Charles A.S. Hall
Kent Klitgaard

Energy and the Wealth of Nations

An Introduction to Biophysical Economics

Second Edition

Springer

Energy and the Wealth of Nations

Charles A.S. Hall
Kent Klitgaard

Energy and the Wealth of Nations

An Introduction to Biophysical Economics

Second Edition

Springer

Charles A.S. Hall
College of Environmental Science & Forestry
State University of New York
Syracuse, New York, USA

Kent Klitgaard
Wells College
Aurora, New York, USA

ISBN 978-3-319-66217-6 ISBN 978-3-319-66219-0 (eBook)
https://doi.org/10.1007/978-3-319-66219-0

Library of Congress Control Number: 2018935116

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To Myrna, my wonderful companion on this and other journeys.

Charles A.S. Hall

To my children, Justin and Juliana Klitgaard Ellis, who have grown from wonderful children into fine adults, and to Deborah York who continues to make me a better person.

Kent A. Klitgaard

Preface

There are four books on our shelf that have the words, more or less, "wealth of nations" in their titles. They are Adam Smith's 1776 pioneering work, *An Inquiry into the Nature and Causes of the Wealth of Nations*, and three of recent vintage, David Landes' *The Wealth and Poverty of Nations*, David Warsh's *Knowledge and the Wealth of Nations*, and Eric Beinhocker's *The Origin of Wealth*. Warsh's book is rather supportive of current approaches to economics while Beinhocker's is critical, but all of these titles attempt to explain, in various ways, the origin of wealth and propose how it might be increased. Curiously, none have the word "energy" or "oil" in their glossary (one trivial exception), and none even have the words "natural resources." Adam Smith might be excused given that, in 1776, there was essentially no science developed about what energy was or how it affected other things. In an age when some 80 million barrels of oil are used daily on a global basis, however, and when any time the price of oil goes up a recession follows, how can someone write a book about economics without mentioning energy? How can economists ignore what might be the most important issue in economics? In a 1982 letter to *Science* magazine, Nobel Prize economist Wassily Leontief asked, "How long will researchers working in adjoining fields ... abstain from expressing serious concern about the splendid isolation within which academic economics now finds itself?" We think Leontief's question points to the heart of the matter. Economics, as a discipline, lives in a contrived world of its own, one connected only tangentially to what occurs in real economic systems. This book is a response to Leontief's question and builds a completely different, and we think much more defensible, approach to economics.

For the past 130 years or so, economics has been treated as a social science in which economies are modeled as a circular flow of income between producers and consumers where the most important questions pertain to consumer choice. In this "perpetual motion" of interactions between firms that produce and households that consume, little or no accounting is given of the necessity for the flow of energy and materials from the environment and back again. In the standard economic model, energy and matter are ignored or, at best, completely subsumed under the term "land," or more recently "capital," without any explicit treatment other than, occasionally, their price. In reality economics is about stuff, and the supplying of services, all of which are very much of the biophysical world, the world best understood from the perspective of natural, not social, sciences. But, within the discipline of economics, economic activity is seemingly exempt from the need for energy and matter to make economies happen, as well as the second law of thermodynamics.

Instead we hear of "substitutes" and "technological innovation," as if there were indefinite substitutes for matter, energy, and the environment. As we enter the second half of the age of oil, and as energy supplies and the social, political, and environmental impacts of energy production and consumption become increasingly the major issues on the world stage, this exemption appears illusory at best. All forms of economic production and exchange involve the transformation of materials, which in turn requires energy. When students are exposed to this simple truth, they ask why are economics and energy still studied and taught separately? Indeed, why is economics construed and taught only as a social science, since in reality economies are as much, and perhaps even principally, about the transformation and movement of all manner of biophysical stuff in a world governed by physical laws?

Part of the answer lies in the recent era of cheap and seemingly limitless fossil energy which has allowed a large proportion of humans to basically ignore the biophysical world. Without significant energy or other resource constraints, economists have believed the rate-determining step in any economic transaction to be the choice of insatiable humans attempting to get maximum psychological satisfaction from the money at their disposal, and markets seemed to have an infinite capacity to serve these needs and wants. Indeed the abundance of cheap energy has allowed essentially any economic theory to "work" and economic growth to be a way of life. For the last century, all we had to do was to pump more and more oil out of the ground. However, as we enter a new era of "the end of cheap oil," in the words of geologists and peak oil theorists Colin Campbell and Jean Laherrere, energy has become a game changer for economics and anyone trying to balance a budget.

In brief, this book:
- Provides a fresh perspective on economics for those wondering "what's next" after the crash of 2008 and the near cessation of economic growth for much of the Western world since then
- Summarizes the most important information needed to understand energy and our potential energy futures

In summary, this is an economics text like no other, and it introduces ideas that are extremely powerful and are likely to transform how you look at economics and your own life.

Charles Hall
Polson, MT, USA

Kent Klitgaard
Aurora, NY, USA
July 2017

Acknowledgments

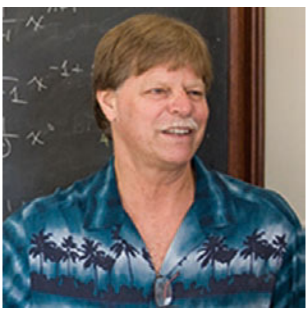
We thank the Santa Barbara Family Foundation, Roger C. Baker Jr., the UK Department for International Development, the Endeavor Foundation, and several anonymous donors for financial support; Steven Bartell, Garvin Boyle, and Jim Gray for excellent editing of words and ideas; Michelle Arnold for assistance in getting the first edition together; Rebecca Chambers, Ana Diaz, Michael Sciotti and Heather Hiltbrand for their able assistance with the data analysis, editing, and graphics for the first edition, and to Tyler Lewis, Sarah Halstead, and Bellina Mushala for assistance with the second edition; and our students over the years for helping us think about these issues. We thank the late, fantastic Howard Odum, who taught us about systems thinking and the importance of energy in everything, and John Hardesty and Norris Clement, who introduced Kent to the limits to economic growth. We thank our colleagues economists Lisi Krall and John Gowdy for providing valuable advice and critique for this and other projects. Their continued collaboration makes our work stronger. We also thank David Packer, Executive Editor at Springer, for believing in us and Myrna Hall and Deborah York for the loving support and infinite patience. We owe a special debt of appreciation to Tina Evans, who provided excellent and valuable suggestions for revising the first edition into what we believe is a much stronger and clearer second edition.

Contents

Authors' Biographies

Charles Hall

is a systems ecologist who received his PhD under Howard T. Odum at the University of North Carolina at Chapel Hill. Charles Hall is the author or editor of 14 books and 300 articles. He is best known for his development of the concept of EROI, or energy return on investment, which is an examination of how organisms, including humans, invest energy into obtaining additional energy to improve biotic or social fitness. He has applied these approaches to fish migrations, carbon balance, tropical land use change, and the extraction of petroleum and other fuels in both natural and human-dominated ecosystems. Presently he is developing a new field, biophysical economics, as a supplement or alternative to conventional neoclassical economics, while applying systems and EROI thinking to a broad series of resource and economic issues.



Kent Klitgaard

is professor of economics and the past Patti McGill Peterson professor of social sciences at Wells College in Aurora, New York, where he has taught since 1991. Kent received his bachelor's degree at San Diego State University and his master's degree and PhD at the University of New Hampshire. At Wells, he teaches a diverse array of courses including the History of Economic Thought, Political Economy, Ecological Economics, The Economics of Energy, Technology and the Labor Process, and Microeconomic Theory and is a cofounder of the Environmental Studies Program. Kent is active in the International Society for Ecological Economics and is a founding member of the International Society for BioPhysical Economics. Recently, his interests have turned toward the degrowth movement, and he has published multiple papers on the subject for *Research and Degrowth*. He has two children and is interested in the outdoors in general: from hiking to beach walking to the occasional round of golf (despite the high energy use of golf courses). Kent is a Californian who still surfs the frigid waters of New England when he gets a chance.

Economies and Economics

Economies exist independently of how we study them. Consequently, there may be significant differences between how an actual economy operates and how we study it. In this section, we will assess how we perceived and taught economics at various times. We begin with the dominance perspective of today, neoclassical economics, in chapter one. In the second chapter, we examine various perspectives from the history of economic analysis from the 18th to the 20th centuries, focusing on the diverse theoretical viewpoints from the past, most of which have little to do with the dominant view of the present. ► Chapter 3 examines the present approach more closely from the perspective of how well that view is related to actual economies. It frequently finds the present view severely wanting, particularly in that it pays little attention to, and indeed is often inconsistent with, the biological and physical world upon which it is necessarily based. ► Chapter 4 introduces a new, different, way in which we can examine economies, one that is in fact based on a proper biophysical underpinning. This approach is called biophysical economics. We emphasize the critical importance of energy here. ► Chapter 5 adds a social perspective that is part of, and consistent with, the essential biophysical framework of this innovative approach.

In general, the entire discipline of economics has paid only a very little attention to energy even though energy was, and remains, the basis of economic activity and growth. Rather economics has treated energy as it treats any other material resource: as a commodity, useful but ultimately substitutable by other commodities. Historically, economists focused their efforts upon capital and labor, and, occasionally, land as the driving economic forces. However, energy issues lay not far beneath the surface of economic reality and many economic concepts. Before the era of classical political economy, English manufacturers had learned to substitute coal for increasingly scarce charcoal to provide heat for their processes. In 1784, James Watt patented a steam engine that could provide rotary motion. The coal-driven industrial revolution was soon to follow. Economies had a new characteristic: growth. Economists now think of growth as a normal characteristic of economies, but this is only a relatively recent phenomenon, and it is highly linked to increasing energy supplies, something that was not characteristic before about 1800.

This book is written by an ecologist and an economist, and part of our objective is to assess where insights and principles from these two disciplines can be combined to understand economies and nature, and their interactions, better. While the two disciplines may appear very different, we believe instead that the phenomena they study are very similar in many ways. From a biological perspective, the economies of cities, regions, and nations can be viewed as ecosystems, with their own structures and functions, their own flows of materials and of energy, and with diversity and stability. Human-dominated systems can exhibit many of the characteristics of natural systems. At the same time, ecology is often referred to as "the economy of nature." There are similarities and differences between organisms in nature and people in modern economies: lions eat gazelles and gazelles eat grasses, trout subsist on insects, and plants exploit nutrients in soils and space in which to intercept sunlight. Individuals and groups find themselves in a relentless struggle to increase their energy gains and decrease energy costs, for their ability to pass on its genes is possible only if it has managed to acquire a large net energy balance. This is also true for humans, but humans are different in that we consciously order the labor process and produce for surplus, rather than for immediate use alone. Producing for surplus dates to the Neolithic transition from hunting and gathering to settled agriculture.

When first encountering the words "biophysical economics," most readers probably asked, "what do those words mean?" The answer is deceptively simple: The word "biophysical" refers to the material world, that which is usually, but not completely, covered by courses in physics, chemistry, geology, biology, hydrology, meteorology, and so on. This can be compared with a "social" or "anthropocentric" (i.e., human-centered) perspective that characterizes modern economics. In this second perspective, which is dominant in our society, humans believe that they can make any world, or set of decisions, or economic systems that they wish – if they can just get the policies right and enough time has passed for new technologies to come on line. The subsequent world becomes our new reality and truth.

But we must ask: How do the powerful, governing physical laws, which we are all prepared to accept in physics, chemistry, and biology classes, operate outside of the scientist's laboratory and the "natural" world? Scientists often think of these laws as imposing constraints on a system. Do these constraints really disappear when human ingenuity is applied to economics and markets? Most economics textbooks would lead you to this conclusion, as growth is just a matter of human actions, technologies, policies, and a healthy dose of ambition. Western culture and its leading commentators (with a few exceptions such as Joseph Tainter and Jared Diamond) do tend to elevate personal and social aspects of a problem, specifically, human actors and their ideas, above any biophysical considerations. Thus, we learn about history as the action of great leaders; wars, if not always battles, are usually won or lost due to the biophysical resources that generals can bring to bear. Napoleon once quipped that "God fights on the side with the best artillery." There is little debate that the South had the better generals in the Civil War, but the North had the industrial might. The North won because of biophysical, not leadership, issues.

Most readers would not argue with the idea that we live in a world that is completely beholden to the basic laws and principles of science. These basic laws include Newton's laws of motion, the laws of thermodynamics, the law of the conservation of matter, the best first principle, the principles of evolution, and the fact that natural ecosystems tend to make soil and clean water while human-modulated systems tend to destroy both. Do economic systems operate outside of these laws? Did the seemingly unconstrained technological and economic expansion of the twentieth century show that these laws were irrelevant or at least insignificant when applied to economics and the satisfaction of human needs and wants?

There is no more important question as we attempt to move beyond the recent financial trauma of the "Great Recession" and the enduring "secular stagnation." Unfortunately, the biophysical laws, particularly as applied to energy, are not understood or appreciated by most people, including most economists. Ironically, our focus on exploiting and investing energy in the economic process has divorced many people from the very biophysical realities that are necessary to sustain them. This includes our ways of building dwellings, living in cites, importing food, being transported and entertained, and so on while isolating our energy using activities in areas generally isolated from people's daily lives. In this book, we examine these issues through an integrated view of economics that emphasizes scientific principles and a more frequent use of the scientific method. Together these chapters provide the beginnings of a powerful new way to think about economics.

Contents

How We Do Economics Today

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_1

1.1 Introduction

We start with a definition of economics: it is derived from the Greek *oikos*, meaning pertaining to the household, so economics is the study of household management Aristotle wrote of this in his *Politics*. He believed wise household management, even if the household comprises the entire state, is part of natural law. But for Aristotle, and those who followed his philosophical lead long into the medieval period, *chrematistics*, or producing and lending money for profit, was unnatural. Economic thought has changed a great deal since ancient and medieval times! Curiously, ecology starts there too, although the ecologist's household can often be much larger. If you think about economics in your day-to-day life, you are probably thinking about providing yourself with the necessities of life (and hopefully a few amenities), that is, the basic stuff you need to survive and, hopefully, be happy. Often you need to think about the trade-offs that exist between the choices you have, a hamburger and no movie vs. Ramen soup and a movie, tuition vs. rent or a vacation, or how to budget whatever financial resources that you have to meet your needs and wants. Many elderly people of limited means must consider the trade-off between food and health care. Hence in very basic perspective, economics is about choice: how much we have and how we should decide among alternatives. Of course, economics pays a lot of attention to money, and a basic starting point for economic thought is that almost everything of concern to humans has a price and can be bought or sold for money. A starting assumption of mainstream economics is that the value of something is represented by its price.

Many people like to talk about the economy. You hear them in barbershops and grocery stores, outside of daycare centers, on the news, at various political functions, and at the park. People wonder whether they should spend their money now or save it for their children's future. Many people are passionate about what the proper role of government should or should not be in the economy. Politicians talk about their economic plans as do journalists and bartenders. These are all legitimate ways to think about the economy, but it is not how academic economics is undertaken. Instead, most mainstream economists build abstract, highly idealized, sets of models. But in this book, we want to do more: we want to capture the essence of how actual economies operate. To do this we need to think deeply about what constitutes actual economies. Many other academic disciplines such as political science, sociology, or even biology do not always have a good command of the basics. We would like to start you off on the right foot when it comes to the study of economics in general and economic theory, reservations. We will do this here by introducing you in minimum space to the main concepts of nearly all basic economics courses, even though later we will address some serious issues we have with that approach.

1.2 Supplying Maximum Human Well-Being Through Markets

Economics courses start with the idea that economics should focus on deriving maximum well-being, as defined subjectively by each individual, and the resources available to each individual. The first question is how should an individual spends his or her money to generate the maximum psychological well-being. The second question is how should the economy as a whole operate to help every individual get as much satisfaction as possible? Although an actual economy is a complex entity with many facets, mainstream economics focuses in large part on what is called "the market." Markets, as places of exchange and trade, have existed since antiquity. However, they were less important in the distant past, because most production of necessities took place in households. It was only in the sixteenth century that markets became a primary way to satisfy daily needs and a place where prices were formed. Adam Smith elevated the study of markets in eighteenth century England to a position of prominence in an era characterized by agriculture and small-scale manufacturing. Here farmers would lay out the leftover vegetables and eggs that they did not use themselves and trade them for money to buy other things such as the products of various smiths or artisans. In these environments purchasers could take their usually hard-earned money and carefully choose what was most needed or desired for their lives without too much in the way of manipulation or compelling authority. Contemporary mainstream economics believes that in an almost magical way "the market" will generate the maximum possible human well-being by generating the largest possible number of most desirable goods and services for each member simply attempting to achieve his or her own self-interest. In the words of Adam Smith: "It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest" [1]. Thus, the basic concept of how economies are thought to work in a "free market" situation is that consumers will purchase goods and services to suit their own conception of the psychological satisfaction each purchase will make and that suppliers will shift to make what people want, for that is where they make their own largest profit. As consumers purchase additional commodities, they will get less satisfaction from the extra one and shift to another commodity.

Market is often imbued with nearly mystical power. Former President Ronald Reagan often spoke about the "magic of the market" [2]. Other ways of thinking about the economy are rarely, if ever considered. Most mainstream economists believe that the basic propositions of economics are true for all places and all times. The economic relations between people and nature that exist today existed tens of thousands of years ago for our hunting and gathering ancestors. The basic assumption of mainstream economists is that there wasn't much difference between the medieval economy and that of the present day. Moreover, the future will be like the present.

1.3 Microeconomics and the Process of Self-regulation

Part of the reason why contemporary economists like this basic worldview is that it expresses the idea that the economy is *self-contained* and *self-regulating*. By self-contained it is meant that the economy is the primary system to be analyzed. It is not a subsystem of something larger such as nature or society. In the mainstream worldview, all human interactions are economic transactions. Nature is external to the system and hardly worth recognition at all. Besides, if necessary, nature can be easily brought within the economy by "internalizing the externalities." Such internalization processes are the subject matter of the emerging field of environmental economics.

In the view of economists, the second concept, that of self-regulation, is very important because it means that an economic system, left to its own devices, will produce outcomes that are efficient and equitable, a very desirable state of affairs. *Efficient* means that resources will flow to their best uses and that no one can be made better off without making another worse off. *Equitable* means that market outcomes are fair. Individuals are rewarded according to their productivity and contribution to society. In other words, the market knows best. If the market forces of competition and flexible prices can unfold without some type of external interference (such as by governments), the result will be that people's needs are met and the economic resources available will be put to the best use, in terms of satisfying human wants and needs. This view of economics is perhaps best exemplified by the words of Voltaire's Dr. Pangloss (based on philosopher and mathematician Gottfried Leibnitz) that "this is the best of all possible worlds." While humans may not be able to fulfill all their wants and desires due to the limits of their purchasing power, at least by making their own free choices they will generate the maximum human psychological satisfaction possible. This has the added virtue, according to market advocates, in democratizing decision-making: society will produce those goods and services that its participants think are best and most desirable rather than what might be advocated by someone who "knows best" for all involved (i.e., centralized planning).

1.3.1 The Primacy of Exchange

In the 1830s, economist Frederick Bastiat declared that "exchange is political economy." By this he meant that the primary subject matter of economics should be the ordinary exchange of money for goods and services. As we will show in the next chapter, in previous times economists, then calling themselves political economists, focused on many processes with a biophysical basis, such as production, distribution, and capital accumulation. Economists still treat these materialistic subject matters, but mostly only within the context of exchange. The basic belief is that one can analyze sufficiently the complex economy just by looking at the processes of buying and selling. This approach comes replete with a definition of economics based on relative scarcity, a conceptual model of the circular flow of exchange value, and the ubiquitous supply and demand diagrams.

1.4 Two Definitions of Scarcity

In the first two centuries of economics, a definition for economics did not exist. During the Great Depression Oxford economist Lionel Robbins, writing on economic methodology, came up with the most widely used definition of economics. He contended that economics was the "study of the allocation of scarce resources among alternative uses." His definition itself needs a little explaining. Allocation means "who gets what." The market allocates this book to a student of biophysical economics, while a pneumatic nail gun may be allocated to a finish carpenter and a tractor to a farmer. Allocation refers to stuff, most often called goods and services. Who gets what money is usually called distribution. The concept of scarcity underlies mainstream economics, although it does not mean exactly what we might first think. It does not relate to the limited availability of things like fish or petroleum or clean water. Mainstream economists rarely deal with such absolute scarcity. If one thinks about it, today is probably the most resource abundant time humanity has ever seen, but still there is scarcity, not to mention enormous poverty. The idea of *relative scarcity* depends upon the assumption that humans have unlimited wants, and any resource would be scarce relative to unlimited wants. Mainstream economists believe relative scarcity has existed in all times and all places. Biophysical economist John Gowdy disagrees. In a marvelous collection of essays on hunter-gatherer economies, entitled *Limited Wants, Unlimited Means*, Gowdy argues that hunter-gatherer bands were quite different than we are today. They were radically egalitarian, had no concept of private property, had few material desires, and faced a cornucopia of nature relative to their very modest wants [3]. This was largely because they were seminomadic, and they had to carry all their possessions on their backs from one water hole to another.

1.5 How the Structure of the Economy Is Perceived

How do economists conceptualize the economy? The most basic model, one found in the first chapters of essentially all economic textbooks, posits two sectors, two markets, and four flows. The model begins from the perspective of the individual. Individuals have but two identities in this view of society. People are either consumers, and are to be found in *households*, or they are producers and can be found in *firms*. Any other identities people may feel affinity toward such as race, ethnicity, nationality, or gender are not considered. And even though real people tend to live in households and work in firms, in this model they are only one or the other. Moreover, they never interact directly with one another. All human activity occurs indirectly through market transactions. People either buy or sell. There are also two markets. The first is the product market where money is exchanged for goods and services. The second is the factor market where the "factors of production," that is, land, labor, and capital, are exchanged for a specific type of money known as *factor payments*. Land receives a rent. Labor gets wages, and capital is remunerated by profit and interest, depending upon whether one is an entrepreneur or a financier. Material goods and immaterial services flow one way, and money flows the other. What is important is the flows of exchange value—things humans perceive as valuable that can be exchanged for money (Fig. 1.1). Value is equated with price and human relations that are not captured in buying and selling, such as the relation between Adam Smith and his mother, are simply not considered. Neither is the human interaction with nature. While all goods

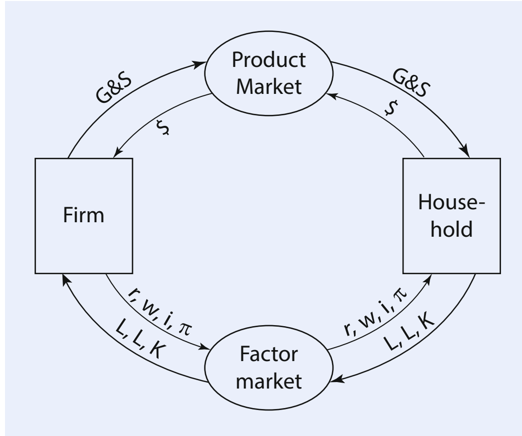


Fig. 1.1 Circular flow model without leaks or injections. In this model the factors of production are: Land (*L*), Labor (*L*), and Capital (*K*). Factor payments are indicated by: Rent (*r*), Wages (*W*), Interest (*i*), and Profit (*π*). The top flow from firm to households represents Goods and Services (*G&S*), while the top flow from household to firm takes the form of Money (*$*)

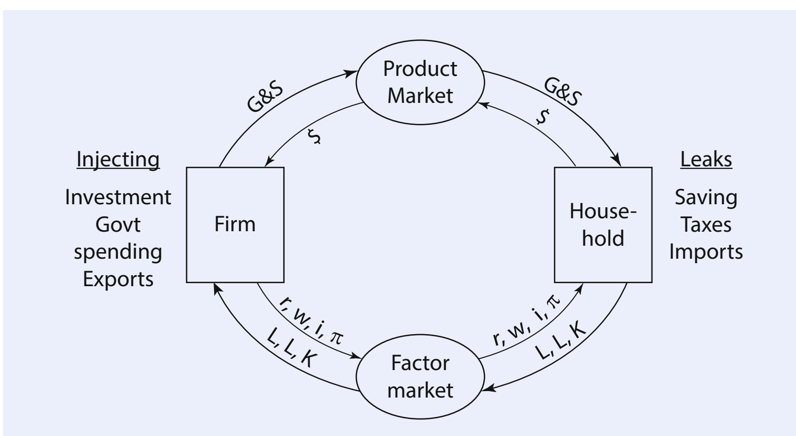
may be relatively scarce, nature imposes no absolute barriers that cannot be transcended by resource substitution, technological change, or entrepreneurial innovation.

The circular flow depicts an economic premise often called "Say's Law of Markets," or simply, "Say's Law." [4]. The economy is self-regulating because the money on the pathway that goes from household to firm exactly equals the money on the pathway from firm to household. This is the simplest, but not the most convincing, explanation of self-regulation. First, it contains no mechanism to translate production, consumption, and spending into forms of income. Second, it requires that everyone spends all their income on current consumption or production. Household members do not save, and businesses do not invest. Nobody buys imports, and firms do not export. No individuals pay taxes, and the government spends no money. But despite these and other problems, Say's Law became an essential cornerstone of economics. After the severity of the Great Depression, economists, especially those following the trail blazed by British economist John Maynard Keynes, began to question the idea of Say's Law. Consequently, when looking at the economy, the circular flow model was augmented to include money that "leaked" out of the spending stream such as saving, taxes, and spending on imports as well as money that was "injected" in the form of investment, government spending, and exports. If more money leaks out of the system than is reinjected, there will not be enough money to buy all the products firms desire to sell. This would cause firms to cut back on production and hire fewer workers. The result is a recession caused by lack of demand. If more money was injected in than leaked out, there would be too much money chasing too few goods, and inflation, or a generalized increase in prices, may result. In addition, economic growth depends upon injections of new investment.

1.6 Supply, Demand, and Their Theoretical Interaction in Markets

A focal point of modern economics is the concept of supply and demand and their interaction to generate price. Supply measures how many units of a good or service sellers want to take to the market, and demand measures how much stuff consumers want to purchase. Both are a little tricky, as, for example, more units would be demanded if the price were lower and fewer if the price were higher. Likewise, if the price was higher, more suppliers would be likely to provide more units to the consumers. A second aspect is that a large number of things that affect the willingness and ability to buy or sell have to be held constant when in the real world they are changing all the time. But if we did not make these assumptions, the model would be very cluttered and difficult to solve even with advanced statistical techniques.

The most important first point to remember is that in the mainstream view the interaction of supply and demand determines simultaneously equilibrium levels of price and quantity. Prices, especially competitively derived prices, are the great regulating mechanism of mainstream economic theory. Equilibrium, also reflected in ◘ Fig. 1.2, is also a useful concept. The idea, appropriated from physics, means a state of rest



◘ **Fig. 1.2** Circular flow with leaks and injections In this model the factors of production are: Land (*L*), Labor (*L*), and Capital (*K*). Factor payments are indicated by: Rent (*r*), Wages (*W*), Interest (*i*), and Profit (*π*). The top flow from firm to households represents Goods and Services (*G&S*), while the top flow from household to firm takes the form of Money (*$*)

where there is no internal tendency to change. If the system is perturbed from the outside, it will return, after adjustment, to the state of rest. It is derived from Newton's third law of motion, that all forces sum to zero, or for every action there is an equal but opposite reaction. In this idealized world of economics, prices, if disrupted (e.g., by an embargo), will return to the original equilibrium situation by means of price competition.

Let us begin with the demand curve and a definition of demand.

Demand measures the willingness and ability of consumers to buy various quantities of goods and services at different prices, with all things that affect this willingness and ability, other than price, held constant.

For those new to mathematical modeling, one benefit of constructing models is the ability to separate cause and effect. This is more difficult if there are multiple causes, so one trick is to pretend that things you know are really changing all the time are constant for the purposes of the model. To lend even more credibility, we give this simplifying assumption a Latin name, *ceteris paribus*, which means all other things held constant. The use of Latin is meant to or is supposed to impress you. The definition of demand is a bit of a mouthful, so let us provide a mathematical shorthand:

$$Q_d = f(p) \text{ ceteris paribus.}$$

This says pretty much the same thing as the long-winded definition in words: the amount you are willing or able to buy depends upon the price. If prices go up, you are willing or able to buy less. If prices go down, you buy more, as long as everything that affects your willingness to buy, other than price, is held constant. Graphically, as in Fig. 1.3 changes in price translate to a movement up and down a stable demand curve.

Note, and it is an important note for those who like technical precision, that a decrease in prices *does not* increase demand, and an increase in prices *does not* decrease demand in this model. It is a technical point that a lot of people get wrong: politicians, newscasters, ecologists, for example. A change in price can change only quantity demanded. Instead, the only change in one or more of our assumed constants (*ceteris paribus* assumptions) can change demand. You should commit the following list to memory if you plan on studying economics formally.

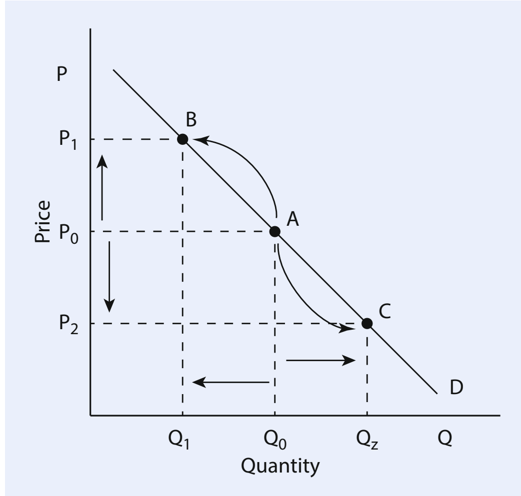


Fig. 1.3 Changes in price lead to changes in QUANTITY DEMANDED. This is graphically depicted as a movement along a stable demand curve. Price (*P*) and Quantity (*Q*)

1.6.1 Assumed Constants for Demand

- Income and wealth
- Tastes and preferences
- Price of related goods
- Consumer expectations
- Number of consumers

If your income increases you are likely to buy more goods and services. If your tastes change, say because of advertising, you might buy more of this and less of that. If the price of a substitute goes up, you would shift to the good in question and buy more. If you expect a sale in the future, you might hold off and buy less now in anticipation. More people with money in their pockets, all other things held constant, will purchase more. A change in demand is depicted as a shift in the demand curve (Fig. 1.4). If it shifts to the right, demand has increased, and a shift to the left signifies a decrease in demand.

Supply looks at the market from the seller's point of view. The definition is remarkably similar to that of demand. If you change consumer to firm and buy to sell, the definition is exactly the same.

Supply measures the willingness and ability of firms to sell various quantities of goods and services at different prices, with all the things that affect the willingness and ability to sell, other than price, held constant. Mathematically:

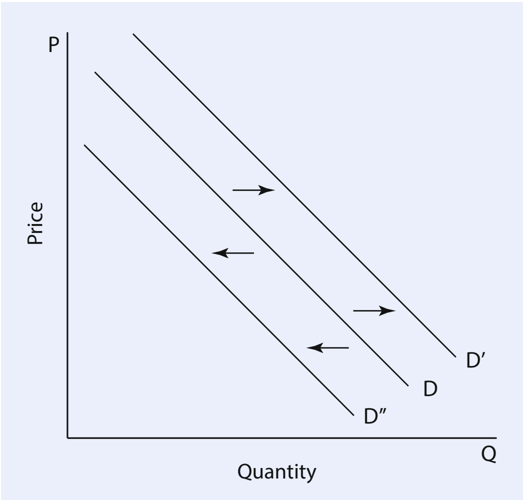


Fig. 1.4 Changes in assumed constants lead to changes in DEMAND. The whole curve shifts to the right for an increase in demand and to the left for a decrease. Price (*P*) and Quantity (*Q*)

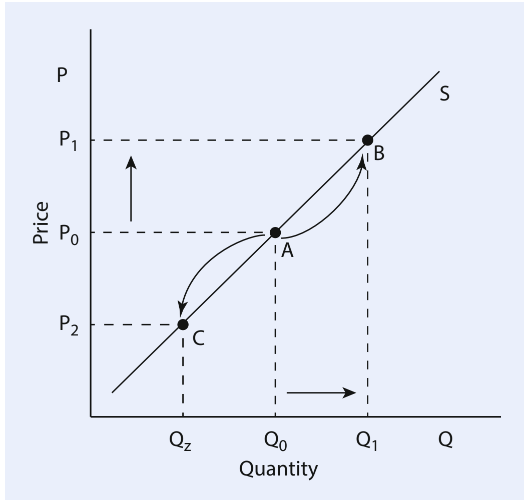


Fig. 1.5 Changes in price lead to changes in quantity supplied. This is graphically depicted as a movement along a stable supply curve. Price (*P*) and Quantity (*Q*)

$Q_s = g(p)$ *ceteris paribus*, where Q_s is quantity supplied, *g* is the functional operator, and *p* stands for price. In this case the supply curve has a positive slope. This means firms will be willing or able to sell more at higher prices, all other things remaining constant. Not surprisingly, the list of assumed constants is different, because it affects the firm's cost of production instead of consumer preferences.

1.6.2 Assumed Constants for Supply

- Technology
- Input or resource prices
- Seller expectations
- Number of sellers

If price changes, the *quantity supplied* changes since at higher prices more suppliers will be interested in selling their stuff (Figs. 1.5 and 1.6). This is depicted graphically as a movement along a stable supply curve. Higher prices bring forth an increased quantity supplied, while lower prices mean that sellers will be less willing or able to bring forth their goods or services, and quantity supplied declines. If one of our assumed constants changes, *supply changes*. Changes in supply are shown as a shift of the curve. If supply increases from technological improvement or lower input prices (e.g., wages, energy, rent), the whole curve

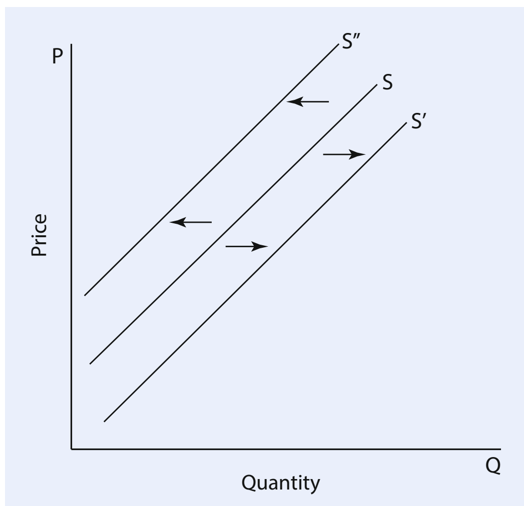


Fig. 1.6 Changes in assumed constants lead to a change in supply. The entire supply curve shifts to the right for an increase and to the left for a decrease. Price (*P*) and Quantity (*Q*)

will shift to the right, so that consumers are willing and able to purchase a higher quantity at the same price. If supply declines because of things like higher energy prices or higher wages, the entire supply curve will shift to the left.

What is the difference between demand and quantity demanded and supply and quantity supplied, and why do we stress it so much? Studying economics is largely about figuring out cause and effect relations on these graphs. Let us summarize these causations:

Causation #1 : change in price leads to changes in quantity demanded (or $\Delta p \rightarrow \Delta Q_d$)

Causation #2 : change in a ceteris paribus condition of demand leads to a change in demand

$$\left(\Delta \text{assumed constants} \longrightarrow \Delta D\right)$$

Causation#3 : change in price leads to a change in quantity supplied ($\Delta p \rightarrow \Delta Q_s$)

Causation #4 : change in a ceteris paribus condition of supply leads to a change in supply

$$\left(\Delta \text{assumed constants} \longrightarrow \Delta S\right)$$

In ending let us say the difference between supply and quantity supplied, as well as demand and quantity demanded, can be seen geometrically. Supply is *all the possible combinations* of price and quantity, given our assumed constants. Quantity supplied is a *single point* on the supply curve. The same goes for demand.

1.6.3 Self-regulations and Changes in the Supply and Demand Curves

The argument for market self-regulation, which was analytically broad and rather unconvincing at the level of the circular flow model, becomes more cogent once the driving force of price competition is added to the process. Remember that a condition of stable equilibrium is that if the state of rest is perturbed, the equilibrium (original conditions where supply and demand are balanced) will be restored by forces within the system. Let us first consider the characteristics of market equilibrium, then two changes that will disturb the state of rest. The first will be a change in prices. We will trace the economic problems involved and show how price competition will restore the original equilibrium state. Next, we will consider one or more changes in our assumed constants (ceteris paribus assumptions) and show how a new equilibrium condition will emerge.

▶ https://www.youtube.com/embed/fcx1sd-9pyKg

We begin our analysis of supply and demand by assuming that the market is in equilibrium.

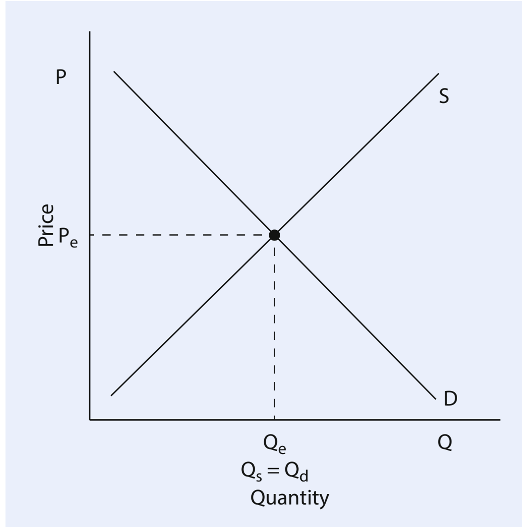


Fig. 1.7 Market equilibrium. At the equilibrium price quantity supplied = quantity demanded. Price (*P*) and Quantity (*Q*)

Here the supply curve intersects the demand curve. The higgling and haggling of the market process has found the one price where the quantity supplied just meets the quantity demanded. At this price, sellers are willing and able to bring forth to the market just the right amount of goods or services that buyers are willing and able to purchase. This does NOT mean that supply equals demand. Since both supply and demand curves represent all the possible combinations of price and quantity, the only way supply would equal demand would be if the curves were superimposed upon one another. This would be impossible because one curve has a positive slope and the other a negative (Fig. 1.7).

Next, assume that prices increase. The same action touches off two effects. The increase in price causes quantity demanded to decline (people will buy less because it costs more) while at the same time it causes quantity supplied to increase (suppliers see more opportunity for profit). At prices above equilibrium ("e" on Fig. 1.8), the quantity supplied is greater than quantity demanded. Economists call this unstable situation a *surplus* and believe that market forces alone will be sufficient to restore the prior equilibrium. A surplus represents unsold goods for sellers. To try to get rid of the surplus, sellers will compete by reducing their prices. If one seller lowers his or her prices, then competitors will be forced to reduce theirs too. The reduction in price increases quantity demanded and lowers quantity supplied, thereby reducing the surplus. If

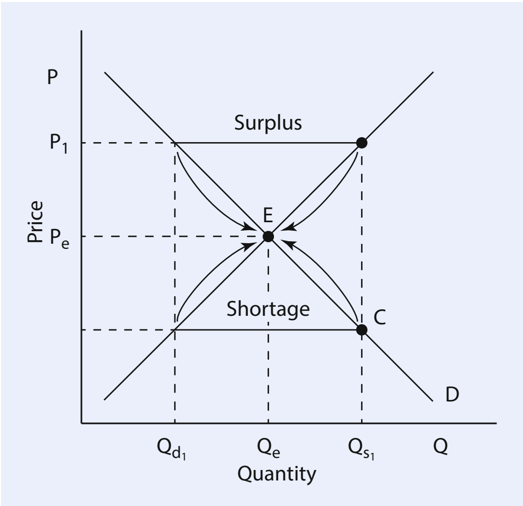


Fig. 1.8 If prices exceed, equilibrium quantity supplied > quantity demanded and a surplus arises. Price competition among sellers to get rid of unsold goods drives the price down toward equilibrium. If price is less than equilibrium, quantity demanded >quantity supplied. A shortage results, and consumers bid up prices to obtain the goods in short supply. Equilibrium is restored by price competition. Price (*P*) and Quantity (*Q*)

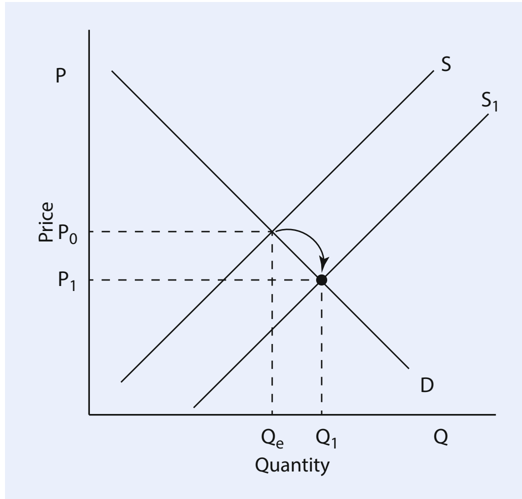


Fig. 1.9 A change in supply drives down equilibrium price. Consumers are willing or able to buy more at the lower price, establishing a new equilibrium. Price (*P*) and Quantity (*Q*)

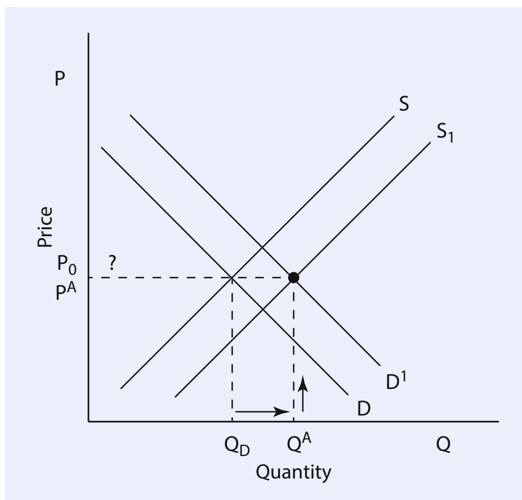


Fig. 1.10 When demand and supply change at the same time, either price or quantity is indeterminate. Price (*P*) and Quantity (*Q*)

the opposite occurs, and prices fall below equilibrium, quantity demanded will exceed quantity supplied. A shortage will ensue causing consumers to compete with one another by offering to pay more for the goods in short supply. The higgling and haggling continue until the market, as if by magic, finds the one price where quantity demanded equals quantity supplied, where the amount firms are willing and able to sell just equal to the amount that consumers are willing and able to buy. At this point, there is no further incentive to change or to raise or lower price. Equilibrium is thereby restored. Nothing other than price competition was needed to restore the balance. At least this is the theory.

Next, assume technological improvement, such as a new and more efficient use of energy. Starting from an equilibrium position, the supply curve shifts to the right because of a change in a ceteris paribus condition of supply. Relative to the starting point, the increase in supply results in lower prices and larger quantities. The resulting price drop leads consumers to increase their *quantity demanded* and purchase more at the new, lower price. This condition, where improvements in efficiency lead to more, not less, resource consumption is known as *Jevons' paradox*. We will consider it in more detail in the next chapter, but it helps to understand the market mechanism behind the process. If consumer incomes were to change along with technology, demand would also rise, pushing both prices and quantities up. The new equilibrium would exhibit some uncertainty because while we could easily say that equilibrium quantities would rise, it would be harder to be certain about price. Because increase in supply would pull prices down, and increases in demand would push them up. We would need more information to tell for certain (Figs. 1.9 and 1.10).

We will not bog you down in any further details. If you find this information interesting, or at least useful, you might want to consult a textbook in *Principles of Microeconomics*. Our favorite is *Microeconomics in Context* by Neva Goodwin and her colleagues at the Global Development and Environment Institute of Tufts University. It is very well written and represents one of the few attempts to introduce the limits of nature and society at the introductory level [5].

1.6.4 Behind the Demand and Supply Curves

Why does a demand curve have a downward slope? Is it just because it seems normal to buy more stuff when it is cheaper or is there something more sophisticated going on? The answer can be found in the theory of *marginal utility*. As we will see in the next chapter, the idea developed over the hundred or so years from the 1790s to the 1890s and depends upon a specific philosophy of human behavior called *utilitarianism*. In the 1790s, British philosopher Jeremy Bentham asserted that humans had but two emotions: we sought pleasure and avoided pain. In the absence of pleasure or pain, we would do nothing. Bentham called the sense of pleasure, or satisfaction, or well-being, utility. He believed that the object of a good society was to provide the greatest good (or level of utility) for the greatest number. Bentham also thought that each individual is the best judge of his or her own pleasure. We can't say ours is better than yours. By strict Benthamist principles, Hall could not claim the joy he feels from his beloved Puccini operas were better than Klitgaard's experience at a Quicksilver Messenger Service concert in San Francisco. Neither of us could claim that our musical preferences are superior to the latest hip-hop compositions that we do not understand completely.

1.7 Margins and Marginal Utility

At this point, we find it necessary to introduce the concept of the margin, which will appear repeatedly in myriad economic contexts. *Margins*, as defined by words, always represent the contribution of the extra, additional, one more, or incremental. *Marginal utility* is the addition to satisfaction that results from consuming one more unit of a product. In terms of mathematics, margins are always the change in the effect divided by the change in the cause, or Δ dependent variable/Δ independent variable. Marginal utility is the change in the result (satisfaction) divided by the change in the cause (consumption). But how much does satisfaction change? Say you have been working clearing brush, and someone gives you something cold to drink. The first few sips would give you a great deal of extra satisfaction. But if you were to drink a gallon, the last few sips would not be as satisfying as the first. This is called *diminishing marginal utility.* The extra satisfaction from one more unit is less when you have a lot than when you have very little. Because marginal utility decreases as we consume more, we would be less willing to pay for the extra consumption of goods. Therefore, the demand curve slopes downward. You might realize there are some problems. If one cannot compare interpersonal utility, how can one aggregate it? This, and many other mysteries, awaits you should you decide to take a course in intermediate microeconomics.

In the 1890s, mainstream, or neoclassical, economists put the theory of supply on a marginal utility basis, changing only the names of the variables but not the analysis. If one input was fixed, say land or capital, as you added more units of a variable input like labor to a fixed amount of land or capital, the amount of extra work would eventually decline. This is known as *diminishing marginal returns*. Each incremental unit of labor produces a little bit less than the last one. As a result, the cost of producing the next unit of output increased on the margin. This rising marginal cost was the basis of a positively sloped supply curve.

1.7.1 Market Structures

The marvel of self-regulation by means of prices alone requires something that may or may not exist in the real world: price competition. Early economic models created an abstract world in which there were so many companies that none could influence the market price, and none had any technological advantage. This structure was known as *perfect competition*, and it rested upon a set of assumptions that must be satisfied simultaneously:
1. A large number of small firms.
2. Each firm is so small as to not affect the market price.

3. Every firm made exactly the same thing (homogeneous product).
4. All firms had perfect knowledge and perfect foresight of market conditions.
5. There are no barriers to entry or exit.

These assumptions result in an economy of tiny, powerless firms that can do nothing except responding to the price dictates of an impersonal market.

In fact, it would be extremely difficult to compete under these conditions. Indeed, all profits other than the maintenance of the entrepreneur (known as normal profit) would be competed away, and all the benefits would accrue to the consumer in the form of the lowest possible prices. All outcomes should be "efficient." Resources would flow to their best use, and individuals would earn their contribution to the total (not surprisingly, known as marginal product), no more and no less.

But businesses want to make profits, retain them and invest them in improved technologies. As far back as the 1500s, British coal companies began to monopolize their markets with the purpose of avoiding price competition, largely by deciding not to overproduce a surplus which would drive down prices. In the United States, merger activity flourished in the years after the Civil War. (We will chronicle the development of monopoly power in ▶ Chap. 9.) Also, in the chaos created by the Great Depression of the 1930s, more and more economists began to question this idealized world of perfect competition. Economists in the two Cambridge's (England and Massachusetts) developed theories of imperfect competition, where firms cooperate for their mutual benefit, rather than compete with one another. We will develop this strand of theory more deeply in the next chapter. Suffice it to say that these models of imperfect competition, while far more realistic, cannot produce outcomes characterized by efficiency and equity. Rather they lead to overproduction, excess capacity, and exploitation. Most conservative economists barely give these criticisms passing notice, although these are crucially important factors in the actual economy.

1.8 Macroeconomics

In the early years of the twentieth century, the ideas of neoclassical economics were extended to interpret the overall, or *aggregate*, economy.

The conservative approach of the 1920s, dubbed "the Classical Model" by John Maynard Keynes, held that the overall economy behaved on the same principles of supply and demand as did an individual firm or industry. It worked as follows, at least in theory: Starting in the labor market, the demand for labor depended upon workers' marginal products, and an individual worker was free to choose the number of hours he or she wished to work by equating the satisfaction of receiving a paycheck (the marginal utility of the wage) with the drudgery of the job (the marginal disutility of the work). The resulting equilibrium assured the economy would operate at full employment. Any unemployment was the result of a surplus of workers, meaning that the price of labor was "too high." Wage cuts could easily restore the balance. Say's Law assured that income translated into spending while the commitment to balanced budgets at home and abroad meant that neither budget nor trade deficits would exist for more than the short while it took markets to adjust. Interest rates, or the price of money, would be set in the market for "loanable funds." Here, the demand for loanable funds was an inverse function of interest rates, which produced a downward sloping curve. The supply of loanable funds would respond positively to increases in interest rates. The resulting market equilibrium created an interest rate that automatically balanced savings and investment. The economy would run like a smooth machine provided that no outside entities like governments or labor unions would disturb the delicate balance.

This explanation held until the collapse of the Great Depression. In the United States, "official" unemployment rates increased to nearly 25% by 1933, and new investment was actually negative. In other words, more equipment was wearing out than was being replaced. The banking system collapsed three times between 1929 and 1933, while international trade dried up under high tariffs and the banner of "America First." Events were worse in Europe and far worse in the poor countries of Africa, Asia, and Latin America. Moreover, the depression lasted for nearly a decade and came to an end only with the spending for Second World War. Out of this chaos, the theories of John Maynard Keynes gained acceptance. Keynes accepted most of the neoclassical economics but rejected Say's Law and the idea that workers could choose their own hours of work based on their utilities and disutilities. Because of these, he concluded that a mature, industrialized, capitalist economy could achieve equilibrium at any level of output, including one at high unemployment levels. Remember that equilibrium means no internal tendency to change, so an economy in this situation would have unemployment rates that persisted. Unchanging unemployment was a factor in the rise of fascism in Europe and in the Bolshevik Revolution. Keynes was determined to save capitalism from itself in the form of a tendency toward politically untenable levels of unemployment and perpetual economic stagnation. Keynes believed that the fundamental cause of a depression was *a lack of demand,* and more specifically a lack of aggregate demand. Aggregate demand is the demand for all goods and service by all economic sectors and consists of consumption, investment, government spending, and foreign trade. If people do not purchase all the goods and services that are produced, surplus inventories will build up. Falling prices will put even more pressure on struggling businesses. They will produce less and perhaps cut wages. But poorer workers spend less money, and the downward spiral begins. This seemed to explain the recurrent recessions and depressions that were characteristic of capitalist societies.

Keynes suggested public works programs, and if those failed he recommended burying money in bottles and paying people to dig them up. He thought anything that put money in people's hand would be part of the solution. He also wanted to abandon the gold standard that kept prices, wages, and profits falling. Keynes also did not believe in the lockstep of a balanced budget. He reasoned that if the government ran a deficit, the economy would expand all the much faster. The cause of the depression was a lack of demand, and a cure for the depression was to boost aggregate demand. In the United States, former Governor of New York Franklin D. Roosevelt was elected president. He implemented a program called the New Deal to relieve the suffering of America's most vulnerable citizens and to begin the recovery. Roosevelt initiated many spending programs but also raised taxes because he also believed in a balanced budget. The recovery was tepid to say the least. Unemployment never dropped below 13% during the entire decade of the 1930s.

1.9 Postwar Macroeconomics

The proof that Keynesian economics "worked" came with the spending for Second World War, and, as Roosevelt put it, "Dr. New Deal was replaced by Dr. Win the War." Nobody complained about big government or deficit spending during the war. As a result, unemployment dropped to nearly 1% by 1944. (We will chronicle the specifics of the postwar experience later in ▶ Chap. 10.) For the time being, it is safe to say that Keynesian economics entered the canon of economic theory, especially for those of us who began studying economics in the late 1960s. But it was a new, more sanitized version of Keynes that developed in the United States. Gone were radical proposals for income redistribution and calls for the "voluntary euthanasia of the rentier class." In its place came a commitment to economic measurement, the "grand neoclassical synthesis," and an obsessive focus on economic growth. Before the Great Depression, the United States had no consistent method of accounting for economic activity. To improve this situation, Congress commissioned Harvard Economist Simon Kuznets to improve the economic statistics available to the nation's policy makers. Even though partial, the statistics proved very useful in the war effort. Fellow Harvard economist Paul Sweezy won the Bronze Star for his statistical work to enable the Normandy Invasion on D-Day, 1944. After the war, the *Survey of Current Business* began to publish the "national income and product accounts." The focus was upon *gross national product* (GNP), or the dollar value of all final newly produced goods and services in the country, as the primary measure of economic success final goods are those purchased by the ultimate consumer, and not sold to someone else. All the components, for example, consumption, investment, government spending, and net exports, added up to equal the GNP. Both economic growth and economic development were defined as increases in gross national product. In 1948, Massachusetts Institute of Technology economist Paul Samuelson created the grand neoclassical synthesis in his textbook, *Economics.* Here he argued that the private sector was best at allocating resources and distributing incomes. The government's participation was needed simply to produce regular and consistent levels of economic growth. This could be done by changing directly the levels of government spending and taxation, known as *fiscal policy*, or by changing the levels of the supply of money and the interest rates, referred to as *monetary policy*. If these policy measures were implemented subtly, economists could "fine-tune" the economy and relegate recessions and depressions to vestiges of the past. If unemployment rose too much to be politically acceptable because of too little aggregate demand, the government could just spend more or tax less. The nation's central bank, the Federal Reserve, could add to the money supply and charge lower interest rates. If prices in general started to rise because of too much aggregate demand, the government could spend less or tax more or make money harder to get and more expensive. The reductions in money and spending would simply bring down prices while full employment was maintained. Life seemed easy, especially in theory. In the real world, attempts to control inflation actually increased unemployment, but British economist A.W. Phillips showed, along with his famous "Phillips curve," how the trade-off could be managed acceptably. The trade-off was a small price to pay for economic growth. This was all predicated on the fact that inflation and unemployment were mutually exclusive events, at least for the time being.

1.10 The Focus on Growth

Keynes himself was not particularly focused on economic growth but upon aggregate demand, economic recovery, and full employment. However, his colleague and biographer Roy Harrod did produce *An Essay in Dynamic Theory* [6] in the last year of the depression. Harrod argued that because of psychological forces, the trajectory of economic growth would be highly unstable. Any deviation from the warranted growth path would touch off unstable oscillations that he compared to a knife edge. Eight years later American economist Evsey Domar published a foundational article that also showed the path of economic growth to be highly unstable [7]. He attributed this instability to "the dual nature of investment." Investment is part of aggregate demand, and its increase leads to growth in GNP. However, investment also produces long-lived fixed capital. If too much capital exists, the overproduction leads to excess capacity which reduces growth. Fine tuning was not as easy as it seemed, although this would not be seen until the 1970s. The best efforts to fine tune the economy in the 1970s were no match for the peak of domestic oil production, and the collapse of the International Monetary Accords. Meanwhile, to the rescue rode Samuelson's MIT colleague, Robert Solow. In his 1956 *Contribution to the theory of economic growth*, Solow made some technical changes to the production function [8]. He accused Harrod of assuming that inputs were used in fixed proportions. Solow constructed a series of equations based on substitutable inputs (also known as the Cobb-Douglas production function) and, viola, the instability disappeared. Solow's analysis did have the problem of a large unexplained residual, and we will address this later in ▶ Chap. 3. In the 1950s, growth theory consisted of the work of Harrod and Domar. By the late 1950s, Solow's approach was given equal footing. By the 1980s, Harrod and Domar's work was relegated to a footnote, and by the twenty-first century all reference to their work had disappeared from the neoclassical literature. What remained was a theory of economic growth in a frictionless, perfectly competitive idealized economy. The model predicted steady growth. Unfortunately, the actual economy produced stagnation, financial collapse, and severe recession.

As we explain in ▶ Chap. 7, the 1970s were a challenging time for Keynesian economics. The international monetary accords, negotiated in 1944 in a grand hotel in Bretton Woods, New Hampshire, no longer functioned. They were predicated on economic power at the end of the war. But post war conditions changed Germany and Japan caught up, and the costs of a failed adventure in Vietnam meant the United States could no longer make good on its promises. Second, high rates of unemployment and inflation occurred at the same time. Attempts to reduce unemployment just raised inflation, while unemployment remained persistently high. Policies designed to reduce inflation were ineffective but raised unemployment. Keynesian economics could no longer "deliver the goods." On top of all that, disruptions in the world supply of oil led to two energy crises in the 1970s. The cheap fuel that postwar Americans had come to see as their birthright was no longer cheap. Moreover, recessions followed every spike in oil prices.

A more conservative approach began to emerge. *Monetarist* economists argued that inflation "always was and always will be a monetary phenomenon." Too much aggregate demand was not the problem, too much money was. Fiscal policy was seen as ineffective, and monetary policy (money supply and interest rates) began to rule the policy roost. Wall Street banker Jude Waninsky devised the idea of "supply-side" economics and convinced the newly elected president, Ronald Reagan, to change policy. According to supply-side economics, inflation and unemployment could be solved by increasing aggregate supply. To do this the cost of regulation and wages needed to fall. The policy also got a boost from the decline in world energy prices. Since then, policies have become more conservative. As we show in ▶ Chap. 7, supposedly liberal Bill Clinton and Al Gore reinvented government by reducing its funding and "ended welfare as we know it." After the 2001 attacks on the World Trade Centers and the Pentagon, President George W. Bush told the Americans to "go out and shop" while increasing military spending and fomenting perpetual war.

In 2008, young Americans came within a hair's breadth of experiencing the same type of depression that their grandparents and great grandparents did. The response of the Obama administration was to implement the equivalent of Herbert Hoover's plan for the economy at the beginning of the Great Depression. The Troubled Assets Relief Program (TARP), patterned after Hoover's Reconstruction Finance Corporation (RFC), poured billions of dollars into the rescue of banks, while leaving millions of everyday working Americans dispossessed from their homes. Government spending for infrastructure projects was part of an overall stimulus program, and military spending continued to grow with active wars in Afghanistan and Iraq. President Obama actively championed a return to Keynesian economics. Efforts at decarbonization were progressive in rhetoric yet small in outcome, as the administration did not see fit to challenge its commitment to economic growth for environmental purposes. Its sustainability program depended largely upon technological change in electrical generation (wind and solar subsidization) coupled with an expansion of hydraulic fracturing of shale gas and tight oil. What lies in store for the United States after the election of Donald Trump remains an open question. One thing is certain, however. "Making America great again" will entail a doubling-down on fossil fuels.

What we have presented so far is an introduction to basic micro- and macroeconomics for those who have not studied economics formally and a brief review for those who have. However, the total discipline of economics does not confine itself to these limited sets of questions. Over the course of history, economics has focused on other questions not usually covered in introductory textbooks. We will end this chapter by posing these questions and answer them in an historical context in ▶ Chap. 2.

Question #1: What Are the Origins of Wealth and Value?

We begin our discussion of the main questions of economics by distinguishing between income and wealth; throughout the ages, the distinction has not always been clear. Wealth has long been seen as an abundance of goods that are available to a society or to an individual. In preindustrial societies, wealth was the stocks of what nature bequeathed us. But as the economy began to grow and develop, wealth began to be defined as the sum of what humans produced, in other words an accumulation of the flows of value extracted from nature. The question as to whether wealth is a stock or a flow has been debated ever since economic theory developed, and the resolution has never been conclusive. The distinction is also complicated by the level of analysis. Most individuals see wealth as a stock of assets that produce a flow called income. Economists of the neoclassical era defined wealth as a stock called *capital*, while "capital" has been extended to describe all factors of production. Ecological economists regularly refer to the stocks of nature as *natural capital*. Mainstream labor economists see their discipline as the study of *human capital*. In the end, questions of capital and income resolve to a discussion of wealth and value.

Question #2: How Are Wealth and Value Distributed?

Some schools of thought find the question of distribution of the rewards of production to be fairly uninteresting. Some find it the focal point of their analyses. In general, classical political economists found questions of production and questions of distribution to be interrelated but analytically separable. Neoclassical economists, however, found them analytically identical. The neoclassical theory of production, known as *marginal productivity*, is also the neoclassical theory of distribution. Marginal productivity theory stated that each "factor of production" would receive exactly its additional contribution to production. John Maynard Keynes, for the most part, accepted the marginal productivity theory of distribution, with a few, but important, reservations. While the theories of distribution are but peripherally related to energy, they are sufficiently important to economics to deserve specific treatment, especially in the neoliberal era.

Question #3 : How Does the Economy Balance Supply and Demand?

Since the late 1700s, most economists have focused on the possibility that the impersonal market forces of competition and flexible prices could balance the needs and desires of consumers with those of firms. Adam Smith wrote first of this possibility although he never drew a supply and demand diagram. His French popularizer, Jean Baptiste Say, codified Smith's vision of the "invisible hand" into "Say's Law," which expressed the idea that the process of producing goods and services simultaneously creates the income to purchase them. This is better known as "supply creates its own demand." Neoclassical economics accepted "Say's Law" as a fundamental part of their system. British neoclassical economist Alfred Marshall provided us with the modern supply and demand schema that we use currently.

Swedish economist Knut Wicksell extended the analysis to the market for savings and investment, concluding that the overall economy would find its equilibrium at full employment. Keynes disagreed fundamentally with this proposition. Rather, he argued, the economy could reach equilibrium at a level of output that was substantially less than full employment and that it exhibited no internal tendency to change from that low-employment equilibrium. Keynes' arguments for governmental intervention in the economy remain hotly debated today, but there is no question that the cycles of boom and bust that followed the publication and at least partial implementation of his ideas have become much more subdued [22].

Question #4 : What Are the Limits to Capital Accumulation?

While the crucially important subject matter of economics from the time of the mercantilists was the accumulation of wealth, the methods of dealing with accumulation and growth changed substantially once the age of abundant and cheap fossil fuels began. All theorists who wrote in the age of solar flow developed theories of self-limiting accumulation. All classical political economists had growth theories that ended ultimately with society in a nongrowing stationary state. But after the introduction of cheap oil, the focus on the stationary state ended, replaced with the idea of indefinite growth as the result of efficiently functioning markets. However, the transition from classical political economy to neoclassical economics also saw a shift from the concept of long-term accumulation to that of static equilibrium. A neoclassical growth theory did not emerge until the 1950s in response to Keynesian views on the internal limits to growth and accumulation. As we enter the second half of the age of oil, we are facing a new set of biophysical limits that interact with the internal limits found largely in the investment process. To address the role of biophysical limits adequately, we first turn to the historical perspectives on the internal limits to accumulation.

Question #5 : What Is the Proper Role of Government?

Classical political economists stood for a limited role of government. These limited roles are embodied, in fact, in the US Constitution. Governments should maintain property rights, enforce contracts, protect the nation from domestic and foreign enemies, and provide public goods. They should not intervene in market processes or regulate prices. Instead the invisible hand of the market would be sufficient to translate individual self-interest into social harmony. Say's Law assured that the overall system would balance at full employment without the need for government direction. Thus, our constitution reflected the dominant economic thought of the time.

Neoclassical economists too accepted this proposition and translated it into mathematical propositions. The Walrasian core of neoclassical economics asserts that individual exchange based on self-interest (in the form of equal marginal rates of substitution among trading partners) will satisfy not only the traders but result in the general equilibrium of the system at a point where no individual can be made better off without harming another. Prices serve as perfect carriers of information, and any intrusion of the government into market processes will distort the markets' price signals and simply make system not work.

Keynesian economics takes a very different position. The private operation of markets periodically produces insufficient demand, and government action is needed to provide sources of demand that the private sector cannot do profitably. Although Keynes himself believed in the necessity for planning in the long term, there is little in Keynesian economics that justifies government intervention in the internal mechanisms of production and profit-making itself. Nevertheless, an increasing number of economists and politicians "bought into" government intervention as Keynes had suggested. For many decades, from roughly 1930 through 1973, Keynesian demand management, or something like it, helped propel a long wave of economic growth that seemed to work extremely well as the US government pumped more and more money into the economy in both war and peace and as the economy grew steadily year after year. Few paid attention to the fact that this was also an era of expanding supplies of cheap oil, which was, according to economics, "just another commodity."

But after the peak of US oil production in 1970, long-term prosperity gave way to long-term stagnation amid rising prices and a disenchantment with Keynesian economics and its attendant requirement for government intervention. This, and other factors, led to a return to political and economic conservatism in the nation accompanied by a conservative resurgence in the economics profession. Neoclassical economists were back in the saddle emptied of Keynesians and legislation reflected their free market orientation. One can argue, however, the long-term result of these "reduce government intervention" policies resulted in the near financial collapse in 2008. The election of 2010 seemed much like a contest between two sets of policies, neither of which worked in the recent past. Subsequent austerity programs such as those in Greece have led not to prosperity but to continued stagnation and increased human misery.

Question #6 : What Is the Role of Money?

What is money and why does it play such a key role over economic activity? Over the course of history, economists and philosophers have look at money from various perspectives. Is it "the root of all evil?" Is money something simple like a medium of exchange, or is money bound up with cultural identity and national sovereignty? Where did money come from, and how has its different uses over time affected how scholars have theorized about it? What is the relation between money and debt? Should money necessarily be backed by some precious metal, or is paper money backed only by the productivity of the economy and the stability of the government's promise to repay its debts? Can one adequately control an economy by adjusting the amount of money that circulates, or does money play a relatively minor role in overall economic performance? Is money merely a lien on energy, or is it far more complicated? Economists have struggled with these questions since people began to use and write about money. Not surprisingly, different schools of thought have different emphases and outcomes.

Historically, money mostly took the form of debt. Cuneiform tablets, one of the first forms of writing, were actually records of debts. Metallic money, stamped with the ruler's image, arose with the military. Precious metals were an effective way to pay soldiers [7]. Today money is mostly debt. Most of the developed world went off the gold standard in the Great Depression of the 1930s, never to return to it. In the post-Second World War era, the US dollar replace gold as the international currency, and gold was demonetized domestically. Currency is now just the debt of the Federal Reserve System, our nation's central bank. Moreover, most of our money supply consists of checks, which are merely the debt of private banks, and monetary policy is simply a matter of a central bank enabling or restricting the commercial bank's ability to create additional debt. But the debate rages on as to whether money drives the economy, or economic activity determines the amount of money in the system.

Over the course of time, money has fulfilled several roles. Money serves as a medium of exchange, as a readily acceptable way to trade goods and services whose use values are not similar. Money can also be a unit of account. When asked "how much is that worth," most people give monetary value, rather than the number of labor hours it took to produce or acquire or the emotional attachments between humans that the good or service represents. Money may be a store of value. That is why many people fear inflation. It reduces the stored value in money. Unfortunately, the different uses of money are not always served equally well by the same currency. Economist

Richard Douthwaite lists several questions that should be asked to figure out how well money functions for its different purposes:

1. Who issued the money? Many beginning students are surprised that most of our money supply is the debt of private banks and not the government.
2. Why did they do so? Most often a bank, like any private enterprise, does so in order to earn profits for their shareholders.
3. Where was the money created? Was it a national currency, a regional money like the Euro, or a local currency like Ithaca dollars or Berkshares?
4. What gives money its value? Is it backed by something like a precious metal or simply the promise to accept the money for payment?
5. How was the money created? Did people go into debt for a central organization like an international bank, or was it a system for debt and credit at the local level?
6. When was the money created? Was it a one-time event or an ongoing process?
7. How well does it work? Does money meet all three goals?

Douthwaite argues that a single form of money does not fulfill all its functions equally well and advocates different currencies for different purposes [9].

1.11 The Need for Biophysical Economics

The ability to live well within Earth's limits calls for fundamental change, and mainstream economics is not designed to guide such system-level transformations [10]. Consequently, economics as is taught today leaves out several crucial factors. It neglects the fact that all work, including economic production, is driven by flows and stocks of energy. Yet energy is not part of the model that sees instead a circular flow of exchange as the primary system. In addition, the turn toward political conservatism and belief in self-regulating markets has caused economists to return to the ideas of Say's Law and perfect competition. But the real world contains monopolies, non-price competition, and great inequities of political power. While economists emphasize growth, the economy produces long-term, secular stagnation, whether or not the theories recognize this. We need a theory that acknowledges both the biophysical limits of nature and the internal limits to economic growth. Much of the rest of this book does that.

References

1. Smith, Adam. 1923. *An inquiry into the nature and causes of the wealth of nations*, 14. New York: Modern Library.
2. Political economist Anwar Shaikh is also an amateur magician. When asked whether he believed in the magic of the market he said "certainly." He then proceeded to make an egg disappear and pull it from the coat pocket of his enquirer. Professor Shaikh then went on to say that magic, like the market, is mostly illusion. What you see is not necessarily what is going on. What we see is the voluntary exchange between equals. What we don't see is the hierarchy of production, the depletion of high quality resources, and the destruction of the environment.
3. Gowdy, John. 1998. *Limited wants, unlimited means*. Washington, DC: Island Press.
4. Notice the quotation marks around "Say's Law." This means we are highly skeptical of the term. For something to meet the criteria of a scientific law, three conditions must be met: the phenomenon always occurs, and there are no counterexamples. Any scientific law must be consistent with other scientific laws. Moreover, the law can be expressed mathematically. Examples that fit are the Ideal Gas Law, Newton's Law of Gravity, and the Laws of Thermodynamics. The existence of periodic recessions and depressions is a counterexample to Say's supposed law. In addition, the fact that the money on the top loop just equals the money on the bottom loop means no waste occurs in production. But the second law of thermodynamics states that some work is always turned into waste heat. Consequently "Say's Law" is contrary to the basic laws of thermodynamics and doesn't really represent a law in the scientific sense.
5. Goodwin, Neva, Jonathan Harris, Julie Nelson, Brian Roach, and Mariano Torras. 2014. *Microeconomics in context*. Armonk/New York: M.E. Sharpe.
6. Harrod, Roy. 1939. An essay in dynamic theory. *Economic Journal* 49: 14–33.
7. Domar, Evsey. 1947. Expansion and employment. *American Economic Review*. 37 (1): 34–55.
8. Solow, Robert. 1956. A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70 (1): 65–94.
9. Douthwaite, Richard. 1999. *The ecology of money*. Devon: Green Bools Ltd..
10. Raskin, Paul, Tariq Banuri, Gilberto Gallopin, Pablo Gutman, and Al Hammond. 2002. *The great transition*. Stockholm: The Stockholm Environment Institute.

How We Got to Where We Are Today: A Brief History of Economic Thought and Its Paradoxes

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_2

2.1 Introduction

This chapter assesses earlier economic theories from an energy perspective, where that is possible. We also make the case that although economics has not dealt with energy very explicitly, the discipline has addressed many other important issues that help us today to understand just how energy operates within economies as well as provide a number of interesting and important perspectives on economies that are not related to energy. The purpose of this chapter and the next three is to utilize the insights and the methods of prior economic schools of thought to build a new theory that explains actual economies much better while addressing energy and biophysical limits to human activity far more explicitly than does mainstream theory.

2.2 Surplus and Scarcity

Economists through the ages usually have commenced their discussions of value, distribution, and growth from two fundamentally different starting points: *relative scarcity* and *economic surplus*. Before the age of fossil fuels, economic theory was based on the premise that nature limited the flow of resources; in other words, there was an *absolute scarcity* of economic goods and services. After the 1870s the physical limits became much less important because of the concentrated power of fossil fuels. Questions of the biophysical means by which the wealth was generated simply fell off the radar screen of economists. The focus of analysis shifted instead to that of *relative* scarcity: that is, of individual subjective choices while facing limited access to money. The theory depended upon the assumption that individual humans were acquisitive and rational beings whose desire for more material goods as the source of happiness could never be satisfied and for whom the desires and preferences of others were irrelevant. No level of output, no matter how abundant, could ever satisfy fully these *unlimited* wants. It is a psychological, not a physical, problem. From this perspective, the clash between limited means and unlimited wants is *the* economic problem. This view of scarcity as the starting point of contemporary economics underlies the usual formal definition of economics as we gave in ▶ Chap. 1.

The discussion of *economic surplus* begins with the premise that society can produce more than it needs for subsistence by organizational and technological means. Access to energy is rarely mentioned but lies beneath the surface. Stated simply, an economic surplus is the difference between society's economic output and the cost of producing it. The *surplus approach* relates to Polanyi's *substantive* definition of economics. In the 1960s, Karl Polanyi wrote and edited a collection of essays on ancient economies, where market-forming prices had little to do with how things were distributed. In *Trade and Markets in Early Empires*, Polanyi and his colleagues realized that markets dated back to antiquity, but that price-forming markets are a contemporary phenomenon. His point was that if one looked at ancient societies through the lens of modern, price-forming, markets, one was likely to miss more than they might discover. In order to understand the ancient economy Polanyi offered a *substantive* definition of economics. We believe that this definition is also an excellent starting point for the integration of energy and human society into economics. "The substantive meaning of economics derives from man's dependence for his living upon nature and his fellows. It refers to the interchange with his natural and social environment, insofar as this results in supplying him with the means of material want satisfaction" [1]. ▶ Chap. 1: In other words, the substantive Definition focuses on how human beings transform Nature… how human beings transform nature to meet their needs. Nature was seen to be abundant. Most economists of the classical period treated it as a "free gift." Economists of the pre-fossil fuel age relied primarily on the economic surplus approach. But by the 1870s came the dawn of the fossil fuel era, the industrial revolution, and the consumer society. For economists, the basic starting point for thinking about economics could be reformulated from *producing* an economic surplus to exchanging commodities that were relatively scarce without thinking much about how products came into being. At the same time, the analytical focus changed from social class to the individual and from an objective accounting of the costs of production to the individual valuation of subjective well-being or utility. The goal of economics became one of figuring out the optimal allocation of resources to best meet human psychological desires. In other words, economic theory was transformed from focusing on obtaining more from nature into an exercise to figure out *who* gets the goods and services and *how* goods and services best enhance subjective well-being. According to the new *neoclassical* economists, the answers were to be found in the magic of self-regulating markets where individual pursuit of self-interest led to social harmony. While this concept was derived from the earlier writings of Adam Smith, it was augmented by mathematical "proofs" appropriated, or better misappropriated, from energy physics. Meanwhile new research in behavioral economics shows that there is little empirical evidence to indicate that human beings actually behave in this "self-regarding" way.

2.3 Economic Surplus as Energy Surplus

Economists of the seventeenth through the nineteenth centuries did not, in fact could not, focus explicitly upon energy as a source of surpluses because the formal concept of energy did not yet exist. Nevertheless, the ability to extract an energy surplus from solar flow or terrestrial stocks forms the basis of economic production and surplus. Contemporary energy analyst Richard Heinberg provides a framework by which to assess the economic roles of such energy surpluses [2]. He argues that throughout history humans have engaged in five strategies to expropriate energy: takeover, tool use, specialization, scope enlargement, and drawdown. *Takeover* was the primary method of early humans, as we appropriated more of the solar energy flow for ourselves by diverting a portion of the Earth's biomass from supporting other creatures to supporting humankind. Our ancestors took over land to grow crops, first as horticulture and later as agriculture, the growing of field crops at the expense of other species. Agriculture turned a complex ecosystem into a simple one. Plants that grew where they were not useful to humans were weeds. Animals that competed for the food were pests. As humans migrated from Africa to the far corners of the world, they took over more and more biocapacity, often disrupting the natural balance. Everywhere humans have gone large mammals have disappeared. The rapid release of chemical energy known as fire aided the process of acquiring energy surpluses. Pioneering biophysical economist Nicholas Georgescu-Roegen termed this a Promethean innovation, which was truly species altering. The only other Promethean innovation was the steam engine. In addition, humans enhanced their abilities to harness the solar flow by domesticating certain animals which could provide more motive power than the biomass necessary to feed them.

Heinberg's second strategy was that of *tool use*. Humans have long-used tools, for tools can augment the takeover of energy from other species and other societies to expropriate ever-increasing amounts of energy from the biophysical system. Specialized tools called weapons aided our ability to concentrate energy in spear points and hunt more effectively, as well as expropriate energy from other societies. Tools have evolved from those that required only human energy for their manufacture and use, such as spear points, to those that use copious amounts of energy and exotic materials from external sources for their manufacture and use, such as the internal combustion engine. As the energy surplus rose to a sufficient level so that not all members of society had to work constantly simply to provide adequate food, humans could begin to specialize on activities such as toolmaking or soldiering. All hierarchical societies that support people who are not immediate producers of crops depend upon this. Increased agricultural productivity could now support classes of artisans, aristocrats, and intellectuals who could better design and build tools and improve social organization designed to capture even greater amounts of energy. All classical political economy, from the French physiocrats to Adam Smith, acknowledged the role that specialization played in determining wealth and value. Howard Odum talks of all kinds of natural and human-dominated systems "self-organizing" to generate "maximum power." From this perspective, humans are not doing anything that other organisms don't do; they are just "good" at it because of their technologies which are now supplemented with the "large muscles" of fossil fuels.

Another strategy of energy appropriation was that of *scope enlargement* or the transcendence of limits. Justus von Liebig found that the limiting factor in the carrying capacity of any biophysical system, especially agriculture, was the factor or input least available relative to the needs of the growing plants or other ecological units. This limit could be pushed back by appropriating the biocapacity of other regions through conquest or trade. Mercantile doctrine rested de facto upon the foundation of acquiring the solar energy surpluses of other regions. The practical aims of traders were later codified by David Ricardo into the doctrine now known as *comparative advantage*. The benefits of trade result from enlarging the scope of the energy would be shared by all traders. Industrial society depended upon the ability of urban industrial centers to appropriate the biomass of rural areas in terms of food and wood for fuel. Unfortunately, many of the nutrients that would have been returned to the soil in the countryside built up as waste in the city. Ecologist Justus von Liebig himself referred to this system of commercial agriculture as "robbery." [4]. Scope enlargement also entailed the stealing of solar surpluses from others through war, exploitation, and colonization largely to provide and enrich the treasuries of dominant powers at the expense of the conquered and colonized.

The last and most successful strategy for increasing carrying capacity that Heinberg describes is that of *drawdown*. Drawdown began occurring when we were able to change from living on steady solar flows to tap nonrenewable stocks of fossil fuels, particularly those of coal, oil, and natural gas. Drawdown was enabled by the development of sophisticated tools and greatly enhanced the previous strategies. With drawdown, humans could appropriate nature sufficiently to support a much higher population at a greater standard of living for a fraction of the population. At the beginning of the age of fossil fuels, around 1800, the world's population stood at approximately one billion. Since then the world supports more than seven times that number. Half of that increase came in the past 50 years following the "Green Revolution" when plant breeders combined hybrid grains with energy-intensive input packages of fertilizers, other agrochemicals, irrigation, and cultivation. While the benefits of increased yields were extended to a broader segment of the world's population, not everyone enjoys food security. There are about 800 million hungry people in the world today.

Heinberg also points out three dangers of the drawdown strategy. First, drawdown of fossil fuels creates pollution. This can take the form of pollutants such as sulfur dioxide and nitrogen oxides that foul the air and acidify the soils and water supplies. Runoff from lands treated with nitrogen and phosphorus fertilizers creates "hypoxic dead zones" in areas such as rivers, lakes, and the mouth of the Mississippi River in the Gulf of Mexico. Secondly, the pollution can take the form of carbon dioxide emissions, whose increasing atmospheric concentration are seen by the broad consensus of scientists as the primary driving force of climate change. Finally, terrestrial stocks of fossil fuels are finite. At the beginning of the twenty-first century, we are at or near the global peak use of these fuels, especially oil. As they become less available and more expensive, societies dependent upon them will undergo dramatic transformation with potentially grave economic as well as social consequences [3].

As we approach the limits to the drawdown of nonrenewable stocks of nature, the idea that we can satisfy human needs and economic priorities by producing and consuming ever greater quantities of material goods should become a matter of inquiry rather than of blind faith. In the post-peak years, we will need to confront the distinct possibilities of absolute scarcity and the diminished capacity to appropriate surplus energy. We need to revisit and reexamine the questions economists have asked for centuries.

2.4 The "Big Es"

Through the ages, different schools of economic analysis focused on a different "Big E." Each of these can be considered as a social construct, a way of thinking about how people integrated with the economy and with nature. The mercantilists organized their thought around a better understanding of *exchange*. How could political economists help change the laws in order to facilitate expansion of commerce and regulation of trade? Classical political economists directed their efforts toward *economic policy.* For them the object of economics was to inform policy makers. The physiocrats wanted to reform French agriculture and encourage large scale production of commercial crops. Adam Smith focused on the elimination of mercantile trade restrictions, while David Ricardo wanted to raise taxes on the aristocracy and facilitate further the trade in food and industrial commodities. Thomas Malthus, on the other hand, wanted to subsidize the aristocracy. Both Ricardo and John Stuart Mill were elected to parliament where they argued effectively to change economic policies. The theories of Karl Marx were grounded in the *exploitation* of labor and the recapitalization of surplus generated by workers and fossil fuel-driven machines.

Neoclassical economic theory revolves around *efficiency and equilibrium*, concepts appropriated from physics. Their fundamental belief is that the economy could be analyzed separately from the rest of the society and that an economy would tend toward a state of balance without intervention from political agencies. The "Big E" for Keynesian economies was *employment*. Keynes genius was to realize that equilibrium could occur at any level of employment, including very high and politically unacceptable levels of unemployment. Keynes advocated the use of the government to insure the economy balances at high levels of employment.

Institutional economics starts with the process of *evolution*. The current professional society of institutional economists is called The Association for Evolutionary Economics (AFEE). Institutionalists such as Thorstein Veblen and John Commons rejected the mechanical analogies of neoclassical economics and looked primarily at the evolution and structural changes in economic and institutions. For ecological economists, the main contribution was that of *embeddedness*. Ecological economists embed a growing economy within a finite and nongrowing ecosystem and sometimes a social system as well. Finally, biophysical economics starts with analysis of the flows of *energy* and analyzes how changes in the quality and availability of energy shape economic activity.

2.5 The Present as History [4]

We should study history as more than just idle curiosity. The lessons of history may provide valuable lessons for the problems of today and tomorrow. Studying history, in our opinion, yields a more sophisticated understanding of how we arrived at our current state of affairs, just as unifying social and natural science approaches allows a better understanding of an economy that is embedded in society and in nature.

We will need a new set of economic theories for the second half of the age of oil: theories that neither treat "Nature's Bounty" as a free gift nor posit resource endowments that appear magically as "manna from heaven." Moreover, we must contend with the problem of limits to growth. In the past humans transcended the boundaries and limits imposed by nature largely by the application of increasing quantities of cheap fossil fuels. But the era we are entering will most probably see the end of this. As high-quality fossil fuels increasingly run short, and the use of all carbonaceous fuels compromises our atmosphere and other natural systems, the specter of living within our means while protecting our home becomes more and more difficult. This is likely to mean the end of the growth economy and will cause us to reconsider the meaning of technological change. As we do this, and as we begin to develop new economic theories appropriate to a new age, we need to consider that many important questions and insights exist in the writings of the economists of the past. Thus, we examine next the most important ideas of earlier economists.

As we trace the origins and development of what we call economics today, we will return to the six questions identified in ► Chap. 1 again and again. We will also introduce the concept that we see each of these questions as being in large part about energy. Even though the questions asked by economists tended to remain the same over time. The theoretical emphases, methods of inquiry and analytical vision were so fundamentally different from one time to another that economic theory can be divided into six distinct periods and "schools of thought."

2.6 Schools of Economic Thought

The different schools of thought often asked similar questions but had very different visions of how the economy worked. They directed their writings toward different purposes and used very different analytical methods. We now ask how each approached the main questions of economics.

2.7 The Mercantilists

Before the European flowering of exploration and commerce in the late fifteenth century, day-to-day life changed slowly in the medieval era. European society was organized around the manor and a strict hierarchy with the church and the landowners at the top, a tiny class or artisans and merchants in the middle, and a landless peasantry constituting most of the population. Church doctrine and economics writings were dedicated to keep life from changing. For the scholastics that shaped the ideas of the medieval days of feudalism, the origins of wealth lie in the land, specifically with the ownership of the land. Those that owned and controlled the photosynthetic capability of the land were wealthy. Those that did not own land were not. The nobility and the church owned the land, and the elaborate principles of medieval law served to concentrate land ownership. Primogenitor demanded that all land was given to the first-born son. Daughters of landowners were expected to marry sons of other landowners. The ban on usury prohibited merchants from acquiring wealth through the charging of interest, and profits by means of trade were limited by the "just price" which covered only the costs of production, transportation, and the return necessary to keep one in "their station in life." Social mobility was a mortal sin. The peasantry, known as serfs, labored primarily in the fields of their feudal lords having but 1 or 2 days per week to till the Commons for their own subsistence needs. All paid taxes to the nobles and tithes to the church.

After what historian Barbara Tuchman calls "the calamitous fourteenth century," the thousand-year stability of feudalism began to fracture, and by the beginning of the sixteenth century, the merchants, much reviled by the nobility and the church, began gaining control of society. Wealth was in new hands that found new uses for it. Art and music prospered and proliferated as did commerce. The age of exploration ushered in the age of long-distance trade as well as the Renaissance. The forests and mines of the "New World" augmented the long-depleted stocks of the old. The writers of the new mercantile period began to redefine the meaning of wealth, from control over land and its biomass to accumulation of "treasure" or stocks of precious metals. This was the essence of mercantilist "economics." By the middle of the seventeenth century, thought on how best to accumulate wealth changed from the treasure itself to the gains made by trade. Treasure, and therefore wealth, would flow to those nations which achieved a positive balance of trade. As much money could be made in control of shipping and customs as could be made mining and refining the treasure itself.

It is important to think about the six key questions from the background of the various dominant economic "schools" of thought as they evolved over the history of economics. The first identifiable school of economic thought was known as *mercantilism*, which was grounded in the economics of long-distance trade. Mercantile doctrine took the form of pamphlets written primarily to justifying the expansion of trade. Although their aims and purposes were practical, mercantilist writers did make advances in questions such as the origins of wealth and value and the accumulation of capital. In many ways, mercantilism was primarily about takeover and scope enlargement.

Mercantilist writers were most often practical business people, not academics. The most famous, Thomas Mun, was a director of the British East India Company. All defined the purpose of the economic endeavor as the accumulation of treasure in the coffers of the nation state. Mercantilists, not surprisingly, took the position that the origin of value, or price, lay in the process of exchange, and they meant to control the terms of that exchange. Their primary mechanisms were colonization, commercial treaty, and war. For most of the sixteenth century, the British battled the Spaniards for control of New World colonies. The seventeenth century was spent engaged in rivalries with the Dutch for control of colonies in the East Indies as well as the Caribbean, while the eighteenth and early nineteenth centuries saw the prolonged conflict between the British and the French. Mercantilists demanded the aid of their governments in determining the terms of trade. The British Parliament passed a series of restrictions (the Navigation and Trade Acts) to assure the positive balance of trade, at the expense of their mercantile rivals and the colonies themselves. By the time that Adam Smith penned the original *Wealth of Nations*, British supremacy was in sight. With the triumphant end of the last mercantile war against Napoleon on June 18, 1815, the world settled down to a long peace but a peace on British terms—*Pax Britannica.*

Mercantilist writers were primarily interested in changing policy to enhance their accumulation of treasure. Few spent any time pondering the historical origins of wealth. Early mercantilists, sometimes known as bullionists, took the position that trade was a pump for wringing gold from a domestic economy. This argument made some sense when a nation exported raw materials, based on the appropriation of solar flow and for which there were many substitutes, and imported finished goods, based on the harnessing of human energy supplemented by the power of wind and water, for which there were few. The terms of trade, or ratio of export prices to import prices, were against the raw material exporter, and they suffered from declining terms of trade. In this case the accumulation of wealth is served well by the restriction of trade.

By the end of the sixteenth century, however, England had become a manufacturing nation and was exporting its products to Europe and to the world. Mercantile thought then turned to crafting an argument that justified the expansion of trade as the primary mechanism to augment a nation's stock of precious metals. The most widely recognized tract of high mercantilism was *England's Treasure by Forraign Trade,* written in 1630 by Thomas Mun and published, after his death, in 1664. Mun's primary purpose was to persuade legislators to abolish the ban on exporting gold. He argued that the export of gold could facilitate the accumulation of treasure if that export led to a positive balance of trade or the excess of exports over imports. To accomplish this goal, Mun and his followers advocated state policies of the regulation of trade. While the mercantilists stood for the expansion of trade, they were not advocates of free trade.

At that time, the ability to extract an energy surplus was limited by the lack of concentrated energy sources. The ability to extract solar flow and turn it into products with economic value could be enhanced only by organizational change, primarily, in the form of plantation agriculture and slave labor. Mercantile doctrine contained no insights as to how to reduce the costs of production, other than the encouragement of the carrying trade, which aided the gains of the trade itself and the accumulation of treasure. Ships were constructed from wood (biomass) and powered by the solar flow of the winds, which may or may not have blown in the desired direction at the desired speed. Yet speed and tonnage improved in the mercantile period. Mercantile doctrine was a matter of scope enlargement by means of expanded trade. As much money was to be made in transportation as was to be made in the initial appropriation of the embodied energy in crops and precious metals. But expanded and speedy transportation was limited by energy availability. Trade, in the mercantile era, was a dangerous and slow endeavor, albeit often a profitable one.

2.7.1 Mercantilist Theory

Mun distinguished between natural and artificial wealth. Natural wealth was what could be spared from domestic use and consisted primarily of agricultural products. Artificial wealth was that derived from trade and manufacturing. Mun thought that acquiring artificial wealth through trade would be more profitable than producing natural wealth domestically. By pursuing a policy of a positive balance of trade, a country without mines would be able to accumulate precious metals. In terms of distribution theory, the mercantilist writer took a very hierarchical position as to where the treasure should flow. Trade was at the top of the scale, followed by manufacturing and then by agriculture. Another mercantilist theorist, Charles D'Avenant, considered a seaman engaged in trade to be worth three farmers [5]. The royal treasury should use the gains of trade to subsidize the carrying trade, and wages should be kept low to restrict consumption, especially consumption of imported goods.

It should not be surprising that the primary advances made by the mercantilists were with regard to the theory of money. While the mercantilists believed in expanded trade, they did not advocate free trade. Rather they believed the government should enforce a strict set of codes known as the Navigation and Trade Acts. Those familiar with American history might recognize that it was the vigorous enforcement of the acts such as the tax on tea and the prohibition of white settler farming west of the Appalachian Mountains (and its large photosynthetic potential) that helped precipitate the American Revolution. The mercantilists believed that treasure would accumulate only if there were to be nonequivalent trade. To accomplish this, the mercantile power needed to run a positive balance of payments. This implied that colonies were to experience balance of payments deficits. The drain of wealth from colonies to augment the coffers of the mercantile powers was a source of social discord in countries other than the United States as well. The mercantilists, while they borrowed heavily to finance their foreign operations, wrote in an era of metallic, or commodity, money. Achieving a positive balance of trade made gold and silver flow into the royal treasury. For them an expansive monetary policy meant acquiring more treasure, not the conscious manipulations of the interest rate and size of the money supply. In fact, the mercantilists argued that the government should simply not take overt action to limit the export of gold or the size of the money supply.

Two important figures represented the transition between mercantilism and classical political economy. William Petty, an English mathematician and physician, began to explore the connection between the costs of production, economic surplus, and the value of commodities by the late 1600s. He was one of the first to express himself mathematically in *Political Arithmetick*, published in 1690, 3 years after his death. Since prices were crucial to the mercantile endeavor, Petty sought to explain the origin of prices and values. His work was the immediate predecessor of the labor theory of value that came to characterize the approach of classical political economy. Yet Petty also stressed the importance of land, reducing all forms of economic surplus to rent. He valued land as the sum of annual rents and was among the first to link the value of land to the rate of interest. Petty also drew upon the political and economic writings of John Locke, who stressed the connection between nature and the products of human labor. Locke believed that nature furnished fundamentally worthless materials, and it took human labor to transform the products of the Earth into something useful. Locke and Petty struggled with the difference between use values, derived from products of nature, and exchange values which resulted from the application of human labor. This distinction would later be clarified in the era of classical political economy. Petty, however, thought along incipient biophysical lines when he reasoned that "Labour is the Father and active principle of Wealth, as Lands are the Mother." [6]. The French precursor to the physiocrats, Richard Cantillon, who was influenced by Petty argued along similar lines. "Land is the Source of Matter from whence all Wealth is produced. The Labour of man is the form which produces it." [7].

2.8 Classical Political Economy: The Physiocrats, Adam Smith, David Ricardo, Thomas Malthus, and John Stuart Mill

By the end of the eighteenth century, mercantilism would give way to *classical political economy*. This era began around 1759 when a French school of natural philosophers called the Physiocrats developed a theory of value that tied the origins of wealth to the photosynthetic capabilities of the land and the agricultural labor that appropriated it. Agriculture in the pre-fossil fuel era transformed solar flow into food by means of land. In 1776 Scottish moral philosopher Adam Smith published *An Inquiry into the Nature and Causes of the Wealth of Nations*, linking a preindustrial and pre-fossil fuel manufacturing process to a general theory of circulation. Smith's book led to the great debates over distribution, population, and, in time, the concept of diminishing returns of Thomas Malthus and David Ricardo and the utilitarianism of John Stuart Mill. These hundred years generated a rich and thoughtful discussion about what the proper focus and moral obligation of economics was and should be.

Classical political economists had an entirely distinct set of purposes. Both the physiocrats and the first important classical political economist, Adam Smith, desired to overturn the mercantilist doctrines of regulated trade. The Physiocrats, who gave us the term laissez-faire ("let us alone"), sought a change from small-scale peasant crop production to large-scale commercial agriculture. One can reasonably assert that Smith's 1776 *Wealth of Nations* was the greatest anti-mercantilist tract ever written. Not only did he believe that state regulation inhibited commerce but also that mercantilist doctrine retarded domestic production. Smith pursued and developed the idea that markets could lead to the expansion of well-being, guided as if by an unseen hand, rather than by the heavy and visible hand of state regulation. Half a century later, David Ricardo would refine the doctrine of mutual benefit from unregulated trade.

The classical political economists, taken as a school, desired to build an economic science and to uncover the origins of wealth. They did this largely through a substantive, and historically specific, study of economic surplus. Their method was essentially a narrative, supplemented by abstract propositions and the occasional recourse to numerical tables. All classical political economists were policy oriented. Adam Smith advocated not only the end of mercantile restrictions but increased expenditures for public education and a high-wage economy; Thomas Malthus and David Ricardo debated the perpetuation or abolition of the Corn Laws limiting the import of food from continental Europe. John Stuart Mill argued in favor of reforms to diminish the gap between those living in wealth and poverty as well as for the emancipation of women.

These political economists grounded their analyses of the origins of wealth and value in the process of *production*, rather than in the process of *buying and selling*, or exchange, as did the mercantilists. Moreover, all used social class as their unit of analysis. The familiar "factors of production" of land, labor, and capital had their origins in the actual, and historically specific, social structure of their days. The primary questions of interest for the classical economists were those regarding the production, accumulation, and distribution of economic surpluses. Their theories of capital were historically specific and related to those of accumulation and value. "Capital accumulation is regarded as a necessity prior to production and production as necessity prior to the exchange of commodities" [8]. Price formation, which has come to dominate modern microeconomics, was of minor concern to them.

2.8.1 The Origins of Wealth and Value

For *classical* economists, who called themselves "political economists," wealth (a stock) and value (a flow) originated in the process of *production*, rather than that of exchange, as the mercantilists believed. Further, the idea that united the diverse classical political economists was that value could be determined objectively by adding up the costs of production. They believed that human labor, augmented by tools, land, and organization of the labor process, was the source of value.

Classical political economists were careful to make two distinctions. They separated use value from exchange value. Unlike modern neoclassical economists, a product did not command a price because consumers found it useful. A commodity commanded a price because the products of nature were transformed by means of human labor. The transition from classical to neoclassical economics represents an epistemological break as far as value theory is concerned. Alongside the distinction between public wealth and private riches, James Maitland, eighth Earl of Lauderdale, wrote in 1819 that public wealth consisted of use values—"all that man desires that is useful or delightful to him." Private riches, however, consisted of exchange values—delightful or useful things that are scarce. So, for Lauderdale a paradox emerged: the enhancement of private riches comes at the expense of public wealth, precisely by making the enjoyments offered by nature scarce so they can command a price [9, 10]. Since the days of the triumph of neoclassical economics, few economists today separate use value and value. They see wealth as merely the accumulation of exchange values, expressed in money form. But as resources become absolutely scarce in the future, a knowledge of the theories that existed, and the theoretical separations that were made, could be a vital component of an economic theory for the second half of the age of oil.

The first classical political economists, the *Physiocrats*, asserted that value originated in the land and the agricultural labor that appropriated the Earth's biomass by planting, harvesting, and transporting food. Only nature created a net product (or *produît net*). Manufacturers were considered sterile in that they only transformed the value created by the land. From their perspective, they added no net product.

In the English-speaking world, in contrast, economic theory extended the creation of value to manufacturing as well as agriculture. The generally acknowledged founder of British political economy was a Scot, Adam Smith. Smith is most often recognized for his belief that the "invisible hand" of the market would transform individual self-interest into social harmony. He began his 1776 opus, *The Wealth of Nations*, by raising the question of value. Smith diverged from both the mercantilists and the Physiocrats. He asserted that the origin of value could be found not in the bounty of nature and agricultural labor but labor in general, specifically in the productivity of labor and the number of productive laborers. Wealth was the accumulation of values generated by producing goods and services for sale on the market. He was writing in the era before fossil fuels were applied widely to manufacturing, and his theory reflected his time. Smith's observations, the most famous being that of a pin factory, led him to believe that the primary method of augmenting the wealth of a nation was to implement the division of labor, where the production process would be subdivided into separate and more productive tasks. Smith, who was a professor of moral philosophy, then had to connect the division of labor to an overall "system of perfect liberty" found in the unencumbered operation of free markets. He did so with a surprisingly simple statement: "The division of labor is limited by the extent of the market" [11]. In order to reap the benefits of the division of labor, a manufactory must have access to a sufficiently wide market to sell the products the division of labor made possible. An important constraint on that perspective, however, barely understood by Smith, was that the market itself was limited by the reliance on solar flow and animal power to transport products of the division of labor.

Smith also deals with the origins of the division of labor. Partly he attributes it to human nature. We all have an ingrained propensity to "truck, barter, and exchange," in addition to possessing a desire to increase the number of necessaries, conveniences, and amusements available to us. Always the historian, Smith addresses the question of how much any particular commodity (known today as a good or a service) was worth in earlier times as well as in his own day. He argues that in the "rude and early stage" of society, before the development of tools and private property, the value of any commodity consisted of the amount of human labor *embodied* in production (meaning the hours of labor that had been used to make something). This was the sole determinant of value or price. Workers could generally fashion their own tools. A distinct tool manufacturing sector would have to wait for the application of more concentrated energy. "Labour was the first price, the original purchase money that was paid for everything. It was not by gold or silver, but by labour that the wealth of the world was originally purchased…If among a nation of hunters, for example, it usually costs twice as much labor to kill a beaver which it does to kill a deer, one beaver should naturally exchange for or be worth two deer." [11]. In this stage of development, the whole product of labor belonged to the producer. But in the eighteenth-century society, characterized by the division of labor, this situation would not hold. At that time "modern" society enhanced the production of each worker through various kinds of equipment, and the owners of capital stock, who provided the equipment and advanced the wages before the crops were harvested, demanded a share of the output. So, too, do the owners of land. Smith argued that the "natural price" or value can be obtained by adding up the natural prices of land, labor, and capital. Smith was not particularly clear about this and had to devote pages upon pages to determining the natural rates of wages, rents, and profits. Moreover, Smith patterned this "rude and early" society after North American Indians, of which Smith knew little to nothing about. Had he been more knowledgeable, he would have realized that the hunter and the trapper would not have made exchanges based on labor hours. They both would have taken their catches to the clan mother. She would have distributed the meat and the fur according to tribal tradition [8].

2.8.2 Smith on Money

This view of the division of labor was crucial to Smith's vision, as his views of money depended upon it. Once the division of labor was established, all people lived by exchange. Money evolved, according to Smith, because the barter system had one important drawback. Exchange could not occur if your trading partner did not desire the use value you possessed and vice versa. Over the years people chose a particular commodity to serve as a currency. Smith lists items, such as cattle, salt, cod, tobacco, and sugar, but argues people eventually chose metals because of their durability. Spartans used iron and ancient Romans copper. Modern commercial nations chose gold and silver, stamped with the image of the ruler to assure weight and quality. Smith does argue, however, that the avarice of all princes and sovereigns led them to debase the currency.

Smith's chapter on money also contains several theoretical positions. In this chapter, he argues for the separation of use value and exchange value and argues that natural price flows only from exchange value. He introduces the idea that natural price is the money expression of the costs of production of land, labor, and capital and prefaces later chapters that will explain why market prices often diverge from natural price. It is also in this chapter that Smith advances the diamond-water paradox and explains the all-important role of relative scarcity in the determination of natural price.

Smith then goes on to explain the original accumulation of stock by the virtuous behavior of those frugal individuals who save. "Capitals are increased by parsimony and diminished by prodigality or misconduct." When the frugal abstain from immediate consumption they add to their capital. They use this capital to set to work industrious persons, and as capital accumulates, the potential productivity embodied in the division of labor rises too. In the end for Smith, the source of the increase of wealth can be found primarily in the increased labor productivity of an increasing population and the virtuous behavior of frugal savers.

The next great English-speaking political economist was David Ricardo, whose 1817 *Principles of Political Economy* [13] represents the definitive statement of classical political economy. Although Ricardo had little to say about the origins of wealth, he made significant contributions to the theory of value. Ricardo was the premier advocate of a pure labor theory of value. He believed Smith to be incorrect when he separated labor *embodied*, the amount of human labor time used in production, and labor *commanded*, or what that labor is worth in terms of purchasing alternative commodities. Ricardo reconciled the two when he declared that capital was simply "dated labor." Most capital at the time was known as circulating capital or the money advanced to purchase labor. Since capital can be reduced to labor, the value of any commodity, or good produced for sale rather than use, was determined solely by the amount of human labor embodied in production.

The problem of dealing theoretically with long-lived fixed capital is an old one, indeed. Ricardo believed that market processes would equalize profit rates. But if one commodity was produced in a more capital-intensive process, problems emerged. If the amount of total capital was the same for two producers, then an equal profit rate meant selling the goods for the same prices, as the market also equalized price. But if, for example, wages increased, it would have a much greater impact on the more labor-intensive commodity. Two goods with unequal amounts of labor would have different prices according to the labor theory of value, as the theory states that the value of a product is a function of the labor put into crafting it, not a function of the use or pleasure derived from the product. But competition in markets would yield the same price. It seemed mechanization was incompatible with the labor theory of value. Ricardo was never able to solve this problem. An unfinished manuscript was found on his desk at his death. His theory did not reflect reality—the less efficient, more costly production would simply be less profitable—as Marx discussed. Ricardo never dealt directly with energy. Nonetheless, he provided two theoretical tools that critically inform energy analysis to this day: the *best first principle* and *diminishing marginal returns*. We will deal with these principles in the next section on income distribution.

John Stuart Mill began his 1848 *Principles of Political Economy* [14] *by* asserting that production was the process of the transformation of natural resources by means of human labor. He began the project of updating and revising Ricardo's *Principles of Political Economy* by expressing an affinity to the labor theory of value. But he ran into the same problem that vexed Ricardo in a mechanized economy. Mill believed the pure labor theory of value applied only when there were equal capital/labor ratios across industries. However, Mill knew this was not accurate depiction of the English economy in the mid-nineteenth century. Instead he fell back onto Smith's adding-up theory of value. In his approach profits were the natural price of capital and a reward for the service the capitalist provided. Mill also relied on an opportunity cost approach. In a phrase taken from Nassau Senior, Mill asserted that profits were also the reward for the "abstinence" of the capitalist who sacrificed by saving and investing instead of consuming.

Mill also rejected the classical doctrine of the wages fund, whereby capitalists advanced only a fixed amount for the payment of wages. If one group organized to increase their wages, it would come at the expense of other wages. This was essentially Malthusian in origin, as the limited ability of pre-fossil energy sources to produce food and the proclivity of the poor to produce children keep wages at bare subsistence. But food production was increasing, and the social order was subject to change in the mid-1800s. Remember, Mill's *principles* were published in the same year, 1848, as the *Communist Manifesto.* Instead, Mill thought wages would be determined in a struggle between workers and capitalists [15].

While claiming some adherence to the labor theory of value, Mill was also a utilitarian. Mill's utilitarianism was rather eclectic and rather different from Bentham's. Bentham, as you may recall, thought that one could not compare the utility of one pleasure to another. Each individual was the best judge of his or her own well-being. Mill, on the other hand, separated higher from lower pleasures. Higher pleasures included those of the Victorian salon: poetry, opera, and philosophical conversation. Lower-order pleasures can be summarized in the modern saying: sex and drugs and rock 'n roll. Mill did not believe, as did Smith and other classicals, that all humans are motivated solely by self-interest. He believed that people are driven by nobler motives then competing with one another to get ahead. In modern terms, a sustainable society had to be a just society. Nonetheless utilitarianism made its way into Mills' value theory in the guise of the separation of use value and exchange value. Recall that Lauderdale had separated public wealth, in the form of use values, from exchange values that commanded a price because of scarcity. Mill concluded eventually that the basis of wealth was not only things that delighted us, or use values, but things that delighted us and were scarce. In other words, wealth could be calculated by summing exchange values or prices. In this sense Mill was the consummate transition figure from classical political economy to neoclassical economics [10].

2.8.3 Classical Political Economy and the Distribution of Wealth and Income

The unequal distribution of wealth was the fundamental problem that had been addressed by the physiocrats. French agriculture yielded little surplus product, as production was on a small-scale subsistence basis with basic wooden (biomass) implements and little application of fertilizer. What little surplus existed was appropriated to support the lavish court in Versailles and to subsidize a set of pampered workshops dedicated to the hand production of luxuries. The Physiocratic program advocated instead the reinvestment of agricultural surpluses on the farm and the creation of large-scale commercial agriculture on the English model. The first economic model ever, the *Tableau Economique*, was designed to illustrate the problem of unequal distribution of wealth. Its modest reforms, however, ran afoul of Louis XVI and were ultimately doomed to failure. The physiocrats ultimate success was the influence they had upon later theorists such as Adam Smith and Karl Marx.

Neither the mercantilists nor Smith focused primarily on the problem of income distribution. Mercantilists, focusing on trade and exchange as the source of wealth, had little to say about the internal order of the domestic economy. This is hardly surprising as the ability to transform fundamentally the process of production by utilizing fossil energy had yet to be developed. Their main focus was the distribution of subsidies. Mercantile doctrine held that a trader was worth several artisans, and artisans are worth many husbandmen. Therefore, subsidies should flow toward those engaged in international trade. Profits were to be made and hence encouraged in the carrying trade and in the exploitation of colonial resources, not by means of reducing the cost of production at home or elsewhere.

Smith, too, wrote relatively little about income distribution, which is surprising given that he was professor of "moral philosophy" and published a lot. Smith did believe that some degree of inequality was natural and that it provided incentives for increased productivity. "Wherever there is great prosperity there is great inequality. For every rich man, there must be at least 500 poor, and the affluence of the few presupposes the indigence of the many." Yet at the same time he believed: "No society can surely be flourishing and happy of which the far greater part of its members are poor and miserable" [18]. Smith truly believed that accumulation of capital would raise living standards for all in the long term, although inequality would persist. In the final book of *The Wealth of Nations*, Smith held out that a commitment to education would also raise the status of the working poor, a position commonly held by many in society today. In his chapter on wages, Smith also wrote at length on the factors contributing to the differences in wages, including the difficulty of learning the trade, constancy of employment, the degree of responsibility, and the uncertainty of success [15]. Smith held a special distaste for the landed aristocracy who loved to reap what they had not sown. He considered rents to be primarily a monopoly extraction on the part of proprietors who did not labor productively. To this day, the term "rent seeker" is one of the most powerfully negative epithets leveled by conservative economists (usually wrongly) at those who do not obtain their incomes by labor or investment.

The next prominent English-speaking political economists writing in the period following the death of Adam Smith in 1790 were Thomas Robert Malthus and David Ricardo. Surprisingly, neither was particularly interested in the origin of wealth. In his 1798 *First Essay on the Principle of Population* [16], Malthus provided a narrative history of the transition from "savagery" (known today as hunting and gathering) to modern societies. Like Smith he favored the (supposedly) virtuous behavior of the parsimonious wealthy classes over that of the prodigal poor. Unlike Smith, he seldom addressed the issues of capital accumulation in his *Essays on Population*. Malthus directed his analysis as to why populations remained stable in early societies and not to why capital accumulated.

David Ricardo subordinated the question of wealth creation to secondary status. For him the real question was one of distribution, and distribution changed according to the specific historical period. Like Malthus he accepted the division of society into classes of landlords, capitalists, and laborers as natural and inevitable. Ricardo believed that the proportions of the whole produce of the Earth which will be allotted to each of these classes, under the names of rent, profit and wages, will be essentially different in different stages of society, depending mainly on the actual fertility of the soil, on the accumulation of capital and population, and on the skill, ingenuity and instruments employed in agriculture. He said, "To determine the laws which regulate this distribution, is the principle problem in Political Economy." [13].

2.8.4 The Origin of the Concepts of Diminishing Marginal Return and Comparative Advantage

Ricardo and Malthus were writing during the late eighteenth and early nineteenth centuries when there was a great rivalry between landowners and emerging capitalists for control of the British economy and society. English Corn Laws, passed in the early 1800s, limited the import of cheaper grains (corn) from continental Europe. This benefited the landed classes by extending the margin of cultivation to poorer quality lands, most of which they owned. Simultaneously the law increased rents and raised wages, since wages were determined by subsistence and ultimately the costs of extracting an energy surplus from poor land. This limited the power and income of the rival capitalists as most of the wealth of society had to go for the necessary food and hence to landowners. David Ricardo and Thomas Malthus undertook great debates concerning the efficacy of the Corn Laws and their effect upon the economy and society. This debate was the genesis of two of the most sacred principles of modern economics—*diminishing marginal returns* and *mutual gains from trade*, technically known as *comparative advantage*. David Ricardo devoted his life to the pursuit of political economy and the repeal of the Corn Laws by crafting myriad arguments in support of the interests of the emerging class of capitalists. His primary aim was to change the distribution of income and wealth from the less productive landed classes to the more productive capitalists, although he himself was a landowner. Malthus argued for just the opposite, the redistribution of income and wealth toward the landowners.

Ricardo enunciated a theory of rent based on the principle of diminishing marginal returns since the price of food (or "provisions") depended upon the costs of production (primarily labor costs) at the no-rent margin (or the land of lowest fertility). The owners of more fertile lands received a rent, so that food grown on more fertile, and less costly, land would sell at the same price as food that was costlier to produce. Ricardo's theory also depended upon the *best first principle*. Farmers, being no fools, would tend to utilize the most fertile and most accessible land first, and poorer lands second. In other words, returns diminished at the margin of cultivation, i.e., the poorest land that was still put into production to meet total food needs. As we shall see in later chapters, this principle is also useful for explaining peak oil and the falling energy return on investment over time. But in the pre-fossil fuel age, the only thing that stood in the way of the redistribution of incomes toward productive commercial farmers and manufacturers was the cumbersome Corn Laws limiting the import of cheap grains. If these laws were repealed, the cultivation of poorer quality lands could be postponed or eliminated.

Ricardo crafted his arguments in the context of benefits to the nation rather than in terms of benefits to a specific class. He reasoned that free trade among nations in finished commodities would result in more goods for a cheaper price than if each nation produced all that they needed on a self-sufficient basis. He also reasoned that capital and labor would be immobile internationally, a proposition subsequently repudiated by advocates of globalization. (We will return to the details of this argument in ► Chap. 8). Moreover, Ricardo believed that such a redistribution of income would enhance the growth of the domestic economy as vibrant profit-seeking commercial farmers would reinvest their returns in improved techniques (what we would call today technology) that would reduce the overall cost of provisions and thereby improve society in general.

Thomas Malthus held the opposite position. He believed that frugal capitalists would over-save and that savings would not automatically find their way into investment. As a result, the economy would lack the demand needed to realize profits, and the economy would fall into a depression. Malthus' solution was the redistribution of wealth to the landed classes who would use it to build monuments and surround themselves with unproductive retainers, ensuring adequate overall demand. We will save the details of the argument for the next section on the balancing of supply and demand, but it is important for the reader to see that many of today's most important economic arguments were developed by Malthus and especially Ricardo as they contemplated the effects of what we would call today free trade.

John Stuart Mill's 1848 book, *Principles of Political Economy*, dominated the discipline until the 1870s but offered little new in terms of value theory. Indeed, he envisioned his own task as little more than updating Ricardo. Mill did offer a unique perspective, however, on income distribution. Production, according to Mill, was subject to natural law (i.e., the limitations of what we would call today resources), as envisioned by Smith, Ricardo, and the other classical economists. But distribution was entirely a matter of the free will of human beings, and humans could change social institutions to accommodate a more equal distribution. Mill therefore showed concern about Irish peasants, industrial workers, and the position of women and supported a series of reforms to increase their share of social wealth and elevate their status. Influenced by his wife, Harriet Taylor, Mill became a tireless advocate of the emancipation of women at work and in the home. Mill wrote that the time of Adam Smith—where the pursuit of self-interest would lead to social harmony—had come to an end as evidenced by the destitution of the working classes and significant social strife. Like Marx, Mill considered the qualitative aspects of social inequality and the future of society. The good life, for Mill, entailed a simpler and more equal society. "I confess that I am not charmed with the ideal of life held out by those who think that the normal state of human beings is struggling to get on; that the trampling, crushing, elbowing, and stepping on each other's heels, which form the existing type of social life are the most desirable lot of human kind, or any but the disagreeable symptoms of one of the phases of industrial progress" [14]. For Mill industrialization brought greater material prosperity, but it also brought many undesirable and unpleasant aspects to the working class that he was interested in overcoming.

2.8.5 Balancing Supply and Demand

Adam Smith's genius lay in his ability to connect productivity increases made possible by the division of labor to events in the broader market. He believed that the natural price of any commodity could be found by the summation of wages, rents, and profits. Smith, however, also contended that commodities do not always sell at their natural prices. Rather, the short-term forces of supply and demand could result in a price that exceeded, or fell below, the natural price. The market price of any commodity was regulated by the quantity that was brought to the market and the willingness and ability of potential buyers to purchase the products. Smith termed this desire, backed by money, "*effectual demand*." If the quantity brought to market falls short of effectual demand, those individuals seeking to acquire the goods will be willing to offer more money for them. Competition among these individuals will result in an increase in market price above the natural price. If effectual demand is less than the quantity brought forth, then market price may fall below natural price. When the quantity brought to the market just equals the effectual demand, market price will equal natural price.

The Physiocrats had not worked out any theory of supply and demand, although the *Tableau Economique* can be thought of as an early circular flow model. What Smith took away from the physiocrats was a confirmation in his belief in liberty. The market provided a mechanism by which the haggling of daily commerce would result in a tendency toward the balance found in natural law. This is most often known as the "invisible hand," and it is greatly admired by many economists today who resent government (or anyone) telling individuals what they should or should not purchase, for example, in response to concerns about climate change [11]. The other side of the coin is that in the absence of government regulation large, powerful corporations have increasing power to regulate markets and impact individual freedoms.

Jean Baptiste Say argued in his *Treatise on Political Economy* that a market characterized by liberty would adjust automatically to produce an equilibrium in which all resources would be fully employed. Say held that every purchase was simultaneously a sale. No one would sell a commodity without the intent to buy another. Money would not be hoarded because it was simply a means of exchange and had no value unto itself. Because of this, supply creates a demand of equal magnitude. Furthermore, the means of purchase are created, in the form of factor payments (wages, rent, and profit) such that there is no shortage of effectual demand. Therefore, according to the principles of Say's law, a general glut of unsold commodities, and a resulting depression due to lack of demand, is theoretically impossible. Say argued that an acute glut is certainly possible, but a glut in one sector would be matched by excess demand in another. Moreover, price fluctuations as described by Smith would assure that price changes born of competition would assure that market price would equalize with natural price. One could say that Say generated an idealized theoretical situation in which the free market would generate the best of all material worlds; many since have believed that to be true.

Malthus rejected Say's law, arguing that a general glut was a defining characteristic of a commercial economy. The years before the publication of his *Principles of Political Economy* were marked by severe depression. The subsequent riots alerted Malthus to the dangerous destabilizing effects of actually existing general gluts. For Say's system to work, every class must spend its entire income. While this was true of the working classes, Malthus realized that the components of price—wages, rent, and profit—were also the incomes of the wealthier classes in England. He argued that capitalists limited their consumption in order to save. This meant that savings must equal investment. But he found that as capitalism progressed, businesses could not find sufficient outlets in which they could receive profitable returns. As investment declined and savings were maintained, a shortage of effectual demand would appear, heralding the onset of a depression due to lack of demand. The Malthusian solution was, as we have already seen, a redistribution of wealth and income to the landed classes. As gentlemen of leisure, they would spend this income on unproductive personal retainers and monuments to themselves which would, according to Malthus, help maintain full demand. They would also patronize the arts, leading to an improvement in the character of society. Servants and artists would consume the material wealth produced by industry but would not produce it. This would negate the cause of an overall lack of demand. Also, as we mentioned previously, the primary mechanism of income redistribution toward the aristocracy and gentry was the continuation of the Corn Laws.

Ricardo defended Say's law and rejected the Malthusian solution of an expansion of unproductive laborers such as servants and retainers. He said that the support of unproductive personal servants would be as beneficial to future production as fires in the warehouses of the business classes. Ricardo believed that market forces would result in the balancing of savings and investment because of the behavior of investors. "No man produces but with a view to consume or sell, and he never sells but with an intention to purchase some other commodity, which may be immediately useful to him, or which may contribute to further production. By producing, then, he necessarily becomes either the consumer of his own goods, or the purchaser and consumer of the goods of some other person" [26]. Ricardo also criticized Malthus for focusing solely on consumption and failing to consider adequately investment itself as a component of effective demand. Ricardo's argument carried the day. His goal of enhancing accumulation by means of redistribution of income and wealth toward capitalists was finally realized in 1846, 23 years after his death, when Parliament repealed the Corn Laws.

2.8.6 Growth, Accumulation, and the Steady State

For Adam Smith, the process of economic growth began with the frugal saving capitalist and the workings of the "invisible hand." The desire to accumulate, which for Smith is innate in the human spirit, manifests itself as saving and investment. Frugal individuals save, invest the capital in expanding the division of labor and employment, and purchase improved equipment. The expansion of employment leads to rising incomes among all sectors of the population providing the means for the extension of the market. Since Smith wrote in preindustrial days, he did not believe that augmented machinery would replace labor. Rather it would expand its employment. But here lies the beginning of the stationary state. As employment and production expanded, so too would the demand for labor. This would serve to raise wages and diminish profits which would hinder further accumulation in the short term. The solution to rising wages and falling profits could be resolved only by the rather cruel operation of nature. Increased wages would lead to a greater number of surviving children. This would increase the supply of labor and result in the subsequent reduction of wages. But the reduction of wages would eventually decrease the labor supply as infant mortality would increase with less money to purchase food. But while nature would operate to regulate the labor market, the long-term tendency was toward decline. When a nation was fully complemented with people with respect to biophysical capacity to support them, wages would fall to the bare subsistence level. As long as food production depended upon limited natural fertilizers and animate power, agricultural productivity would remain low and wages would tend toward subsistence. When the nation was fully stocked with all that the low level of wages could support, profits would fall as new investment opportunities vanished. Thus, the fate of a vibrant system of perfect liberty was the stationary state. Smith saw this as unfortunate, as the quality of life in the progressive state was vibrant but life in the stationary state was melancholy. Life in the declining state was tragic. But for Smith, no nation was close to achieving its full complement of labor and capital, so the stationary state was a prospect for the distant future [28]. Smith's analysis of accumulation gave economists two methodological lessons that are strong still today. The lack of economic growth was stagnation which was to be avoided at all costs. Moreover, the tragedy of the end of accumulation was found in the distant future. Today economists, politicians, and citizens alike tend to follow Smith's logic. Growth is the primary goal of most economic policy now, and many believe that the environmental consequences of growth will not occur for at least a hundred years.

Less than a decade after Smith's death his optimism, or that of his followers, was dashed. The arrival of the steady state seemed imminent instead of not distant. British philosopher Thomas Carlyle surveyed the debate over the end of accumulation between Thomas Malthus and David Ricardo and dubbed political economy "the dismal science." The primary limit to accumulation for Ricardo was the existence of diminishing marginal returns. Given the existence of the Corn Laws, the extension of cultivation onto poorer lands resulted in reduced harvests and increasing rents accruing to the landowners. The increases in rents and wages would diminish profits, resulting in the cessation of productivity increasing investments as soon as potential profits fell to the prevailing rate of interest. Only a suspension of the Corn Laws could remove the limit to growth.

Malthus saw the primary impediment to long-term accumulation in the increase in human population at a rate that would soon overwhelm the ability to provide sufficient food, resulting in mass starvation. Malthus advocated not only measures to limit population by "courting the return of the plague" but a transfer of wealth to the morally restrained landed classes. But Malthus too saw internal limits to accumulation. Capitalists tended to over-save, thereby limiting effectual demand needed to extend the market and justify the increased level of production. He advocated the redistribution of wealth to the aristocracy who would spend the income on retainers and monuments to themselves, eliminating the shortage of effective demand and perpetuating all that is good in modern society. For both Malthus and Ricardo questions of accumulation ultimately resolved the questions of distribution of wealth.

2.9 Proper Role of the Government

The theory of classical political economy follows directly from the political theory of Enlightenment philosopher John Locke. Locke's basic idea was that the reason for government was the protection of private property and that government works best when it limits itself to this function. The familiar dictum of Thomas Paine, that "the government that governs best governs least," is consistent with this Enlightenment view. It should surprise no one that the two most important Enlightenment documents in the English-speaking world, the American *Declaration of Independence* and Adam Smith's *Wealth of Nations,* spoke directly to the proper role of limited government.

Smith argued that mercantile restrictions, especially the granting of royal charters and high rates of tariffs, favored the large trading corporations, limited competition, and reduced the benefits for the public. Smith clearly spoke of the mercantile monopolists and their government benefactors when he said: "People of the same trade seldom meet together even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices." [16] Smith was optimistic that with a government that supported mercantile monopolies out of the way the "system of perfect liberty" could flower and the pursuit of self-interest could be channeled into social harmony by means of price competition, "as if by an invisible hand."

Contrary to the common wisdom, Smith did not oppose government in all its forms. He simply did not believe that a government representing mercantile monopolists should meddle in the market process and distort the workings of the system. Since the system of natural liberty depended upon the increase in productivity that resulted from the division of labor, Smith was aware of, and sensitive to, the plight of detail worker in the manufactories based upon the division of labor. He thought that repeating the same unvarying tasks would render a worker as "rude and ignorant as it is possible for a human creature to become." He therefore recommended strongly public expenditure on education. In Book V of the *Wealth of Nations,* Smith declared that a sovereign had three fundamental obligations: to provide for the common defense, to maintain an independent system of courts to adjudicate property rights, and to construct public works necessary for the smooth functioning of commerce. Smith also believed that poverty relief was another proper role for the government. According to Robert Heilbroner [17] there were a million and a half paupers out of 12 million people in Smith's Britain. Other than paltry poor relief, a welfare system simply did not exist. Smith believed the expansion of markets could relieve poverty. He called for the repeal of the Poor Laws that tied the poor to their local parishes as a condition of receiving their meager subsistence. Sometimes, he thought, the best thing the government could do was to stay out of the way.

This idea of laissez-faire, "leave us alone," did not originate with Smith. Rather it was the brainchild of the aforementioned Physiocrats. Their program consisted of the removal of onerous taxes like the quitrent (having to pay rent on a property you own), sharecropping, and forced labor. By such means the process of the economic surplus flowing away from the productive classes on the farm and into the lavish court and the luxury workshops could be stopped. To expand French agriculture from small-scale subsistence to large-scale commercial necessitated investment on the farm. Only if the taxes were repealed could these investment funds be generated.

Classical political economists shared some fundamental ideas, despite their differences. All believed that nature, in the form of land, played a role in the creation of value, along with the role played by human labor in transforming the products of nature into sellable commodities. As long as land was a fixed factor of production with a highly concentrated ownership, expanding food production would be difficult. From this observation came the principle of diminishing marginal returns. In addition, the classicals shared the perspective that a fixed quantity of lands, in conjunction with high fertility rates, would reduce wages to a paltry subsistence level. All classical political economists grounded the analysis in terms of social class, and all were focused upon economic policy. The coming of the fossil fuel era would bifurcate economic theory into two distinct approaches by the 1870s, one based upon economic surplus and the other based on relative scarcity. The first political economist to understand the power of fossil hydrocarbons in transforming the productivity of human labor was Karl Marx. Within 3 years of the publication of the first volume of *Capital* [18], the first theories of neoclassical economics would appear.

2.10 Karl Marx

The German philosopher turned political economist Karl Marx was probably the first political economist using the labor theory of value to comprehend fully the industrial revolution and the role that mechanization and fossil energy played in its development. Marx is seen by many in the environmental community as an economic determinist. This comes largely from his treatment of the biophysical world as "a free gift of nature." As we have already seen, this was a customary practice among the most prominent political economists, especially David Ricardo, whose works Marx admired. Another oft-quoted passage comes from an early work *The Poverty of Philosophy*, Marx's critique of utopian socialist Pierre Joseph Proudhon's *The Philosophy of Poverty.* In this book, he said: "the hand-mill gives you society with the feudal lord; the steam-mill society with the industrial capitalist." For Marx, this was no simple mechanical relation but a set of complex dynamics between humans, energy, and machinery. Marx was both fascinated and admiring of the increased output made possible by the application of fossil fuels to production. "The bourgeoise, during its rule of scarce 100 years, has created more massive and more colossal productive forces than have all preceding generations together" [19]. According to Adam Smith, 10 men in his time, using the system of the division of labor, made 48,000 sewing needles every day. A single needle-making machine, however, makes 145,000 needles every hour. One woman or one girl superintends four such machines and so produces nearly 600,000 needles in a day or over 3,000,000 in a week [20]! Marx thought that this was a marvelous means of making labor more productive, and he clearly understood, but did not dwell upon, the role of energy in this process. According to contemporary political economist Andreas Malm steam power engenders and extends the role of the division of labor, transcends strength, skill, and endurance, and allows for substantial increases in labor productivity. In his more mature work, *Capital,* Marx realized that changing machinery and energy led to a different mode of production, which led to changing social relations. Improvements in energy and machinery change the economy by working through the agency of human labor [21].

2.10.1 The Origins of Value and Wealth

Marx began the first volume of *Capital* with a chapter entitled "The Commodity." The basic reality of capitalist society, the commodity, possessed a "two-fold nature." It possessed both use value and value. Commodities were produced for sale, rather than for personal use but could not be sold if they had no use value. This distinction between use value and value was crucially important for Marx, as it was for earlier political economists. While use value was the origin of wealth, exchange value was the sole basis for price or, simply, value. In his later political commentary, *A Critique of the Gotha Programme*, Marx chided other socialists for claiming that labor was the source of all wealth and therefore labor deserves the entire product. Marx's position that wealth, as a use value, also has its origin in nature and that capital also plays a part in its creation [22]. Exchange value or, simply, value depended upon the average amount of socially necessary labor that was embodied in its production, a similar, if more refined, version of the labor theory of value of Ricardo.

This distinction manifests itself as an analysis of circuits. The first Marx called "simple commodity circulation." An independent artisan entered the market possessing title to a commodity. He or she would sell that commodity for money and use the proceeds to purchase another commodity. The goal here was to obtain a different use value of the same value (say 10 hour of labor). Like previous adherents to the labor theory of value, Marx began *Capital* by assuming all commodities exchanged at their value. With the goal met, the circuit self-extinguished, although the owner of another commodity may make another exchange. Money, for Marx, was a medium of exchange. In simple commodity circulation, if C represents a commodity, and M represents money, the circuit can be depicted as C-M-C. The value at the end equals the value at the beginning.

Marx contrasts this with the circuit of capital. Money, in this case, was the object of desire, not just a medium of exchange, or as he called it, "the universal equivalent of commodity values." The capitalist begins with money, buys commodities, and sells them for more money. The additional money is then reinvested and the system becomes self-perpetuating. Unlike most economists, who viewed capital as a thing, Marx saw capital as *a process of self-expanding value!* Schematically this is represented as M-C-M', with M' > M. But how is this possible if all commodities exchange at their value? The answer is found in the types of commodities capitalists buy as capitalists. As wealthy individuals capitalists may purchase expensive transportation, elegant housing, and fancy clothes. But as capitalists they purchase means of production (machines and energy) and labor power. Marx made special efforts to distinguish labor from labor power. Labor power, or work per unit of time, was a commodity with an exchange value. The value of labor power was the cost of reproducing the worker or the subsistence wage, with subsistence defined culturally and historically as the average bundle of wage goods consumed by the working class, not a minimum biological subsistence in terms of calories. Labor power was also the potential to work. Labor was a use value and part of the human essence, as expressed by Marx's collaborator and benefactor Frederick Engels in his essay, "The Part Played by Labour in the Transition from Ape to Man" [23]. Since labor is the essence of humanity, a capitalist does not purchase either labor or the human being. Rather he purchases a worker's ability to work for a specific amount of time. If a capitalist can get a worker to produce more in a day's work than the cost of subsistence, then the extra value produced, or surplus value, accrues to the capitalist. This surplus value is the basis of profit.

In the era before the widespread use of fossil fuels, the only means of increasing surplus value was to either lengthen the working day or increase the intensity of the labor process. Both measures had physical and social limits. Both increasing the time workers remained on the job without increasing wages and implementing harsh supervision provoked absenteeism, high quit rates, political Factory Acts to limit the work day, and many, many strikes. Marx called this method *absolute surplus value*. Although working hours have fallen from the daily average of 12–14 hours in Victorian England, profits have not disappeared. This means another method must have been successful. Marx called this *relative surplus value*. The basic premise of classical political economy was that workers were paid at their value. Reducing wages below the costs of subsistence was not a long-term option for capitalists. However, if capitalists could reduce the costs of wage goods, they could reduce the money wages of workers but maintain their real wage, which was the value of labor power. Mass production, powered by fossil fuels, accomplished this goal. Moreover, mechanization augmented the possibilities of intensifying the labor process. Coal-driven steam engines could provide continuous power, and they and steam-powered machines could be run faster than machines driven by other sources.

Marx's analysis was qualitative as well as quantitative and focused upon the quality of work life as well as wages. Economists who focused only on the quantitative aspects of lower prices and higher productivity overlooked the changes in the process of labor. Marx's critique of the existing political economy was grounded in terms of both qualitative and quantitative approaches to value. He believed that qualitative relations among people undergird the quantitative relations between people and things. The accumulation of capital depended upon the extraction of surplus value from immediate producers (i.e., workers), and the profit rate depended upon increasing the rate of surplus value or labor productivity. To accomplish this, the character of work became stripped of its meaning. The mental work was separated from the manual work, first by organizational means such as the division of labor and later by the application of fossil fuels to machinery. These changes had many social impacts. The worker became an appendage to the machine, no longer directing its application for the improved quality of the product, but rather the worker had to follow the dictates and pacing of the machine. The intellectual unity of head and hand was severed for all but a very few workers whose skills were sufficiently unique such that they could not easily be replaced by machines. The resulting alienation that the worker felt from the products and processes of production would drive social change. Marx believed it was likely that wages could rise with economic growth but that the changes in production and the degradation of the labor process could not be overcome with more money. This qualitative aspect formed a crucial part of Marx's theory of income distribution and inequality and the inevitability of social revolution.

2.10.2 Supply and Demand

Marx chided Ricardo for defending the automatic balance between supply and demand ("the childish babble of a Say, but hardly worthy of the Great Ricardo"). Marx argued that Say's law was applicable only to the stage of simple commodity circulation where an independent artisan enters the market with a commodity and sells it for money to purchase a different commodity. It was not applicable to an industrial capitalist society. The possibility of such an equilibrium occurring in a simple economy did not imply its inevitability in a modern one. Marx's writings on the balance of aggregate supply and demand in a modern economy can be found in the little-read Volume II of *Capital*, where Marx discussed the process of exchange. Here Marx begins with the abstract and highly unlikely possibility of a nongrowing capitalist economy, where the entire surplus value is consumed and the economy goes on year after year at the same level and composition of output. He calls this "simple reproduction," as opposed to a growing economy that he terms "extended reproduction." To begin the analysis, Marx divides the economy into two sectors or "departments." Department I produces means of production, known today as the capital goods industry. Department II produces means of consumption. In both sectors, the total value (V) is composed of the sum of constant capital, variable capital, and surplus value. Equilibrium necessitates that the output of these two sectors is balanced [24].

In plain English, Marx believed that the combined demand of workers and capitalists in the department producing capital goods had to balance the demand for capital goods in the consumption goods sector. The formula for this is $c_2 = v_1 + s_1$, where c stands for constant capital or means of production, v stands for variable capital or the money advanced for wages, and s equals surplus value. This is highly unlikely and highly abstract. It is a mathematical equilibrium condition. The reason for the low likelihood is that capitalism is a dynamic system of self-expanding value. The driving force of capitalist competition is technological change to increase labor productivity. Capitalists simultaneously restrict their own consumption while paying workers no more than the value of the subsistence wage to accumulate capital. Therefore, there is no reason to believe that this abstract equilibrium condition will occur in an actual economy. If the conditions of simple reproduction are not met in an actual economy, crises can occur for a variety of reasons. The pace of technological change may result in a capital-labor ratio that increases faster than does labor productivity, precipitating a tendency for the rate of profit to fall. Slowly growing wages and technological unemployment may lead to insufficient effectual demand, and disproportionalities may develop as the capital goods and consumption goods sectors grow at different rates. For Marx, sectoral imbalances are the norm, while the possibility of a balance in aggregate supply and demand is but a highly unlikely theoretical possibility that contradicts the very essence of capitalism [28].

In the first volume of his 1867 opus, *Capital*, Marx turned to the accumulation that occurred prior to the emergence of industrial capitalism [13]. His chapters on "the so-called primitive accumulation" chronicle the process by which former artisan producers and independent farmers—even before the evolution of industrial capitalism—were forcibly "stripped of the means of production" by those with more financial or political power and left with only their labor power to sell. Furthermore, Marx analyzes the effects of mercantile strategies where fortunes were built on colonization, slave labor, and war. Unlike Smith, who attributes the origins of wealth and capital to the virtuous behavior of the frugal saver, Marx declares "If money…comes into the world with a blood-stain on its cheek, capital comes dripping from head to toe, from every pore, with blood and dirt" [25]. Thus, Marx added, or continued to add, a moral dimension to how economies worked under different systems.

Accumulation

Marx did not have a theory of the stationary state. Unlike his classical predecessors, Marx wrote in an era of fossil fuels where a fixed supply of land was no longer the limiting factor. Rather, he believed the internal contradictions of the capitalist system could result in its passage into socialism before the physical basis of the end of accumulation arrived. For Marx only human labor created new value, although it was augmented to an unprecedented extent by the application of coal to large scale mechanization. Such mechanization reduced the per unit labor content of commodities resulting in the reduction of their prices. Capitalists competed by means of mechanizing to reduce the price of their individual commodities below the social average. But as the expansion of constant capital increased faster than the increase in productivity, profits would fall. This touched off an economic crisis, which could not, in the long run, be overcome by the mere addition of more fossil fuel-driven equipment. Marx termed the tendency for the rate of profit to fall a "law of motion of the capitalist system." A second law of motion was the tendency toward monopolization, as during the crisis better capitalized and better managed companies would acquire their less fortunate rivals, creating bigger operations that were owned by fewer capitalists. The resulting depression "solved" the tendency for the rate of profit to fall by decreasing the level of capital to labor, as bad debts were written off and factories shuttered, as well as by increasing the productivity of labor when desperate workers would work harder for less. Before the stationary state set in the increasing severity of periodic crises and a socialist political party would transform society by instituting rational planning into the investment process resulting in an end to economic crises and the true beginning of human history.

2.10.3 Marx and the State

Human history has not actually worked out as Marx envisioned. His vision of socialism was one where workers would use the state to humanize the labor process and to distribute incomes more evenly. It was a system, unlike capitalism, that was not crisis prone and crisis dependent. Growth and accumulation would serve the needs of the populace, rather than being the *sine qua non* of the system. Communism would arrive when the state was no longer needed and workers could manage the economy by themselves. Socialism and communism in the real world tended to be characterized by strong rather than by withering states, and worker alienation remained high. After the fall of the Soviet Union in the late 1980s and the transformation of the People's Republic of China into highly centralized state capitalism, few examples of socialism in the real world still exist. The prospects for the future, whether some form of democratic socialism could still emerge, remain unknown at this point. But the future of capitalism is also unknown. In ▶ Chap. 23 we will discuss a series of planetary boundaries and biophysical limits, some of which we have already exceeded. We do not know how a system in overshoot can grow its way into sustainability. Neither do we know how a nongrowing capitalism can exist in the absence of stagnation and high unemployment.

Marx on Money

As mentioned previously in the analysis of circuits, money took different forms for Marx. It could be a simple means of exchange or it could be money capital. This money capital could be used to purchase means of production (constant capital) or labor power (variable capital). Surplus value was the basis of profit and accounted for in monetary units. Like his classical predecessors, Marx wrote in an era of commodity money or money that was backed up by a precious metal. This meant the amount of money could not be expanded at will, as is the case today. However, Marx was also aware of the extension of credit and that in a time of economic crisis, financial factors themselves could exacerbate the crises caused by a tendency for the rate of profit to fall.

The Metabolic Rift

By the third volume of *Capital*, Marx was deeply concerned about the fate of the Earth, arguing that capitalism systematically undermines the material conditions of its very existence: human labor and the soil. He was profoundly influenced by Justus von Liebig, telling Engels that the work of the agricultural chemists was more valuable than that of the political economists. In his chapters on ground rent, Marx tried to incorporate new understandings of energy and entropy. Ricardo based his principle of diminishing marginal return on the "original and indestructible powers of the soil." Through his careful study of Liebig, Marx realized that the powers of the soil are not indestructible. Rather, large-scale commercial farming (British high agriculture) according to Liebig was a "generalized system of robbery." Nutrients would be shipped from the rural agricultural districts in the form of food and not returned to the soil. Unfortunately, because matter and energy are not destroyed, these missing nutrients, which we now know to be nitrogen and phosphorous, emerged as pollution in large cities such as London. We will see in ▶ Chap. 23 on planetary boundaries that the disruption of such biogeochemical cycles remains a problem in the modern day. The appropriation of the land by large-scale agricultural monopolies created a metabolic rift between humans and nature, and the abolition of these monopolies would be essential to create the kind of society we now call sustainable [10].

2.11 The Origin of Neoclassical Economics

This period of classical economics lasted through the early 1870s. Then the discipline underwent a profound transformation in questions of value, production, and distribution. This shift in emphasis and analysis led soon to the emergence of *neoclassical* economics, based on the concept, or perhaps faith, that mechanical details of the market economies are based on the "invisible hand" of Adam Smith. Furthermore, that markets are self-regulating by means of competition and flexible prices and that these could be well represented by analytical models borrowed from physics. The originators of this idea came from the French-Swiss Léon Walras, Englishman Stanley Jevons, and Austrian Carl Menger and focused much less on production and much more on "marginal value," that is that the value of something became less the more of it you had. Neoclassical thought derived from this "marginal revolution" was fully synthesized by the early years of the twentieth century and remained the primary mode of thought until the Great Depression of the 1930s. Then system-wide economic collapse rendered the prevailing orthodoxy incapable of understanding the depth of economic decline or formulating policies to improve it. In this climate of dislocation, the theory advanced by the British economist John Maynard Keynes provided an alternative that soon dominated the profession.

Keynes visions and methods are in sharp contrast with neoclassical economics, which was enunciated in the 1870s and continually refined until the present day. The neoclassicals were interested in the development of universally applicable theory, modeled after physics and independent of historical context. Nobel Prize winning economist Robert Solow stated this clearly if somewhat tongue in cheek:

> My impression is that the best and brightest in the profession proceed as if economics is the physics of society. There is a single universal model of the world. It needs only to be applied. You could drop a modern economist from a time machine—a helicopter maybe, like the one that drops money—at any time in any place, along with his or her personal computer; he or she could set up business without even bothering to ask what time and what place [26].

British economist G.L.S. Shackle stated that the principle around which neoclassical economics was organized, the principle that served as the equivalent of gravity in celestial mechanics, was self-interest [7]. But neoclassical economists focused not on the pursuit of self-interest, as did Smith and the classical school, but upon the *maximization* of personal self-interest through the mechanism of people buying what they want in markets. Their approach was mathematical and abstract and based upon relative scarcity as a universal principle. In short, neoclassical economics was the marriage of differential calculus with utilitarian philosophy. The classical focus on social class as the unit of analysis was replaced with that of the individual, and the role played by accumulation gave way to a stress upon static equilibrium and allocative efficiency. A neoclassical analysis of growth was not to appear until the 1950s when it was enunciated by Robert Solow.

2.11.1 Value and Wealth

Perhaps the greatest break with classical political economy came in the area of *value theory*. Classical political economists all commenced their analyses from the viewpoint that value and wealth were created in the process of production and that value could be calculated *objectively* from the costs of production. Neoclassical economics was, and continues to be, grounded in the proposition that value, like beauty, is in the eye of its beholder—that is a matter of *subjective* well-being or utility. Their overall objective was not to pursue the origins of wealth as much as to show, under ideal theoretical conditions, that market economies are self-regulating by means of small, or marginal, fluctuations in prices driven by competition on the individual level. The result of voluntary trades, based solely on the maximization of self-interest, leads us to a situation of Pareto efficiency (named after its originator, Vilfredo Pareto) where no one individual can be made better off without making another worse off. this state is called *Pareto Efficiency*. According to neoclassical doctrine, government intervention could do no good, and much harm, as it would distort the signals of the market, which is seen as a perfect carrier of information [27].

An important problem facing economists in 1870 was what is often called the "diamonds vs. water" paradox. Water was, and remains, essential for human life. But since it was abundant and often available for the taking in rural areas, it did not command a high price. In the parlance of classical political economy, water had great use value but little exchange value. Diamonds, on the other hand, had little use value, except as ornaments, but a very high exchange value. Classical political economists would attribute this to the great amount of human labor that had to be expended in mining the stones, cutting them, and polishing them for the market. Water, on the other hand, took little labor to harvest from the ground.

The newly evolving neoclassical economists saw this as a "paradox." But from our perspective, the reason was not some fundamental problem with the classical view but was because the neoclassicists did not separate use value from exchange value. Unlike classical economists, who saw exchange value as independent from use value, the early neoclassical economists viewed use value, now called *utility*, as the source of exchange value. Thus, the relative prices of water and diamonds now became a paradox to them because how could something so useful be so cheap, while something of little use, like diamonds, command such a high price? Their resolution was to make exchange value *subjective*. Diamonds were costly because people liked them, they were not especially abundant, and people were willing to pay a lot of money for them. Water was mundane but abundant. Scarce commodities carried a higher price.

The change that neoclassical economics brought to this problem was a change in the conception of value. Classical economics believed that humans generated value objectively by transforming the products of nature into things humans wanted through the actions of labor. Neoclassical economists, on the other hand, thought that the origin of value was subjective. Value was determined by human preferences, and these preferences were revealed by what humans chose to purchase in the marketplace. This was, at least in theory, a very democratic process in that any consumer is as important as any other in that his or her purchases will "send a market signal" to the whole economy about what that economy should be producing.

This subjective approach became the jumping-off point for neoclassical economics, whose practitioners began a "new" approach to economics in the early 1870s—and still dominate the profession today. They differed from the classical economists in not being particularly interested in the origin of wealth, other than to agree with Smith that the origins of wealth could be traced to the virtuous behavior of individuals. By and large they accepted the idea that wealth was a stock. From the beginning Swiss economist Léon Walras, one of the originators of neoclassical economics, saw the study of economics as the transformation of stocks of natural resources into human-satisfying utilities, with production relegated to a rather irrelevant intermediate position [28]. Thus, neoclassical economists changed the focus of the discussion from an *objective* theory grounded in economic surplus and the (labor) costs of production to a *subjective* utility grounded in psychological scarcity which ultimately was translated into willingness to pay. To create the core of neoclassical economics, this idea was married to utilitarian philosophy, based on the propositions that individuals rationally endeavor to increase their happiness, also married to differential calculus as well. If a commodity provided utility (greater happiness), more of that commodity would provide more total utility.

The focus of early neoclassical thought was also upon *marginal utility* or the extra utility received from consuming one more unit of the good. Neoclassical economists believed that it was marginal utility, also known as the final degree of utility, or rareté, that determined value or price. Marginal utility declines as more of a commodity is consumed. Thus, the first liter of water that might be consumed would have a nearly infinite value, and each subsequent liter was less valuable to the subjective tastes of the consumer. Since water was abundant, it was not worth too much. Theoretical "rational consumers" were thought to continue to trade with one another until the marginal utilities of the two traders equalize. At that point neither party will benefit from additional trading. No individual consumer can be made better off by trading without making another worse off. This is the genesis of what is called *Pareto efficiency*. The reader should note the irony that although the neoclassical concept of value is based on "economic scarcity," this is only *relative*, *not absolute* scarcity. Even though industrialization made possible an abundance of goods, neoclassical economists spoke about scarcity only from the perspective of an individual's infinite wants.

The theory of neoclassical economics assumes that in a money economy, consumers will continue to purchase a "set" of two or more "commodity bundles" even though they have less and less additional value to them. Therefore, they experienced *diminishing marginal utility*. The consumer will cease buying when the ratio of marginal utilities equals the ratio of prices, resulting in "consumer equilibrium." Graphically this

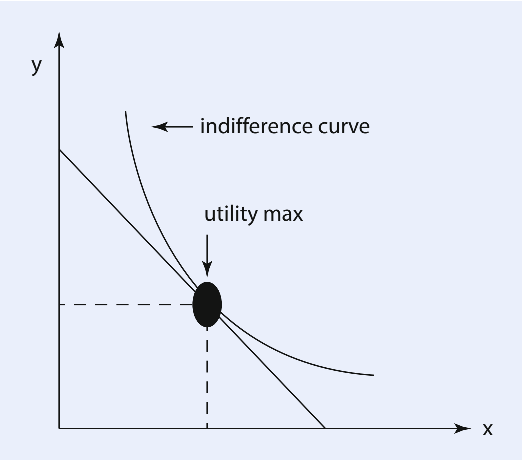


Fig. 2.1 Consumer equilibrium is reached when the slope of the indifference curve (or the marginal rate of substitution) just equals the slope of the budget constraint. At this point the consumer values the tradeoff at the same rate as does the market

can be depicted in Fig. 2.1. Utility is constant along any indifference curve, labeled U_0. Its slope is the ratio of marginal utilities. A budget constraint, B_0, exhibits a slope that is equal to an exogenous price ratio.

In other words, when a consumer trades good x for good y at the same rate that the market trades them off, she or he will be in the best possible position. As prices change, so, too, does the equilibrium position, with lower prices generally resulting in higher quantities purchased. Although the initial assumptions require that interpersonal utilities cannot be compared, they can be aggregated mathematically. The standard "rite of passage" for every student of intermediate microeconomics is to decompose these changes into income and price effects and derive a downward sloping demand curve, despite the complete unreality of the assumption.

For a consumer-based price theory to replace the classical value theory based on costs of production and social classes, let alone come to dominate economic thinking, a reasonably large cohort of consumers must exist. This consumer class was first created by means of low food prices, enabled by the application of fossil fuels to economic production. The industrialization of agriculture began to drive food prices down by the early 1830s, and the increase in productivity made possible by the application of coal to machinery drove down the price of wage goods sold to workers. Moreover, mechanization was accompanied by an increase in the ranks of supervisory employees who enlarged a nascent middle class whose incomes allowed the expansion of consumption and the expansion of the market [29].

By the late years of the nineteenth century, neoclassical economists expanded their early marginalist roots by extending the marginal utility approach to the analysis of production. They believed that production functions mirrored utility functions and that efficient production begot equilibrium in utility as well. Factor price ratios (such as the ratio of wages to profits) were substituted in their equations for the price ratios of utility theory, while ratios of marginal productivities, or the change in output with respect to the addition of one more factor, took the place of ratios of marginal utilities. Producer equilibrium occurs when the two ratios are equal. Moreover, the theoretical distinction between production and distribution found in classical political economy simply vanished. The theory of production and the theory of distribution are one and the same in neoclassical economics. The neoclassical theory of production does not deal explicitly with energy, but the very functions themselves are built upon pre-thermodynamic energetics [30]. The typical production function is simplified to include only capital and labor as the independent variables that produce output.

2.11.2 The Marginal Productivity Theory of Distribution

The neoclassical vision of distribution could not have been more different from that of Mill. Rather than separating the mechanisms undergirding production and distribution, as Mill had done, the neoclassical theories of production and distribution are virtually identical. For 20 years, following the 1870s marginal revolution, neoclassical economics, based on scarcity and utility, was solely a theory of demand. Production was still based on classical principles of cost. But classical theory utilized an economic surplus approach, which entailed the possibility of exploitation—value created by one class is appropriated by another. Marginalism would become neoclassical economics only when production was placed on a marginal utility basis, and the possibility of exploitation was eliminated (at least in theory).

The fundamental idea is that each factor of production [land (T), labor (L), and capital (K)] earns its marginal product (or incremental contribution to total output), no more and no less, as rational individuals follow the price signals of the market. The result is equitable—one's reward depends solely upon one's contribution to society. The marginal product of labor therefore equals the wage rate (w), profits (π) are equated with the marginal product of capital, and rents (r) are determined by the marginal product of land. This can be added up to generate the total output (P), with MP_L being the marginal contribution of labor, capital, and land.

$$P = MP_L \bullet L + MP_K \bullet K + MP_T \bullet T.$$

- P = Total output.
- MP_L = Marginal product of labor.
- L = Labor.
- MP_K = Marginal product of capital.
- K = Capital.
- MP_T = Marginal product of land.
- T = Land.

Unfortunately, this equation can be true only under a limited set of mathematical conditions. British economist John Hobson showed that if marginal product of labor exceeded the average product (or the output elasticity is positive), the product of $MP_L \bullet L$ can exceed the total output to be distributed. But this is possible only if one or more of the factors (e.g., labor, capital) are not paid their marginal contributions. Economists in the neoclassical tradition arrived at some elegant solutions in the subsequent years. However, they depended upon two conditions being met. Equations of degree one are linear, either because they have no exponents, or because the exponents add to one, as in a Cobb-Douglas production function. The function had to be linear homogeneous, and of degree one, and equal to zero. Furthermore, production had to exhibit constant returns to scale [49].

Constant returns do exist when output expands proportionately with the increase in all inputs. In 1928 mathematician Charles Cobb and economist Paul Douglas published an article on long-term trends in income distribution. They were most interested in why the distribution of income remained stable, despite momentous changes in industrial structure and the position of the United States in the world economy. However, the paper is most famous for the *Cobb Douglas production function*.

$$Q = aK^{\alpha} L^{1-\alpha}$$

This equation says that the quantity of production (Q) equals the product of capital (K) and labor (L). The Greek letter α represents capital's share of the income distribution, while the remainder ($1 - \alpha$) was labor's share. Cobb and Douglas estimated capital's share to be 25% of national income, with labor receiving 75%. The fact that the two exponents added to one assured constant returns to scale and the substitutability of resources. Land, which symbolizes all natural resources and which had been used in most previous assessments, was simply left out of the equation, as was energy. Both were subsumed, inappropriately, under the category of capital, as capital, as a productive asset, is essentially useless without energy. But if all inputs are substitutes, the theory implies that society can maintain, and even increase, its level of output in the virtual absence of resources or energy, even were these included explicitly. This failure of neoclassical economics to include energy in their basic equations of production has bothered many biophysical scientists greatly, including Nobel prize winning chemist Frederick Soddy, anthropologist Leslie White, ecologist Howard Odum and his students Robert Costanza and Charles Hall, physics trained economist Phillip Mirowski, and other economists including Nicolas Georgescu-Roegen. Nearly a century after the formulation of these neoclassical equations, Cleveland et al. [31] and Reiner Kümmel [32] showed that 90% of productivity increases can be attributed to increases in net energy, that the productivity of labor is principally determined by the energy used to subsidize labors' muscles, and that capital is important because it is the means of using energy. More explicitly when energy is inserted into Cobb Douglass type functions, it is a far more important determinant of changes in production than is either capital or labor. Why this basic and empirically incontestable concept has escaped incorporation into general economic thinking is astonishing to us and to the distinguished scientists mentioned above.

The marginal productivity theory can be shown, mathematically at least, to produce equity, or fairness, but only under conditions known as *perfect* *competition*. As seen in ▶ Chap. 1, this hypothetical market structure entails creating an abstract model in which equally powerless firms meet perfectly rational consumers in an impersonal market. In addition, firms must be willing to accept zero economic profit in long-term equilibrium. In this model entrepreneurs earn only a "normal" profit, which is what they could earn in wages working for someone else. In 1934 economist Joan Robinson demonstrated that such outcomes are equitable *only* under conditions of perfect competition. In perfect competition workers are paid what they are worth. The pay of a worker is a combination of their individual productivity and the value of the extra production that the firm sells. Technically, this is known as the value of the marginal product, or MRP. It is the production of the marginal product of labor ($MP_L = \Delta Q/\Delta L$) time marginal revenue ($MR = \Delta TR/\Delta Q$), where total revenue (TR) equals all the money the firm brings in by selling its products. $TR = P \times Q$. The Value of the Marginal Physical Product is marginal Revenue × Price. Since in perfect competition, and only in perfect competition, marginal revenue equals price, the marginal revenue product and the value of the marginal physical product are identical. However, in imperfect competition, with any degree of control over price on the part of the firm, marginal revenue is less than price. This means the marginal revenue product (what workers are paid) is less than the value of the marginal physical product (what workers are worth.) Mrs. Robinson referred to this as exploitation. She and we believe this to be the normal, not exceptional, situation [33].

In summary, neoclassical economists built a mathematically elegant structure establishing a nonexploitation theory of distribution. The functions that explain distribution are the same as those that describe production. The two theories are indistinguishable. However, the theory depended upon structures that do not occur in the real world: perfect competition, unlimited and reversible input substitution, and constant returns to scale. In addition, they do not give energy any special role—it is just another commodity. Nonetheless students of economics are trained routinely and often exclusively on such models of perfect competition. It is the only market structure that has been conceptualized in which distribution is equitable and exploitation cannot exist, but it is contradictory to the reality in which humans operate.

2.12 What Most Economists Missed: The Impact of Industrialization

The production and accumulation of wealth have been a central issue of economics since its earliest days, but the concept that energy is a critical factor in that production was (and is) generally treated only peripherally, if at all. The physiocrats understood that land was the origin of wealth, but they had little or no explicit understanding of land as the way that solar energy was captured and turned into things of economic value such as crops or wood, often from photosynthesis of nutrient-starved plants on long-depleted soils. Not surprisingly, most early economics focused upon understanding and explaining the primacy of land in overall production. But since resources took substantial amounts of human labor to extract and the rewards were distributed unequally, they rightfully thought labor was important. Malthus thought that the meager agriculture of his time would limit human populations to something like what were present in his day.

But once humans discovered coal, and later oil, our ability to do economic work, including agricultural production, soared. The energy density found in these new resources led to the rapid transformation of the human condition. Population, which had barely grown for a thousand years, reached one billion in the early 1800s and has soared to nearly seven and one-half billion by 2017. A very few economists, such as William Stanley Jevons and Karl Marx, did address energy explicitly. Jevons found in 1860 that "all economic activity leads back to coal." Marx understood that large-scale economic production was not possible without coal and emphasized the role of coal and machinery in increasing labor productivity. We know now that energy is central to all economic issues and is likely to have serious influence upon, and even limit, the economist's usual goal of economic growth. But even Marx and Jevons mentioned energy only peripherally in their most important writings [34, 35].

As time went on humans constructed an economic infrastructure of factories, refineries, bridges, automobiles, suburban homes, and shopping malls and could now exert a greater degree of control over nature than at any time in the past. However, working conditions for those laboring in nineteenth century textile factories were often horrid and degrading, as are the contemporary conditions for most textile workers in Africa, Asia, and Latin America. The prosperity of the fossil fuel era has not been equally visited upon the world's diverse population. Yet the economic situation, in terms of access to material goods, for most people has not been better than today [36]. This is largely because each human can generate vastly more wealth per unit time than in the past because of the subsidy by fossil fuel.

Thus, humans have increased their ability to acquire and accumulate resources through the use of fossil fuel. Although we have been trained to think about the economy as something run by money, from our perspective, money is just our means of keeping track of debt, facilitating exchange and serving as a lien on the acquisition of surpluses of energy and labor. The fossil fuel-based economy has given each of us in the industrialized countries the equivalent of 60–80 of what futurist Buckminster Fuller called "energy slaves," and the more money you have, the more energy servants you can have. Why economists mostly missed the importance of the industrial revolution as they developed their theories is rather a mystery.

2.12.1 Supply and Demand

Most of what is taught as introductory microeconomic theory in English-speaking colleges and universities today is but an updated version of the neoclassical theory enunciated by Alfred Marshall in his 1890 *Principles of Economics* [37]. Marshall was among the first to link marginal utility with demand, and he aggregated market demand curves from individual ones. By linking supply and demand, Marshall reasoned that equilibrium in the labor market occurred when individuals decided to supply hours to the market up to the point where the marginal utility of the wage equaled the marginal disutility of the work. This unrealistic point, although rejected by Keynes, still forms the theoretical core of modern labor economics.

Marshall also based his analysis on the substitutability of resources. Consumers will substitute one good for another based on the ratio of extra utility to the price that must be paid. Rational consumers substitute the relatively cheaper good for the more expensive one, as long as happiness or utility remains nearly constant, and substitution ends when the ratio of marginal utility to price is the same for all commodities considered. This is known as the equimarginal principle. Marshall's mode of analysis applied equally to the firm as it did to the consumer. The theory of the firm began with the "representative firm" that exhibited no marketing, energy, or technological advantage over any other. He divided his analysis into periods. The short period was one in which one factor (capital) was fixed but labor would be allowed to vary. This period was ruled by diminishing marginal productivity. If a firm applies increasing quantities of a variable input to a fixed input, eventually the rate of increase in output begins to decline. Ricardo first enunciated this idea in his debate with Malthus, but Marshall formalized it.

The onset of diminishing marginal returns implied an increase in marginal cost. If each additional laborer produces less output then, once diminishing returns have set in, a firm would need to hire additional workers, at additional costs, to produce the same incremental increase in output. The marginal cost curve, above the minimum point of average variable cost, becomes the supply curve. This became the basis of profit maximization for the individual firm. Profits would be optimized at the point where marginal cost equals marginal revenue or the extra income derived from selling an additional product. Since the marginal cost curve was equivalent to supply and marginal revenue could be equated with demand, this point also represents the intersection of supply and demand. Any profits in excess of the normal rate, which Marshall termed "quasi rents," would be eliminated by price competition among firms (Fig. 2.2).

In Marshall's long period, all factors of production are variable. Consequently, diminishing marginal returns, which require the application of variable inputs to fixed inputs, cannot operate. Long period costs were regulated by economies of scale. Traditionally classical political economists had posited that capitalists would add fixed and circulating capital up to the point of constant returns to scale, where output expanded proportionately with the application of inputs, and the investment of additional resources would yield no more than the value of the investment. But Marshall saw no a priori reason to assume constant returns. In the era where land played the

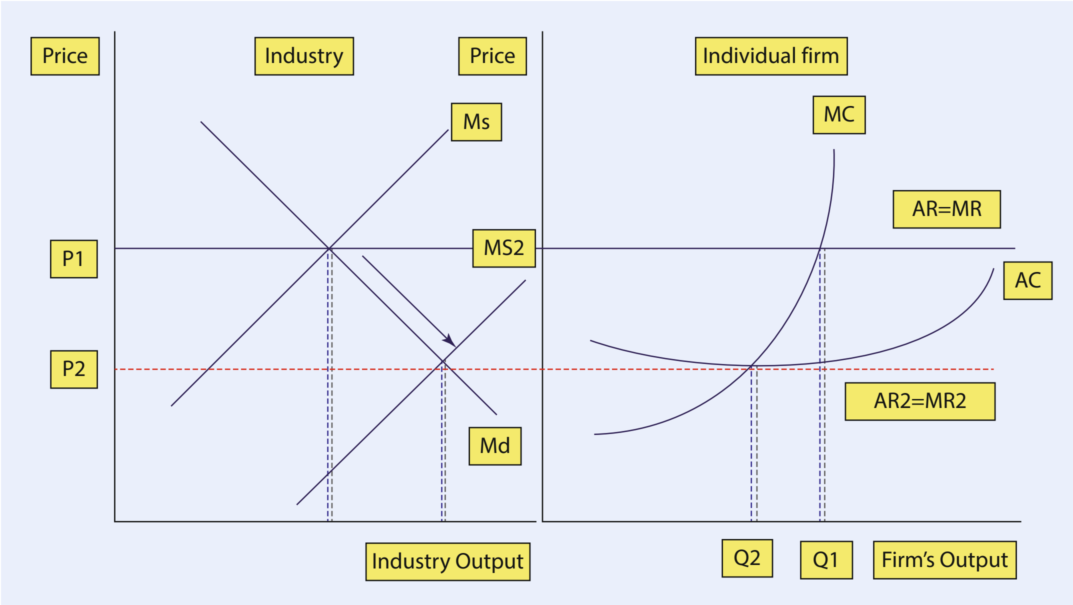


Fig. 2.2 Any profits in excess of the normal rate, which Marshall termed "quasi rents," would be eliminated by price competition among firms

primary role in production, Marshall, following Ricardo, believed there was a tendency toward decreasing returns to scale. But in the era in which the restrictive role of nature was diminished, and the application of fossil energy could increase productivity dramatically, the tendency was toward increasing returns [29]. This rendered a parabolic long-run average cost curve where constant returns to scale represented the minimum achievable cost.

In Marshall's neoclassical synthesis, the market will self-regulate to generate long-period equilibrium where marginal revenue = marginal cost = price = the minimum short-period average cos = long-period average cost. At this point profits are forced to the "normal" level, and the outcome is *allocative efficiency*. Allocative efficiency occurs when the market price covers fully all underlying incremental costs and resources flow to their most lucrative use. Firms unable to achieve constant returns to scale can produce only at above average cost and will be forced into bankruptcy by price competition. Since the firm-level supply can be aggregated into market supply and market demand is simply the summation of individual demands, the market-level balance of supply and demand is the most efficient allocation of resources. The idea that markets allocate efficiently is a deeply held belief of almost all economists, including most ecological economists.

By the 1920s, supply and demand analysis had been extended to describe the workings of the primary sectors of the overall economy. According to Marshall the supply of labor, set by the disutility of the work, would come into balance with the demand for labor, which was determined by marginal productivity by means of subtle adjustments in the price of labor, or the wage rate. If wages were below the equilibrium rate, shortages would occur, causing competing employers to offer a higher wage in order to attract workers. Unemployment was a surplus of labor, caused by workers demanding wages in excess of equilibrium. The solution for unemployment was therefore a reduction of wages. In many ways, the neoclassical or market model provided logical rationales for management to pay labor as little as possible.

Economist Knut Wicksell offered an analysis of the market for savings and investment, called loanable funds, based on the idea of the self-equilibrating market. Savings were specified as a positively sloped function of the interest rate (the price of money). Those with enough income to save would be induced to augment their savings by an increase in interest rates. Investment was negatively related to interest. At higher interest rates, the costs of borrowing rose, and less profitable investment projects would be curtailed. The market would find its own equilibrium interest rate, and savings would equal investment.

2.12.2 The Neoclassical Perspective of Indefinite Growth

Neoclassical economists held a very different opinion on the future of economies. The potential of continued dramatic increase in productivity (made possible by fossil fuels, although that was not mentioned) relegated questions of accumulation and growth to secondary status. Consumption was limited only by a budget constraint. However rational consumers would maximize their well-being by substituting cheaper for more expensive goods, so that consumption could increase indefinitely. A similar process worked on the production side. Initially the optimal situation was the point at which supply balanced with demand. While the possibility that that reinvested profits could lead to economic growth was considered, the focus was clearly on static equilibrium. Only later, in the profound depression of the 1930s, did a neoclassical theory of growth begin to emerge. Sir John Hicks developed the idea of the *elasticity of substitution*—which meant in practice that expensive, unreliable labor could be substituted by cheaper, more reliable capital. He believed that a progressive society necessitated a positive elasticity. In other words, the price of progress was the redistribution of wealth from labor to capital. This would allow growth to continue indefinitely.

One conspicuous exception exists in the work of William Stanley Jevons. Before Jevons solidified his reputation as a marginalist, he produced the previously mentioned empirical work, *The Coal Question*. Jevons theory was based on that of Malthus, but he argued the limiting factor had switched from corn to coal. He had no particular interest in sustainability or resource conservation. Rather he wanted to maintain England's industrial and imperial domination of the world. These depended upon the development of mass production industry, especially textile manufacturing, and industry depended upon an adequate supply of cheap coal [10]. But Jevons believed there was no prospect for a reliable and cheap substitute for coal and that England's mines were slowly becoming exhausted. This would render much of England's population superfluous (and perhaps incapable of being fed) and essentially create the conditions for the return of the stationary state. While Jevons offered no satisfactory solutions, his essay represents the initial exercise of the economic consequences of the absolute scarcity in the age of fossil fuels [29]. For Jevons, England's greatness depended upon lavish use of a declining resource, and he looked toward the future with trepidation. "We have to make the momentous choice between brief but true greatness and longer continued mediocrity" [34]. Many today still fear this prospect, and this will make the transition to living within nature's limits all the more difficult.

Jevon's Paradox, today called the rebound effect, shows that an increase in resource efficiency increases resource use. If you recall the supply and demand diagrams from ▶ Chap. 1, the rebound effect becomes less paradoxical. Increased resource efficiency increases supply and relative to stable demand, drives down prices. The lower prices increase quantity demanded and resource use. Savery's inefficient steam engine used very little coal because almost nobody could afford to use it. Watt's engine led to the expansion of coal use because the engine was efficient enough to compete with, and eventually dominate, water power.

2.12.3 Accumulation and Growth

Neoclassical growth theory emerged as a critique of the Keynesian economists Roy Harrod and Evsey Domar, who proposed separately that the growth path of a capitalist economy would be unstable because of the system's internal dynamics. We will review their work in detail in the next section. In 1956, Robert Solow argued that the flaw in the Harrod-Domar approach was in the way they specified their equations. According to Solow, the Harrod-Domar model used fixed proportions between labor and capital. When he replaced these fixed coefficients with a Cobb-Douglas function the instability disappeared and the functioning of markets would lead to stable growth trajectories. Solow managed to turn a social problem into a technical one and maintained the neoclassical ideal of self-regulating markets over the long term [37].

Unfortunately, Solow's model suffered from a large unexplained residual. As we stated previously, Reiner Kummel explained the residual satisfactorily by adding energy to the production function. Solow's explanation was that the residual was due to technological change that could increase output without increasing the quantities of labor and capital [38]. In this approach, technological change was exogenous, appearing as "manna from heaven" rather than being determined within the parameters of the model. In the mid-1980s, following the most severe recession since the Great Depression in 1981–1982, neoclassical economists sought to model technological change as endogenous to the process of accumulation and growth. Economists such as Paul Roemer and Robert Lucas theorized that investments in innovation and "human capital" were important determinants of economic growth. These models are often termed "AK" models because all inputs were specified as a form of capital. No longer were there land, labor, and capital. Now there are natural capital, human capital, physical, capital and money capital. Expenditures on education and training are therefore important for the future and within the domain of proper government activity, and government policies should focus on innovation and competition. The model assumes that marginal productivities are constant at the aggregate level, so no declines occur due to the addition of capital. The models also tend to utilize perfect competition as the basic market structure. Although short-term monopoly profits might fall to those in research and development, but free entry into the market will equalize these profits in the long term.

The current state of the art of neoclassical growth theory is known as dynamic stochastic general equilibrium theory or DSGE. Dynamic refers to change over time, which is the very essence of growth. Stochastic is used not just as being probabilistic but in the sense of the economy being subject to random errors. If all errors are random, then policy prescriptions are essentially irrelevant. The model is cast within Walrasian general equilibrium theory. If you recall, general equilibrium holds that individual agents will trade amongst one another, with accurate knowledge and foresight of prices until no trader can be made better off until another is made worse off. This is referred to as Pareto efficiency. Since all agents have the same perfect information of the present and future and the same reasoning process, they can be treated as exactly the same, and the entire economy can be reduced to a single representative agent. Technological changes are a random error, and treated as frictionless, despite the fact that technological changes benefit some in the real world and hurt others. Moreover, in the words of James K. Galbraith, capitalism is treated as a perfect or nearly perfect system, the analog of a frictionless physical system, that adjusts to random shocks and results in a steady-state growth trajectory [38]. The two main variants of DSGE are real business cycle theory, based on perfect competition, and new Keynesian economics, which allows for some monopolistically competitive price setting.

2.13 Thorstein Veblen and the American Institutionalists

Institutionalism as a school of economic thought focuses on the structural transformation of social institutions over time, and not price formation, as the key to understanding how an economy functions. Institutional change affects human behavior and human behavior affects institutional change. Institutionalism's main proponent, Thorstein Veblen, can be classified more as a social critic than as an economist, for he read and wrote widely in science, politics, anthropology, philosophy, and history, as well as economics. Veblen is most known for being a vociferous critic of neoclassical economics, taking on the giants of the day such as monetary theorist Irvin Fisher and his own mentor, John Bates Clark who was the originator of marginal productivity theory. Veblen's critique of neoclassical economics, a phrase which he coined and often used interchangeably with "the hedonistic approach," stemmed from his active study of science, especially Darwinian evolution. Veblen adopted Darwin's descent with modification based on random variation and natural selection in a way that was far different than other "social Darwinists" such as Herbert Spenser and William Graham Sumner who focused on the competitive nature of humanity and "the survival of the fittest." Veblen juxtaposed the competitive, or predatory, side of humanity with the altruistic, that took the form of the "instinct of workmanship" or the parental bent. Veblen's adaptation did contain one important difference from Darwin's ideas on evolution in the nonhuman world. Humans can adapt to changes in the biophysical world by changing their behavior within one generation. Within this context, much of Veblen's work was rhetorical in the strict sense of the word: the art of persuasion. Veblen urged his readers to adapt to the pecuniary exploits and fraudulent behaviors of the "captains of industry" in the Guilded Age.

In one of his first articles, the 1898 "Why Economics is not an Evolutionary Science," Veblen took on the utilitarian theory of human behavior and classified it as "pre-Darwinian." He asserted that the utilitarian perceptions of the rational, self-absorbed *homo economicus* were incorrect, as they allowed for neither the adaptation of the individual nor the institution of the market, which shapes individual behavior and is shaped by behavior in return. For Veblen, the economic life of the individual was a process of cumulative adaptation, with the economic agent and the social environment being the result of the last adaptation. This is a far cry from the unchanging individual with self-regarding preference sets who is unchanged by the institutions of the market. In Veblen's words:

> The hedonistic conception of man is that of a lightening calculator of pleasures and pains, who oscillates like a homogenous globule of desire of happiness under the impulse of stimuli that shift him about the area, but leave him intact. He is an isolated, definitive human datum, in stable equilibrium except for the buffets of the impinging forces that displace him in one direction or another. Self-imposed in elemental space, he spins about symmetrically about his own spiritual axis until the parallelogram of forces bears down upon, whereupon he follows the line of the resultant. When the force of the impact is spent, he comes to rest, a self-contained globule of desire as before [39].

According to Rick Tilman, editor of *A Veblen Treasury,* Veblen's work was centered around duality and conflict. Some of the main conflicts included those between superstition and science, between business and industry, and predatory exploit and warlike animus vs. peaceable congeniality and workmanlike efficiency. These conflicts appeared in all of his major works [40].

Veblen's best-known work was his 1899 *Theory of the Leisure Class* [41]. It was here where he coined the term "conspicuous consumption." Veblen historically and anthropologically, analyzed the role played by a growing economic surplus (based on an energy surplus) in the development of a class that did not have to work. Veblen's analysis began with hunting and gathering societies and the emergence of settled agriculture (which, in the parlance of the day, Veblen called savagery and barbarism). Predatory activities such as war and sports led to the highest of social statuses, and people emulated these upper classes to improve their own senses of well-being. Veblen believed that the utility preferences of the common man could not be understood in absence of understanding the preferences of the upper classes. In this work, he began to utilize his concept of instincts that would appear throughout the remainder of his works. Veblen used instincts differently than would an animal behaviorist. For Veblen, instincts were more like propensities. They were purposive, learned behaviors.

In his 1904 *Theory of Business Enterprise* [42], Veblen refined the distinction between business and industry and the instinct of predation and the instinct of workmanship. Pecuniary activity (making money) was grounded in the instinct of predation while making products found its base in the instinct of workmanship or doing a good job for the sake of doing a good job. Veblen called the process of consciously denying efficiency for the sake of pecuniary gain (think of the recent revelations regarding Volkswagen) to be sabotage. It was also in the *Theory of Business Enterprise* that he enunciated his theory of the business cycle. Veblen was among the first to incorporate analyses of monopoly concentration and finance into his analyses, stating that the cause of economic instability lay in excessive capitalization and credit inflation. Fundamentally, there is a tendency for firms to borrow too much based on overestimation of their future earning power. When banks and creditors realize this, they call their loans which set off a chain of bankruptcies and liquidations. When expectations of future earnings coincide, once again, with reality, the bankruptcies stop until the next round of speculative excess drives the cycle once again. Veblen thought that the growth of monopolies and wasteful government spending might stop the cyclical instability but did not express a great deal of optimism.

Veblen never enunciated a theory of income distribution, although he spent a great deal of time criticizing the marginal productivity theory of his mentor, John Bates Clark. He thought that the assumption that compensation equals effort was wholly untenable, for there was no reliable way to measure an individual's contribution in a social setting, especially when pecuniary activity, based on the instinct of predation, was at the base of the process. Veblen ridiculed theories of abstinence and waiting that justified the appropriation of economic rents by those absentee owners who did their best to avoid hard work. He also realized, because of the pecuniary processes, that wages were administered rather than being the result of supply and demand in competitive markets. Veblen was an ardent supporter of unions, most notably the Industrial Workers of the World, and advocated for democracy in the workplace instead of dictatorial control by the agents of predation.

Economist Lisi Krall asserts that this distinction between business and industry and the concept of administered prices is crucial for understanding the dynamics of the oil industry in the second half of the age of oil. The Saudi-led cartel realizes that if prices are maintained at too high a level for too long, the oil-dependent industrial nations of the global North will find alternatives and find them quickly. Historically, the business strategy has been to limit production to maintain the "correct" administered prices, just the process that Veblen termed "sabotage." Yet despite price fluctuations and peak oil, the motives and the power of the oil industry have not been negated [43]. With the advent of new technologies such as hydraulic fracturing, one cannot fully understand the future prospects without looking at the institutional structure of the oil industry. The theories of Thorstein Veblen are a good place to start.

2.14 Keynesian Economics

The beginnings of *Keynesian economics* date to 1936 with the publication of *The General Theory of Employment, Money, and Interest* [44]. In this work Keynes was mostly interested in how uncertainty led to declines in capital investment and an imbalance with aggregate savings. He concluded that periodically the overall level of economic activity would fall as a result of falling investment, leading to an overall decline in the level of (aggregate—or total national) demand for goods and services. The economy could come to rest at an equilibrium point that was characterized by elevated levels of unemployment unless the economy was stimulated by an outside force. Keynes attributed the depression to a market economy's inability to sustain sufficient demand for goods and services over the long period, as well as the misguided policies of neoclassical economics that reduced consumption demand as they advocated wage cuts to reduce business cost. Keynes believed in a mild redistribution of income from rich to poor, primarily by means of job creation, and government stimulation of demand during recessions. Keynes was somewhat of an advocate of economic planning and restricted trade.

A more "business-friendly" although perhaps somewhat corrupted Keynesian economics was synthesized, primarily in the United States, in the 1950s. Most students of economics learn that Keynes was mostly about the government's use of its power to tax and spend (known as *fiscal policy*) and its control over the price and quantity of money (*monetary policy*) to keep the economy on an even keel. For decades, it appeared to many that Keynesian economics was the longed-for antidote to periodic business downturns until, in the 1970s, it itself fell victim to the prolonged economic stagnation following the peak of US oil production and the subsequent "energy crises." Keynesian economics was no longer able "deliver the goods" of economic growth with stability. Neoclassical economics made a strong comeback from the 1980s until the global financial collapse of 2008 and the subsequent recession. Recently Keynesian economics has seen somewhat of a revitalization, but there also has been a great deal of resistance to Keynesian measures that exists in the circles of economic policy as well as in economic theory. As of 2017 there is no clear agreement of what kind of economics works and what kind does not.

Economies in general, and capitalist economies in particular, suffer from strong cycles of expansion and recession. Recessions tend to bring enormous hardships to people as workplaces close, and fewer people are employed. John Maynard Keynes had, unlike his neoclassical predecessors, developed a theory that these cycles were caused by internal conflicts. The market as a system was not self-regulating. In his 1936 work, *The General Theory of Employment, Interest, and Money,* Keynes showed that a mature market economic system could reach equilibrium at considerably less than a full employment level. Consequently, the market could not be left to its own devices to restore balance, especially if was already "balanced", but by doing so through substantial numbers of people unable to find work. Keynes considered himself a "moderate conservative" and was primarily interested in saving the market economy from its own worst feature of periodic depressions accompanied by high rates of unemployment. Instead of believing that market forces of competition and flexible prices would correct the ills of depression, Keynes thought that the imbalance of savings and investment led to a deficiency of *aggregate demand* that is for goods and services. Rather than wishing to replace capitalism with another form of organization and governance, Keynes believed that judicious use of government policy could boost the overall level of demand to reduce unemployment during recessionary times. In the 1950s a new generation of economists calling themselves Keynesians would attempt to "fine-tune" the economy by spending more when the economy was contracting and less when it was expanding too rapidly as to make prices rise. These actions would, they thought, tend to smooth out economic fluctuations over time. One can argue that in fact it worked, as the proportional fluctuations in the US economy decreased to much less than before the general acceptance of Keynes's ideas. We will explore this period in our chapter on the postwar economic order.

2.14.1 Keynes and the Taming of Economic Cycles

John Maynard Keynes, who influenced the application of economic theory to day-to-day economics more than nearly anyone else since Adam Smith, had little to say about wealth and value or price formation. He accepted, on face value, utility theory and marginal productivity theory and was relatively uninterested in price formation. He did base his critique of the labor market on the proposition that wages were "sticky" and did not fall as workers attempted to protect their standards of living. This, however, was not original to Keynes, as his neoclassical mentor Arthur Cecil Pigou had worked on this topic.

John Maynard Keynes had little to say about income distribution and what he did offer was contradictory. In ► Chap. 2 of *The General Theory* he stated that the classical theory of employment rested upon two premises. First, the wage equaled the marginal product of labor. This established the demand for labor as capitalists would hire labor only up to the point where the marginal product of labor equaled the prevailing equilibrium wage. At that point, they would cease hiring additional workers. Second, neoclassical theory asserted that the marginal utility of the wage equaled the marginal disutility of the work or the pleasure obtained from the wage earned equals the displeasure of the work done. In other words, the prevailing wage is sufficient to bring forth the needed amount of labor. While he rejected premise number two, Keynes accepted marginal productivity theory without reservation. But this implies that a reduction in wages can expand employment. Unfortunately, this was inconsistent with much of Keynes' main point that the economy can balance at full employment only if the population has enough money to spend purchasing the products they have manufactured.

In ► Chap. 10 of *The General Theory*, Keynes discusses the relation of savings vs. spending in stimulating the economy. Specifically, he examined the role played by the propensity to consume (or the fraction of additional income that is spent). Keynes utilized R.F. Kahn's multiplier principle when he considered overall investment and employment, which states that income is expanded by an amount that equals propensity to consume, that is, the amount of consumption changes with respect to the rise and fall of income. Mathematically, $k = \Delta C/\Delta Y$, where. C symbolizes consumption and Y stands for aggregate income. But Keynes realized that savings came primarily from the wealthy, which he called "the saving classes." If the poor saved a smaller proportion of their incomes than do the rich, then a redistribution of wealth would result in greater total spending and a greater multiplier effect and a more rapid expansion of income and employment. But Keynes never came out for a policy of income redistribution. Rather he addressed the issue indirectly, calling for an expansion of public works [44].

Overall many economists, especially classical economists, thought deeply about the questions of distribution of wealth between the different classes of society. We can say that their discourse, and others like it, had a great deal of effect on the actual implementation of economic policy, at least until the last two or three decades. This was because tax and other government policies based on their thinking tended to result in a much greater equity in the distribution of the great wealth made possible by the industrial revolution, especially in the United States and Europe.

These two conclusions about the functioning of aggregate markets served as the backdrop for John Maynard Keynes' critique of neoclassical economic policy. For John Maynard Keynes, the question was not one of whether overall, or aggregate, supply would balance with aggregate demand, but whether or not the balance would occur at full employment. Keynes began his 1936 opus, *The General Theory of Employment, Interest, and Money*, by accepting all the neoclassical postulates except two. He rejected Say's law and Marshall's idea that the supply of labor is determined by the interaction of the marginal utility of the wage and the marginal disutility of the work. Whether this change in two initial propositions constituted a revolutionary change in the profession or was a matter of "moderate conservatism," as Keynes himself believed, has been and probably will continue to be a matter of considerable debate. But Keynes' conservatism was not about domestic spending. He saw the enterprise economy of the 1930s as being limited by internal and external factors. The internal factor was the persistence of severe unemployment and social dislocation that characterized the depression. The external factor was the presence of two alternate systems, Fascism and Bolshevism, which Keynes found highly distasteful. Keynes conservatism came from his desire to save and perpetuate the free enterprise system. His moderation came from a belief that leaving the economy to its own devices and awaiting the triumph of market forces would be insufficient to solve the problems created by the Great Depression.

The prevailing orthodoxy in the middle third of the last century was grounded in the notion that savings determined the level of investment. Furthermore, the balance of saving and investment was needed to achieve the overall balance of supply and demand. A simplified version modifies the circular flow model, (which is essentially a depiction of Say's law), to accommodate the reality that not all firms and household members spend all their money in current consumption. Money "leaked" out of the system flow when individuals saved a portion of their income, when taxes were levied on income, and when purchases of foreign goods were made. On the other hand, income flowed into the system when businesses made investments, when the government purchased goods and services, or when an economy sold goods in foreign markets and received the incomes from doing so. Consequently, the traditional circular flow model can be augmented with both leaks and injections.

Given the conventions of the early twentieth century of a political commitment to a balanced budget that equated government spending and taxation, along with an international gold standard that balanced imports and exports, the main question facing Keynes was to what degree would savings balance with investment? Unless savings and investment balanced the aggregate supply of products (which were increased by investments) and the effectual demand for them (which were increased by consumptive expenditures) would not balance at full employment. He believed that finding adequate investment outlets for surplus savings, and not wage reductions, was the key to finding a macroeconomic equilibrium at full employment. The prevailing orthodoxy, on the other hand, was to treat the market for savings and investment as a market for loanable funds. Competitive market forces would lead savers and investors to vary the amount of funds with the price, leading the market to find an equilibrium rate that balanced savings and investment. Keynes disagreed vehemently that this was how it worked. Savings, in his analysis, depended upon income, and savings would increase only as income rose. Investment depended upon expected profit and the rate of interest. Savings and investment were functions of different variables. Keynes believed there were no reasons for planned (or *ex ante*) savings to equal realized (or *ex post*) investment.

Keynes' greatest concern was not a shortage of savings but savings that exceeded investment. The orthodox method of increasing savings was to increase the interest rate. This had the unfortunate effect of simultaneously depressing investment, thereby reducing the level of aggregate output and employment. As investment fell, so too did employment. Workers with less money buy fewer products, forcing business to reduce investment once more. The economy spiraled into depression, and when it came to a balance, the equilibrium was at a low level of output and a high level of unemployment. But if the interest rate is not determined in the loanable funds market, where is it determined? For Keynes, interest was a monetary phenomenon. The amount of money in the system depended upon the interaction of the supply of money (determined politically by monetary authorities) and the preference investors have for holding their money as cash (called transactions demand) or as balances to be invested in financial securities (called speculative demand). Money plays an essential feature in a modern economy, and the economy could not run without it. For Keynes, the fundamental problems of investment were those of uncertainty. The present, when investments are made, lies between an unchangeable past and an unknowable future. Despite efforts of economists and mathematicians, the uncertainty posed by investment over the long term makes the rational calculations of neoclassical microeconomic theory essentially impossible. The future is sufficiently uncertain that the self-regulatory capacity of the laissez-faire economy is unlike that posed by neoclassical theory. Keynes believed that the object of the accumulation of wealth entailed investing now to receive rewards in the distant future. But our knowledge of the future is uncertain. In an oft-quoted passage from his 1937 *Quarterly Journal of Economics* article entitled "The General Theory of Employment," Keynes declared:

> The calculus of probability, tho mention of it was kept in the background, was supposed to be capable of reducing uncertainty to the same calculable status of certainty itself…By "uncertain" knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty; nor is the prospect of a Victory bond being drawn. Or, again, the expectation of life is only slightly uncertain. Even the weather is only moderately uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth owners in the social system in 1970. About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know [45].

The use of money allows for a method to avoid all of one's assets being fixed in permanent and unchangeable assets. This ruled what Keynes called the speculative demand for money. But speculation is subject to waves of pessimism and optimism. While the primary driving force of output, and therefore employment, was investment, the level of consumption was also important in determining the level of aggregate demand. The amount of consumption, like savings, was dependent primarily upon the level of income. The fraction consumed (the marginal propensity to consume) was subject to multiplier effects. Since the poor spend a greater fraction of their income than do the wealthy Keynes believed that some augmentation of income growth could be affected by a redistribution of income. Given the uncertainty of investment, and the limitations of expanding the economy by means of money creation when interest rates are low, Keynes allowed for the state to spend to assure sufficient aggregate demand for the economy to balance at the full employment level of income. We will return to his methods in the final question of this chapter.

What should we conclude about this main question of economics, about how economists view whether supply can possibly balance demand, and lead the economy away from the troubling boom and bust patterns that have characterized capitalism? The optimist might point out that most economists believe that firm-level supply is aggregated into market supply and likewise market demand is simply the summation of individual demands. Together these forces operating at the market level balance supply and demand well enough and in a way that is the most efficient allocation of resources. The idea that markets allocate efficiently is a deeply held belief of almost all economists. But cycles remain, although much less as a percent of GDP following the publication of Keynes magnum opus and its partial implementation [31]. Even so, today Keynes, as represented by arguments as to whether, or to what degree, governments should undertake deficit spending to restore ailing economies, is very much hotly contested. The cynic might say "economists throughout the history of economics often held strongly held beliefs that were in fact often contradictory to each other. Today we have little or no better idea than in the past as to which is correct." This is hardly a surprise to anyone who follows economics today.

What is missing from this and other economics questions is a consideration of what has become what is likely to be the most critical issue of economics today: issues of energy and other resources. And environmental degradation. The issue was always how to take nature's abundance and mobilize forces to turn that into wealth and employment. We can perhaps understand how this came about since economics was mostly developed before the appropriate science, but the roots of economics have hardly budged with the new information we have now on resources and the environment, and probably most economists today do not think there is any particular reason to worry too much about resource or environmental limitations.

2.14.2 Accumulation and Growth

Keynesian economics gained prominence in the failed growth economy and Great Depression of the 1930s, but it was, perhaps surprisingly, not particularly oriented toward growth. Rather it focused on an explanation of the role of inadequate demand and uncertainty in producing depression, as well as the futility of relying upon markets alone to produce sufficient demand to end the depression. A Keynesian growth theory was not to emerge until the very end of the depression in 1939. Roy Harrod, Keynes' collaborator and biographer, began his "Essay on Economic Dynamics" with the conflict between what he termed "the actual growth rate" (G) and the "warranted growth rate" (G_w). The actual growth rate is the percentage change in output from year to year, $x_1 - x_0/x_0$. The warranted growth rate follows from the Keynesian tradition of psychological theories of the trade cycles. This is best remembered as Keynes' idea of the role of "animal spirits" in the investment process. It is the growth rate that leaves all parties satisfied that they have produced neither more nor less than the correct amount or the growth rate that will lead them to produce just enough to maintain the rate of growth. The warranted growth rate is determined by the ratio of the propensity to save or the change in savings relative to the change in income ($s = \Delta S/\Delta Y$) and the value of capital goods needed to produce one unit of output (C). Stated mathematically: $G_w = s/C$. The instability emerges from this fundamental equation. If there is excessive output and G exceeds G_w, the actual increase in capital goods per unit of output falls below the desired level and will lead to an undesired reduction of the capital stock by means of inventory depletion. Investors will then increase their capital stock even more, causing a further movement of G from G_w. The larger the gap, the greater the stimulus for further expansion. If the actual growth rate falls below the warranted growth rate, excess capacity will emerge, resulting in a decline in the incentive to invest. This creates a positive feedback loop and economic instability due to the internal dynamics of the investment process. In Harrod's own words:

> A departure from equilibrium, instead of being self-righting, will be self-aggravating…A unique warranted rate of growth is determined jointly by the propensity to save and the quantity of capital required by technological and other considerations. Only if producers keep to this line will they find that on balance their production in each period has been neither excessive nor deficient. On either side of this 'field' in which centrifugal forces operated, the magnitude of which varies as the distance of any point from the warranted line. Departure from the warranted line sets up an inducement to depart farther from it. The moving equilibrium of advance is a highly unstable one [46].

On the 17th page of a 22-page essay, Harrod introduces the concept of the natural rate of growth (G_n). We mention the pagination because Solow's aforementioned critique of Harrod in his 1956 paper advanced the proposition that Harrod's conflict was between the warranted and the natural growth rate, not the warranted and actual growth rate as Harrod contended. Population, work/leisure preferences, capital accumulation, and technology determined the natural growth rate, which was defined as the maximum growth rate allowed by these factors. Furthermore, there is no inherent tendency for the warranted and natural growth rates to coincide. If the warranted rate were to exceed the natural rate, a depression (or stagnation) would result as the social and economic forces are limited by the systematic biophysical limits found in the natural rate. The warranted rate must fall to the natural rate, and this can be achieved only by chronic unemployment. Harrod's policy recommendations were to "manipulate the proper warranted rate [by means of public works, fiscal and monetary policy] so that it would be equal to the natural rate" [Harrod 1939: 32].

Harrod's natural growth rate can be interpreted on a biophysical basis by adding the quantity and quality of energy sources, as well as the assimilative capacity of the atmosphere and the oceans as independent variables. The fundamental problem remains that the warranted growth rate that would lead to maximum profits exceeds the natural rate but now by a greater fraction. Stagnation and unemployment will still result, and stimulative measures, which may be successful in the short term, will not rectify the long-term problem. This problem would reverberate through the economy as a whole, including labor and financial markets. Combined with the structural changes enabled by information technologies that reduce the need for human workers in the manufacturing and service sectors, short-term stimulative policies may be successful in increasing the growth rate but will not lead to full employment [38]. The impact of the structural shift in labor markets will not be measured fully in the unemployment rate but rather in slow rates of growth of wages and labor force participation and in long-term underemployment. Reducing the warranted rate to the natural rate will be a difficult problem in the absence of significant social restructuring. Energy prices will eventually rise, as the undulating plateau created by the interaction of supply and demand is transcended by geophysical realities. Most probably before we lack access to sufficient fossil fuels, growth-dependent financial markets may fluctuate wildly before declining. The debt-based global economies will find it difficult, to say the least. In a nongrowing economy, capitalist societies may well tear themselves apart with distributional conflicts in the interim. The problem of living within nature's limits is considerably more difficult than technological optimists believe. What is needed is to decouple employment from economic expansion.

Seven years after Harrod's paper was published, Evsey Domar enunciated a similar theory of growth and instability, although he never read Harrod's work until after his own papers went to the publisher. He made explicit connections between economic growth and employment in two papers published in the immediate postwar period [47, 48]. The expansion of employment depends not just upon the growth of national income but upon the *rate of growth* of national income. Job growth necessitates that national income and effective demand (consumption + investment) grow perpetually at an *increasing rate*. After making a set of simplifying assumptions including no time lags, the use of net savings and investment, and a constant price level, but not fixed proportions of labor and capital, Domar set out a model word add dynamic elements to the static Keynesian system. New investment is simply capital accumulation. It increases national income but also increases the productive capacity of the economy. Unfortunately, the national income that produced full employment would not be sufficient to produce it in the next because of increases in technology, the labor force, and access to new resources. Domar criticized the mainstream (neoclassical) approach of increasing income by reducing prices, as price decreases were a rare occurrence in the monopolized economy that he observed.

The essence of Domar's argument lies in the *dual nature of investment*. New investment is simply capital accumulation. As a form of spending, investment increases aggregate demand and national income. However, on the supply side, investment also increases productive capacity. The instability comes from the fact that the stimulation of demand is short lived, while the expansion of capacity is long lived. Excess capacity reduces the demand for new capital formation. From this simple realization, Domar developed a model that included the growth of investment on both the supply side and the demand side.

If Y = national income, α = the marginal propensity to save, then $1/\alpha$ will equal the multiplier (k) which shows the degree to which an increase in spending will translate into an increase in national income. Domar also posited that σ represented the productive potential of the economy or, more precisely, the average social productivity of investment. σ measured the dollar amount of capital needed to produce a dollar increase in national income. From the supply side perspective, σI represented the total output that an economy can produce. From the demand side, $\Delta I\ 1/\alpha$ indicates the total aggregate demand. In equilibrium:

$$\Delta Y = \Delta I 1 / \alpha = \sigma I$$

To maintain a constant state of full employment, investment and national income must grow at a constant percentage rate $\alpha\sigma$, which equals the rate of compound interest. To expand employment to keep up with resource availability, technology, and labor force, growth investment must grow perpetually at an increasing rate. This is unlikely, if not impossible, because the buildup of excess capacity stunts the rate of investment growth. Domar's model was in the tradition of multiplier-accelerator models. Balanced growth is difficult because to have a high multiplier, one must have a high marginal propensity to consume. To have an equally high accelerator, one must have a high propensity to save. Since the sum of marginal propensity to save and the marginal propensity to consume = 1, it is impossible for this seemingly simple mathematical condition to exist in the real world. Domar concluded by stating that excess capacity would not be a problem in a competitive economy, as those firms with too much capital would go bankrupt. Yet in a monopolized economy, excess capacity would be a chronic problem that the private sector could not solve on its own. According to Domar, the government needed to assume the role of investment banker to keep the funds for expansion flowing.

2.15 Biophysical Economics

Most of the economic schools mentioned so far were growth oriented to greater or lesser degrees. The main disagreement then, as now, was how would growth be best achieved? Classical political and neoclassical economists tended to focus upon market processes in achieving accumulation and growth. Karl Marx explored the internal contradictions that inhibited the accumulation process. Keynesian economics relied on the role of the government to provide the growth stimulus when private economy could not. In the absence of growth, employment would stagnate and human well-being would decline. In the early classical era, growth could be achieved principally by organizational means; the capacity to increase material output by means of technological change barely existed. It was only in the later stages of classical political economy, neoclassical, and Keynesian economics that the ability to increase output dramatically by means of harnessing energy-dense fossil fuels was possible.

What, then, should be the purpose of biophysical economics, the approach we are advocating in this book? Clearly it must deal with a world that is increasingly dependent upon stocks of fossil fuels, the depletion of those stocks, and the increasing difficulty of achieving growth as depletion occurs. Unlike the utilitarians, biophysical economics considers and encourages the possibility that humans can achieve happiness by means other than the acquisition of ever-increasing quantities of material goods—goods that cannot be produced with declining resources. As such, it calls back to the center stage the question of distribution: for generations that question has been suppressed. If the pie has been getting larger then everyone can get a larger piece. But if the size of the pie is not growing, who should get how large a piece?

Biophysical economics serves as a wake-up call to the impending and inevitable end of the economy based on high-quality fossil fuels and with it the end of growth economics. It also provides important caveats as to which of the many alternatives proffered has a good chance of succeeding by providing guidelines for the assessment of alternative sources of energy. How we can live well within nature's limits is a question we can no longer afford to postpone or subsume to a series of equations unconstrained by reality. But to answer this whole new set of questions, we must first assess how economists have addressed the age-old ones, for these questions remain as relevant for these new conditions as they were for the circumstances when they were asked. In other words, for a relatively few decades—a century and a half at most—in the most favorable situations has a year by year increase in general affluence been the normal condition. It was not true back when early economists were writing and it appears no longer true. So, we must pay attention once again to their questions—but we need to do that while including an energy perspective.

2

2.16 Summary

In this chapter, we chronicled the development of economic thought over the ages focusing, when we could, on the role played by energy. We also tried to emphasize the major transitions that occurred in the actual economy and explored how they affected the course of economic thought. Economic thinking and writing in the ancient and medieval world tended to justify the prevailing social order of a small elite controlling the society through land ownership. Collective sets of privileges and obligations were codified as natural law, and individual self-advancement was castigated as a mortal sin. By the early 1500s, individualism emerged in the age of exploration, the Renaissance, and the Enlightenment.

The first recognized school of thought were advocates of expanded trade known as the mercantilists. Their basic theory held that the origin of wealth could be found in the process of exchange. Buy cheap, sell dear. The real money was to be made in colonial exploitation and control of the carrying trade. By the mid-1700s, the idea that wealth and value could be determined by adding up the costs of production, rather than by counting sales, began to emerge. The first school of thought, the Physiocrats, held that all value came from "the natural bounty of the land" and the agriculturalized labor that transformed nature's bounty. By the late 1700s, the idea of value being produced by labor in general became the norm, as enunciated by Adam Smith and David Ricardo. Value, and price, could be determined objectively by adding the costs of production, especially labor costs.

Smith and Ricardo, along with Thomas Malthus, lived in an age of the solar flow. Animate power, biomass, and water served as the primary, and limited, energy resources. Energy was embodied in a fixed supply of land and that fixed amount of land gave rise to diminishing returns and pressures of a growing population upon the limited capacity to grow food. Limited energy densities helped account for the small-scale nature of production. Although all the classical economists advocate policies of capital accumulation and economic growth, all believed the eventual fate of an economy would be a nongrowing stationary state. Ricardo and Malthus, especially, engaged in great debates about the distribution of society's income, and each advocated a policy to redistribute income to their favored classes, as class was the primary unit of analysis. Ricardo favored putting money in the hands of the newly emerging capitalist class, who would invest the money to drive economic progress, while Malthus favored the landed aristocracy who would spend the money on comforts and personal servants, ensuring adequate spending and keeping the economy from stagnation and depression.

John Stuart Mill was a transition figure. He started out as an advocate of the labor theory of value but popularized the principles of utilitarianism that would come to characterize neoclassical economics. Mill still believed that the fate of the economy was in the stationary state, but unlike his predecessors believed such a state could be superior to a growing economy where individuals stepped on one another's backs in order to get ahead.

Karl Marx was the first economist of the industrial revolution. He realized the productive power of fossil fuel-driven machinery to enhance labor productivity and to augment wealth and income. But Marx's analytical method looked for contradictions. The same economic forces that increased wealth and income expanded the exploitation of labor. The economic process of recapitalizing the surplus labor of workers sets the condition for an internally generated decline in the rate of profit and an economic crisis. The process of capital accumulation that resulted in increased wealth also undermined systematically the very material conditions of its existence: the worker and the soil. Unlike prior classical political economists, Marx was interested in the transition to the next society, rather than the perpetuation of the existing form of capitalism.

Within a decade of the publication of the first volume of Marx's *Capital*, a fundamental epistemological break occurred in economic theory. The marginal revolution occurred, and the determination of value was to be found in the sphere of exchange rather than in the process of production. In addition, value now depended upon the subjective well-being of the individual rather than upon an objective counting of labor hours. Social class ceased to be a proper category of analysis, and the historical specificity of classical political economy gave way to universal theory. The focus of accumulation gave way to a search for static equilibrium. The marginalism of the 1870s became the neoclassical economics when the process of production was placed on a marginal utility basis. Supply and demand graphs made their appearance, and the purpose of economics became price determination. By the 1920s, the neoclassical approach expanded beyond the well-being of the individual and began treating the economy as a whole as if it were an individual market. Competition and flexible prices became the method of self-regulation, not just for individual markets but for the entire economy.

At the turn of the nineteenth to the twentieth centuries, institutional economists such as Thorstein Veblen criticized strongly the neoclassical ideas about human behavior, ideas on perfect competition, and the very idea of self-regulation. For Veblen and his followers, price formation should not be the focus of economics. Rather economic evolution by means of structural and institutional change was the path to a deep understanding of how an economy operated. Veblen, like Marx, based his theory on conflict and contradiction: the conflict between business and industry and the conflict between the ethic of workmanship and that of predation.

While the dominance of neoclassical economics could survive the ideas of a Veblen, it could not so easily survive the devastation of the Great Depression. The idea of self-regulation fell into disrepute in an era of 25–50% unemployment rates and a collapse of industrial production. This social dislocation provided a fertile backdrop for the ideas of John Maynard Keynes who argued that economic equilibrium could occur at any level of output, even levels that produced high unemployment. Keynes advocated reducing unemployment by expanding overall, or aggregate, demand. He discounted the idea that the private economy could produce sufficient demand, so he advocated the role of government spending as a solution. Yet Keynesian economics was not about producing economic growth, it was about recovery and stability. If anything "proved" Keynesian economics worked, it was the economic recovery, especially in the United States that accompanied the Second World War. Little concern was displayed, even by the most conservative legislators, for budget deficits when it came to defeating fascism.

After the war both a Keynesian and a neoclassical growth theory emerged. Keynesian growth theory emphasized the instability of the economy, while neoclassical growth economics stressed the idea that substitution of resources would result in a steady-state growth path. Keynesian and neoclassical debates characterized the 1960s, but Keynesian economics fell into disrepute when its policies could not solve the problem of simultaneous recession and inflation. Neoclassical economics reemerged as the dominant mode of economic thought in the 1980s and has remained the primary approach by which today's students are taught about the economy.

Yet there is need for a more comprehensive theory, as the mainstream Keynesian-neoclassical synthesis excludes the crucial role of energy and discounts the disruption of the Earth's biophysical systems. This is the void that biophysical economics seeks to fill. Fortunately, many lessons can be learned from economic analyses of the past, especially those of classical political economy. We hope you have gained a better understanding and appreciation that the role of history plays in shaping the future.

? Questions

1. Do you think that combining natural science and economics is a good idea? Why or why not?
2. How is a city like a natural ecosystem? How is it different?
3. What ideas did you get in this chapter from earlier economists that you think might be important for understanding our current situation?
4. Can you think of a "peak oil" situation that occurred 150 years ago? Does that have any relevance today?
5. Why do you think economists have tended to ignore energy in their basic equations? Were they justified in doing that?
6. Where did the early group of economists known as the physiocrats believe that wealth came from?
7. Define relative vs. absolute scarcity.
8. What is economic surplus?
9. What are Heinberg's "five strategies for obtaining energy?"
10. Discuss one of the four main economic questions.
11. List four major schools of economics over time and one idea associated with each
12. What is natural capital?

13. What is the source of wealth for a physiocrat? A classical economist? A neoclassical economist? Yourself?
14. What was the "Wealth of Nations?" How does that relate to the title of this book?
15. Give one of the great economic ideas derived by David Ricardo.
16. What was the "diamonds vs. water" paradox? How was it resolved?
17. How did Keynes think we could diminish the large swings in the capitalist economy?
18. What did classical political economy have to say about the distribution of wealth?
19. Discuss comparative advantage.
20. What is the "best first principle?"
21. Was Karl Marx principally interested in communism?
22. Did Mill think about the distribution of wealth?
23. What important factor did the Cobb-Douglass production function leave out?
24. What are the main two views as to whether economies can balance supply and demand?
25. What earlier economist probably had the largest impact on what is taught today in basic economics courses?

References

1. Polanyi, K., C. Arensberg and H. Pearson (eds) 1957. *Trade and Market in the Early Empires: economies in history and theory*. Glencoe: Free Press.
2. Heinberg, R. 2003. *The party's over*. B. C. Canada: New Society Publishers, Gabroiola Island.
3. Foster, John Bellamy, and Fred Magdoff. 2009. *The great financial crisis*. New York: Monthly Review Press.
4. This chapter heading is taken from Sweezy, Paul. 1953. *The present as history*. New York: Monthly Review Press.
5. Rubin, I.I. 1979. *A history of economic thought*. London: Pluto Press.
6. Rubin, p. 73.
7. Rubin, p. 75.
8. Graeber, David. 2014. *Debt: The first 5000 years*. Brooklyn/New York: Melville House.
9. Daly, Herman. 1998. The return of the Lauderdale paradox. *Ecological Economics* 25: 21–23.
10. Foster, John Bellamy, Brett Clark, and Richard York. 2010. *The ecological rift*. New York: Monthly Review Press.
11. Smith, Adam. 1923. *An inquiry into the nature and causes of the wealth of nations*, 17. New York: Modern Library.
12. Adam Smith. 1923. p. 47.
13. Ricardo, David. 1962. *Principles of political economy*. Cambridge: Cambridge University Press.
14. Mill, John Stuart. 2008. *Principles of political economy*. Oxford: Oxford University Press.
15. Dobb, Maurice. 1975. *Theories of value and distribution since Adam Smith*. Cambridge: Cambridge University Press.
16. Malthus, Thomas. 1960. *Principles of population*. New York: Dutton. p. 128.
17. Heilbroner, Robert L. 1953. *The Worldly Philosophers: The Lives, Times And Ideas Of The Great Economic Thinkers*. Touchstone.
18. Marx, Karl. 1976. *Capital*. Vol. I. London: Penguin Books.
19. Karl Marx and Frederick Engels. 1848. *The communist manifesto.*
20. Tucker. 1978.
21. Malm, Andreas. 2016. *Fossil capital*. London: Verso.
22. Marx, Karl. 1941. *Critique of the Gotha program*. London: Lawrence and Wishart.
23. Engels, Frederick. 1940. *The dialectics of nature*. New York: International Publishers.
24. Sweezy, Paul. 1942. *Theory of capitalist development*. New York: Monthly Review Press.
25. Marx. 1976. pp. 925–926.
26. Solow, quoted in Perelman. 2006. *Railroading economics*. New York: Monthly Review Press.
27. Michel DeVroey. 1975. The transition from classical to neoclassical economics. *Journal of Economic Issues. 9:415–439*
28. Passinetti, Luigi. 1979. *Lectures on production*. London: Palgrave Macmillan.
29. Klitgaard, Kent. 2009. Time for a change in economic theory. *Proceedings of the 61st Annual Convention of the New York State Economics Association I: 54–63.*
30. Mirowski, Philip. 1989. *More heat than light*. Cambridge: Cambridge University Press.
31. Cleveland, CJ, R. Costanza, C.A.S. Hall, and R. Kaufmann. 1984. Energy and the United States economy: A biophysical perspective. *Science* 225: 890–897.
32. Kummel, R. 1989. Energy as a factor of production and entropy as a pollution indicator in macroeconomic modeling. Ecological Economics 1(1): *161–180*.
33. Robinson, J. 1939. *Introduction to the theory of employment*. London: Macmillan.
34. Jevons, W.S. 1865. The coal question: An inquiry concerning the progress of the nation and the probable exhaustion of our coal mines. MacMillan and Company, London and Cambridge.
35. Marx, Karl. 1993. *Capital*. Vol. III. London: Penguin Books.
36. Roser, M. 2017. 'The short history of global living conditions and why it matters that we know it'. *Published online at OurWorldInData.org*. Retrieved from: https://ourworldindata.org/a-history-of-global-living-conditions-in-5-charts/ [Online Resource]
37. Alfred Marshall. 1890. *Principles of economics.*

38. Galbraith, James K. 2014. *The end of normal*. New York: Simon and Schuster.
39. Veblen, Thorstein. 1919. Why economics is not an evolutionary science. In *The place of science in modern society*, ed. Thorstein Veblen, 73. New York: B.W. Heubsch.
40. Tilman, Rick. 1993. *A Veblen treasury*. Armonk: ME Sharpe.
41. Veblen, Thorstein. 1965. *The theory of the leisure class*. New York: Augustus M. Kelley.
42. ———. 1965. *The theory of business enterprise*. New York: Augustus M. Kelley.
43. Krall, Lisi. 2014. An institutionalist perspective on the economy and the price of oil. In *Peak oil, economic growth and wildlife conservation*, ed. J.E. Gates et al. New York: Springer.
44. Keynes, John Maynard. 1964. *The general theory of employment, money and interest*. New York: Harcourt, Brace, and World.
45. ———. 1937. The general theory of employment. *Quarterly Journal of Economics*. 51: 209–223.
46. Harrod, Roy. 1939. An essay on dynamic theory. *The Economic Journal*. 49: 14–33.
47. Domar, Evesy. 1947. Expansion and employment. *American Economic Review*. 37 (1): 34–55.
48. ———. 1946. Capital expansion, rate of growth, and employment. *Econometrica* 14: 137–147.
49. Robinson, Joan. 1934 Euler's theorem and the problem of distribution. *The Economic Journal*. 44 (175): 398–414.

Problems with How We Do Economics Today

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_3

3.1 Introduction [1]

The first chapter of this book summarized how we undertake economics, and our explanation for that approach in the modern Western world. The second chapter introduced the idea that this contemporary view of understanding economics is just one of many ways that humans have understood how the economy operates. The last century has seen the ascendancy, indeed intellectual dominance, of neoclassical economics (NCE, also known as Walrasian economics). The basic NCE model represents the economy as a self-maintaining circular flow among firms and households, driven by the psychological assumptions that humans act principally in a materialistic, self-regarding, and predictable way. Unfortunately, the NCE model violates a number of physical laws and is inconsistent with actual human behavior, rendering it to be an unrealistic and a poor predictor of people's actions. Recently, an array of experimental and physical evidence and theoretical breakthroughs demonstrate the disconnect between evidence and neoclassical theory. Despite the abundance and validity of these critiques, few economists seriously question the efficacy of the neoclassical paradigm that forms the foundation of their applied work, although behavioral economists such as George Ackerlof and Richard Thaler have received Nobel Prizes precisely for questioning the assumptions about rationality. This is a problem because policy makers, scientists, and others turn to economists for answers to important questions. The supposed virtues of "privatization," "free markets," "consumer choice," and "cost-benefit analysis" are considered to be self-evident by most practicing economists, as well as many in business and government. In fact, the evidence that these concepts are correct is rather slim and contradictory. Thus, this chapter is a strong critique of economic theory, in this case NCE.

We offer a review and synthesis of NCE, paying particular attention to the lack of connection of NCE to biophysical reality and its inadequate characterization of human behavior. When all the criticisms are taken as a whole, it is clear that the NCE framework stands on an untenable foundation and that some other basis for interpreting economic reality must be found. NCE is very limited in its usefulness and cannot guide us in our attempts to deal with the most critical issues of our time, such as the depletion of oil and gas, climate change, financial crises, inequality, and the destruction of much of nature. We end by sketching alternative characterizations of human behavior and economic production.

3.1.1 Economic Issues Appropriately Assessed with Conventional Economics

Before we begin we wish to emphasize that there are any number of conventional "economic" questions about which we believe that conventional economic procedures are accurate and appropriate. For example, we have no argument with cost and gain accounting procedures used by businesses and individuals. One must balance one's own accounts using just dollars, although one can think about the meaning of those dollars in terms of their energy backing. Our issue is with the theory that forms the basis for economic thinking. This theory is the basis for more complex economic thinking.

3.2 Some Fundamental Myths of NCE

The edifice of NCE is built on myths and based on an outdated worldview. These myths are not merely harmless allegories because they provide the foundation upon which economic policy is made and cultural attitudes are distilled. Thus, the worldview and policy prescriptions of most economists can only be described as "faith based" because many fundamental tenets of NCE are inconsistent with economic reality.

3.2.1 Myth 1: A Theory of Production Can Ignore Physical and Environmental Realities

Real economies are subject to the forces and laws of nature, including thermodynamics, the conservation of matter, and a suite of environmental requirements. NCE does not recognize or reflect the fact that economic activity requires the inputs and services of a finite biophysical world which is usually diminished and degraded by that activity.

3.2.2 Myth 1a: The Economy Can Be Described Independently of Its Biophysical Matrix

NCE begins with a model depicting abstract exchange relations considered only as goods and services and money within a world unrealistically limited to markets, firms, and households. Real economies also require material and energy from the natural world to allow that exchange and are limited by the material and energy transformations necessary for economic activity. Students are introduced to the *circular flow model* of the economy in the first days of principles of economics. This conceptual vision of the economy is one of a self-contained and self-regulating system independent of the biophysical system and its laws. There are but two sectors, households and firms, with goods and services going from firms to households, and productive inputs (land, capital, and labor) going from households to firms. As seen in ► Chap. 1, all human interactions take place in markets. Firms acquire the property right for land, labor, and capital in the factor market by payment of rents, wages, interest, and profit. Consumers in the household receive goods and services in exchange for money. All exchanges are seen as voluntary and made in the pursuit of self-interest. For the model, at this basic level, to be self-regulating the money that flows from firm to household (the sum of factor payments) must equal the total expenditures on goods and services. No money is saved, and no profits are retained by business for reinvestment. But more importantly from a thermodynamic point of view, the material and energy inputs required for production are simply left out of the model.

Neither monetary value nor physical materials are lost to heat or erosion as inputs are transformed into goods and services. Thus, the circular flow model represents an abstract notion of an economic system that cannot exist.

The NCE notion of scarcity is disconnected from biophysical reality for it is never absolute but only relative to unlimited wants. If we are confronted by the limits of one resource, the imaginative human mind, driven by the proper set of monetary incentives and protected property rights, we will always create a substitute. No input is critical, therefore neither absolute scarcity nor the need of any particular resource is a problem in the long run. Thus, in the NCE world the economy can simultaneously experience relative scarcity and infinite growth. Competitive prices, formed in markets, assure that resources flow to their best use.

Nicholas Georgescu-Roegen and his student Herman Daly were among the first to point out the absurdity of this depiction of production. Real economies cannot exist outside the global biophysical system, which is essential to provide energy, raw materials, and a milieu within which it can operate and assimilate wastes [2, 3]. Their first step to make an economic model consistent with reality is to put the economy *inside* the global biophysical system. Some natural scientists have gone several steps further. Several writers [4–7] demonstrate clearly that the NCE model is unacceptable because (1) its boundaries are drawn incorrectly and (2) the model is de facto a perpetual motion machine because it has neither energy inputs nor entropic loss. Many economists today, including many recent Nobel Prize winners (e.g., Paul Krugman, Amartya Sen, Joseph Stiglitz, George Akerlof, and Elinor Ostrom) have very serious reservations with the contemporary model. Most of the authors referenced in this paragraph, the authors of this book, and many other physical and social scientists are not interested in simply making corrections to the basic NCE models. Instead these scientists and others believe that the NCE model is incorrect at its core. For starters, while money may cycle seemingly indefinitely among goods and services, the real economic system cannot survive without continual inputs from, and outputs to, nature.

3.2.3 Myth 1b: Economic Production Can Be Described Without Reference to Physical Work

The neoclassical economists' model of production does not require any specific physical inputs but is solely an exchange of existing entities among firms and households. The economic process is driven not by the availability of physical resources, but rather by human ingenuity as depicted in the still widely used Cobb-Douglas function. The quantity of output produced (Q) is a function of only capital (K) and labor (L).

$Q = AK^{\alpha}L^{\beta}$ where α represents capital's share of output, β stands for labor's share, and $1 > \alpha > 0$. Moreover, $\alpha + \beta$ must add to one, so $\beta = 1 - \alpha$. The product of capital and Labor is also multiplied by some constant A, considered "pure technological change," or total factor productivity.

3

In this model technology is independent of the inputs of land and capital and is calculated as a "residual" left when the contributions of the measured factors (i.e. capital and labor) are subtracted from the growth rate of total economic output [8]. Not surprisingly the residue tends to increase over time. Thus, most economists believe that technology is difficult to measure but can increase the productive power of the economy without limit. With the assumption that there are no diminishing returns to technology, there is no need to worry about physical work or the scarcity of any productive input.

The preoccupation with pure technological change as the driver of economic growth has caused earlier neoclassical economists to virtually ignore the critical importance of energy in powering the modern economy [8]. In contrast, many natural scientists and some economists have concluded that the explosion of economic activity during the twentieth century was due principally to the increase in the ability to do work through the expanding use of fossil fuel energy. In fact, the neoclassical economist's technology residual disappeared when energy was included as an input. Energy as a factor of production was *more* important than either capital or labor for Germany, Japan, and the United States in recent decades [6]. Further Ayers and Warr [9] found that most improvements in "technology" have been simply an increase in the quantity of energy used or the efficiency of getting it to the point where the work is done. Although NCE models purport to show that technology alone has driven the industrial economy, historically, it has been a technology that mostly has found new sources of, and applications for, energy.

There are a number of additional, more specific, criticisms that the natural scientist can level against the basic neoclassical model as summarized in Hall et al. 2001 [6]. These criticisms are devastating to the fundamental approach taken by neoclassical economics and taken together mean that there is no possibility that we can assign any validity to the basic neoclassical model.

3.2.4 Specific Criticism 1: Thermodynamics

Contemporary economics and its fundamental household-firm-market model (Fig. 3.1) pays

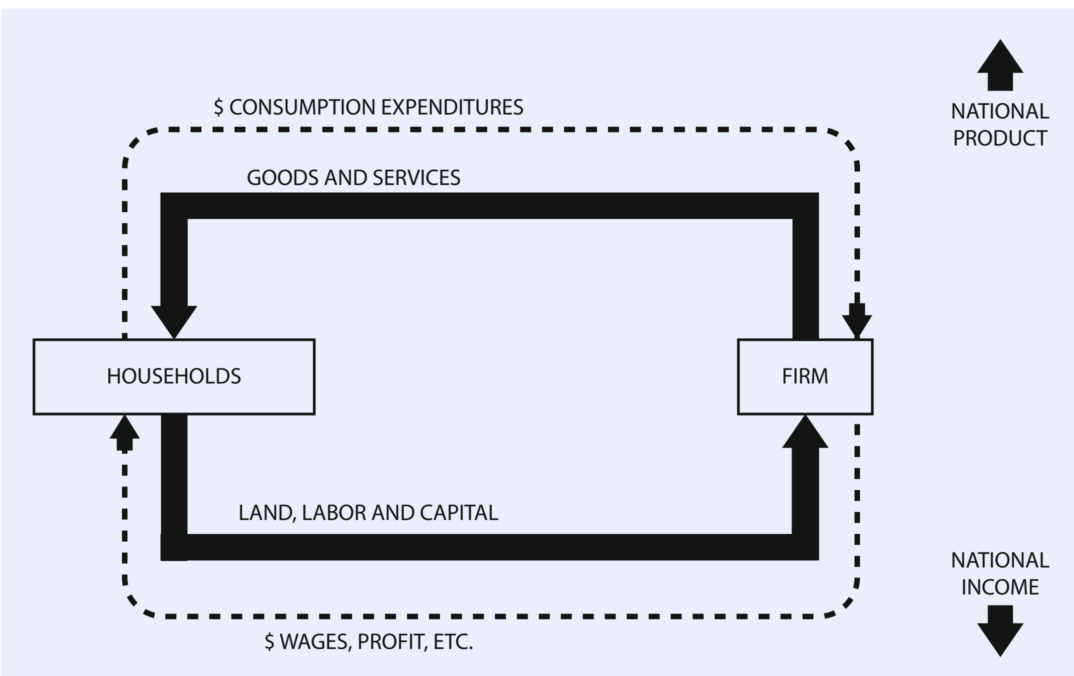


Fig. 3.1 The neoclassical view of how economies work. Households sell or rent land, natural resources, labor, and capital to firms in exchange for rent, wages, and profit (factor payments). Firms combine the factors of production and produce goods and services in return for consumption expenditures, investment, government expenditures, and net exports. This view represents, essentially, a perpetual motion machine. See also Fig. 1.1

only minimal attention to the first law of thermodynamics, and none at all to the second. In fact, the second law is completely incompatible with the conceptual model known as the circular flow. In the circular flow diagram, there is never any value lost to waste or entropy. Specifically, there is no dissipation of the useful work of energy as it is used, and hence no requirement in that model for an input of new energy. This is a serious conceptual flaw and an obstacle to designing economic policies that can meet the challenges of pollution, resource scarcity, and depletion successfully. In effect, the two laws of thermodynamics say, "Nothing happens in the world without energy conversion and entropy production." The consequences are: (1) Every process of industrial and biotic production requires the input of energy. (2) Because of the unavoidable entropy production, the valuable part of energy (called exergy) is transformed into useless heat at the temperature of the environment (called anergy), and usually matter is dissipated, too. This results in pollution and, eventually, the exhaustion of the higher-grade resources of fossil fuels and raw materials. (3) Human labor, powered by food, can be, and was, replaced by energy-driven machines in the course of increasing automation. This has allowed an increase in the productivity of labor, as each worker can do more real work. But it also makes much of labor increasingly superfluous.

Although the first and second laws of thermodynamics are among the most thoroughly tested and validated laws of nature and state explicitly that it is impossible to have a perpetual motion machine (i.e., a machine that performs work without the input of exergy), the basic NCE model *is* a perpetual motion machine, with no material requirements and no limits (◘ Fig. 3.1). Most economists have accepted this incomplete model and have relegated energy and other resources to unimportance in their analysis. Rather than placing the economy within the confines of nature, this approach relegates all the limits of nature to a minor position within a system of self-regulating markets. This attitude was cemented in the minds of most economists by the analysis of Barnett and Morse [34], who found no indication of increasing scarcity of raw materials (as determined by their inflation corrected price) for the first half of the twentieth century. However, their analysis, although cited by nearly all economists interested in the depletion issue, was seriously incomplete. Cutler Cleveland showed that the only reason that decreasing concentrations and qualities of resources were not translated into higher prices for constant quality was because of the decreasing price of energy [10]. Thus, it is only because of the historic abundant availability of many natural resources that economics can assign them low monetary value despite their critical importance to economic production.

3.2.5 Specific Criticism 2: Boundaries

The basic model used in neoclassical economics (◘ Fig. 3.1) does not include boundaries that in any way indicate the physical requirements for, or effects of, economic activities. We believe that at a bare minimum ◘ Fig. 3.1 should be reconstructed as ◘ Fig. 3.2 to include necessary resources and generation of wastes. Taking this assessment one step further, we believe that something like ◘ Fig. 3.3 is the diagram that should be used to represent in more detail the physical reality of an economy's working. It shows the flow of energy and matter across the boundary separating the reservoirs of these "gifts of nature" from the realm of cultural transformation within which sub-boundaries indicate the different stages of their further transformation into the goods and services of final demand. Such a diagram should be presented to every student in an introductory economics course so that the ways the economic process operates in the real world are properly understood. Another way of reflecting the necessary changes is that ◘ Fig. 3.4 shows the standard economist's view of one person's role in the economy, while ◘ Fig. 3.5 gives the biophysical perspective of what biophysical materials are actually needed to operate the economy for one person for 1 year. Superior, and more detailed conceptual models of the biophysical perspective will be found in ▶ Chap. 5.

3.2.6 Specific Criticism 3: Validation

Natural scientists expect theoretical models to be tested before applied or developed further. Unfortunately, economic policy with far reaching consequences is often based on economic models that, although elegant and widely accepted, are not validated. Economists test regularly many hypotheses. Topics such as the effects of income on consumption or the tax rate on economic output are regularly subjected to the rigors of linear regres-

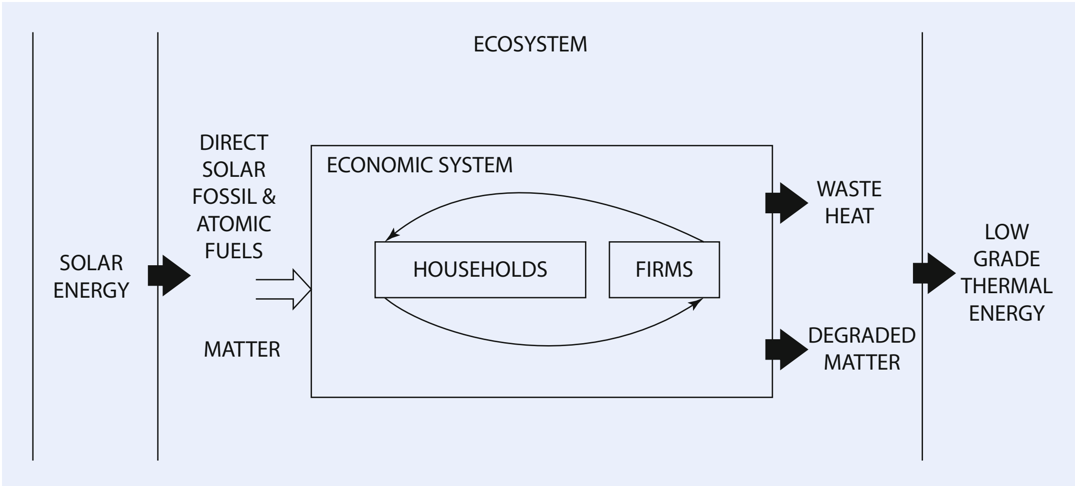


Fig. 3.2 Our perspective, based on a biophysical viewpoint, of the minimum changes required to make Fig. 3.1 conform to reality. We have added the basic energy and material inputs and outputs that are essential if the economic processes represented in Fig. 3.1 are to take place (Source: Daly [3])

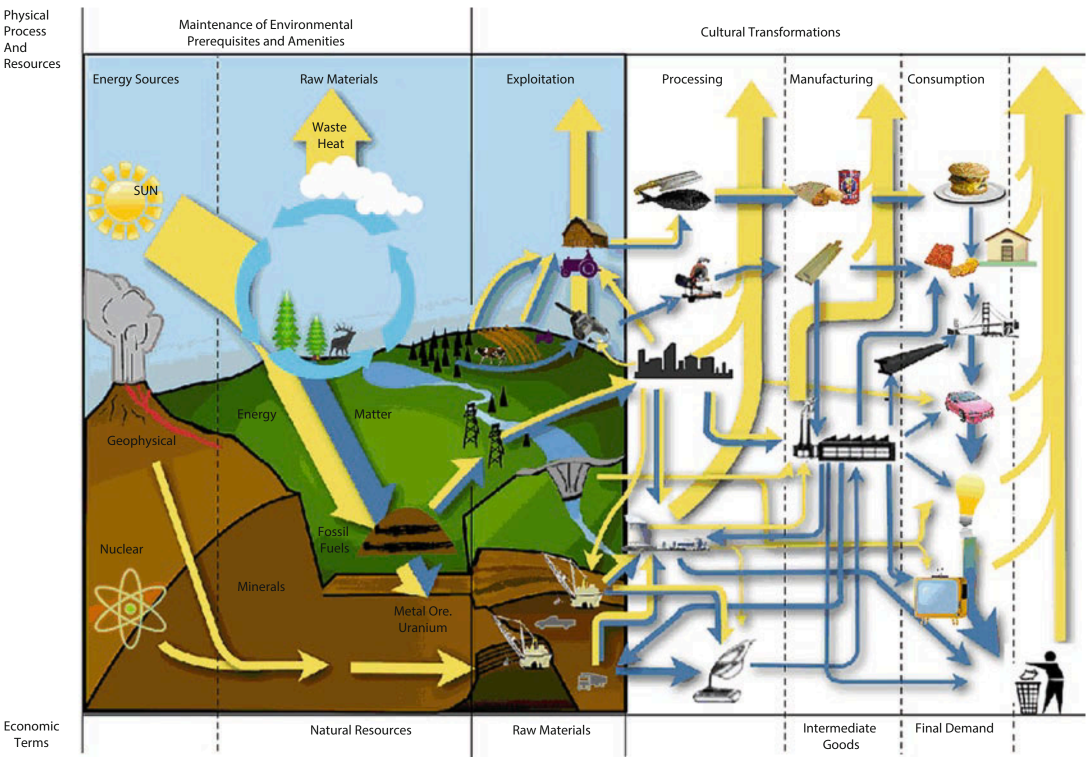


Fig. 3.3 A more comprehensive and accurate model of how real economic systems work. This is the minimal conceptual model that we would accept to represent how real economies actually work. Natural energies drive geological, biological, and chemical cycles that produce natural resources and public service functions. Extractive sectors use economic energies to exploit natural resources and convert them to raw materials. Raw materials are used by manufacturing and other intermediate sectors to produce final goods and services. These final goods and services are distributed by the commercial sector to final demand. Eventually, non-recycled materials and waste heat return to the environment as waste

Fig. 3.4 A conventional economist's view (or perhaps a caricature of that) of one person's inputs and outputs to the process of economic production for 1990 (Source: Hall et al. [32])

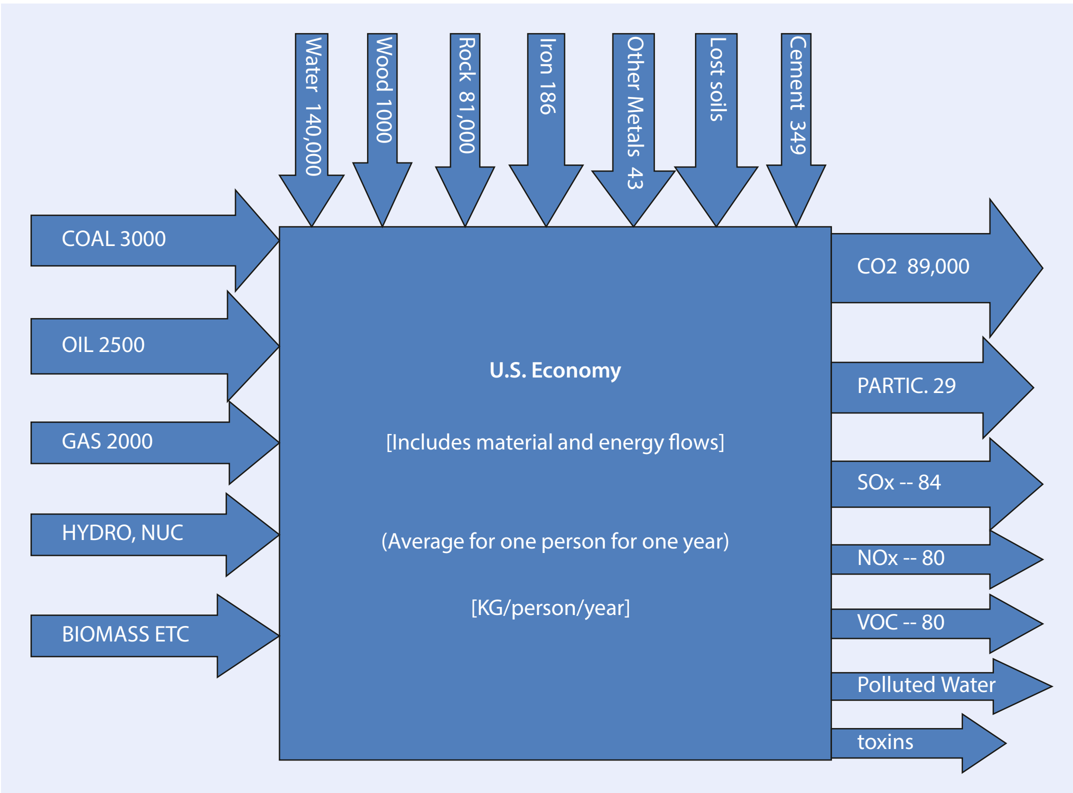


Fig. 3.5 The actual material and energy flows associated with one person's involvement in the economy for the same year (Source: Hall et al. [32])

sion, and even nonlinear statistical methods. However, questions about the ideological worldview of NCE are not often tested. Behavioral assumptions such as rationality, self-regarding preferences, and the connection between higher levels of material consumption and happiness are not always, if ever, tested. Neoclassical economists consider them to be "maintained hypotheses" that do not require empirical verification. Validation also proves difficult or impossible because both classical and neoclassical theories were originally developed using concepts of production factors as they existed in preindustrial and agrarian societies [14]. These theories have been transferred more or less unchanged to applications in the modern industrial world. No provisions have been added to the basic theory for industrialization and its consequences. As the Nobel Laureate in Economics Wassily Leontief noted [12], many economic models are unable "to advance, in any perceptible way, a systematic understanding of the structure and the operations of a real economic system"; instead, they are based on "sets of plausible but entirely arbitrary assumptions" leading to "precisely stated but irrelevant theoretical conclusions."

While we have no argument with the development of theoretical assumptions or models, they normally should be put forth as hypotheses, that is, as a good assumption or guess as to how something operates. This is how the scientific method works, and this is the most powerful way that we have to find out how the world actually works. Then the hypotheses can be tested, and if it stands up well can be advanced to a theory or perhaps eventually a law. But although there are some economists who appropriately use hypotheses, there has been no attempt to build up the main theoretical model of economics as a set of testable and tested hypotheses. Instead economics is constructed as a series of logical constructs that make a certain sense (from a limited perspective), but hardly encompass how real economic systems operate. We believe economists should adopt this perspective and test the supposedly maintained hypotheses, instead of treating belief in the self-regulating market as a matter of faith!.

Most noneconomists do not appreciate the degree to which contemporary economics is laden with arbitrary assumptions. Nominally objective operations, such as determining the least cost for a project, evaluating costs and benefits, or calculating the total cost of a project, normally use explicit and supposedly objective economic criteria. In fact, such "objective" analyses, based on arbitrary and convenient assumptions, produce logically and mathematically tractable, but not necessarily correct, models.

The authority economists often assign to their "physics-based" models, starting with the basic neoclassical model of the economy, are somewhat curious. In neoclassical production theory, the price vector is given by the gradient of the output in the space of the production factors just as the vector of a conservative physical force is given by the gradient of potential energy in real space [13]. The quite imperfect economic analogy should not be confused with the thermodynamically rigorous model in physics, and unavoidably fuzzy economic models should not become more precise simply because they distantly share concepts.

3.2.7 Myth 2: A Theory of Consumption Can Ignore Actual Human Behavior

The second main way that conventional neoclassical economic models are unrealistic is that the model assumes that humans behave as individuals and do not care what others think of them. These are referred to as "self-regarding preferences." Yet, we have known since the time of Aristotle that humans are social animals. Few of us would want to live in total isolation, no matter how many creature comforts we might possess. Interestingly enough, most economists pay homage to Adam Smith, but few have ever read him in the original. If you chose not to follow this path, we suggest you read Smith's *Theory of Moral Sentiments,* in which he spends hundreds of pages detailing how social approval governs our behavior, and that humans have an altruistic side as well as an individualistic one. But just as NCE production assumptions violate principles of physics, its assumptions about human behavior are inconsistent with both a large body of psychological and neurological research and even everyday human experience. It is well established that real human beings are other regarding, that is, how one person values a certain economic outcome depends on how much it is valued by others. It is also well established that the consumption of market goods cannot be equated with an individual's happiness. Nevertheless, the fundamental behavioral assumptions of NCE require self-regarding consumers whose happiness depends essentially, or even only, upon their consumption of market goods. The cultural context of behavior is deemed irrelevant to neoclassical economic analysis as the emphasis is entirely on the behavior of the isolated individual.

3.2.8 Myth 2a: *Homo Economicus* Is a Scientific Model That Does a Good Job of Predicting Human Behavior

At the heart of standard neoclassical economic theory is the model of human behavior embodied in *homo economicus* or "economic man." Economic texts usually begin with a very general statement about human nature that is soon codified into a set of rigid mathematical principles resting upon the idea that "people maximize their well-being by consuming market goods according to self-regarding, consistent, constant, well-ordered, and well-behaved preferences." However, the assumption that people are entirely, or mostly, self-regarding has been shown to be false by considerable contemporary work in behavioral economics, neuro-economics, and game theory [15–17]. For example, Henrich and colleagues, after examining the results of behavioral experiments in 15 societies ranging from hunter-gatherers in Tanzania and Paraguay to nomadic herders in Mongolia concluded: "[T]he canonical [NCE] model is not supported in any society studied." In experimental settings and under real-world conditions, humans consistently make decisions that favor enforcing social norms over ones that lead to their own material gains [18]. Gintis describes several experiments showing that humans are both far more altruistic and far more vindictive than the NCE "rational" actor model allows. They will make decisions to punish persons they will never again encounter if those people "cheat" in experimental transactions, even if this means considerable monetary loss to themselves. Rather than humans being simply self-regarding, they have a high regard for seeing that others "follow the rules" and treat other people decently.

The centrality of the behavior of isolated individuals is reflected in the notion that consumers are sovereign, meaning independent in their behavior, in a market economy. Ackerman and Heinzerling [19] point out that the rise of economic orthodoxy put consumers at the center of analysis. The idea is that producers respond to consumer preferences rather than the reverse. Yet we all know that, in fact, consumer tastes are both grossly and subtly manipulated and that firms barrage us with advertising to increase their market share. Nonetheless, the centrality and preeminence of the individual in orthodox economic analysis precludes any analysis or emphasis on the context of individual behavior.

3.2.9 Myth 2b: Consumption of Market Goods Can Be Equated with Well-Being and Money Is a Universal Substitute for Anything

Most economic texts simply equate utility with happiness and assume that utility can be measured indirectly by income without any substantive or formal discussion of the matter [20]. The higher the income, the better off an individual (and hence society) is supposed to be. Yet there is considerable evidence that, past a certain point, income is a positional good; that is, if everyone's income goes up there is little or no long-term gain in social well-being. This implies that policies designed merely to increase per capita income may have a negligible effect if the goal is to improve social welfare.

Psychologists have long argued and documented that well-being derives from a wide variety of individual, social, and genetic factors. These include genetic predisposition, health, close relationships, marriage, and education—as well as income [20]. It is generally true that people in wealthier countries are happier than people in poorer countries, but even this correlation is weak, and the happiness data show many anomalies [21]. For example, some surveys show that people in Nigeria are happier than wealthier people in Austria, France, and Japan [22–24]. Past a certain stage of development, increasing incomes do not lead to greater happiness. For example, real per capita income in the United States has increased sharply in recent decades, but reported happiness has declined [25].

When economists equate utility with income in the NCE model, this affects the policy recommendations of economists which in turn impact the natural world. According to Arrow and colleagues [26], "sustainability" means simply maintaining the discounted flow of income over time. Leaving future generations with the same or greater real income than the present leaves them at least as well-off no matter what happens to specific features of the natural world. By this reasoning if the present discounted value of a rainforest is $1 billion in ecosystem services if left intact, but can generate a discounted investment flow of $2 billion if it is clear cut and sold, then it is the moral responsibility of the present generation to cut down the rainforest. With $2 billion the future generation could buy another rainforest or something of equal value and have $1 billion left over. This is the logic used by some economists to justify the destruction of a substantial portion of the planet's ecosystems and species [27].

3.2.10 How the Neoclassical Model Fails to Deal with Distributional Issues

A different but extremely important and pungent critique of the neoclassical model comes from recent work by John Gowdy [27, 28]. Gowdy takes as his starting point the welfare model of John

Rawls. (Here "welfare" is the same as "utility.") The basis of *welfare economics* is that each individual gains welfare proportional as his or her real disposable income increases. Thus, a given individual will be "better off" by a factor of two if he or she has 2000 rather than 1000 dollars to spend (or if prices are half as much). This concept also uses the idea of Pareto optimization. Both the Rawlsian and the Pareto approach assume that there is a linear relation between individual welfare and money. Thus, if one individual becomes five times wealthier (say from $1000 to $ 5000), that is as great a social good as five people becoming twice as wealthy (say from $1000 to $2000 each). This is an important concept that lies behind welfare economics and has been used incessantly as a logic for developmental plans that tend to pay most attention to increasing GNP and relatively little attention to *who* gets the proceeds. This of course avoids the contentions within the developing world that development tends to enrich those who have, while doing little, or even impoverishing, those that have not. By the Rawlsian-Pareto logic, or at least as employed by most contemporary neoclassical economists, if the total wealth is increased the distribution is not important, or at most is quite secondary. The entire economic perspective is often associated with social notions that people are well-off or not in accordance to their own efforts rather than due to factors outside their control.

Gowdy argues against the economists' position that distribution is not an important issue by summarizing considerable recent psychological investigation that shows that human welfare and happiness does not increase linearly with income, but rather is curved downward. Hence supplying poor people with the basic necessities of life generates a greater deal of happiness and welfare with a given amount of money compared to much less happiness or well-being generated by the same amount of money in the hands of someone who is well-off. Curiously this is a conclusion also reached by thinking about the concept of marginal value—that the first units of something have much more value than additional units—a fact conveniently ignored by marginalist neoclassical economists! Instead the marginal utility of money is assumed to be constant. If it were not, the neoclassical theory of income distribution could not produce efficiency and equity. Finally, according to Gowdy and Gintis, the extensive social research done in recent years has completely undermined the "value neutral" assumptions that are the base of welfare and neoclassical economics and calls into question all the basic tenets of neoclassical economics.

3.2.11 What Economists Think of These Ideas

Mostly conventional economists do not think at all about these problems with conventional economics but stick very closely to the accepted neoclassical model. But there are some partial exceptions. The Nobel Laureate in Economics Robert M. Solow considered the possibility in 1974 that "The world can, in effect, get along without natural resources" because of the technological options for the substitution of other factors for nonrenewable resources [11]. More recently, Solow stated "It is of the essence that production cannot take place without some use of natural resources." Clearly, there is need for more analytical and empirical work (some of which we provide in later chapters) on the relation between economic production and natural resources, especially energy, and how much of the resources are actually needed. Many economists today, including many recent Nobel Prize winners (e.g., Akerlof, Krugman, Sen, Stiglitz) have very serious reservations with the contemporary model, although none has explicitly endorsed the biophysical alternative.

We might ask why economists pay so little attention to the biophysical alternative. The conventional neoclassical view of the low importance of energy and materials goes back to the early days of neoclassical economics. Initially, the focus was not so much on the generation of wealth but rather on the "efficiency of markets" and the distribution of wealth. The model of pure exchange of goods starts without considering their production. With a set of mathematical assumptions on rational consumer behavior, it was shown that through the exchange of goods in markets, an equilibrium situation results in which all consumers maximize their utility. This benefit of (perfect) markets is generally considered the foundation of free market economics. It shows why markets, where greedy or at least "self-regarding" individuals meet, work at all. Later, when the model was extended to include production, the problem of the physical generation of wealth had to be inseparably coupled to the problem of the distribution of wealth. In the neoclassical concept of equilibrium, the activity of profit maximizing entrepreneurial behavior generates the situation where factor productivities (e.g., the respective contribution of capital, labor, and energy) equals factor prices. This means that in conventional economic analysis, the weights which the production factors contribute to the physical generation of wealth are determined by, and evaluated by, the factor cost shares. Thus, energy's importance is assumed by most economists to be equal (only) to its cost, which typically is small, only 5–10% of the cost of all goods and services.

Unlike their classical predecessors, neoclassical economists do not even bother to include the process of how things are actually made in their analyses. They just take the input prices and put them into a function, and the price and quantity of output are automatically generated. Here lies the historical source of the economists' underestimation of energy as a production factor, because in industrial market economies energy cost, on the average, is only 5–6% of the total factor cost (and of GDP). Therefore, economists either neglect energy as a factor of production altogether, or they argue that the contribution of a change of energy input to the change of output is equal only to energy's small cost share of 5–6%. This has led to a long-lasting debate on the impact of the two energy price explosions in the years 1973–1975 and 1979–1981 when the cost of energy increased to 14% of GDP even while supplying less physical energy. As we show below, and more explicitly in Hall et al. (2001), energy is more important in production than either labor or capital, although all three are needed. Curiously energy's low price is the reason for its *importance*, *not its unimportance*. For 200 years the economy has received huge benefits from energy without having to divert much of its output to get it. This is because basically we do not pay nature for energy, but only the cost of exploiting it. Likewise, the finite emission absorption capacity of the biosphere is more important to future economic growth than its present (nearly vanishing) price seems to indicate.

Neoclassical models built on the assumptions of ◘ Fig. 3.1 cannot explain the empirically observed growth of output by the growth of the factor inputs. There always remains a large residual (i.e., a statistical "leftover" that is not explained by the factors used in the analysis, in this case, capital and labor). This is formally attributed to what economists call either "technological progress" or improvements in "human capital," which are long-term increases in skill and education of workers. Even Robert Solow stated, "This ... has led to a criticism of the neoclassical model: it is a theory of growth that leaves the main factor in economic growth unexplained" [11]. As we will argue below, weighting a factor by its cost share is an incorrect approach in growth theory.

In fact, the human economy uses fossil and other fuels to support and empower labor and to produce and utilize capital. Energy, capital, and labor are then combined to upgrade natural resources to useful goods and services. Therefore, economic production can be viewed as the process of upgrading matter into highly ordered (thermodynamically improbable) structures, both physical structures and information. Where the economist speaks of "adding value" at successive stages of production, one may also speak of "adding order" to matter through the use of free, or unbounded and available, energy (exergy). The perspective of examining economics in the "hard sphere" of physical production, where energy and material stocks and flows are important, is called biophysical economics. It must complement the social sphere perspective.

3.2.12 Why Theory Matters

It is in the policy arena that the ideological nature of NCE reveals itself most completely. Most economists substitute the mythical NCE world of rational agents, certainty, and perfect information for the complex reality and uncertainty of real economies. Where reality and the neoclassical model disagree, reality is increasingly forced through policy to conform to the neoclassical model [29]. Neoclassical economists generally assume that people always respond rationally and consistently to price signals; therefore, the goal of economic policy is to assign property rights and "get the prices right." The corollary assumption is that things of value to people have a price, and anything without a market formed price must lack value. Prices are theoretically capable of reflecting all the relevant attributes of any good or service and all that people value. The rest of us are asked to take the validity of these assumptions and analyses on faith and to turn our complex decision-making increasingly over to barely regulated markets and cost-benefit analyses. This emphasis frequently leads to fundamental policy-related failures and problems that include the following:

1. The ultimate policy goal of NCE is not to correct any particular problem directly but rather to correctly value the problem in terms of everything else so that the "calculating machine" of the market can establish the pecking order of priorities. The focus on establishing "general market equilibrium" frequently means neglecting essential details of the policy problems under consideration, especially those for which it is difficult or impossible to determine a price (i.e., oil depletion, environmental degradation, and global climate change). Hence when we purchase a gallon of gasoline, we pay only for getting that gallon to the pump, not for finding a new gallon to replace it, or something else if oil depletion makes replacement impossible.
2. The NCE model makes no qualitative difference between needs and wants, or among commodities produced, or among specific productive inputs, including energy. Everything we find useful is treated like an abstract commodity substitutable for and by anything else. Absolute scarcity does not exist nor, within certain broad limits, are any specific conditions deemed necessary for human existence. Value is a relative matter expressed in relative prices. Because no single thing is essential, substitution among resources and commodities will occur until the marginal value of a commodity is the same for all commodities. At this point, rational individuals have made optimal choices, and the sum of all optimal choices leads us to the "best of all possible worlds." Thus, the tastes of affluent teenagers in malls for unnecessary but heavily advertised clothes or gadgets are given as much weight per dollar spent as health care or education for the less affluent.
3. The model assumes that aggregate income is a complete and sufficient measure of well-being. Operationally this means that total costs and benefits of policies can be determined by merely adding the monetary changes in the incomes of all isolated individuals affected. This implies that relative income effects don't matter to the individual—for example, a loss of $1000 to a poor person can be more than compensated for by a gain in $1100 to a billionaire. Similarly, neoclassical economists consider preferences to be exogenous to social context. Yet numerous studies have found that relative income effects matter and sometimes these effects can completely cancel out increases in total income which is always the primary goal of NCE. How much one person values a gain or loss depends on what others get, the income of each person relative to others, the "fairness" of the income change, and a variety of other social factors which are not included in the NCE model.
4. "Sustainability" in the NCE model means sustaining only the discounted flow of per capita income, not anything else such as biodiversity, oil stocks, human health, or social cohesiveness. This is known as weak sustainability. However, to live within nature's limits, we need to arrive at the conditions of strong sustainability, which requires that the profits from the depletion of a resource or degradation of an ecosystem are reinvested in developing alternatives or restoring degraded systems. This entails looking at the bigger picture of how market systems function and interface with the biophysical world [29–32]. Consequently one cannot arrive at a social decision to achieve an optimal macroeconomic scale by merely aggregating many separate efficient market outcomes.
5. Perhaps most importantly the neoclassical model has nothing to say about the relative power of diverse groups of people to influence the "free market" through influencing politicians with expensive contributions, through supporting advertisements in the media, or simply through their own massive purchasing patterns. The consequence has been to increasingly make the rich richer and the poor poorer. The advertising campaign against the role of government has undercut many programs that have helped alleviate somewhat the difference between the rich and the poor. There is a rich literature on this subject [33], much of it extremely critical of the neoclassical model, but much of the public still believes that markets are the best way to distribute economic goods and services

despite the lack of compelling evidence that this is true. For example, Sekera has demonstrated clearly that government can deliver services more efficiently than private entities, but few citizens seem to understand that. The work of Piketty and Sekera, along with that of other income distribution scholars, will be developed in more detail in ▸ Chap. 23.

NCE dominates policy making yet provides an inadequate toolbox for confronting the major problems of the present world: global climate change, biodiversity loss, oil depletion, loss of wilderness, and the recalcitrant problems of poverty and social conflict. It has been used as the basis for "the Washington Consensus" which has been and continues to be exported to the developing world with essentially no assessment of its effectiveness or basis in reality and with enormous social and environmental problems [30, 31]. We are led to believe that our most pressing environmental and social problems can be dealt with by simulating efficient market outcomes as if this alone provides the elixir for all that ails us. Yet we know that the concept of market efficiency rests on an untenable and faulty foundation and that the real market economy is not best described in this framework. The perpetuation of neoclassical economics, usually to the exclusion of other possible approaches, is essentially the substitution of faith for reason, science, and empirical testing in many areas of economics. We must move beyond this "faith-based" economics and find a more illuminating way of understanding economic activity and informing decision-making so that our policies will amount to something more than window dressing for the status quo.

? Questions

1. What are some of the "myths" of neoclassical economics? Do you agree that these are myths? Why or why not?
2. Why is the circular flow model of the economy inconsistent with the laws of thermodynamics? Is that possible?
3. Nicholas Georgescu-Roegen and his student Herman Daly are economists. Why are they such critics of conventional economics?
4. Economic productivity in neoclassical economics is usually represented as a function of capital and labor. Do you agree with that perspective? Why or why not?
5. What, in your opinion, should be the proper boundaries to be used in economic analysis? Can you draw a picture of how you would represent these boundaries?
6. What does validation mean? Why is this often difficult for economic models?
7. What are thought to be (within conventional economics) the main characteristics of *homo economicus* (or "economic man")?
8. Do you think that having greater amounts of money to spend will make you happier? Why or why not? Do you think wealthier people that you know are happier than poorer people?
9. Does an increase in income of, say, 1000 dollars have the same meaning for a wealthy person as for a poor person? How does that relate to the usual economist's position on Pareto optimality?
10. Why have neoclassical economists attempted to generate a "value neutral" approach to economics? To what degree have they succeeded, in your opinion?
11. Why does theory matter in economics?
12. What does sustainability usually mean within conventional economics? What might be some problems with that definition?

References

1. This chapter is derived in part from Gowdy, J., C. Hall, K. Klitgaard, and L. Krall. 2010. The end of faith-based economics. *The Corporate Examiner* 37: 5–11, and Hall, C., D. Lindenberger, R. Kummel, T. Kroeger, and W. Eichhorn. 2001. The need to reintegrate the natural sciences with economics. *BioScience* 51: 663–673.
2. Georgescu-Roegen, N. 1975. Energy and economic myths. *Southern Economic Journal* 41: 347–381.
3. Daly, H. 1977. *Steady-state economics*. San Francisco: W. H. Freeman.
4. Cleveland, C., R. Costanza, C. Hall, and R. Kaufmann. 1984. Energy and the U.S. economy: A Biophysical perspective. *Science* 225: 890–897.
5. Hall, C., C. Cleveland, and R. Kaufmann. 1986. *Energy and resource quality: The ecology of the economic process*. New York: Wiley Interscience.
6. Hall, C., D. Lindenberger, R. Kummel, T. Kroeger, and W. Eichhorn. 2001. The need to reintegrate the natural sciences with economics. *BioScience* 51: 663–673. Also: Kummel, R., J. Henn, and D. Lindenberger. 2002. Capital, labor, energy and creativity: Modeling. *Structural Change and Economic Dynamics* 3 (2002): 415–433.

7. Wilson, E. 1998. *Consilience: The unity of knowledge*. New York: Alfred Knopf.
8. Denison, E.F. 1989. *Estimates of productivity change by industry, an evaluation and an alternative*. Washington, DC: The Brookings Institution.
9. Ayres, R., and D. Warr. 2005. Accounting for growth: the role of physical work. *Change and Economic Dynamics* 16: 211–220.
10. Cleveland, C.J. 1991. Natural resource scarcity and economic growth revisited: Economic and biophysical perspectives. In *Ecological economics: The science and management of sustainability*, ed. R. Costanza, 289–317. New York: Columbia University Press.
11. Solow, R.M. 1974. The economics of resources or the resources of economics. *American Economic Review* 66: 1–14. Also Solow, R.M. 1994. Perspectives on growth theory. *Journal of Economic Perspectives* 8: 45–54.
12. Leontief, W. 1982. Academic economics. *Science* 217: 104.
13. Mirowski, P. 1989. *More heat than light*. Cambridge: Cambridge University Press.
14. McCauley, J.L., and C.M. Kuffner. 2004. Economic system dynamics. *Discrete Dynamics in Nature and Society* 1 (2004): 213–220.
15. Gintis, H. 2000. Beyond Homo economicus: Evidence from experimental economics. *Ecological Economics*. 35: 311–322.
16. Camerer, C., Loewenstein, G., 2004. Behavioral economics: past present and future. In: Camerer, C, Loewenstein, G. Rabin, M (Editors), Advances in behavioral economics, Princeton/Oxford, Princeton University Press 3–52.
17. Henrich, J., et al. 2001. Cooperation, reciprocity and punishment in fifteen small-scale societies. *American Economics Review* 91: 73–78.
18. http://www.youtube.com/watch?v=u6XAPnuFjJc
19. Ackerman, F., and L. Heinzerling. 2004. *Priceless: On knowing the price of everything and the value of nothing*. New York/London: The New Press.
20. Frey, B., and A. Stutzer. 2002. *Happiness and economics: How the economy and institutions affect well-being*. Princeton: Princeton University Press.
21. Diener, E., M. Diener, and C. Diener. 1995. Factors predicting the well-being of nations. *Journal of Personality and Social Psychology* 69: 851–864.
22. Brickman, P., D. Coates, and R. Janoff-Bulman. 1978. Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology* 36: 917–927.
23. Blanchflower, D., and D. Oswald. 2000. *Well-being over time in Britain and the U.S.A.*, NBER working paper no.7481. Cambridge, MA: National Bureau of Economic Analysis.
24. Lane, R. 2000. *The loss of happiness in market economies*. New Haven/London: Yale University Press.
25. Meyers, D. 2000. The funds, friends, and faith of happy people. *American Psychologist* 55: 56–56.
26. Arrow, K., P. Dasgupta, L. Goulder, G. Daily, P. Ehrlich, G. Heal, S. Levin, K. Goran-Maler, S. Schneider, D. Starrett, and B. Walker. 2004. Are we consuming too much? *Journal of Economic Perspectives* 18: 147–172.
27. Gowdy, J. 2004. The revolution in welfare economics and its implications for environmental valuation. *Land Economics* 80: 239–257.
28. Gowdy, J., and J. Erickson. 2005. The approach of ecological economics. *Cambridge Journal of Economics*. 29 (2): 207–222.
29. Makgetla, N., and R. Sideman. 1989. The applicability of law and economics to policymaking in the third world. *Journal of Economic Issues* 23: 35–78.
30. Hall, C. 2000. *Quantifying sustainable development: The future of tropical economies*. San Diego: Academic Press.
31. LeClerc, G., and Charles Hall, eds. 2008. *Making development work: A new Role for science*. Albuquerque: University of New Mexico Press.
32. Hall, C.A.S., R.G. Pontius, L. Coleman, and J.-Y. Ko. 1994. The environmental consequences of having a baby in the United State. *Population and the Environment* 15: 505–523.
33. Sekera, June. 2016. *The Public Economy in Crisis*. New York: Springer; Piketty, Thomas. 2014. *Capital in the Twenty-first Century*. Cambridge, Massachusetts: The Belknap Press.
34. Barnett, H., and C. Morse. 1963. *Scarcity and Growth*. Washington, D.C.: Resources for the Future.

Biophysical Economics: The Material Basis

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_4

In our first chapter, we provided a review of how modern (neoclassical) economics operates as a social science-based means of allocating "scarce" resources, including the philosophy behind the assumptions that govern the operation of that model of how economies work. In this approach markets are seen as especially important as a means of making economic decisions and guiding the allocation of productive resources.

▶ Chapter 2 reviewed earlier approaches to economics, many of which were based on a more explicit understanding and acknowledgment of the biophysical basis of real economies. Thus, while most people who do think about economics today probably believe that the conceptual model (neoclassical circular flow model) that dominates economics is the only possible and proper way to think about economics, there are many alternatives. In fact, as was obvious from ▶ Chap. 2, there are many very different ways we can think about economics that accurately describe at least some important aspects of what is going on in real economies. We just happen to live in a time when there is a dominant form that excludes other world views of what constitutes economics.

Many criticisms have been leveled at this dominant "neoclassical" model. ▶ Chapter 3 undertakes a thorough and damning review of the many problems that exist with the intellectual basis of conventional economics. It focuses specially on the conceptual and logical problems that arise from assuming that economics should be based only on the social sciences even while the basis for actually existing economies is the production and transport of goods and the provisions of services, all of which must take place in the real world of matter and energy and hence are best studied using the natural sciences. Professional economists as a group tend to be uninformed about, or uninterested in, the criticisms that have been leveled at their discipline. In a sense, they have been successful at circling their wagons to protect their core beliefs, ignoring the criticisms, and proceeding with their craft, independent of the criticisms or the degree to which it is or is not successful in describing reality or making predictions.

In the next two chapters, we introduce the reader to another equally or, we believe, more appropriate and accurate approach to economics—biophysical economics. The concept has a very old history, starting with the recognition by whatever might have passed for economists in the Stone Age that one's material well-being depended upon nature and those things that humans might be able to extract from nature and the difficulty or ease in doing so. As humans eased into the first stages of agriculture, we know that they paid a great deal of attention to the material conditions of their economic life due to the large part of whatever wealth they had that was "invested" into observatories, temples, and activities that attempted to understand and beseech their gods to provide rains, good harvests, and so on. The people may not have understood well the forces that generated or not their economic production, but they knew them as important. The work of Anthony Aveni, for example, has led to an entire new discipline of archeoastronomy. He has shown convincingly that entire cities (such as the area around the temples of the sun and moon in Mexico) were constructed to determine the movement of the sun relative to the Earth, leading to a better understanding of planting times.

While we cannot interview such people, as they are long dead, we can examine (or could until recently) the various cultures around the world that are little touched by industrialization to see how they operate. Do they in fact operate in a way consistent with the assumptions of modern neoclassical economists? When the anthropologist Karl Polanyi undertook an examination of a large series of preindustrial "folk" societies, he found that while market transactions had always existed, most people traded their surplus goods [1]. Things were not produced specifically for sale, and markets did not form prices. Societies allocated what we now call goods and services by means of trade, reciprocity, and redistribution. In other words, economics was based more on their material basis than on money. We call this material basis "biophysical economics."

4.1 Background to Biophysical Economics

As we stated in the first sentence of the introduction to this first section, economies exist independently of how we perceive or choose to study them. Also, we noted that economists have chosen over the past 150 years—for more or less accidental reasons—the social sciences and an inappropriate and overly simplified analytical model borrowed from physics as the essential conceptual base for undertaking our definitions and analyses of economies and economic systems. This is the case, even though actual economies are as much about biophysical as social activities. Actually existing economies must be based on many things, including the physical materials and the energy required to provide goods and services as well as the NCE-sanctified market interactions that transfer these goods and services from firm to household or household to firm. This was well understood by earlier physiocratic and classical economists. Curiously, starting in the 1870s or so, economics somehow became *only* a social science, and it has remained that way for the most part. In this social science-based model, the material world is represented only by the prices of things within the material world.

4.2 What Is Biophysical Economics?

Biophysical economics is a system of economic analysis that is based on the biological and physical properties, structures, and processes of real economic systems as its conceptual base and fundamental model. It has two components: biophysical science itself and economic analysis that is consistent with science and other social sciences such as psychology and anthropology. It acknowledges that the basis for nearly all wealth is nature and views most human economic activity as a means to increase (directly or indirectly) the exploitation of nature to generate more wealth. As such, it focuses on the structure and function of real economies from an energy and material perspective, although it often considers the relation of this structure and function to human welfare and to the money (i.e., dollar) flows that tend to go in the opposite direction to energy [2]. From a biophysical perspective, one's job is viewed as trading one's time at work (the monetary value of which is related to the energy flows of society controlled by the individual) for access through wages and salaries to the energy flows of the general economy. This "general economy" contains goods and services created from the extraction of energy from the earth in anticipation of some demand for them. At present, each dollar we spend requires roughly 5 megajoules (about half an 8 oz. coffee cup's worth of oil or equivalent energy) to generate the good or service purchased. With economic inflation, the energy per dollar decreases over time so that in 1970, one could receive about ten times more energy (as used to generate goods and services) per dollar than he or she can today. The ice cream that fueled Hall's paper route in 1954 cost only 5 cents, but required for its production roughly the same amount of energy as today. A biophysical economist might ask "how many minutes of labor did you have to put in to earn that nickel? At your current salary, do you put in more or fewer minutes for that ice cream cone? If your salary is high, is it commensurate with the energy flow in society that you control?" Or perhaps "when you spend your salary, how much of the world's nonrenewable resources are depleted, and how much did you contribute to changes in the atmosphere?"

Figure 3.1 is the "firm and household" diagram said to represent the economy in most introductory economics textbooks. We find this model, which represents the basis of most economic theory and teaching, to be less than useful in representing the real things that must occur within a real economy. As developed in Hall et al., this representation violates the laws of thermodynamics (which nothing real can do), has completely inadequate and incorrect boundaries, and has not been tested using the scientific method [3]. Our perception of the simplest diagram that one could use to represent a real economy, which is far more complex and infinitely more accurate than Fig. 3.1, is Fig. 3.3. This diagram, and real economies, includes (from left to right) (1) energy sources (principally, the sun) that are essential for any economy; (2) the material that circulates upon the earth's surface through natural and seminatural ecosystems; and (3) the human-dominated steps of exploitation, processing, manufacturing, and consumption. Blue and yellow arrows show the transfer of materials and energy through the economy. Raw materials are refined by human activities using fossil fuels until the heat is dissipated and the materials are either released as wastes to the environment or recycled back into the system. From this diagram, one could argue that the most important activity of the economic process is the proper functioning of the hydrological cycle, since virtually all economic production and manufacturing are extremely water intensive. From the standpoint of a traditional economist, the hydrological cycle is not important because we pay next to nothing for it. A biophysical

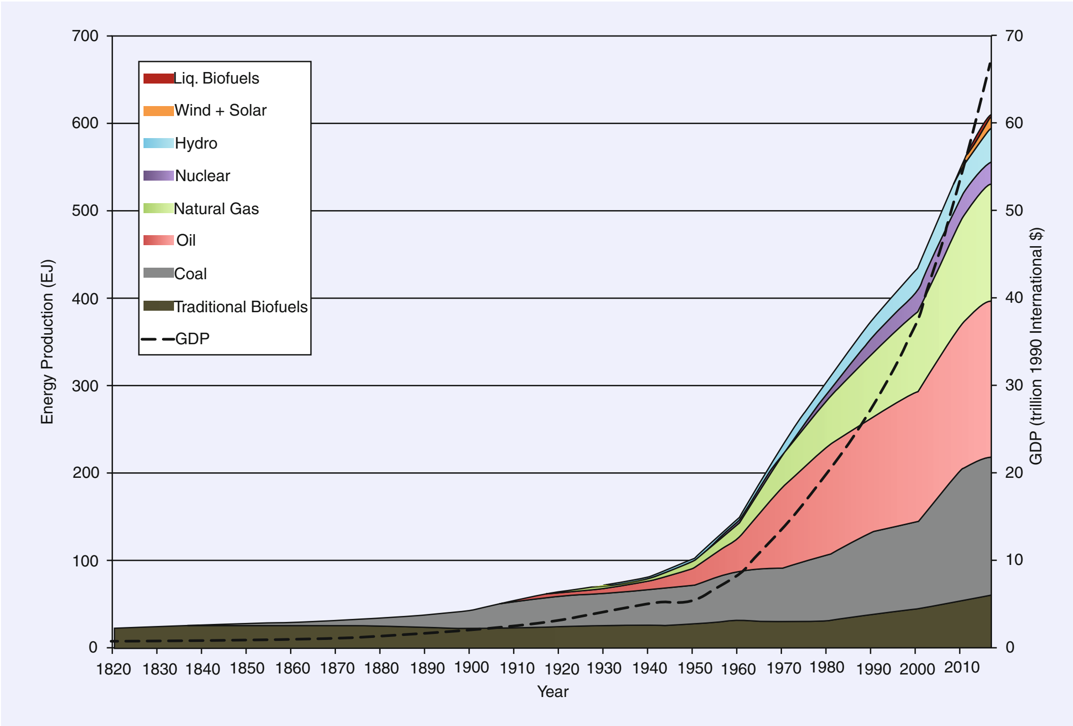


Fig. 4.1 Increase in fossil energy use and economic activity for the world

economist, on the other hand, would argue that it is critical for many reasons and that it is only because we can extract its services from nature at little direct monetary cost that we can have the high generation of wealth within today's economy.

A fundamental premise of biophysical economics is that wealth is produced basically by the application of energy, initially human muscles, draft animals or wood, and increasingly fossil fuels, to the resources of nature to generate wealth (Fig. 4.1). This can readily be seen from several pictures of agricultural harvesters (Figs. 4.2 and 4.3). Studies of the cost of energy to society show that energy has become much cheaper over past centuries as fossil fuels were exploited (Fig. 4.4).

4.3 Conceptual Sources of Biophysical Economics

Biophysical economics derives from three main sources of ideas: (1) earlier thinking by economists, such as François Quesnay and the eighteenth-century physiocrats, who called themselves "Les Economistes," and a few economists of the latter part of the nineteenth and the beginning of the twentieth centuries; (2) conceptual thinking about how ecosystems operate; and (3) scholars from various disciplines at the end of the twentieth century who introduced a new perspective about the limitation of the Earth to support an increasing human population. All of these ideas first came together under the word "biophysical" in a 1984 cover article by Cleveland et al. in *Science* entitled "A Biophysical Analysis of the United States Economy." These concepts gained a further following by the great interest in the "peak oil" movement of the first decade of the twenty-first century and a series of meetings on BioPhysical Economics in Syracuse, New York, starting in 2008. The participants formalized the International Society of BioPhysical Economics in 2015, and the organization continues to meet on an annual basis. These ideas are developed in more detail below.

Fig. 4.2 33 Horse-power combine in about 1900

Fig. 4.3 200 Horse-power combine in 2000

4.4 Biophysical Basis of Early Economists

The present social science focus of economics and economists was not particularly the case with earlier economists, who were more likely to ask "where does wealth come from?" than are most mainstream economists today. In general, these earlier economists started their economic analysis with the natural biophysical world, probably simply because they had common sense but also because they deemed inadequate the perspective of earlier mercantilists who had emphasized sources of wealth as "treasure" (e.g., precious metals) derived from mining or trade. The first formal school of economics, the French Physiocrats, focused on land as the basis for generating wealth [4]. The biophysical perspective continued with Thomas Malthus' famous "Essay on the Principle of Population," (there were six of them) which assumed that human populations would grow exponentially—because it seemed unlikely that anyone, other than the well-born, would control the "passion between the sexes"—unless somehow "checked" by factors that either reduced the birth rate or increased the death rate. Since Malthus had little faith in the "moral restraint" of the working classes and believed that birth control was "vice," he recommended a rather draconian social policy to increase the death rates among the poor. In Malthus'

4

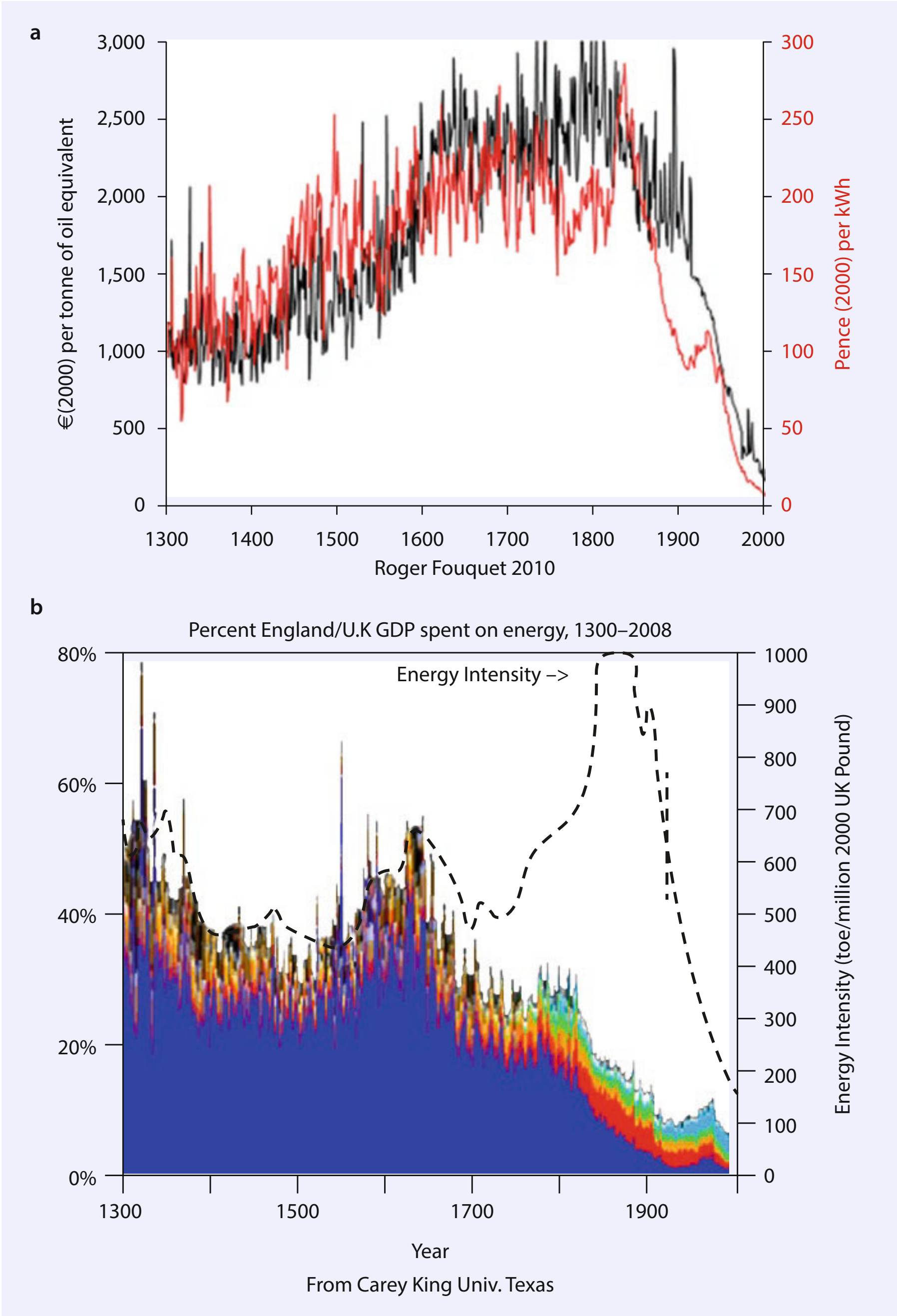


Fig. 4.4 **a** Efficiency of the global economy, as determined by the ratio of GDP produced to energy used (Source: Fouquet). **b** Percent of GDP spent on Energy for UK (From King)

view the agricultural production needed to feed this exponentially increasing human population could grow only linearly, i.e., less rapidly than the number of humans. He also opposed the importation of cheaper continental grains, as a limited food supply assured increasing rents for his patrons, the landed aristocracy, and squeezed the profits of the rival capitalists. It was this view that the human prospect was limited by inadequate food supplies and that class conflict was inevitable, which led the Victorian philosopher Thomas Carlisle to give economics the label of "the dismal science."

As chronicled in ▶ Chap. 2, Adam Smith and other classical economists focused on land

and especially labor as a means of expropriating the resources generated by the natural world, and then transforming them into materials that we perceive as constituting wealth. Later, David Ricardo made important observations about the general need to use land of increasingly inferior quality as populations (and hence total agricultural production) expanded. Karl Marx, who focused on the part played by labor in creating value, realized the crucial role played by nature in creating wealth. He was keenly interested in the long-term adverse effects of large-scale agriculture on soil quality and firmly believed that capitalism exploits the land as it does labor, and the process of capital accumulation creates a metabolic rift in the organic connection between humans and nature.

Thus a number of economists made important conceptual and philosophical advances that formed the basis upon which biophysical economics has been built. Early economists Quesnay, Malthus, Carlisle, Smith, Ricardo, and Marx all were aware of, to varying degrees, the importance of biophysical inputs and processes to the economy. Additionally, Kenneth Boulding in his paper "The Economics of Coming Spaceship Earth" focused on the impossibility of continued economic growth on a finite planet: "Anyone who believes that exponential growth can go on forever in a finite world is either a madman or an economist." Nicholas Georgescu-Roegen was a Harvard-trained economist who found the intellectual structure of conventional neoclassical economics enormously misleading and wrote extensively detailing the failures of conventional economics. His most prominent contributions were *Energy and Economic Myths* and *The Entropy Law and the Economic Process*. But the real foundation for biophysical economics was laid by his student Herman Daly, who through a series of excellent books and presentations examined the biophysical requirements for modern economies. His main emphasis was that we could not grow indefinitely and that any growth at all would cause unacceptable damage to the Earth. His main vehicle for thinking about this was the development of "steady-state economics," that is, an economics not based on growth. In addition he was among the first, and certainly the most thoughtful, in criticizing the intellectual underpinnings of conventional economics because it did not begin with the biophysical reality of the physical systems that are essential for supplying the materials and energy required for any economic activity. Nor did it consider the limiting effects of entropy. Daly extended Karl Polanyi's idea of the embedded economy with a focus upon the economy as a subsystem of the planetary ecosystem. His thoughtful and gentle personality allowed him to deliver very sharp criticisms to the economic community from one of their own. Nevertheless, most of Daly's many followers came from outside, not inside, the discipline of economics. Other economists who made important contributions to biophysical economics include John Gowdy and Lisi Krall, especially as regards their work on humans as an ultrasocial species and the crucial role played by the production of surplus at the dawn of the Neolithic era.

4.5 Ecology as a Source of Ideas

Ecology as a concept for understanding nature dates back to at least Theophrastus in ancient Greece, and the economic importance of properly functioning natural systems was well recognized by various scientists in, for example, Ukraine and Russia during the first half of the twentieth century. But ecology as a self-understood academic discipline hardly existed before the middle of the last century. One key event was the publication of Eugene Odum's *Principles of Ecology*, and another was the publication of Howard Odum's (Eugene's younger brother) *Environment, Power, and Society*. The latter was an ambitious attempt to show commonalities among various natural ecosystems and human societies using energy flow diagrams. Thus we can consider ecosystems such as natural streams, forests, or grasslands as economic systems ([2, 5]; ◘ Fig. 4.5). These systems have "economic" structures for production (photosynthesis), consumption (grazing, predation, respiration), and transfer of "goods" (food, minerals) through exchange processes (e.g., transfer of materials and energy between the physical environment and organisms through processes of plant uptake of nutrients and capture of energy, plant and animal uptake of water, and transfer through food chains). They are different in that the human economy is the result of conscious effort by humans and their expenditure of energy to change nature into what humans want.

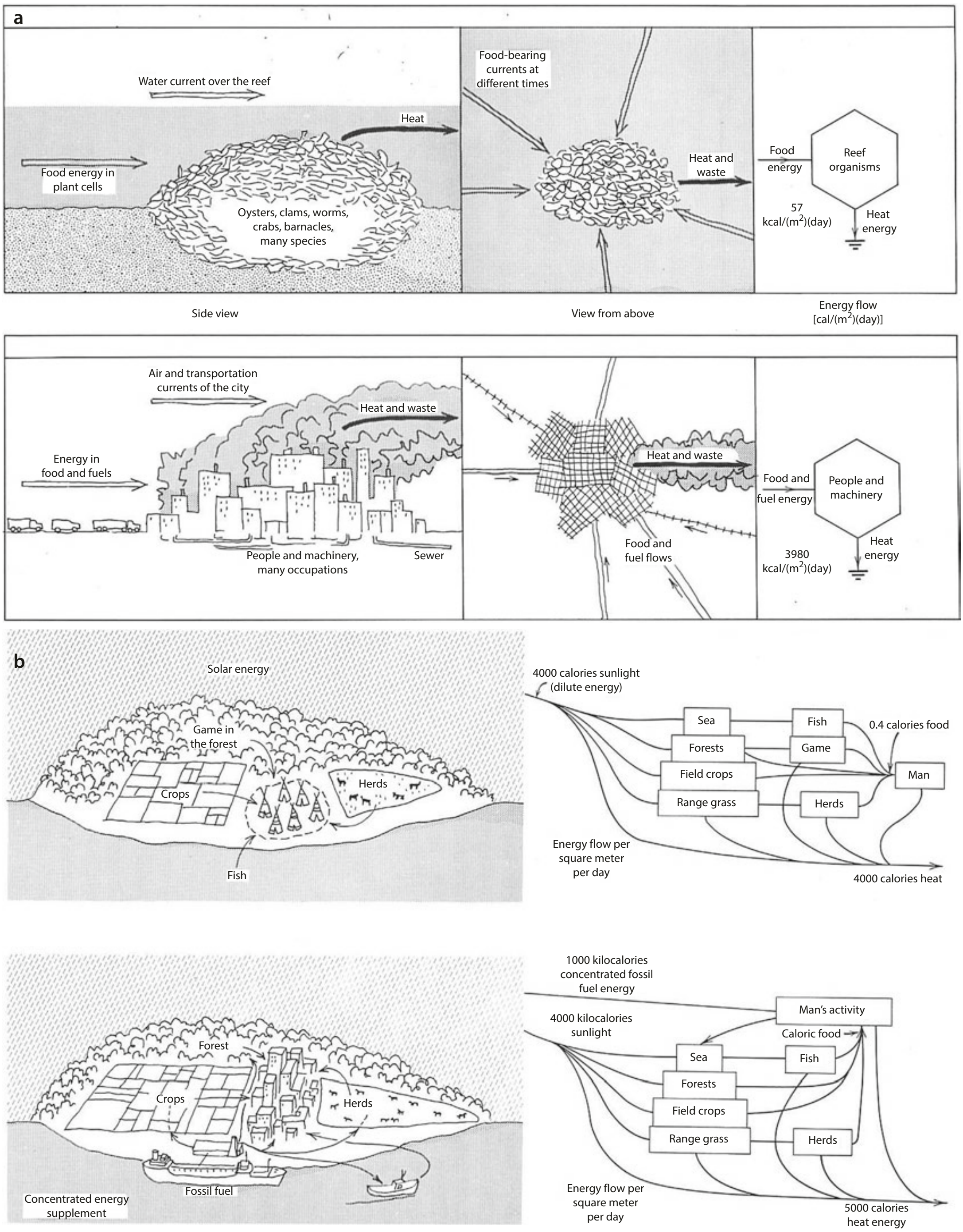


Fig. 4.5 **a** Similarities of oyster reef cities in nature and a human city (From Odum [2]). Both are necessarily subsidized by energies ging food in from elsewhere and removing wastes. **b** Comparison of agrarian system with industrial system (From Odum [2])

A critically important insight that has been gained from the study of these natural ecosystems is the importance of a thermodynamic perspective. While the importance of energy for biology was well recognized by, for example, Ludwig Boltzmann as early as 1880, energy as a concept was not well understood during the time period when modern economics was being developed.

Whether economics as a discipline would have been built on a foundation of thermodynamics, as were physics, chemistry, and ecology, if thermodynamics had been better developed during modern economics formative stages is impossible to answer but seems likely [6].

As thermodynamics was developed, ecologists began to understand that in the absence of a continual input of energy, the highly ordered molecules within an ecosystem will, over time, degrade into completely random assemblages. It is only the continual input of energy from the sun, the capture of this energy by green plants, and the effective transfer of energy to other components of the system that allows ecosystems and their components to fight the general tendency of all things toward randomness (often called a tendency for disorder or entropy to increase). Ecosystems have often been called "self-organized" entities; organisms within the ecosystem, and perhaps ecosystems themselves, interact to build a biological structure that best captures and utilizes available energy [7]. The blueprints laid down in an organism's DNA are fine-tuned through natural selection so that energy may be used to capture, reorder, and maintain both additional energy and the molecules in that organism in the otherwise extremely unlikely patterns that we call life. Energy is captured and used to generate biological structure, which in turn maintains, replicates, and sometimes changes itself through natural selection. It does not take too much imagination to transfer this concept to human economies, as both are equally biophysically based, although additionally powered by fossil fuels (◘ Fig. 4.5). This is what Howard Odum initiated in *Environment Power and Society* and other publications [8].

A particularly interesting and important concept developed for both natural and economic systems by Howard Odum was "the maximum power principle" (MPP). Odum and his colleague physicist Richard Pinkerton started with simple physical systems such as Atwood's machine (a simple pulley with two baskets that allowed differential loading of energy inputs and backforce [9]). They then change the ratio between the force (the weight in the heavier, descending basket) and the backforce (the weight of the load ("goods") in the ascending basket). They found that the maximum useful work that could be undertaken per unit of time (i.e., power) was when the heavier basket was twice the weight (gravitational force) of the ascending basket. If the baskets were more equal in weight, more "goods" could be delivered per trip and per unit input energy, but the machine worked more slowly even if more efficiently. Conversely, when the baskets were very different in weight, the goods were delivered very rapidly, but not much per trip, and most of the input energy went into heat as the basket hit the ground. The maximum useful work per unit time was when the ratio between the two weights, and hence forces, was 2:1. Odum and Pinkerton went on in that paper, and many more, with many other examples including economic analogies. The basic idea was that in a competitive world, one cannot be too efficient, for otherwise one's competitors would exploit resources before you were able to. The consequences are not always comfortable. For example, if the United States chose to use Middle Eastern oil more slowly, would that open up the resource for additional exploitation by the Chinese?

Most of us would not consider systems of nature as "real" economies because that term tends to be reserved for systems that include humans, human processes, market transactions, money, and/or other human-directed activities. Nevertheless, actual economies (including those of the city of Syracuse, NY, or of the country of Costa Rica) are, in fact, subject to the same forces and laws of thermodynamics as natural ecosystems and have much in common with them—structure, function, energy requirements, material cycles, and so on. In our view modern cities, agricultural systems, and even entire nations are indeed industrial ecosystems. Since the structure of many human-constructed systems (e.g., cities) contains so much more abiotic and animal mass than that of natural systems, the energy requirement to construct and maintain them is much larger and must be supplied from outside the system. Today this requires not only the usual input of solar energy but also the concentration of massive quantities of fossil fuels and energy-intensive materials, which in turn generate enormous "ecological footprints" on the rest of the world. Hence these "real" economies are as much about the movement of materials and the use and dissipation of energy as they are about the social or human-involved transactions.

As a consequence of studying natural systems from this perspective, many ecologists, led by Howard Odum [e.g. 2, 8], were quite ready to begin to look at economies from an energy and material perspective, and they already had the conceptual, measurement, and modeling tools to do so. They found ready acceptance and collusion with Herman Daly and a very few other economists, but no interest at all from traditional neoclassical economists.

4.6 Limits to Growth

The third main source of ideas that led to the development of biophysical economics was a series of quite startling and, many would say, pessimistic scientific reports about the future that took place in about 1970. The most important, or at least the ones that received the most attention, were the "Club of Rome's" *Limits to Growth* [10] and *The Population Bomb* [11] by Paul Ehrlich. *The Limits to Growth* was the result of a computer model generated at MIT, initially by Jay Forrester, and then refined and promulgated by his students Dennis Meadows, Donella Meadows, Jorgen Randers, and William Behrens III. The model was a very basic projection of human population, including birth rate and death rate, per capita industrial production, per capita food production, pollution, and nonrenewable resources, which is modeled as one entity that is depleted over time. The results of the "standard run" generated growth in population, food production, and industrial production for a while but eventually serious complications due to pollution or resource depletion that led to serious population decline. The investigators varied their assumptions in many ways to see if they could generate a scenario that was stable. Counterintuitively, increasing investments or controlling pollutants only delayed the negative impacts. In the model, it was only by extreme population control and eliminating *all* investments that a stable future could be derived. The model and its critics are discussed further in ▶ Chap. 12. *The Population Bomb* discussed the growing human populations and the many problems associated with humans using more and more of the Earth's resources to support them. This too predicted quite dire implications of the continuing population growth. While the most extreme of Ehrlich's predictions did not come true, there are many ways that in fact his predictions have come to pass [12].

These reports added concerns about human population growth and pollution to the existing concerns based on the predictions by Shell Oil Geologist M. King Hubbert [13] of the inability of both the United States and the world to keep increasing the production of petroleum. Hubbert assumed that over time the use of a nonrenewable resource would grow and then decline in approximately a normal-shaped (bell-shaped) curve, initially increasing rapidly and then reaching one (or several) peak when about half of the resource was consumed. Again, the results of this projection are somewhat ambiguous: oil is still abundant and relatively cheap (although more expensive than in the past), but globally conventional oil has ceased growing or nearly so. But a country-by-country analysis shows that oil production in most countries does follow fairly closely a Hubbert curve [14]. As we discuss later, the timing of some of these predictions for the globe may be a little off, while the fundamental pattern projected is very much on target.

These reports implied in various ways that the human population appeared to be becoming very large relative to the resource base needed to support them—especially at a relatively high level of affluence—and that it appeared that some rather severe "crashes" of populations and civilizations might be in store. Meanwhile, many new reports in scientific journals were published about the many environmental problems such as acid rain, global warming, pollution of many kinds, loss of biodiversity, and the depletion of the Earth's protective ozone layer. The oil shortages, the gasoline lines, and some electricity shortages in the 1970s and early 1980s all seemed to give credibility to the point of view that our population and our economy had in many ways exceeded the world's "carrying capacity" for humans, that is, the ability of the world to support humans and their increasingly affluent lifestyle.

Universities hired many new people in the previously obscure disciplines of ecology and environmental sciences, and there was a great surge of interest by students in issues related to resources and the environment. Although courses in environmental economics were added to some college catalogs, most economists ignored these issues or or, if anything, modeled nature as part of the economy and added in environmental factors to the list of things that would be regulated by rational individuals responding to price incentives. The notion of external limits to growth, based on biophysical constraints, got a chilly reception from the community of mainstream economists, although the idea of an economy limited by nature began to develop following among political economists in the early 1970s [15] . Although economists have written about the *internal* limits to growth since the eighteenth century, these new works raised a new possibility: our futures would be limited by nature as well. Historically, humans have been able to transcend nature's limits by employing increasing amounts of energy to the problems at hand. But were we nearing those limits, either in supplies of energy or in the consequences of using it? If so, the age of convenience and growth of affluence and human well-being, primarily in the global North, [15] might be replaced by living within our means or even degrowth. The message was not popular. President Jimmy Carter discussed on television the need for Americans to conserve and even installed solar collectors on the White House roof. He said that the American people should view the energy crisis as "the moral equivalent of war." For many people, it did seem like humans had reached the limit of the abilities of the Earth to support our species.

Most economists, however, did not accept the absolute scarcity of resources or the concept of limits to growth. The return to growth, they said, was just a matter of implementing a series of proper incentives and market-based reforms, as well as dispensing with the dangerous ideas of absolute limits. A series of scathing reports appeared directed at those scientists who wrote articles with the "limit" perspective (e.g., Passell et al. [16]). They argued that economies had built-in market-related mechanisms that would deal with short-term (relative) scarcities. Technical innovations and resource substitutions, driven by market incentives, would solve the longer-term issues. Critics of the early antinuclear movement belittled the idea that using less electricity or generating it from less dangerous sources was remotely viable. For them it was generate more nuclear power or "freeze in the dark."

These three lines of thought converged more formally in mid-1985 with the development of the International Society for Ecological Economics, along with national affiliates, and the journal, *Ecological Economics*. There was a sense by many that this society and journal, while undertaking important research, focused too much on putting a dollar value on environmental goods and services while mostly missing the issues of depletion, the institutional context in which economic decision-making takes place [17], and a continued commitment to neoclassical models and analysis. Consequently about 20 years later, the International Society for Biophysical Economics was formed. Starting in 2016, the initiation of the journal *Biophysical Economics and Resource Quality* devoted itself to publishing papers exclusively grounded in biophysical economics. Some economists, increasingly, are agreeing with biophysical economists on the need to reform basic concepts in economics to reflect the importance of energy [e.g., 18].

4.7 What Can One Do with Biophysical Economics?

Biophysical economics thus far has focused on five major issues:

1. The inadequacy of neoclassical economics (see ▶ Chap. 3)
2. The need to incorporate biophysical realities into economics (see ▶ Chaps. 3, 4, and 5)
3. The importance of the fossil fuel revolution for economic growth (see ▶ Chaps. 4, 6, 7, 8, 9, 10, and 11)
4. Limits to growth as a real (if complex) issue (▶ Chap. 12)
   - This includes peak oil, declining energy returns on investment (EROI).
   - Can renewables substitute for fossil fuels?
5. The need to improve and generate better estimates for EROI and equations for biophysical economics (see ▶ Chap. 18)

Money is of great practical concern in day-to-day life as we get paid or exchange money and so gain access to food, fuel, or housing. In addition, many financial practices that seem closely related, such as accounting, bookkeeping, and simply balancing one's checkbook, are based on money and have great practical importance and apparent reality. In fact, those of us who advocate biophysical economics realize that money is useful as a medium of exchange. And of course there is no substitute for proper economic bookkeeping and the normal everyday use of money as a medium of exchange to obtain needed goods and services. But what about the theory behind economics? Is that the best way to understand the routine use of money in our economy? We think not, as is obvious. Others disagree. Conventional economics is useful to those at the top, because it justifies the prevailing economic order. It treats continuing consumption of more and more stuff as the key to happiness and holds that a private enterprise economy is the best society that has ever existed. Most importantly conventional economics is based on money as the key to valuing and acquiring stuff. Thus even the most arcane economics has a certain appeal because it uses money, which translates easily in most people's mind.

4

But what is important to understand is that money in its modern form is fiat money or "money by decree." As such it has no intrinsic value but value only in representing the willingness of society or its representatives to accept it for payment. The government accepts the money for payment of taxes. We also use money to acquire energy or energy-derived products to generate the good or service that will then be made available to the bearer of the money. *Thus, money is a lien, meaning a legitimate claim, on energy (or energy that has been spent), as well as labor, commodities, or money itself.* As we saw in ▶ Chap. 1, money can serve as capital and as a medium for speculation as well. Money creation is how banks make profits, but ultimately money can be best understood as a lien on energy.

Let us give an example. One might buy a wonderful, high-quality bagel in New York City for a dollar (more with cream cheese and/or lox). Behind that dollar lies many biophysical activities each of which must occur. Natural gas must be used in Louisiana to make fertilizer, which is then barged up the Mississippi River to Nebraska where a tractor uses oil to spread the fertilizer on the land and plant the wheat seeds. Later a tractor uses more diesel to harvest the wheat, which is then shipped by railroad to New York and ground into flour. Electricity is used to mix the flour and boil the water it is cooked in. Energy is used to make the fertilizer plant, the barge, the tractor, the railroad, and so on. It is the physical expenditure of energy to do all the work necessary to generate the bagel that is necessary, and the dollar is how we keep track of that. Part of the dollar goes to pay for the energy used at each step, and part of it is used to pay for those who directed the work (i.e., labor and management). Energy is also needed to give meaning to the worker's paycheck or to the proceeds given to capitalists through their profits. Roughly 5 megajoules (one-seventh of a liter of oil) of energy was used to make that one bagel.

In addition to, or overlapping with, the list given above, we can ask how biophysical economics has been applied by its practitioners. The first, most general application is visualizing how the economy operates and understanding how energy is required at each step. This can be seen by studying ◘ Fig. 3.3 carefully. This leads to a number of nonobvious implications. For example, if one wishes to live within nature's biophysical limits, then people in the rich nations of the world need to think mindfully about the energy used and embodied in their day-to-day lives and act accordingly.

A second major application of biophysical economics is to evaluate realistically the potential for economic expansion in poorer countries. Most of the world is quite poor, and there are many efforts directed at improving the lot of the poor, some successful and some not so much. Overall there has been a considerable improvement in the lot of the poor in the past six decades, and in fact since at least 1820, when according to one study, 90% of people were then living in extreme poverty compared to only 10% now living in extreme poverty, mostly in the global South [15]. Some might consider with Roser's numbers rather optimistic as about a third of the Earth's human population live on less than two dollars a day. Where has this increased wealth come from? The principle cause has been the continuing industrial revolution, which used fossil fuels to replace animate power and, most importantly, increase food production (see ◘ Figs. 4.1, 4.2, and 4.3).

Nevertheless, there is a dominant viewpoint among many of our agencies responsible for financing development, such as the World Bank, that what is important is not spending money in the public sector but turning as much as possible of the financial workings of a country over to the private sector. It is commonly held that the private sector is more efficient than government. Empirical analysis of this question does not clearly show that to be the case and in fact often the reverse [19]. The extreme view of the concept that it is better to privatize functions commonly undertaken by governments is called the "Washington Consensus," and it was used to guide development in Latin America in recent decades, often with disastrous results [20]. Biophysical economics approach has been used to examine these often misguided policies and alternatives that may offer hope for the poor (e.g., [21]).

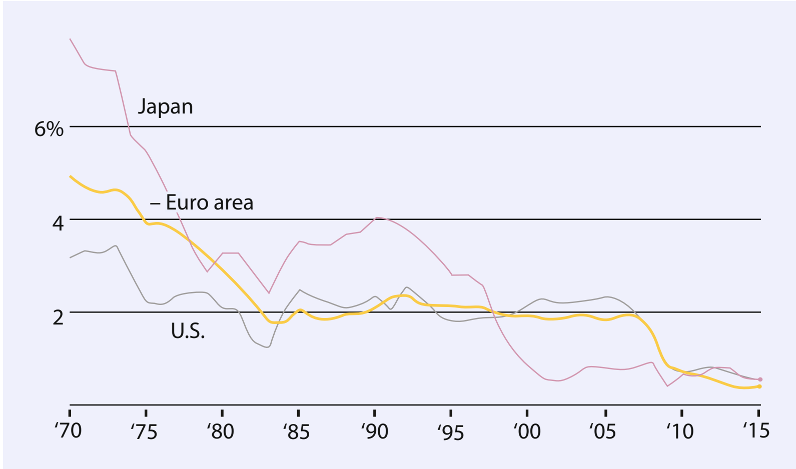


Fig. 4.6 Economic growth of Japan, Europe, and the United States, showing the general decline over time

A third application of biophysical economics is to understand some important trends in the world today and to help us prepare for a future that might be quite different. These trends include secular stagnation, peak oil and declining EROI. There are also a whole suite of issues around reducing carbon release and whether renewable energy can replace some important proportion of fossil fuels.

4.8 Secular Stagnation

The global North, from the end of Second World War until the 1970s, once Europe had recovered from the war, experienced historically unprecedented growth in energy use, and to some degree energy efficiency and economic output increased substantially (Fig. 4.1). But the growth rate of most of the world's industrialized economies has slowed since the 1970s. The economies of the United States, Europe, and Japan have experienced declining growth since the 1980s, and most of the current world economic growth is driven by China and India, although they too may be declining, (Fig. 4.6). As of mid-2017 the GDP of countries in Europe and Japan had been essentially stagnant for a decade or two. During the past decade, the United States had a GDP growth rate of about 1%, about half of the historical standard of 1.9% since the Civil War [22] and about the same as the population growth—hence there was no increase in average per capita wealth. Among economists there is considerable discussion and controversy about these essentially stagnant economies [e.g., 23]. Much of this focuses on factors internal to the economy: consumer spending, debt, banks, deficit spending, and Keynesianism—whether or why governmental deficit spending, which has been used extensively in the past to "jump start" economic growth, no longer works as it once did. We will discuss more extensively mainstream and heterodox views of secular stagnation in the next chapter.

Biophysical economics may provide such an explanation [24]. Most adherents to biophysical economics believe (as do many others) that conventional (neoclassical) economics is fundamentally flawed (see ▶ Chap. 3). Biophysical economics believes that there is a general relation between the declining abundance of resources, as reflected in lower production and EROI for oil and other important fuels, and the decline and cessation of growth (Fig. 4.7). Murphy and Hall put forth a model that gave a biophysical economics-based explanation of economic cycles that seems consistent with the actual behavior of economies(Fig. 4.8). The case for this was stronger up to mid-2015, when oil was trading at $100 a barrel. At the time of this writing, it is about $50 a barrel, still high by historical standards and relative to, e.g., the 1990s when growth was still strong. The OECD country with highest growth, although still low, is the United States. In the United States, natural gas, not quite as valuable as oil but still an excellent fuel for industry, was at a very low price, about a quarter of the long-term price, reflecting over production from fracked areas in, e.g., the Marcellus Shale in Pennsylvania. This could be the reason for the slightly higher growth of the

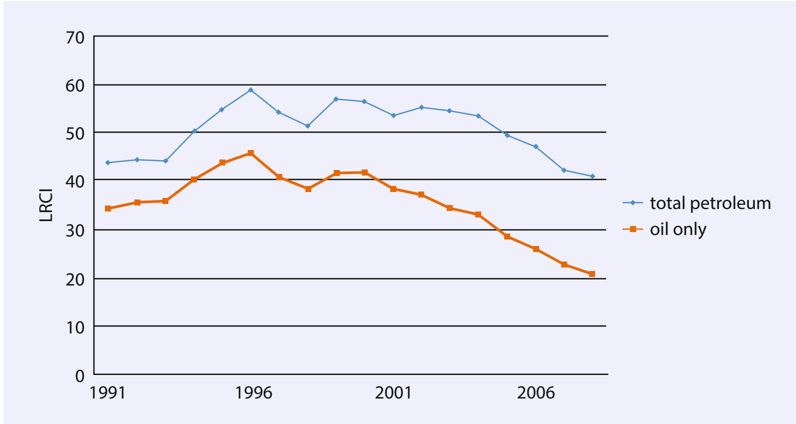


Fig. 4.7 Example of decline in EROI (Norwegian oil and gas; From [37])

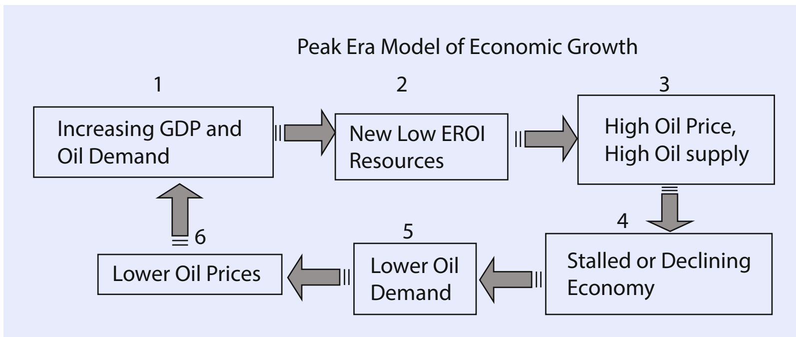


Fig. 4.8 Model of cycling of oil prices (From [38])

US economy compared to other OECD countries. Curiously the oil and gas companies are still drilling new wells even while they are losing money!

There may be another useful biophysical concept from ecology. Eugene Odum in 1969 wrote a good paper representing the behavior of ecosystems over successional time, that is, from the establishment or colonization of life at a site, such as an ecosystem that develops on a bare patch of land or a newly filled aquarium until the ecosystem reached "climax," when it no longer accumulates biomass. At first, as biomass became established, production (the capture of energy from sunlight) and respiration (the use of energy for maintenance metabolism by all living things) each increased rapidly, with photosynthesis being larger than respiration (Fig. 4.9). The difference between the two represented the energy absorbed by the increasing biomass. But then at some point, the respiration of the increasing biomass equaled the production of the plants, and the system stopped accumulating biomass. The relatively constant biomass remaining at steady state was limited by the respiration (i.e., for maintenance) energy costs being as large as the gain from the capture of energy from the incoming solar energy, and the system adapted to that. This takes place around the world as most ecosystems are limited by the incoming solar radiation (or water) and rather than growing indefinitely they reach a steady-state biomass level. Odum believed that human societies too would initially grow rapidly (i.e. new construction exceeding maintenance requirements, resulting in the accumulation of infrastructure) but then would approach equilibrium as energy costs to maintain infrastructure became very large. This is very different from the indefinite growth of economies expected by most economists. So have modern highly developed economies with enormous infrastructures (think roads and cities) reached a stage where all of the available energy is used for "maintenance metabolism" to support the infrastructure that exists and little is available for net growth? Could this be an explanation for secular stagnation? Will our existing growth-oriented economic models still be appropriate (Fig. 4.10)? Or is it sufficient to say that the growth of economies simply reflects the growth of the energy that allows that to happen, and that energy, once easy and cheap to get, is no

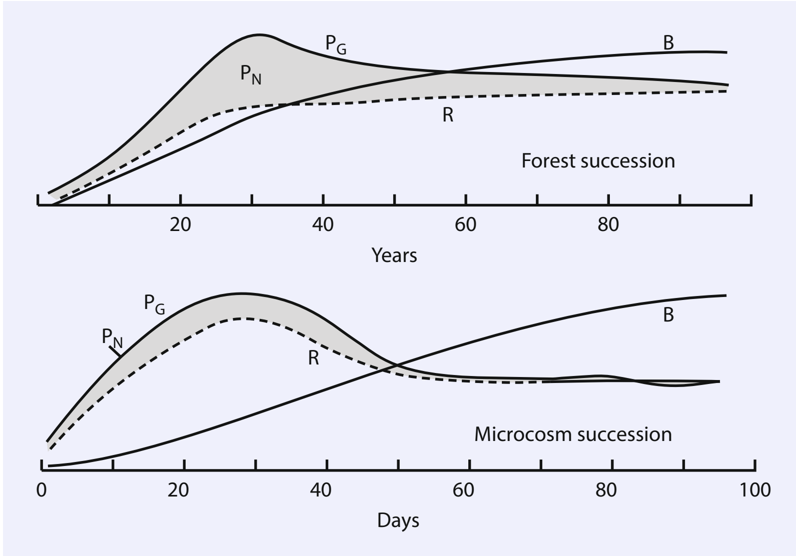


Fig. 4.9 Ecological concept of succession following disturbance (or initiation of ecosystem succession). As time passes in a novel ecosystem (such as in a new forest clearing), both photosynthesis and respiration increase, although initially photosynthesis increases more rapidly, resulting in the production of more biomass. Eventually the increasing biomass uses more and more energy for respiration, so photosynthesis and respiration are equal (called "climax") (From [39])

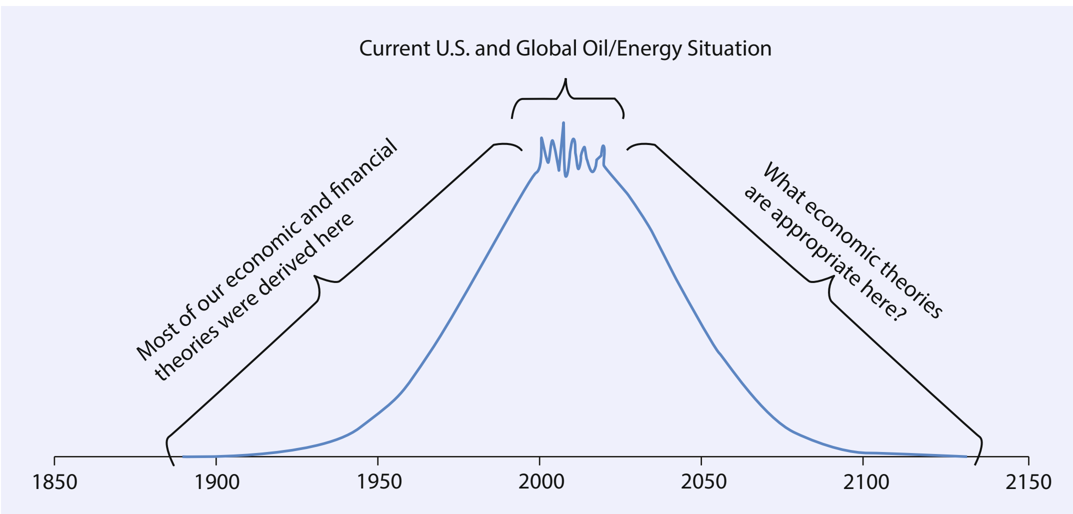


Fig. 4.10 Relation of our development of economic models relative to the general pattern of increase, stagnation, and eventual decline in production and availability of fossil fuels

longer so? Either way biophysical economics has approaches that appear very useful to understanding secular stagnation which should be explored much more than they have been so far.

Meanwhile, the main problem that we face with regard to fossil (and other) fuel supplies is not their total quantity on Earth (there are enormous supplies remaining) but their quality. To survive and thrive, all species must balance the relation between the energy cost of getting needed resources, including additional energy, and the energy (or other attributes) in the resource exploited. This applies to predators hunting for food who must compare the energy expended in the chase and the chances of success with the energy obtained from the prey. Likewise, human hunter-gatherer societies, if they are to survive, must generate substantial surpluses of energy relative to their own investment energies. It also applies to modern human industrial economies, although they are different in that the energy invested and gained is not metabolic but exosomatic (outside the body) energy. Thus, a critically important issue for examining our energy future is what is called energy return on investment (EROI) of fuels. Investments are required to get fossil and other fuels out of the ground and into society. These investments are in terms of energy as well as dollars, and just as we need a profit from a financial investment, we need a net energy profit from our energy resources for society to continue. We will develop these concepts more fully later in ► Chaps. 18 and 19.

Perhaps the most important policy question for us now is how we should make our energy investments. Huge investments will be needed if we are to maintain the enormous human infrastructure that we have built—simply to fight the inevitable generation of entropy which nature dictates will occur. As we move into the future, EROI is a critically important component of the decisions we have to make but hardly sufficient by itself. The main problems that we face at this time with respect to understanding our situation are as follows:

1. The apparently incessant decline in EROI which will greatly limit our options for investing in new energy technologies, whatever they might be.
2. The need for professional, objective means of gathering the needed data and evaluating the alternative energy sources and claims. It would seem that such an evaluation would have to come from peer-reviewed or government-sponsored program.
3. The total inadequacy of conventional economics for the job.

4.9 Why Have Not Most Traditional Economists Paid More Attention to Biophysical Economics?

Economists have had to pay little attention to biophysical economics largely because there seems to be no crushing limits to growth as of yet. Perhaps the basic question is whether Malthus has been put to rest by the evolving technology of modern society. Most economists would answer "Yes, Malthusian concerns have been put to rest, mostly by continuing technological advancement." According to Bridge [25], "there is a post scarcity narrative – a postindustrial (market-generated) resource triumphalism – in which resource scarcity no longer poses a limiting factor to economic development… Neoliberal prescriptions for marketization and privatization have come to dominate nearly all areas of public policy over the last two decades."

But there have been at least three biophysical factors that seem to be at least as important: the opening of the Americas to immigration by the surplus Europeans, their virtual extinction of the Native American potential competitors there, and the industrialization of agriculture, which generated an enormous increase in food production and removed the stranglehold of land as a fixed factor of production, upon which Malthus' entire theory depended. A fourth factor might be considered technology by itself, although most technologies were associated with industrialization so that we might consider them as a force working together. With this increasing creation of economic surplus, economics, starting with Keynes, focused increasingly on consumption and became more and more intertwined with the social sciences. Simultaneously, concepts of economic production have focused increasingly on capital as an abstract but critical notion, while labor has been reconstituted as "human capital," and land has simply been omitted. Recently, in an attempt to give value to nature, ecological economists have christened biophysical stocks as "natural capital" that subsequently produce flows known as natural resources. But the continuing abundance of energy, food, and disposable income by at least a large part of the developed world has tended to relax any concerns that economists might have about resource limitations and hence the need for biophysical economics. An additional issue is that economics as a discipline tends to be "hermetic," meaning completely enclosed within itself.

Many economists argue that since energy costs are equivalent to only some 5% of GDP, then they are trivial in importance compared to the rest of the economy and that we need not be too concerned about future possible energy shortages. But what if this cheap energy declines in abundance, as seems inevitable to many of us? When energy and minerals increased to 12% of GDP in the oil-constrained and economically devastated decade of the 1970s, as is likely to occur again, perhaps soon, the economic consequences were enormously adverse. Hamilton [26] has found that whenever the cost of energy approaches 10% of GDP, a recession will occur. One can argue that if the present 5% of GDP energy cost is subtracted from the current economy, most of the other 95% of GDP will cease to exist. In other words, we are extremely lucky that we must pay only the extraction costs, rather than the full-value production, value to society, or replacement costs that Mother Nature might charge if there were mechanisms for her to do so. The full price would have to include the costs of natural capital depreciation, including both the fuel itself and the nature destroyed by its extraction, shipping and use, as well as the military costs of assuring resource availability. These we are hardly paying at present. If and when we run out of luck, and these costs come due, as will likely be the case, economics will become a whole new ball game in which the focus will return to production and which will result in a new way of thinking about monetary and energy investments. Thus, there are good reasons to examine economies from a biophysical and energy perspective as well as from a social- and market-based perspective. This may be a difficult leap initially, but the shift in perspective should become obvious and desirable once the idea is broached.

4.10 Are We Becoming More Resource Efficient?

The material demands for societies continue to grow despite very little empirical data to support the popular idea that economies are becoming more efficient in turning resources into economic production. In fact, considerable empirical data suggests that many economies are becoming less efficient (see, e.g., 27–29) even while total consumption increases nearly everywhere. (One partial exception to this statement is seen in the US economy since 1980, where GDP appears to be increasing somewhat more rapidly than the increase in resource use—although about half of that supposed increase in efficiency is through the increased proportional use of higher-quality fuels such as primary electricity). An extremely important question becomes: is petroleum a transition element along the energy source road from slaves to draft animals to water power to coal to oil and gas to…something else? Or are liquid and gaseous petroleum a one shot, extremely concentrated, relatively environmentally benign, high energy return on investment (EROI) premium fuels that we will never see again at such a large scale? We suspect the latter. A second critical question to which we do not yet know the answer is "which will win the race between innovation/substitution and depletion." In the case of petroleum from the United States, Mother Nature seems to be winning, as the EROI has declined from at least 30 to 1 in 1970 to 18 to 1 in the late 1990s [30, 31]. When Cleveland made appropriate corrections for the fact that increasingly we are investing higher-quality energies (e.g., electricity) into producing oil over time, the "quality-corrected" EROI has declined much more sharply, to about 11 to 1 [32]. The EROI for our legacy giant oil fields continues to decline even as new oil becomes more difficult to find [33]. Likewise, the energy cost of getting a ton of pure copper in the United States has increased despite massive increases in technology because the best ores are long gone [31, 32].

Essentially no resources today can be viewed as truly sustainable at present rates of production, consumption, and growth because all are subsidized by cheap petroleum. "Sustainability" projects such as those of ecotourism and, indeed, the entire economy of "sustainable" places such as Costa Rica [20] are not sustainable at all due to their ever-increasing dependency on petroleum and the debt that implies. The assumptions of growth-oriented economists have resulted in enormous economic and energy investments in developing tens of thousands of expensive resorts in many lovely but otherwise poverty-stricken tropical areas that are based on the assumption that the people that live there can and should live indefinitely on the crumbs that fall off the tables of the industrial world's momentary wealth. As the supply of cheap petroleum is exhausted through the increased exploitation of the Earth's highest-quality and most accessible energy resources while demand for its products continues to grow, the world will likely be in for some very rough sledding ahead. We as a society must recognize the need for a more biophysically based economic system, which includes a focus on material things such as land, water, soil, food, timber, other fibers, and, most importantly, energy. The economy must focus once again on the most fundamental issues of providing food, clothing, shelter, basic transportation, and other necessities. It must come up with real solutions to the critical problems we face (e.g., energy depletion and impacts, soil erosion, over fishing, water management, massive inequity in the distribution of wealth, etc.) that have been neglected thus far due to our temporary patch-up "solutions" of cheap oil. We must rethink very carefully what any increase in efficiency might bring because of Jevons' paradox, for example [33]. We must think about the critically needed international development assistance in entirely new ways, and we cannot allow an unjustified faith in the supposed virtues of neoclassical economics mask where it is used to sanctify the massive neocolonialism sweeping the less developed world [34]. If in fact the grim results of the *Limits to Growth* do come to pass, do we castigate those politicians who for "moral" reasons removed population from the agenda of the US Government? How about those economists who argued foolishly against that model's utility or, more generally, a biophysical approach to the Earth's problems? Do we put them in jail for the lives lost and for encouraging us to make investments in the wrong places?

4.11 Biophysical Economics as a Means of Synthesizing Traditional Approaches to the Generation of Wealth

One can summarize the three most important approaches to economics as the physiocrats (with their focus on land), classical economics (with its focus on labor), and neoclassical economics (with its focus on capital). These seem to be completely independent conceptual approaches to economics. Yet all can be understood from a biophysical perspective as an appropriate focus on the main energy sources of their time. Land was important when the main energy input to economies was the sun: farmers redirected the solar energy of ecosystems to human and draft animal mouths, and wood provided the most important source of heat so that land became a source of wealth as emphasized by the physiocrats. During the late eighteenth and nineteenth century, workers were increasingly concentrated in factories where their physical efforts were important to the productive process. Over time, the landed gentry who owned large solar-collecting estates were replaced at the top of the financial and social ladder by the new mill owners and then industrialists who directed new production systems using the more concentrated energies of coal and then oil. Therefore, the physiocrats, such as Quesnay, were correct for the time and place in which they lived, when the land-derived capture of solar energy generated the most wealth. Adam Smith, a contemporary of Quesnay but living in England, was correct for the time and place in which he lived, when craft labor was increasingly the main way to generate wealth. Ninety years later, when artisans were replaced by unskilled factory operatives, and landed aristocrats were displaced by industrial capitalists, Marx was able to contribute penetrating insights into the relations of the new classes of people who controlled different types and quantities of energy flow. Perhaps today neoclassical economists are partially correct to put the focus on capital—which is the means of utilizing fossil energy. Unfortunately, when all inputs are considered capital, it is more difficult to see this than in the days when capital primarily signified means of production.

What these "mainstream" production functions fail to emphasize is what every biophysical economist knows to be the truth: it is the energy that does the work of producing wealth and is essential for its distribution as well, whether that energy is derived from land, labor, or capital-assisted fossil fuels. Ayres and Voudouris [35], Kummel [36], and Hall and Ko [29] have shown that the production of wealth in industrial societies is almost perfectly a linear function of the energy use in those societies and that the correlation gets tighter and tighter when proper corrections are made for the quality of the energy used (e.g., coal vs. electricity) and for the amount of energy actually applied to the process (e.g., electric arc vs. Bessemer furnaces). Much, perhaps most, technology is ultimately about these things. It may seem obvious now that wealth is generated by the application of energy by human society to the exploitation of natural resources. Nature generates the raw materials with solar and geological energies, and human-directed "work processes" are used to bring those materials into the economy as goods and services. These processes have been made enormously more powerful over time through technologies that are mostly ways to use more or higher-quality energies to do the job. Energy would be the first element to be considered by most natural scientists if they were asked to construct a production function because they are trained to think that way and because it is statistically the most important factor—more important empirically than either capital or labor [36]. Where neoclassical economics treats production as just another case of the maximization of individual preferences, biophysical economics treats production as scientists treat work—the transformation of inputs into outputs using energy while subject to the laws of thermodynamics.

4.12 Summary

Our expectations for our lives for the past several hundred years have been based on an expanding universe of lands (e.g., the Americas), energy, and energy returns on energy invested. This has generated in the minds of most of the world's affluent people the expectation that there was at least the possibility of their bettering their own material lot, and for many this indeed occurred. We hear it in the pronouncements of economists from all sides, how we are facing a situation where many young people no longer have an expectation of more than their parents. In fact, the issue often enters the political arena as a failure of this or that government to make the economy grow and some other governmental philosophy having some magic power to reinstate the growth that they perceived as normal, as a birthright. To some degree this decline in economic growth for average people is clearly because the reins of power have increasingly passed into the hands of the wealthy, who tend to look out for their own. Most of us no longer live in a democracy but rather a plutocracy, meaning the rule of the wealthy. But something else is happening too: Malthus is finally catching up with us, if not exactly now (it probably is) then it is likely to come on in spades soon. The global population and its affluence can no longer be supported without piling up enormous debt, in monetary terms but also energetically and environmentally. Everywhere we look there are serious environmental issues starting with the potential impacts of climate change. There is certainly a lot of attention paid in some circles to climate change. But we believe the potential impact on our future society from issues related to energy supply and EROI are likely to be as large or even larger. It is likely that the effects of both will occur in the same time frame, probably the next few generations or perhaps sooner. Hopefully the understanding and use of EROI in analyses and public media will help soften the hard landing ahead of us, as the high-quality fossil fuels that have allowed many of the world's 7.3 billion people to live in relative luxury by the standards of old are increasingly depleted.

But neither our economists nor our politicians have the conceptual base or mental models to deal with this and still rely on mental models where the only operational levers for society are within the economy and are often some kind of untested political ideology. Rather economists must understand that much of what has determined human history and is likely to continue to do so comes from outside the immediate economy and is far less susceptible to internal manipulation (Fig. 3.3). Biophysical economics is one antidote to this but hardly sufficient. We need an entirely new approach to education, including how we can work together to face a world with increased constraints on our energy and economic growth (Fig. 4.10). Our present economic conceptual and mathematical models are not only inappropriate but hugely misleading if and as we enter this future.

Additional appropriate references can be found at a supplement to our paper "Hydrocarbons and the Evolution of human culture" (Nature 426 no. 6964. p. 318–322). ► www.nature.com/nature/journal/v426/n6964/extref/nature02130-s1.doc

References

1. Polyani, K. 1957. Trade and market in early empires in 1957.
2. Odum, H. T. 1971. *Environment, power and society*. Wiley Interscience (2nd edition 2005 Columbia University Press).
3. Hall, C.A.S., D. Lindenberger, R. Kummel, T. Kroeger, and W. Eichhorn. 2001. The need to reintegrate the natural sciences with economics. *Bioscience* 51 (6): 663–673.
4. Quesnay, F. 1766. Analyse de la formule arithmetique du Tableau Economique de la Distribution des dispenses annuelles d'une nation agricole. Journal de agriculture, du commerce et des finance; Christensen, P.P. 1989. Historical roots for ecological economics – biophysical vs. allocative approaches. *Ecological Economics* 1: 17–36; Christensen, P.P. 2003. Economic thought, history of energy. In *Encyclopedia of energy*, ed. C. Cleveland, 117–130. Elsevier.
5. Ricklefs, R.E. 1976. *The economy of nature*. Chiron: Portland.
6. Hall, C.A.S. 2017. *Energy return on investment: A unifying principle for biology, economics and development*. New York: Springer.
7. Odum, H.T. 1988. Self organization, transformity, and information. *Science* 242: 1132–1139; Odum, H.T. 1995. *Ecological and general systems*. Boulder: University Press of Colorado; Odum, H.T. 1995. Self organization and maximum empower. In *Maximum power: The ideas and applications of H.T. Odum*, ed. C. Hall, 311–330. University Press of Colorado; Brown, M.T. and C.A.S. Hall (eds). 2004. Through the macroscope: An H.T. Odum primer. *Ecological Economics* 178(1–2): 1–294.

8. ———. 1973. Energy, ecology and economics. Royal Swedish Academy of Science. *AMBIO* 2 (6): 220–227.
9. Odum, H.T. and R. Pinkerton. 1955. Time's speed regulator. American Scientist.
10. Meadows, D., D. Meadows, J. Randers, and W.H. Behrens. 1972. *The limits to growth: A report of the Club of Rome's project on the predicament of mankind*. New York: Universe Press.
11. Ehrlich, P. 1968. *The population bomb*. New York: Ballantine Books.
12. Ehrlich, P.R., and A.H. Ehrlich. 2016. Population, resources, and the faith-based economy: The Situation in 2016. *BioPhysical Economics and Resource Quality* 1 (1): 1–9.
13. Hubbert, M.K. 1969. Energy resources. In *The National Academy of Sciences–National Research Council, Committee on resources and man: A study and recommendations*. San Francisco: W. H. Freeman.
14. Hallock, J., P. Tharkan, C. Hall, M. Jefferson, and W. Wu. 2004. Forecasting the limits to the availability and diversity of global conventional oil supply. *Energy* 29: 1673–1696; Hallock, Jr., John L., Wei Wu, Charles A.S. Hall, Michael Jefferson. 2014. Forecasting the limits to the availability and diversity of global conventional oil supply: Validation. *Energy* 64: 130–153.
15. Roser, M. 2017. The short history of global living conditions and why it matters that we know it. Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/a-history-of-global-living-conditions-in-5-charts/ [Online Resource] May 10, 2017.
16. Passell, P., M. Roberts, and L. Ross. 1972. Review of limits to growth. *New York Times Book Review*.
17. Klitgaard, Kent, and Lisi Krall. 2012. Ecological economics, degrowth, and institutional change. *Ecological Economics* 84: 158–196.
18. Steve Keen. 2017. podcast. https://www.youtube.com/watch?v=BAjN6bG7XzM; http://www.eclectications.com/#post3.
19. Sekera, J. 2017. *The public economy in crisis. A call for a new public economics*. New York: Springer.
20. Hall, C.A.S. 2000. *Quantifying sustainable development: The future of tropical economies*. San Diego: Academic Press.
21. McKibben, W. 2017. The race to solar-power Africa. *The New Yorker* June 26, 2017. New York.
22. Gordon, Robert. 2016. *The rise and fall of American economic growth*. Princeton: Princeton University Press.
23. Galbraith, J. 2014. The end of normal. Simon and Shuster N.Y.; Irwin, N. 2016. We're in a low growth world; how did we get there? The upshot. August 6 2016; Hans Despain. 2016. Secular stagnation: Mainstream versus Marxian traditions. *Monthly Review* 67(4).
24. Hall, C.A.S., and K. Klitgaard. 2006. The need for a new, biophysical-based paradigm in economics for the second half of the age of oil. *Journal of Transdisciplinary Research* 1 (1): 4–22.
25. Bridge, Gavin. 2001. Resource triumphalism: Postindustrial narratives of primary commodity production. *Environment and Planning* 33: 2149–2173.
26. Hamilton, J.D. 2009. Causes and consequences of the oil shock of 2007–08. *Brookings Papers on Economic Activity* 1(Spring): 215–261; Hamilton, J.D. 2011. Nonlinearities and the macroeconomic effects of oil prices. Macroeconomic Dynamics, 15, 472–497.
27. Ko, J.Y., C.A.S. Hall, and L.L. Lemus. 1998. Resource use rates and efficiency as indicators of regional sustainability: An examination of five countries. *Environmental Monitoring and Assessment* 51: 571–593.
28. Tharakan, P., T. Kroeger, and C.A.S. Hall. 2001. 25 years of industrial development: A study of resource use rates and macro-efficiency indicators for five Asian countries. *Environmental Science and Policy* 4: 319–332.
29. Hall, C.A.S., and J.Y. Ko. 2004. The myth of efficiency through market economics: A biophysical analysis of tropical economies, especially with respect to energy and water. In *Land, water and forests in the tropics*, ed. M. Bonnell, 40–58. UNESCO. Cambridge University Press, Cambridge UK.
30. Hall, C., C. Cleveland, and R. Kaufmann. 1986. *Energy and resource quality: The ecology of the economic process*. New York: Wiley Interscience.
31. Guilford, M.C., P. O'Connor, C.A.S. Hall, and C.J. Cleveland. 2011. A new long term assessment of EROI for U.S. oil and gas production and discovery. *Sustainability* 3: 1866–1887.
32. Cleveland, C. 2005. Net energy from the extraction of oil and gas in the United States. *Energy* 30: 769–782.
33. Hall, C.A.S. 2004. The myth of sustainable development: Personal reflections on energy, its relation to neoclassical economics, and Stanley Jevons. *Journal of Energy Resources Technology* 126: 85–89.
34. LeClerc, G., and C.A.S. Hall, eds. 2007. *Making world development work: Scientific alternatives to neoclassical economic theory*. Albuquerque: University of New Mexico Press.
35. Ayres, R., and V. Voudouris. 2014. The economic growth enigma: capital, labour and useful energy? *Energy Policy* 64: 16–28; Ayres, R., and B. Warr. 2005. Accounting for growth: the role of physical work. *Structural Change and Economic Dynamics* 16: 181–209.
36. Kummel, R. 2011. *The second law of economics: Energy, entropy, and the origins of wealth*. New York: Springer.
37. Grandall, L., C.A.S., Hall, and M. Hook. 2011. Energy return on investment for Norwegian oil and gas in 1991-2008: Sustainability: Special Issue on EROI. 2011. Pages 2050–2070.
38. Murphy, D.J, Hall, C.A.S. 2011. Energy return on investment, peak oil, and the end of economic growth. Annals of the New York Academy of Sciences. *Special Issue on Ecological economics*. 1219: 52–72.
39. Odum, E.P. 1971. The strategy of ecosystem development. Science 164: 262–270.

Biophysical Economics: The Economics Perspective

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_5

5

5.1 Introduction

From its inception, biophysical economics has been dedicated to the unity of the approaches of natural and social science as a way to understand the interaction of humans and nature. This chapter explores methods of better accomplishing this goal. We believe that such a unity of methods of inquiry will produce a deeper understanding of the present and future problems that face us in a potentially energy-short and climate-compromised world. Usually the study of natural and social sciences is approached separately. Ecology and biophysics are studied as natural science while economics and social history as social sciences. Upon what grounds can we unify them?

This unification is not as simple as it may seem, as the methods differ greatly, not only between natural and social sciences but among social sciences themselves. Unlike many natural sciences, social sciences seldom rely on controlled laboratory experiments. If one wanted to test the proposition that lack of adequate nutrition adversely affects learning outcome in young children, it would be ethically suspect to set up a control group and feed them well and simultaneously deprive an experimental group of enough to eat. That would be considered morally reprehensible.

Along with consistent use of the scientific method, natural sciences have principles which all practitioners believe. All natural sciences must be consistent with the laws of thermodynamics and basic principles of evolution. The field of quantum mechanics unified physics and chemistry. All natural sciences share these starting points. The case is not the same with social sciences. Mainstream economists start with the idea that individuals are rational actors who respond to incentives. The goal is to find the right set of incentives to make people with self-regarding preferences cooperate with one another. However, many psychologists are reluctant to accept the rational actor model. Moreover, some try to prove what economists simply assume. Anthropologists study mainly small-scale societies and focus upon culture. Sociologists give primacy to large-scale modern societies. Political scientists believe that political processes determine behavior, while neoclassical economists place very little value on studying social or political institutions [1]. In addition, social sciences usually have a sense of purpose. But a sense of purpose implies delving into the messy world of actual human behavior and fundamental differences in what constitutes a good society. Is a good society one in which market principles are sufficient, or is it a collective society in which the government is the agent of the common good? Does voting imply democracy, or is a large-scale participation of all social classes necessary? Is a good society an equal society or one that rewards individual effort with great riches? Social sciences have debated these topics for ages, and they continue in the present day without being fundamentally resolved. We do not have definitive answers to these questions, but rather the perspective that the role played by energy in providing the material basis for our society needs to be part of the debate. Without understanding the crucial role of energy in transforming the way we do work, the way we consume, and the way we interact with one another, we doubt these debates can ever be resolved.

We cannot speak for all social sciences in this text, although as time passes, the incorporation of more social scientists and their varied approaches into biophysical economics is an important goal. A theoretically and methodologically diverse approach is capable of understanding more dimensions of complex problems than is a single, disciplinary approach. Rather we will focus on the integration of biophysical sciences with economic and historical analysis as we produce biophysical economics. Access to high-quality energy is crucially important in determining what can be produced and how. Discovering and exploiting energy resources takes place in an economic context. For example, coal is abundant and has an energy return on investment (EROI) that exceeds many forms of alternative energy such as wind and solar. Yet, at the same time, coal mines are shutting down, laying off workers, and filing for bankruptcy. Understanding this complex interaction between the possibilities found in nature and those in the human economy is the goal of biophysical economics.

In the previous chapter, we outlined a set of potential principles for biophysical economics from the perspective of biophysical science.

1. The inadequacy of neoclassical economics.
2. The need to incorporate biophysical realities into economics.
3. The importance of the fossil fuel revolution for economic growth.

4. Limits to Growth is a real (if complex) issue: peak oil and declining energy returns on investment (EROI). We must ask the question: Can renewables substitute for fossil fuels?
5. The need to improve and generate better estimates for EROI and equations for biophysical economics.

As we turn, in this chapter, to biophysical economics from a social science approach the list needs to be somewhat modified. The first principle is that biophysical economics needs to include economics. We need to study the economy in its own right and not reduce all human economic activity to questions of access to energy. The most important point is that social and economic principles *must be consistent* with the basic laws of science and with research in other behavioral sciences! With that said, let us now augment the above list.

1. Biophysical economics must go beyond a critique of neoclassical economics.
2. We have done that in many books and articles, including ▶ Chap. 3, already. Fundamentally, we need to transcend the orthodox ideas of the rational, self-interested, hedonistic, individualistic *homo economicus* operating in the idealized world of perfect competition and replace it with a broader understanding of actual humans who behave in social and institutional contexts. Moreover, biophysical economics needs to transcend the idea that the study of price formation should be the fundamental goal of economics.
3. Building upon point number one, we need to incorporate *economic reality* into biophysical economics.
4. We live in a world economy characterized by large-scale multinational corporations, which maximize profits in the long run and are often more powerful than the governments of the nations in which they operate. In addition, industrial corporations, and not just banks, are often complex financial institutions. General Motors' major profit source comes from its financing operations, General Motors Acceptance Corporation, which also sells home mortgages. We can no longer accept the notion that capital simply denotes means of production in an era where the share of gross domestic product claimed by the Finance, Insurance, and Real Estate (FIRE) sector has increased from about 30% in 1970 to more than 90% now [2]. The economics in biophysical economics must reflect the globalized, financialized, monopolized nature of the actual economy in which we live.
5. One can easily observe the connections between fossil fuel use and economic activity, but biophysical economics must explain the connection more deeply.
6. In other words, what are the causal mechanisms that link increased fuel use to overall economic performance and labor productivity? In ▶ Chap. 8, we develop such a link. The fundamental question is why did coal-driven steam engines displace water as a power source when coal was expensive and water power was essentially free, once the water wheel was constructed? The answer lies in the connection between human labor and fossil fuels. Water-powered factories were located primarily in rural areas where adequate and disciplined labor supplies were difficult to obtain. In urban centers, such as Manchester, England, large numbers of workers were ready and able to work for mere subsistence wages. The consistent power of a fossil fuel power source allowed manufacturers and inventors to produce self-acting machinery, replacing not only large numbers of workers but also reducing the need for skilled labor in the production process [3, 4]. Business leaders do not want to give up their plans for profit making, capital accumulation, and economic growth simply because high-quality energy is less available or energy prices higher. They will often reorganize the labor process spatially and in terms of skill requirements and number of workers to accomplish their goals. The causal link between fossil fuels and economic performance runs through human labor, and the transition from the solar flow to the terrestrial stock enabled a veritable revolution in labor productivity.
7. Limits to growth are economic as well as biophysical, in the sense that the process of capital accumulation is self-limiting, even without biophysical constraints.
8. There are many limits to growth built into the internal dynamics of the capital accumulation process, resulting in prolonged

periods of slow economic growth. Since the 1930s, economists have called this phenomenon "secular stagnation." Most schools of thought, from the most conservative to the most progressive, have weighed in on the causes of slow economic growth. We would like to assert that the best starting point for the development of a biophysical economic theory of stagnation, one which ties the biophysical limits to the internal dynamics of investment, lies in heterodox political economy and institutional economics, rather than in neoclassical theory. In this context, we will explore a more realistic explanation of the role that technological change plays in economic development based in epoch-making or Promethean innovations that fundamentally reorder economy and society. Answering questions about the adoption of alternative energy sources will require a realistic theory of technological change grounded in actual historical practices and power relations.

9. Developing a more formal analysis for biophysical economics is a good long-term goal, and a great deal of current research is directed toward this goal. However, there is a certain danger in a quest for a formal set of analytical tools in the absence of a solid conceptual model. Nobel Laureate Paul Krugman captured the essence of this problem when he stated: "the economics profession went astray because economists, as a group, mistook beauty, clad in impressive-looking mathematics, for truth" [5]. There are many equations in economics that constrain actual human behavior to fit the equations. If the results do not fit the idea that market economies produce efficiency, equity, and the maximum of human well-being, then the actual behavior is simply dismissed. Inclusion of concentrated industries (or monopoly power) does a great deal of harm to neoclassical theory. That is why monopolization is treated as an afterthought. On the other hand, mathematics can be very useful. It can allow one to put the barrage of data with which we are confronted into recognizable patterns that can be more easily analyzed. Mathematics alone, however, cannot substitute for a theory grounded in economic and biophysical reality. For that, we need to develop a rigorous conceptual model, whose pre-analytical vision is grounded in biophysical and economic reality. We will propose just such a set of models later in this chapter.

5.2 A Selective History of Biophysical Economic Thought

In this section, we will review prior writing by the present authors and their colleagues on biophysical economics. Since biophysical economics has a historical approach as part of its core methodology, many of the articles mentioned contain reviews of a broader literature. The term biophysical economics was first used explicitly in the 1980s by Charles Hall and his colleagues Cutler Cleveland, Robert Costanza, and Robert Kaufmann. In a paper entitled "Energy and the U.S. Economy: A Biophysical Perspective" [6], the authors test several hypotheses relating energy use to economic activity and find that gross national product (GNP) and labor productivity are correlated closely with energy use, especially when corrected for energy quality.

5.2.1 Energy and the US Economy

The economic goals of stable prices, full employment, and increasing per capita wealth were met during the long expansion from 1940 to 1970. After 1973, however, these goals became incompatible. Increased spending produced not stable prices, full employment, and prosperity but simultaneous recession and unemployment. Keynesian tools no longer worked well, and Keynesian theory fell into disarray. The present authors propose alternative explanations to the beginning of the long period of post-1973 stagnation by introducing biophysical factors such as oil consumption, the energy return on investment, and improvements in resource quality into the argument. The paper lists several goals and hypotheses. The approach was to approach macroeconomics from a thermodynamic and production perspective rather than from the traditional neoclassical view of creating well-being by exchange of goods and services for money according to human preferences. In their view, productions upgrade the organizational structure of matter and energy into lower entropy goods and services. Production is a work process which necessitates available free energy. A comprehensive analysis of economic production needs to include the thermodynamics of work. Furthermore, changes in resource quality affect the ease, and cost, of extracting energy and the economic throughput of matter and energy.

They argue that economic policy must incorporate the physical properties of resources, lest the predictions and policy recommendations be less accurate and less effective. They examined the relations between fuel use, economic output, and labor productivity over the 90 years preceding the publication of the article by the method of ordinary least squares using both time series and cross-sectional analyses. They found a high coefficient of determination (R^2) of 0.98 for time series and cross-sectional estimates. This showed a strong link between economic output and fuel use. They found that a large degree of the increase of labor productivity was due to an increase in direct energy use as well as the indirect energy use embodied in capital equipment. This is an excellent example where an understanding of the biophysical basis (increasing energy used per worker per hour) of a social process (increasing labor productivity) is facilitated by a biophysical assessment. Furthermore, they found changes in the price level correlated with changes in the money supply relative to the physical supply of energy but expressed concern that the energy costs of locating, extracting, and refining fuel have risen despite significant technological changes. Technology made previously inaccessible resources economically feasible, but at the expense of increasing energy intensity of extraction. Economic output per unit of fuel use fell 60% since 1939. Oil discoveries peaked in the 1930s, and oil production peaked in the 1970s. Since then, energy returns on investment have fallen from 30:1 in the mid-1960s to 18:1 in 1977 and to 10:1 in 2007. They conclude that if the nation wishes to sustain economic growth, alternative fuels with the same EROI as fossil fuels must be found. In the absence of such discoveries, energy availability and quality will be a limiting factor in continued economic growth.

5.2.2 The Ecology of the Economic Process

Hall, Cleveland, and Kaufmann followed the 1984 *Science* paper with a book-length monograph called *Energy and Resource Quality: The Ecology of the Economic Process* [7] in 1986. In this book, they use the principles of systems ecology to analyze economic processes, defining in the economy how energy is used to transform natural resources into goods and services to meet society's material needs. Energy and economic systems comprise a fundamental, interacting, ecosystem whose mechanism cannot be understood by viewing ecosystems and economies in isolation. Understanding the role of energy in human affairs is tied to virtually all environmental and economic questions, so energy should be an analytical focal point. They state their motivations were a fascination with human-dominated ecosystems based upon fossil fuel consumption and a dissatisfaction with the state of current economic theory. They argued that the energy basis of economic activity, is not all that determines economic phenomena but should be a crucial component to supplement standard economic analysis. This book provided detailed analyses of thermodynamics, the energy requirements of human activity, and the concept of the energy return on investment, then expressed in kilocalories.

The book introduced the method of careful examination of schools of economic thought that could serve as alternatives to the inadequately developed neoclassical economics, including the well-known figures from classical political economy such as the Physiocrats, Adam Smith, David Ricardo, and Karl Marx, along with lesser known luminaries such as Sergei Podolinsky, Fred Cottrell, Frederick Soddy, and Wilhelm Ostwald who included biophysical phenomena in their social and economic analyses. They also included a series of diagrams and conceptual models, such as the economic activity as a continuous process from solar energy to extraction, to production, to consumption and waste, along with a model of an economy embedded within a flow of energy from the sun through the ecosystem through the economy, and finally to waste heat. These models appeared in other articles by Hall and colleagues, as well as in the pages the first edition of this book (Figs. 5.1, 5.2, and 5.3). At this early stage of theoretical development, they simply inserted a circular flow of exchange value into the biophysical flows of energy and materials. Later in this chapter, we will present some more sophisticated approaches. The book provides detailed analyses of the availability of many energy sources, from conventional oil to solar power, and very detailed studies of the EROIs of agriculture, imported oil, gas, coal, and nuclear power, including the then-unanticipated costs. The book continues with a section on the general impacts of burning fossil fuels, including changes in atmospheric concentrations of carbon dioxide, human effects upon the carbon cycle, ocean acidification, and crop production. The book ends with an editorial on the fading beacon of economic growth and the stark choices that face society as the availability and quantity of oil decline.

5.2.3 Historical Perspective and Current Research Trends

The next year Cutler Cleveland produced an essay on the historical perspective and current research trends in biophysical economics [8]. Here he augmented themes that first appeared in the 1984 and 1986 works with Hall, Kaufmann, and Costanza. Ignoring physical laws has prevented standard economics from understanding fully the economic significance of changes in energy quality upon basic support service and waste assimilation. Furthermore, economic factors of production such as labor and capital depend upon low-entropy matter and energy. Neither capital nor labor, alone or in combination, can create natural resources. Cleveland provides a more detailed history of economic analysis starting with the Physiocrats and classical political economists. He then integrates the laws of thermodynamics into economic theory in a more detailed manner with the work of Podolinsky and concludes that the ultimate limits to economic growth lay not in the relations of production but in physical and economic laws. Later in this chapter, we will argue that a more complete understanding of the growth process will stem from a fuller understanding of the interaction of biophysical and internal limits found in the relations of production. Cleveland expands his analysis by focusing on Frederick Soddy, who developed an economic analysis on biophysical first principles, Alfred Lotka who initiated the discussion of Maximum Power, and the technocratic movement, who advocated a society run by technocrats rather than politicians and businessmen. Special mention is reserved for M. King Hubbert, who first enunciated the theory of peak oil and asserted that the industrial and fossil fuel era is just a transitory phase in human history. The work of Hubbert and Lotka were reflected in the work of pioneering systems ecologist Howard T. Odum who developed a systematic methodology for using energy laws to analyze the combined system of humans and nature. The *Ecological Modelling* article shows how biophysical economics was enhance by the empirical work of Energy Resources Group at the University of Illinois who developed an input-output model based on energy flows from which to calculate direct and indirect energy costs.

Particular praise is heaped upon Nicholas Georgescu-Roegen and his student, Herman Daly, for formally incorporating the laws of thermodynamics, especially the entropy law, into economic theory. Georgescu-Roegen asserted that thermodynamics was the physics of economic value and that it was the most economic of all physical laws, as it came from Sadi Carnot's experiments upon that human creation: the steam engine. This rendered the economic process unidirectional and not circular, as low-entropy energy and matter are transformed into high-entropy waste in the process of production. But the steps in between are what interests humans. Yet human agency is required to produce happiness in the human world. Low-entropy matter and energy are necessary but not sufficient.

The article ends by enunciating the principle that the absence of biophysical principles renders economic growth theory unable to make viable predictions about long-term trends, given the large and unexplained statistical residuals that are attributed to a vague and simplistic notion of exogenous technological change. From a biophysical perspective, standard economic theory needs to pay attention to the economic impacts of how changes in resource quality affect humans.

5.2.4 The Need to Reintegrate the Natural Sciences with Economics

In 2001, Charles Hall and colleagues published an article in *BioScience* calling for less isolation among academic disciplines related to economics [9]. They begin by asserting that wealth that is distributed in markets must be produced in the natural world. As part of the natural world, production must obey the laws of physics, chemistry, and biology. Unfortunately, standard economic models disregard key aspects of production. This was not always the case, as the theories of classical political economy were more fundamentally grounded in nature. Physiocratic theory gave primacy to the land as the most fundamental component in the production of wealth. For classical political economists like Adam Smith, Thomas Malthus, and David Ricardo, land as a fixed factor of production gave rise to diminishing marginal returns and the tendency of wages to drop toward subsistence level. Karl Marx was probably the first political economist of the fossil fuel era, and he fully understood the potential of coal-driven machinery to augment, and sometimes displace, human labor, resulting in rising productivity. By the 1870s, classical political economy was supplanted by neoclassical economics, so today most of the world's economic decisions are based on models that are inconsistent with nature. Hall et al. argue neoclassical theory is inadequate because (1) it is not grounded in the biophysical world (2) the basic principles of economics are logical posits, not tested hypotheses. They do not contain a flow of energy through the system but focus entirely on markets and exchange. They suggest that the circular flow model be replaced by a model that embeds the economy in the necessary energy flows, the model first introduced in *Ecology of the Economic Process*. They also critique mainstream theory for its validation processes. The fundamental assumptions about human behavior, such as acquisitiveness, rationality, and self-regarding preferences, are never put to statistical testing. Economists assert these are "maintained hypotheses" that do not require testing. However, from the biophysical point of view, these assumptions should be subjected to empirical verification.

The authors then ask the question, why does neoclassical economics assign such a low value to nature? Conventional economists do so because advanced industrial economies spend only 5–6% of their economic output on energy, which therefore gives energy a low value by the economists' monetary criteria. Although fossil-derived energy gives each of us 70–80 "energy slaves" to do the hard physical labor of yesteryear, energy is usually not included whatsoever in neoclassical production functions. The article then extends the prior work of one of the coauthors, Reiner Kümmel, who inserted both energy and creativity into the basic production function postulated by Robert Solow in his famous 1956 article "A Contribution to the Theory of Economic Growth." Solow used capital and labor as the sole independent variables in his equation, and the equations were structured to allow ample substitution of inputs (by using a Cobb-Douglas production function). While the model produced a steady-state growth path in place of the volatility of earlier models by Roy Harrod and Evsey Domar, it also produced an unexplained residual of up to 70% which Solow attributed to technological improvement (this was called the "Solow residual"). When Kümmel and colleagues included energy and creativity in the list of independent variables, after taking the elasticities (or %Δ in result/%Δ in cause) of all independent variables, and testing them with a LINEX function, the residual virtually disappeared. Energy explained nearly all of the "Solow residual" and was more powerful than either capital or labor! The social implications include the prediction that expensive labor will continue to be replaced by cheap capital and energy. Price does not always reflect scarcity and importance, and the goal of sustainable development must be reconsidered carefully in the light of energy and materials requirements. In less developed nations, policies based on neoclassical economics may lead to an overexpansion of debt, and humans tend to seek political explanations for events precipitated by biophysical causes. Biological implications of the analysis are based on the fact that agriculture, medical technology, wildlife management, and conservation all require energy. Human well-being stems from the redirecting of energy from natural food chains and processes to human ends. Finally, overpopulation, groundwater pollution, and changes in the carbon cycle and composition of the atmosphere are not externalities but part of the fossil fuel system.

5.2.5 The Early History of Modern Ecological Economics

In 2004, Ingrid Røpke authored a review article on "The Early History of Ecological Economics" [10]. She raised several methodological issues about the social and internal process of research in a quest to trace the approach to intellectual rigor. She also included institutional contexts and political factors, as well as diffused social influences. She contends that early ecological economics was quite open to diverse ideas in its conference and the pages of its journal, *Ecological Economics*. Early ecological economics conceptualized the economy in terms used to describe nature, and the focus on thermodynamics revealed half-forgotten authors that not even Nicholas Georgescu-Roegen was aware of when he wrote *The Entropy Law and the Economic Process*. These many authors showed the groundbreaking work in the 1970s through the 1990s to ground economics in biophysical reality. This is, perhaps, best summed up by a quote from English economist Mick Common. "You can't understand the last 200 years of human history without understanding energy. We could have accumulated vast amounts of capital, but it wouldn't have done what it has done for us, had it not exploited fossil fuels. Energy is what you need to do work, and doing work is what economics is all about."

Røpke listed many important themes and observations, most of which have to do with energy quality. They included the ideas that the decline in energy used per unit of gross domestic product was the result of using higher-quality fuels. Regarding labor productivity, technological change relies on capital that uses more fossil fuels per laborer, so increased labor productivity can be attributed to fossil fuels. Agriculture captures solar energy, but modern, fossil fuel-driven agriculture is far less efficient. Empirical models, such as input-output, along with distribution theory analyze the effects of energy taxation. She also asks the questions: do prices correlate with direct and indirect energy inputs? Does embodied energy provide a good measure of the value of goods and service? Røpke discusses the role played by systems theory, especially that derived from the work of Ilya Prigogine, as well as by institutional economists such as the French Regulationist School, although she recognizes that environmental analyses played only a small part in many institutionalist journals. Early ecological economics was a meeting place for researchers committed to the idea that environmental issues and biophysical limits needed to be taken seriously.

5.2.6 A New Biophysically Based Paradigm

The 2006 publication of "The Need for a New, Biophysically-Based Paradigm in Economics for the Second Half of the Age of Oil" [11] marked the first scholarly collaboration between Charles Hall and Kent Klitgaard. The paper began with the familiar critique of neoclassical economics and indicated the skepticism of the basic conceptual model among prominent economists and Nobel Laureates. But the article also introduced a new critique of ecological economics. As seen in the Røpke survey article mentioned above, ecological economics began with a call for transdisciplinary research and a commitment to methodological pluralism. By 2006, according to the authors, ecological economics had abandoned its roots and has become, in essence, a branch of mainstream environmental economics specializing in putting a monetary value on ecosystem services and natural capital.

Hall and Klitgaard then reiterate the transition from a more biophysically-based approach that characterized classical political economy to the abandonment of biophysical reality with the transition to neoclassical economics that limited its research agenda to the study of the exchange process, based on hedonistic human behavior and perfectly competitive markets. The article criticized the use of neoclassical production functions for the exclusion of energy and energy quality as an independent variable, showing that such models did not produce accurate results or predictions. The authors asserted that economics should not be solely a social science at the expense of biophysical science; stated that the object of biophysical economics was to study the biological and physical properties, structures, and processes to the actual economy; and advocated the methods of systems ecology as a starting point. The paper ends by asking the question "are we optimistic or pessimistic?" The authors expressed optimism that there are far superior ways of using resources than those of the present but pessimistic in that the decisions are too often left to market processes.

5.2.7 EROI, Peak Oil, and the End of Economic Growth

The year 2011 saw the publication of 21 articles that showed the theoretical diversity that constitutes biophysical economics. David Murphy and Charles Hall collaborated on a paper in "Ecological Economic Reviews" asserting that the causes of the long-term economic slowdown and the financial crisis of 2008–2009 could be attributed to changes in energy supply and fuel prices [12]. For the past 40 years, increased use of fossil fuels has driven economic growth. The ability to increase the global supply in the future is doubtful, given the depletion of cheaper, and easier to access, conventional oil and their replacement with lower-quality unconventional, and more expensive, sources such as deepwater wells and Canadian oil sands. This situation creates a series of feedbacks that the authors term the economic growth paradox. Further increases in oil use, given the depletion of low-cost fuel sources, will require rising energy prices; the higher prices reduce quantities demanded of fuel which dampens economic growth. Consequently, the economic growth of the last 40 years is unlikely to continue without some remarkable change in how we manage the economy.

Historically, there has been a tight correlation between oil consumption and economic growth, and aside from a few interruptions, oil supplies have kept pace with demand. Since 1970, oil consumption has increased by 40% and GDP has tripled. However, since US oil production peaked in 1970, every oil price spike has been followed by a recession. The article shows that the 1973 oil shortage produced four effects.

1. There was a decline in oil consumption.
2. The capital stock and existing technologies became too expensive to operate at higher-energy prices.
3. Marginal cost increased for manufactured goods.
4. The cost of transport fuels rose.

Expansionary periods showed the opposite trends. Lower oil prices and higher consumption were indicative of a growing economy. During times of economic expansion, oil prices averaged $37 per barrel, while they averaged $58/barrel in times of recession. Oil consumption rose 2% per annum in expansionary years and declined by 3% during recessions. According to the authors, rising oil prices are not compatible with long-term expansion. Evidence for this proposition includes the fact that production now exceeds discoveries, oil production is flat despite rising oil prices, and most of the easy-to-find oil has already been found. Much of the increase in the world's oil supply in the period of 2004–2008 came not from increases in new sources but from a drawing down of Saudi spare capacity, which fell from 6% to 2% over these years. Oil production has leveled off despite higher prices, which is an empirical phenomenon in conflict with standard economic theory. Murphy and Hall respond to critics of peak oil theory by saying to the critics who believe sufficient substitutes will be forthcoming, given the correct price incentives, by stating "you can't produce what you can't find." There is no substitute for conventional oil at the same price and the same quality.

The paper then turns to an analysis of energy returns on investment. The authors cite energy analyst Nate Gagnon's research who states that the EROI for oil from all publicly traded international companies fell from 36:1 in the 1990s to 18:1 in 2004. That was due to the fact that new sources of oil are more energy intensive to produce than are old ones and that enhanced recovery techniques that boosted production for 4 years had a short life. Oil production in fields such as Mexico's Cantarell fell precipitously in this time period. The authors predict that the production of conventional oil will continue to decline in the coming years. This renders the business-as-usual strategy of pursuing economic growth untenable because of the economic growth paradox. The causes of economic stagnation and recession can be found in the biophysical explanation of this paradox. Economic growth spurs oil demand. Increased oil production can only be met from lower EROI sources. As extraction costs rise, so do oil prices. The price increases stall economic growth, and the contraction reduces the demand for oil. The reduced demand results in lower prices. Peak oil is likely to take the form of an "undulating plateau" instead of a nicely formed Gaussian maximum. But, in the end the higher prices of more costly, lower EROI, fuels will dampen future economic growth.

5

5.2.8 Ecological Economics and Institutional Change

Lisi Krall and Kent Klitgaard also published, in 2011, a biophysical critique of mainstream ecological economics and its inability to understand the changes necessary to achieve economic justice while living within the Earth's finite limits [13]. They begin by recognizing the importance of the embedded economy as a conceptual model. Yet they criticize ecological economics for allotting too much effort to finding the right price for nature in the form of valuing natural capital and ecosystem services, and too little to understanding the foundational underpinnings and internal logic of a capitalist economy This is due largely to the affinity of many ecological economists to neoclassical methods and to their reluctance to consider fundamental social and institutional change as necessary to achieve sustainability. This leads to a cursory understanding of the systems dynamics. For example, in Costanza et al.'s *Introduction to Ecological Economics*, the authors survey, as did Cleveland, earlier economic thought. However, the reader sees Smith without the division of labor, Malthus and Ricardo without the Corn Laws, and Marx without crisis theory, the essence of these authors analyses. The early ecological treatment of the history of political economy focused primarily on enunciating the biophysical principles found in classical political economy, but did so without a broad understanding of the political and economic conditions under which these theories were advanced. Furthermore, Daly provides the neoclassical criterion of setting marginal benefits equal to marginal costs to determine when to stop producing at the macroeconomic level. However, the problem is not just *when to stop, but how to stop.*

Krall and Klitgaard contend that ecological economics has split into two branches, one focusing on valuing natural capital and the second on developing steady-state economies, and that both flow from the original work of Herman Daly. Daly contends that an economy, when it is working well, does three things. It allocates goods and services, distributes income, and determines macroeconomic scale. He proposes standards and methods for evaluating these goals. Daly also asserts that these three categories can be separated analytically. The criterion for allocation is efficiency, which can best be left to markets. Distribution should be based on justice, and macroeconomic scale should be based on sustainability or living well within Earth's limits. These last two features need to be planned. But how does one plan for justice and the absence of growth in a system that produces inequality along with goods and services and depends upon growth, without subjecting the population to increased poverty, unemployment, and lack of opportunities? Moreover, in the actual economy, allocation, distribution, and macroeconomic scale are united in the process of the reinvestment of society's economic surplus. Herman Daly was not the first to separate these categories analytically. Paul Samuelson did much the same in 1947 with his "grand neoclassical synthesis." The differences between Samuelson and Daly were that Samuelson believed that income distribution problems could be solved by the market, as could allocation, and that the government should be responsible for promoting economic growth. Daly, instead, was a proponent of a steady-state, no-growth economy, where well-being and development could be divorced from economic growth by limiting the throughput of matter and energy to the economic system, while increasing its efficiency.

However, as business historian Alfred Chandler points out in *The Visible Hand*, the efficiency improvements of the industrial revolution came by means of *increasing throughput!* This creates a conflict between the firm's need to grow and the biophysical need to reduce growth. Moreover, the purpose of a capitalist enterprise, from the smallest entrepreneur to the largest multinational corporation, is to reduce costs, expand market share, and plow the profits into increased scale of operations. Krall and Klitgaard assert that the logic of profit making at the firm level is incompatible with eliminating growth at the macroeconomic level. To achieve a steady state and any hope of sustainability, the fundamental logic of the system must be brought to the fore. The authors make the case that ecological (and biophysical) economics would be best served by abandoning neoclassical ideology as soon as possible and build a better theory based on heterodox political economy and institutional economics. They give a brief introduction to the main heterodox and institutional schools that prevail today: Social Structure of Accumulation, the Monthly Review School, and the Development without

Growth approach. The emphasis is on the compatibility of the logic of capital accumulation and the social institutions that enable it. The article ends with a quote from Thomas Jefferson. "Laws and institutions must go hand in hand with the progress of the human mind. As that becomes more developed, more enlightened, as new discoveries are made, new truths disclosed, and manners and opinions change with the change in circumstances, institutions must advance also, to keep pace with the times."

5.2.9 Ecological Economics, Degrowth, and Institutional Change

The next year, Klitgaard and Krall followed their 2011 article with a more comprehensive explication of heterodox and institutional theories [14]. They present evidence, in the form of US and global rates of investment, profits, productivity, and gross domestic product, that the age of economic growth is coming to an end. They attribute this decline to both biophysical constraints of declining energy quality and rising cost and to the internal dynamics of the capital accumulation process. Evidence shows that the economic output has been increasing at a decreasing rate since the 1970s and that employment is linked to the percentage growth rate of investment and final demand. At the same time, total output has tripled since 1970. It is the absolute accumulation of the effluents of this growth that is pressuring the environment. This creates the dilemma that we are growing both too slowly and too fast at the same time. Economic growth rates are not sufficient to support increasing employment but are too fast to live within nature's biophysical limits. The authors contend that if ecological and biophysical economists do not pay adequate attention to the social dimensions of unemployment and economic stagnation, their valuable insights on living within the planet's biophysical limits will be ignored or rejected by the population as a whole. This creates a difficult situation in that, if the economic system reaches its internal limits at the same time the biophysical limits are reached, a transition to a sustainable economy will be exceedingly difficult. To understand the possible trajectories of transition at this historical moment, we must understand the interaction of the economy and the biophysical world as a complex system and understand the boundaries, inputs, outputs, and feedback mechanisms. Mainstream, neoclassical, and Keynesian economics do not provide an adequate basis for systematic analysis in the modern era. The authors reiterate their call for the adoption of models based in heterodox political economy and institutional economics as the basis of a viable model of the social component of biophysical economics. Neither mainstream Keynesianism nor neoclassical theory recognizes sufficiently the existence of internal limits to growth that accompany the biophysical limits to growth. Heterodox political economy and institutional economics build the social limits to growth into the core of their theories and are therefore more compatible with a biophysical approach than are mainstream analyses.

Political economists have been writing about the economy as a system since the 1700s. Smith, Ricardo, John Stuart Mill, and Marx all presented comprehensive, systematic expositions of how the economy works. In the late 1930s and immediate post-Second World War period, Keynesian economists such as Evsey Domar, Alvin Hansen, and Roy Harrod presented analyses as to how the internal dynamics of the investment process led to cyclical instability and long-term stagnation. Political economists Paul Baran and Paul Sweezy surveyed the work of these economists, plus the writings in the Austrian, Marxist, and institutional traditions to produce a theory that because of the ability to produce a surplus, the problem was one of how that surplus could be spent. If not enough ways to spend the surplus could be found, the result would be chronic stagnation or low growth rates. In the 1980s, a school of thought called the *Social Structure of Accumulation* evolved from studies of how changes in the institutions of the labor process and labor markets impacted the long swings of prosperity and stagnation. By the 1990s, this analysis was elevated to include more macroeconomic variables. They recognize the advent of neoliberalism, based on privatization, remilitarization, and the distribution of wealth from labor to capital which heralded the emergence of a new Social Structure of Accumulation in the top tiers of society. The neoliberal era was grounded in growth-oriented policies that could not produce growth. The average growth rate in the decade of the 2000s, when many neoliberal policies were implemented, was a mere one-tenth of 1% higher than was the growth rate of the depression decade of the 1930s. Neoliberals call for a return to market principles of price competition to restore economic growth and stability. Historically, however, the regulatory mechanism has been one of the periodic depressions rather than subtle price adjustments driven by competition. Klitgaard and Krall end their article with a call for a new economic framework that focuses on the interaction between the internal and biophysical limits to growth. They question whether the present institutional arrangements of globalized and monopolized multinational corporations and governments that serve their interests can provide enough employment while sustaining the biophysical integrity of the planet.

5.3 Hydrocarbons and the Illusion of Sustainability

In 2016, Kent Klitgaard published an article entitled "Hydrocarbons and the Illusion of Sustainability" [15] in the special issue of *Monthly Review*, commemorating the 50th anniversary of the publication of Baran and Sweezy's *Monopoly Capital.* Klitgaard contends that although energy issues played but a minor role in Baran and Sweezy's *opus*, they presented an excellent method by which to analyze current energy dilemmas and biophysical limits, within a context of the limits found in the dynamics of the capital accumulation process. He chronicles recent declines in resource quality, the economic effects of oil price spikes, and the recent bankruptcies of coal companies. After a brief summary of the theory of monopoly capital, Klitgaard goes on to argue that the formation of monopolies went hand in hand with hydrocarbon development, from the mid-1500s when the London Hostmen's guild gained control of the British coal trade in order to restrict output and maintain prices to the role of Standard Oil in forming a domestic monopoly and becoming the world's first powerful multinational corporation. He incorporates the theory of fossil capital to argue that without access to coal to power industrial machinery, the industrial revolution would probably never have occurred. It was the switch from the solar flow to the terrestrial stock that allowed early industrialists to discipline labor adequately, drive down wages, and reduce the price of wage goods.

If, as Baran and Sweezy argue, the normal stage of monopoly capitalism is economic stagnation, what accounts for periods of prosperity? The authors of *Monopoly Capital* provide evidence that war and its aftermath and epoch-making innovations propel periods of above-normal growth. Klitgaard points out that all the epoch-making innovations that drive prosperity, the steam engine, the railroad, and the automobile, were fossil fuel intensive. He also referred to a letter from Sweezy to Nicholas Georgescu-Roegen (graciously given to Kent Klitgaard by John Gowdy, Georgescu-Roegen's PhD student) that showed not only the close personal and professional connection between Sweezy and Georgescu but also the close connection between epoch-making innovations and the species-altering "Promethean Innovations" developed by Georgescu-Roegen. In one of Sweezy's last *Monthly Review* articles, entitled "Capitalism and the Environment," Sweezy attributed growing environmental destruction to not only the increase in fossil fuel consumption but to the dynamics of capital accumulation itself. Capitalism depends upon capital accumulation, and degrowth and the steady-state economy needed to achieve life within biophysical limits are incompatible with a system that needs to grow forever. We need a system based upon decent work, equitable distribution, and respect for nature's limits, not one based on inequality and endless expansion.

5.4 Toward an Economic Theory for Biophysical Economics

A biophysical economic theory must be consistent with the principles of biophysical science. Such a theory must also be grounded in a solid historical understanding of how an actual economy works. The economic arguments of biophysical economics to date have dwelled mostly with the shortcoming of neoclassical economics and with a search for elements of greater understanding in classical political economy that preceded neoclassical economics. As seen in ▸ Chap. 2, classical political economists mostly lived in a world that either predated the world of fossil energy or was written at the formative years of the fossil economy. For them, land was a fixed factor of production that begrudgingly yielded its output. The transition to the tremendously productive power embodied in the chemical bonds of hydrogen and carbon allowed economists to stop thinking about the constraints of absolute scarcity, subsistence wages, and the inevitable arrival of the stationary state. Now all scarcity was relative to individuals' supposedly insatiable need for material comforts. Economics became the study of exchange processes and price formation. This critique has appeared regularly in the biophysical economic literature since the 1980s. The boundaries of the neoclassical system are drawn incorrectly as they do not include inputs of high-quality energy nor heat waste. Neoclassical economics ignores the second law of thermodynamics. The neoclassical framework is dominated by negative or self-canceling feedback mechanisms. Without these, self-regulation would be impossible. Moreover, neoclassical analysis ignores positive feedbacks. Positive or self-perpetuating feedbacks potentially produce tipping points and the need for fundamental, systemic, change. Maurice Dobb [16] makes the point that all new theories begin with a critique of the old. Yet it is now time to start building a biophysical theory on new methods and new ideas. In short, it is time to link theoretically the internal limits to real-world economic systems with the biophysical limits. While it is certainly possible that contemporary neoclassical economists could contribute to biophysical economics, or that the techniques of the paradigm may be useful, biophysical economics tends to reject the dominant neoclassical framework due to its inconsistencies with biophysical science. We, for example, do not advocate a rejection of all standard approaches to the quantification of money. Where then can one find a sophisticated framework by which one can make the causal link between energy quality and availability and economic outcomes?

It is now time to begin constructing such a theory. We propose that the theory starts with the actual economy that we experience today. The economy is global, concentrated, and driven by the needs of finance. It is time to abandon the unrealistic abstraction of perfect competition. A viable biophysical economic theory must be consistent with the known laws of science and the current level of research in other social science disciplines such as anthropology, political science, psychology, and sociology. It includes the notion of the embedded economy, in which the economy is a subsystem of both society and nature. The idea of an economy embedded in a larger society dates back to Karl Polanyi, and the notion of an economy embedded in a biophysical system and its energy flows traces back to at least Nicholas Georgescu-Roegen and his student Herman Daly. These ideas are abstractions, but much more realistic and complete abstractions than are those of the pure exchange economy. Biophysical economics should also include a theory of technological change, whereby changes in technology are both embodied in the economy, rather than appearing as "manna from heaven" as in much neoclassical growth theory, and can result in profound social and geographical reorganization of the economy and of society. A biophysical economic theory should also realize that the slow growth of the past four decades is not simply an aberration nor the result of poor policy choices. Rather, secular stagnation is as embedded in our current system as the economy is in nature. Slow growth is the result of changes in the accumulation process. These changes began to occur even before the age of declining resource quality and falling EROI. The economy has its own internal dynamic that operates in conjunction with biophysical constraints. It is crucial to understand both sets of limits to growth to address the problems of providing reasonable incomes and decent work to the majority of the world's population as we approach the world of the future that is likely to be slow growing, energy short, and climate compromised.

5.5 Secular Stagnation, the Theory of Monopoly Capital, and the Institutions of Accumulation

The term "secular stagnation" was coined by Alvin Harvey Hansen in his 1938 book *Full Recovery or Stagnation*, meant to explain the second crash of the Great Depression and extend Keynes' idea of an underemployment equilibrium to the long term [17]. US unemployment in 1937 rose from a level of 14% that year to 19% in 1938 and not falling into "single digits" until the Second World War began. In Hansen's terminology, the Recession of 1937 commenced long before "full recovery" occurred. Hansen believed that a mature economy, whose basic industrial infrastructure had long ago been "built from scratch," would face limited investment opportunities in the future. The epoch-making innovations of the past, such as the railroads and automobiles, were unlikely to provide for vibrant investment in the future. Furthermore, the geographical frontiers of the country had been reached, and population growth was in decline. Parenthetically, after the war when the "baby boom" began, Hansen wrote an article in *Life* Magazine declaring that kids were a built-in tool to fight recession, as the spending to support them would increase aggregate demand. He argued that stagnation was caused by shortfalls in investment and these could be caused by a number of reasons including income inequality that limits purchasing power and consumption demand, excess capacity, and market saturation. With investment opportunities vanishing, Hansen called for policies of constant and large-scale deficit spending on the part of the government to provide the demand that the private sector could not. The long postwar expansion and a new automobile boom seemed to relegate Hansen's theory to an interesting theory of the past until the economy began to stagnate in the 1970s. Forty-five years later, Lawrence Summers, former Harvard president, vice-president of the World Bank, and architect of neoliberalism, told the Federal Reserve Board that the country was, once again, in a state of secular stagnation. The response of mainstream economists ranged from dismissal to embrace. Mainstream critiques, dubbed Mainstream Ideas of Secular Stagnation (MISS), fell into two camps. Conservative economists tended to blame the slowdown in growth on exogenous, supply-side factors that would limit productivity growth such as an antibusiness climate and government regulations that raised business costs, a dysfunctional labor market where workers' skills were mismatched with available jobs, a lack of infrastructure spending, and stasis in retailing. None mention declining energy quality as a supply constraint. Liberal economists tended to favor demand-side explanations such as a reduction in capital investment associated with the digital economy (a server bank and internet connection requires fewer investment funds than does a steel mill or power plant), a debt overhang from the previous financial explosion, and credit markets that are insufficiently flexible to allow an interest rate that is low enough (essentially negative) to enable monetary policy to produce full employment [18]. Hans Despain contends that neither liberal nor conservative mainstream approaches capture the essence of the problem: that secular stagnation is built into the dynamics of the capital accumulation process.

Scholars of the left have understood this connection since the early 1930s. Michael Kalecki, a contemporary of Keynes who had published Keynes' entire system, and more, in Polish 3 years before the publication of the *General Theory*, asserted that the natural outcome of competition is monopoly concentration. The degree of monopoly could be calculated by measuring the ability to mark up prices over prime costs such as labor, machinery, and energy. This is a crucial element for a biophysical economic theory as it means in the modern economy prices are administered and not set by supply and demand. If biophysical economics is to be more than just another branch of mainstream economics, it needs to develop a sophisticated theory of administered pricing, especially as regards energy. Kalecki also recognized that the great tragedy of investment was that it was useful and could be easily overbuilt. He also realized that business cycles, in the age of demand management and fiat money, are political and can be manipulated by government policy. Josef Steindl, following in Kalecki's footsteps, asserted that endogenous factors, especially the concentration of oligopolies, were the root cause of long-term stagnation. In a competitive economy, falling profit margins due to unused productive capacity would mean bankruptcy. But in a concentrated economy, large corporations adjust to market conditions by reducing quantity not reducing prices. The increase in monopolization thereby raises profit margins but also increases excess capacity. Although gross profits may rise, excess capacity reduces net profit margins and investment stagnates because investors do not see sufficient profits forthcoming by building new capital equipment when they can utilize what they already have [18].

Paul Baran and Paul Sweezy also analyzed the mature capitalist economy in their 1966 work, *Monopoly Capital*. Their book provoked considerable controversy among political economists because they argued that Marx's observation of the tendency for the rate of profit to fall was driven by price competition. But once Marx's prediction that competitive firms were replaced by concentrated and centralized industries (now called oligopolies), the tendency for the rate of profit to fall should be replaced by the tendency of the economic surplus to rise. Starting from the classical notion of economic surplus, or the difference between the value of the output and the sum of subsistence consumption and replacement investment, they argued that modern capitalism is dominated by giant corporations (or oligopolies) which maximize long-term profits by administering prices, avoiding price competition, extending market share, and reducing the cost of production. This hypothesis of the 1960s is backed up by considerable evidence in the second decade of the twenty-first century. The number of industries in which the top four firms control 50% or more of the market has risen from 5 to 185 since the 1950s. Gross profits of the top 200 US corporations have risen from about 14% in 1950 to approximately 30% in 2008.

As a result, the economic surplus tends to rise and needs to be absorbed by finding adequate spending outlets. If it is not, production will decline and chronic stagnation will appear. Baran and Sweezy stated that there were three methods of absorbing this rising economic surplus: it could be consumed, invested, or simply wasted. To analyze the increase in consumption to levels sufficient to avoid stagnation, Baran and Sweezy chronicle the development of the "sales effort." Mass consumption was not the result of rational consumers maximizing their subjective utilities subject to limited incomes, but a conscious effort on the part of profit-seeking corporations and the state to assure that consumption levels are adequate to absorb economic surplus by creating needs that did not exist in the past and products to fulfill them. Investment directly absorbs the economic surplus but simultaneously creates more surplus to be absorbed in the next period. Waste such as planned obsolescence or excessive military spending could also serve as a potential absorber as well as war itself. Baran and Sweezy show that a market economy would succumb to long-term stagnation in the absence of waste. If they are correct, moving toward sustainability by reducing waste may exacerbate the economic stagnation that is already occurring within our current economic structure. If the economy depends upon ever-growing consumption, then it will be quite difficult to live well within nature's limits, especially as the fossil energy needed to produce the goods and services is declining in quality. It is certainly possible to see the overextension of credit in our present era in the same vein. Certainly, in a rationally planned economy, employment could be boosted, and the environment improved, by large-scale public investment in nonfossil transportation and the construction of a nonfossil infrastructure. However, Baran and Sweezy argue that large-scale public investment would not absorb sufficiently the economic surplus generated by the economy because of the power relations of monopoly capitalism. Public investment that competed effectively with the private sector would be kept within limits. Their argument seems to have contemporary relevance, as the role of the government as a demand manager is being debated both in the United States and in Europe at the present time.

Because of the chronically unabsorbed surplus, the normal state of a concentrated industrial economy is slow growth, or secular stagnation, not the assumed steady-state growth path of neoclassical economics. In fact, the economic literature also refers to secular stagnation as the "Sweezy normal state." However, if stagnation is the normal state of the economy, how would one explain periods of prosperity such as those that occurred in the 1960s? One biophysical explanation is low oil prices for a prolonged period that allowed for the increase in labor productivity. Yet the theory of monopoly capitalism adds a different dimension. Baran and Sweezy attributed prosperity to either war and its aftermath or epoch-making innovations. The end of the Second World War saw the United States rise to the position of global hegemon. It controlled the world's financial system, had sole possession of nuclear weapons until the late 1940s, and had the world's only viable industry after the war. By the 1970s, the international monetary accords had fallen apart, Germany and Japan had caught up industrially with the United States, and the United States spent billions of dollars fighting wars in Southeast Asia. Epoch-making innovations that stimulate demand and employment, absorb vast quantities of investment capital, create myriad peripheral industries, and result in large-scale geographic shifts are few and far between. Baran and Sweezy list only three: the steam engine, the railroad, and the automobile. All these innovations were propelled by cheap and available fossil fuel. Without the automobile, we would not have the shopping mall, suburban housing, fast food, nor the soccer mom. In the era of declining energy quality and availability, will there be an alternative vehicle by which to absorb surplus? Certainly, the Internet and social media have provided nowhere near the same levels of employment and investment, although they are a ubiquitous part of the lives of many today. Biophysical economics would be well served by developing a theory that links fossil fuel use to the institutions of accumulation and the needs for employment.

In the 1980s, Sweezy and Harry Magdoff turned their attention to the rise of financial institutions in the pages of their journal, *Monthly Review*. They argued against the mainstream proposition that the exploding number of financial instruments were dragging down real investment. Rather they asserted, and backed with considerable statistical evidence, that investment funds were flowing toward Finance, Insurance, and Real Estate (FIRE) precisely because the real economy was stagnant and profitable investments were not forthcoming, especially in the second half of the age of oil. The share of GDP accruing to the FIRE sector increased from about 30% at the start of the second half of the age of oil (1970) to more than 90% by 2010 [2]. To put the matter bluntly, the economy was kept from even more serious stagnation by a combination of military spending, financial speculation, and conspicuous consumption. How is sustainability to be accomplished without a fundamental reorganization of society's institutions when these are the primary drivers of even sluggish growth?

5.6 The Social Structure of Accumulation

Further explorations in political economy and institutional economics have focused upon the interaction of short-term business cycles, long-term trends of expansion and stagnation, and the institutional structure in which economic activity takes place. One of the most fruitful of these explorations is the work of the Social Structure of Accumulation theorists. The many economists writing in this tradition define a Social Structure of Accumulation as the institutional context in which profit making occurs. Unlike Baran, Sweezy and Magdoff, who came of age during the Great Depression, they represent a new generation who came to academic maturity in the long post-Second World War expansion and questioned the idea of secular stagnation. Instead, they embraced Nikolai Kondratieff's theory of long waves and began to link their phases of expansion and contraction to changes in the conditions of labor. Kondratieff's theory was embraced by Harvard economist Joseph Schumpeter as an alternative explanation for long-term decline to that of Hansen, his great intellectual rival. Although Schumpeter was himself very conservative, he nurtured and supported young scholars of all political inclinations, including Paul Sweezy, Paul Samuelson, and Nicholas Georgescu-Roegen. The roots of liberal neoclassicism, neo-Marxism, and biophysical economics all trace in a way back to Schumpeter.

The institutional revival of the 1970s and 1980s showed that the functioning of markets is embedded within a context of social institutions. Just like embedding the economy in a finite and nongrowing biophysical system forces us to think about the limits of the primary system, embedding the functioning of markets within a social system forces us to think about the interaction of markets with the broader set of institutions. David Gordon and colleagues termed this interaction of macroeconomic cycles and the institutional context the Social Structure of Accumulation (SSA). An SSA is the institutional context in which capital accumulation occurs. In some historical eras, the institutions are broadly supportive of profit making, and the SSA enters an expansion phase. The economy enters a long swing of growth [19]. At some point, however, the institutional conditions change, and the SSA collapses, leaving a decline of roughly 20–25 years in its wake. Phillip O'Hara summarizes this position succinctly when he states: "The system requires certain 'public goods' or systems functions to promote accord, agreement, organization, communication, and information to moderate conflict and instability that so-called 'free markets' would otherwise largely be without" [20].

SSAs go through distinct phases of exploration, consolidation, and decay. A long wave with an undertone of stagnation coincides with collapse of an SSA, for example, the SSA of the early twentieth century industrial revolution collapsed in the Great Depression. Progressive capitalists explore innovative ways of conducting production and marketing. As they become successful, a new set of institutional arrangements are consolidated and become the basis of a long period of growth. Eventually, after 20 some years, changes in variables such as technology, world power arrangements, and labor organization cause the SSA to decay. In the decay, the world economy begins to stagnate and a new long wave with an undertone of stagnation ensues.

Bowles, Gordon, and Weisskopf [21] extended the determinants of a Social Structure of Accumulation from the conditions of labor to broader categories of world relations and domestic considerations. A postwar SSA was constructed based on US hegemony, the recognition of unions in a limited capital-labor accord, the limitation of price competition among large firms, and a capital-citizen accord based on the politics of economic growth. The postwar SSA could not survive the early 1970s with the collapse of the Bretton Woods accords, the peak and beginning of the decline of US oil production, and the era of stagflation. The conflict was between a system that needed growth for economic and political purposes but simply could not produce it. After a period of impasse, a new, neoliberal SSA began to be constructed upon more conservative goals of (1) removal of international barriers to the free movement of commodities and capital, (2) the withdrawal of the state from regulatory activity, (3) privatization of state enterprises and public services, (4) a shift to regressive taxation, (5) the end of the capital-labor accord, (6) the replacement of corespective oligopoly behavior by renewed competition, and (7) a faith in entrepreneurial spirit and free market ideology [21]. The most recent SSA is coming to an end as inequality, stagnation, increasing resource scarcity, and the exaggerated positive feedback loops, exacerbated by speculative finance, create untenable conditions for the long-term stability of the system [22].

The SSA is supposed to provide the institutional framework for long-term sustainable growth, at least until it breaks down. Yet if, as seems likely, every scientific measure of human impact upon nature indicates that we are in overshoot, then there is no possibility of configuring a new Social Structure of Accumulation based on renewed growth. Rather degrowth is demanded, a social structure of deaccumulation. But at the same time, the main power structures of government and corporations and their supporting institutional structures believe that growth is needed to achieve a stable prosperous economy, with the absence of growth seen as economic crisis. Wolfson and Kotz state the matter forthrightly: "Capitalism does indeed display a powerful accumulation drive. That drive is one of its central features. It is doubtful whether capitalism could survive without the accumulation of capital—it would be torn apart by the conflict without an 'expanding pie'" [23].

The fundamental differences between the Monthly Review School and the Social Structure of Accumulation approach are secular stagnation vs. long waves and, epoch-making innovations vs. institutional restructuring. The Social Structure of Accumulation school believes that the global economy is seeing renewed competition where the Monthly Review school sees another form of oligopolistic rivalry. The SSA school also believes that the right set of social institutions can produce another period of long-term growth. That is harder to believe in the age of declining resource quality, but there are many important lessons to be learned from both approaches. Most importantly, these examples ground their theories in actually existing economies that change historically and within the context of social institutions. We believe that they could serve as a good starting point, although not the definitive ending point, of a viable theory of growth for biophysical economics.

5.7 Equations and the Conceptual Model

Before we rush headlong into formalizing a set of equations by which to describe biophysical economics, we should first establish a solid conceptual model. The equations of mainstream economics are derived from the pre-analytical vision that the economy is self-contained and self-regulating by means of price competition. We reject both of those notions. Rather than reproduce equations based on a faulty conceptual model, it is time to advance candidates for a better starting point.

In ► Chap. 3, we presented a model in which the economy was embedded in a larger biophysical system that was dependent upon a flow of solar energy, entering as visible light and exiting as waste heat. However, that model, first advanced in *Energy and Resource Quality: The Ecology of the Economic Process*, places a simple circular flow model within the economy which is also embedded within the environment. Subsequent research has shown that the circular flow is an inadequate way to model the complex interactions of a biophysical economy grounded in solar flow, fossil

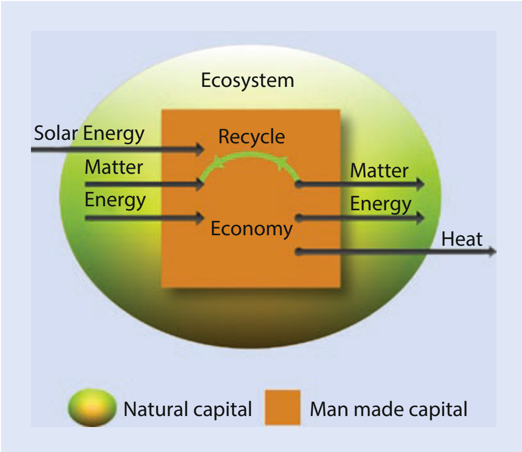


Fig. 5.1 This figure is a depiction of the basic ecological economics model developed by Herman Daly. It shows that the economy is embedded within in the ecosystem, and also shows the transformation of low-entropy solar energy into high-entropy heat

fuels, extraction, production, distribution, and waste. There is no role for institutions or actual human behavior in this model. We must do better.

We would like to advance three candidates for the conceptual starting point of biophysical economics. The first is an early visual model that was advanced by ecological economist Herman Daly [24; Fig 5.1]. The modeling of an embedded economy is one of his greatest accomplishments. Daly puts a growing, open economy inside a finite and nongrowing ecosystem. He then differentiates between an empty world, filled with natural capital but largely devoid of human-made capital, and a full world that is abundant with human-made capital but in which the products of nature have become seriously depleted. The primary purpose of the model was to show the need for a steady-state economy that operates within nature's finite limits. This model has been also developed by Hall et al. as given in Fig. 3.3.

The second was another visual model developed by Neva Goodwin, Jonathan Harris and their colleagues at the Global Development and Environment Institute (GDAE) associated with Tufts University in Medford, Massachusetts, USA [25; Fig. 5.2]. The model embeds the economy

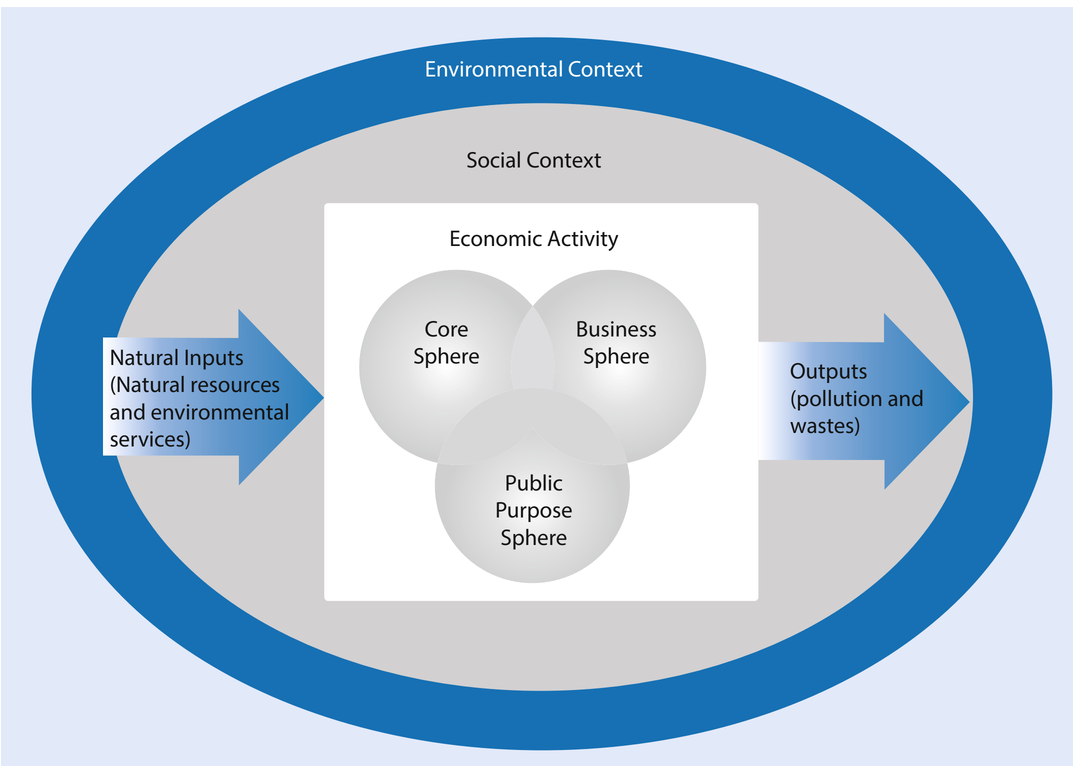


Fig. 5.2 The model embeds the economy not only in an ecological context but also in a social context (Neva Goodwin [25])

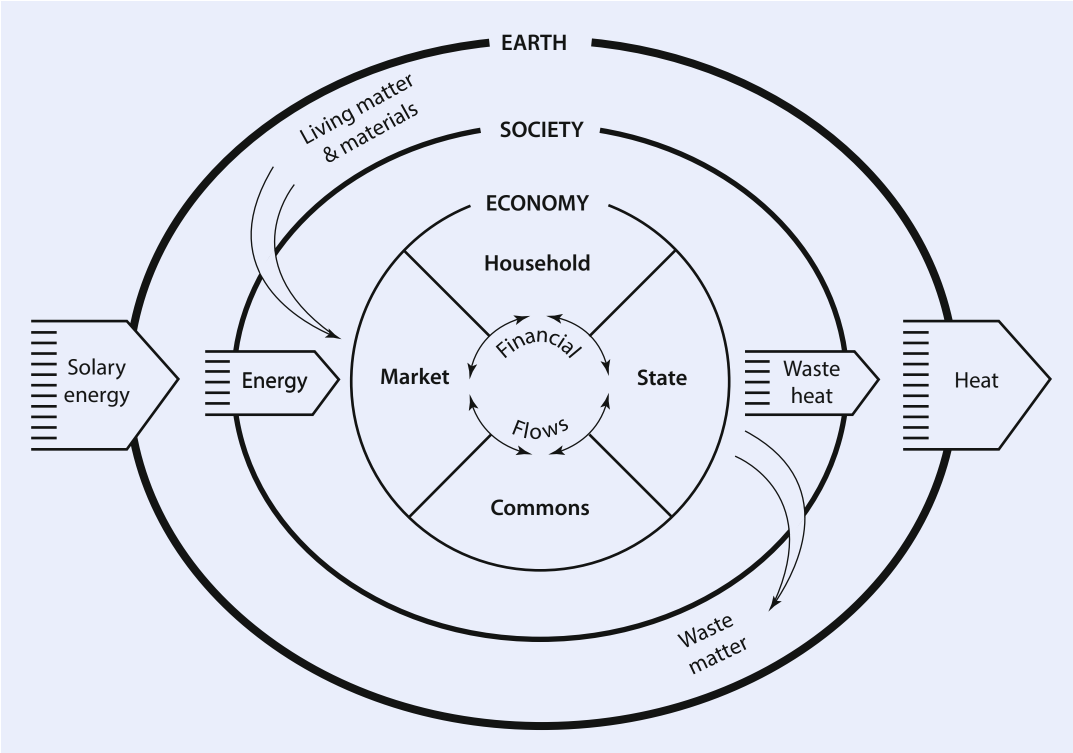


Fig. 5.3 Shows an economy that includes not just the market but also household production, the government, and the all-important commons (Kate Raworth [26])

not only in an ecological context but also in a social context. Not all human interactions are exchange relations. In the real world, there are interpersonal interactions that do not involve the transfer of money. This part of the economy is termed the core sector. The part of the economy modeled by mainstream economics is called the business sector, while the model adds a public purpose sector of governments, nongovernmental organizations, and not-for-profit enterprises. The use of Venn diagrams shows direct, personal interaction among the sectors, not just indirect interaction mediated by markets.

A third, but similar, approach is the brainchild of development economist Kate Raworth and is used to model her commitment to a safe and just operating space for humanity [26]. She asserts that the visual pictures of neoclassical economics are all wrong and need to be replaced by images that see the big picture, nurture human nature, and show skepticism about economic growth. Her conceptual model shows an economy that includes not just the market but also household production, the government, and the all-important commons. The recent work of GDAE-affiliated public policy analyst June Sekera [27] shows clearly how the very notion of public service and the commons have taken a beating in the neoliberal era. Restoring the commons to a prominent place in the pre-analytical vision is a welcome addition in our opinion.

None of these models fits the exact needs of biophysical economics. All are rather vague about the role of energy. Yet they are a much better starting point than is the circular flow model based on hedonistic human behavior, perfect competition, and pure exchange. When we get the conceptual model specified sufficiently, a set of equations will be forthcoming.

References

1. Gintis, Herbert. 2007. A framework for the investigation of the behavioral sciences. *Behavioral and Brain Sciences*. 30: 1–16.
2. Foster, John Bellamy, and Fred Magdoff. 2012. *The endless crisis*. New York: Monthly Review Press.
3. Malm, Andreas. 2015. *Fossil capital*. London: Verso.
4. Braverman, Harry. 1974. *Labor and monopoly capital*. New York: Monthly Review Press.
5. Paul Krugman. 2009. How did economists get it so wrong? *New York Times*. September 2.

6. Cleveland, Cutler J., Robert Costanza, Charles A.S. Hall, and Robert Kaufmann. 1984. Energy and the US economy: A biophysical perspective. *Science* 225: 890–897.
7. Hall, Charles A.S., Cutler J. Cleveland, and Robert Kaufmann. *Energy and resource quality: The ecology of the economic process*. New York: Wiley Interscience.
8. Cleveland, Cutler J. 1987. Biophysical economics: Historical perspectives and current research trends. *Ecological Modelling* 38: 47–73.
9. Hall, Charles A.S., Dietmer Lindberger, Reiner Kümmel, Tim Kroeger, and Wolfgang Eichhorn. 2001. The need to integrate the natural sciences with economics. *Bioscience* 57 (8): 663–673.
10. Røpke, Ingrid. 2004. The early modern history of ecological economics. *Ecological Economics* 50: 293–314.
11. Hall, Charles A.S., and Kent Klitgaard. 2006. The need for a new, biophysically-based paradigm for economics in the second half of the age of oil. *International Journal of Transdisciplinary Research*. 1 (1): 4–22.
12. Murphy, David J., and Charles A.S. Hall. 2011. Energy return on investment, peak oil, and the end of economic growth. In *Ecological economic reviews*, ed. Robert Costanza, Karin Limburg, and Ida Kubiszewski, vol. 1219, 52–72. Boston: Blackwell on behalf of the New York Academy of Sciences.
13. Krall, Lisi, and Kent Klitgaard. 2011. Ecological economics and institutional change. In *Ecological economic reviews*, ed. Robert Costanza, Karin Limburg, and Ida Kubiszewski, vol. 1219, 247–253. Boston: Blackwell on behalf of the New York Academy of Sciences.
14. Klitgaard, Kent A., and Lisi Krall. 2012. Ecological economics, degrowth, and institutional change. *Ecological Economics* 84: 185–196.
15. Klitgaard, Kent A. 2016. Hydrocarbons and the illusion of sustainability. *Monthly Review* 68 (3): 77–88.
16. Dobb, Maurice. 1975. *Theories of value and distribution since Adam Smith*. Cambridge: Cambridge University Press.
17. Alvin Hansen. 1938. *Full recovery or stagnation?* New York: W.W. Norton and Company. Alvin Hansen was Kent Klitgaard's "academic grandfather," the dissertation supervisor of Sam Rosen, who was Klitgaard's dissertation supervisor. Other PhD students of Hansen included Evsey Domar and Hyman Minsky.
18. Despain, Hans G. 2015. Secular stagnation: Mainstream versus Marxian traditions. *Monthly Review* 67 (4).
19. Gordon, David M., Richard Edwards, and Michael Reich. 1982. *Segmented work, divided workers*. Cambridge: Cambridge University Press.
20. O'Hara, Phillip. 2006. *Growth and development in the global political economy*, 17. London: Routledge.
21. Bowles, Samuel, David M. Gordon, and Thomas Weisskopf. 1990. *After the wasteland*. Armonk: M.E. Sharpe.
22. Kotz, David. 2009. The financial and economic crisis of 2008: A systematic crisis of neoliberal capitalism. *Review of Radical Political Economics*. 41: 395–317.
23. Wolfson, Martin, and David Kotz. 2010. A reconceptualization of the social structure of accumulation theory. In *Contemporary capitalism and its crises*, ed. T. McDonough, D.M. Kotz, and M. Reich, 79. Cambridge: Cambridge University Press.
24. Daly, Herman. 1996. *Beyond growth*. Boston: The Beacon Press.
25. Neva Goodwin. J. Harris, J. Nelson, B. Roach, and M. Torras. 2014. *Microeconomics in context*. Armonk: M.E. Sharpe.
26. Raworth, Kate. 2017. *Donut economics*. White River Junction: Chelsea Green.
27. Sekera, June. 2016. *The public economy in crisis: A call for a new public economics*. New York: Springer.

Energy and Wealth: An Historical Perspective

The first five chapters focused mostly on economics, that is, the procedures by which we study our economies. It included reviews of the main ways we use today and in the past, and critiques of the dominant forms today. It offered an alternative perspective based on including natural as well as social sciences in the consideration of economics. The next three chapters focus more on the economies themselves, including their historical and biophysical basis. We believe that these reviews reinforce the virtues of using a biophysical approach to understanding real economies.

Contents

The Evolution of Humans and Their Economies

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_6

The first four chapters focused mostly on economics, that is, the procedures by which we study our economies. This chapter focuses more about the economies themselves, including their historical and biophysical basis.

6.1 The History of Formal Thought on Surplus Energy

Many natural and social scientists from different disciplines have thought deeply about the long-term relation of humans and wealth production. Many of them have concluded that the best general way to think about how societies evolved over time is from the perspective of surplus energy. Human history, including contemporary events, is essentially about exploiting energy and the technologies to do so. This is not the perspective taught in our schools, and the role of energy is essentially missing from our dominant history. Instead, human history usually is seen in terms of generals, politicians, and other personalities. But the options and successes and failures of these generals, politicians, and others are extremely dependent upon the energy and other resources available to them for undertaking whatever they undertake.

This chapter will develop the alternative perspective that the fates of past civilizations and other events of the past can be better understood from the perspective of energy availability and in particular surplus energy. *Energy surplus* (or net energy) is defined broadly as the amount of energy left over after the costs of obtaining the energy have been accounted for. The energy literature is quite rich with papers and books that emphasize the importance of energy surplus as a necessary criterion for the survival and growth of many species, including humans and the development of science, art, culture, and indeed civilization itself. While each acknowledges that other issues such as human inventiveness, nutrient cycling, and entropy (among many others) can be important, each is also of the opinion that it is energy itself, and especially surplus energy, that is key. The issue is not simply whether there is surplus energy but how much, what kind (quality), and at what rate it is or was delivered. The interplay of those three factors determined the flow of net energy and hence the ability of a given society, whether modern or ancient, to divert attention from growing sufficient food or the attainment of water toward trade, warfare, or luxuries, including art and scholarship. Indeed, humans could not possibly have made it this far through evolutionary time, or even from one generation to the next, without there being some kind of net positive energy, and they could not have constructed such comprehensive cities and civilizations or wasted so much in war, without there being substantial surplus energy in the past.

6.2 The Prehistory of Human Society: Living on Nature's Terms

Human populations must first feed themselves and after that generate sufficient net energy to survive, reproduce, and adapt to changing conditions. While a moderately small percentage of people in industrial societies today worry about getting enough to eat, for many in the less industrialized global South, getting sufficient food is still a major concern. The focus on food acquisition has also occupied much of humanity's time throughout history. For at least 98% of the 2 or so million years that we have been recognizably human, the principal technology by which we as humans have fed ourselves has been that of hunting and gathering. Contemporary hunter-gatherers—such as the !Kung of the Kalahari Desert in Southern Africa—probably live as close to the lifestyle of our long-term ancestors as we will be able to understand. Studies by anthropologists such as by Lee and Rapaport confirmed that indeed present-day (or at least recent) hunter-gatherers and shifting cultivators acted in ways that appeared to maximize their own energy return on investment.

Richard Lee studied the energetics of the !Kung while they were relatively unaffected by modern civilization [1]. A charming, although romanticized, view of their culture is readily accessible in the movie "The Gods Must be Crazy." Life for a hunter-gatherer is basically about taking nature as it is found and finding ways to survive on those resources. The key challenge was gaining the needed food energy. For the !Kung, this was undertaken by women gathering mongongo nuts and men hunting antelope and other animals. Mongongo nuts are the most abundant resource that provides the largest part of the energy and

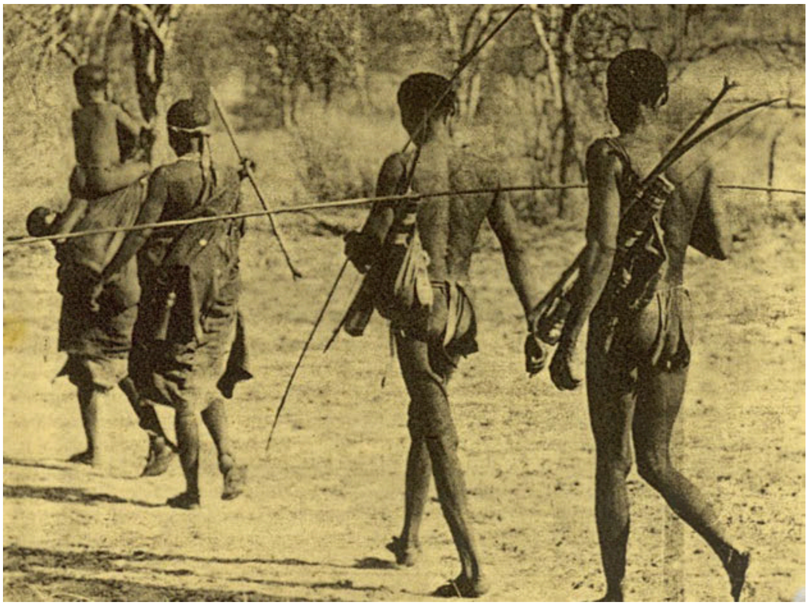


Fig. 6.1 !Kung people, modern-day hunter-gatherers, probably represent how all of our ancestors lived their lives for far more time than even the time since the start of agriculture (Source: Science Magazine)

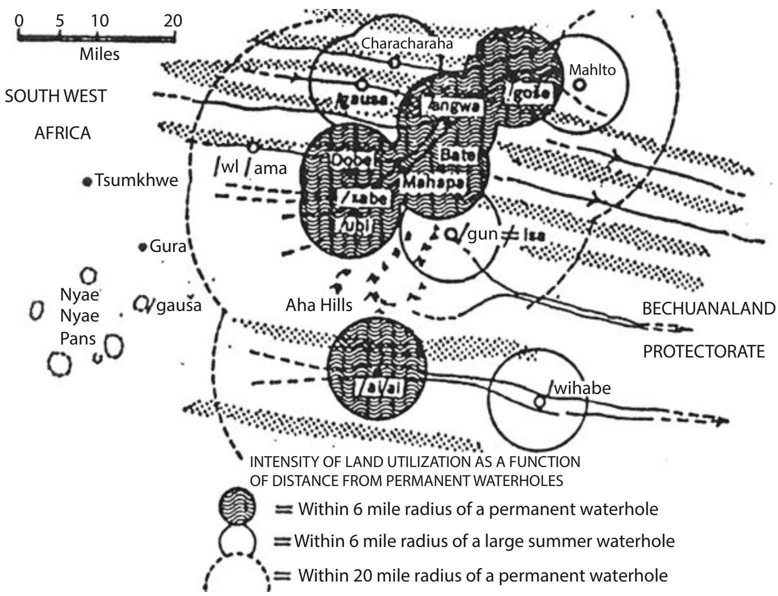
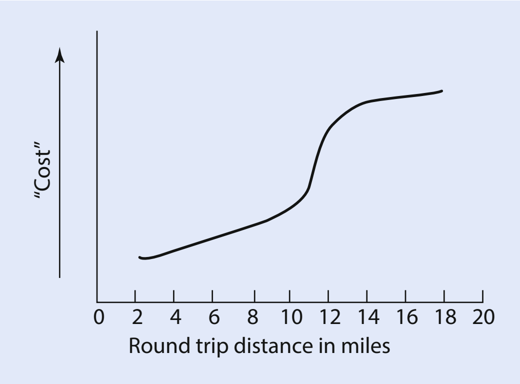


Fig. 6.2 Map of the various waterholes in the Kalahari desert that the!Kung migrate from and to over the seasons. The exhaustion of easy food in the region of one waterhole necessitates movement to another (Source: Lee 1973)

protein consumed by the !Kung, although game is very much appreciated and gives needed additional protein to the diet. Life is good for the !Kung, at least before their major contact with civilization (Fig. 6.1). According to Lee's studies, the !Kung spent far fewer hours working each day than do people living in industrial societies; a lot of their time was spent in leisure activities. Young women tended to be sexually active (which was considered normal) from as early as age 9, but tended not to get pregnant until about age 18, when they had sequestered enough body fat so that the pregnancy was possible (i.e., it appears that the human body protects young women from pregnancy when they do not have enough energy surplus to carry a fetus) [2]. Life for the !Kung was not always simple, however, for they lived in a desert and were constrained by their need for water and food. In their homelands of Botswana, there are only a relatively few waterholes, and it is essential to set up camp near one of these waterholes (Fig. 6.2). Mongongo trees are spread ran-



6

Fig. 6.3 Determinants of !Kung EROI. At a distance of about 11 miles, energy cost increases greatly because an additional day is needed. When the !Kung have exhausted the mongongo nuts within 1 day's walk, they have to make a substantial investment in walking 2 days to get a new supply of nuts (Source: Lee 1973)

domly around this part of the Kalahari Desert so initially the !Kung can derive all the food they need from relatively short excursions from their camp. As time goes on, they deplete the nuts (and game) within easy reach, so that each day they must make a longer and longer trip to gather enough mongongo nuts to feed their families. At some point, they have gathered all the mongongo nuts within a day's hike. Then they must make much longer, overnight trips to get them. Since they eat a lot of food both going and coming back, they consume a substantial portion of the food they went out to get! This greatly increases their energy investment and lowers what we call their energy return on investment or EROI (Fig. 6.3; see ► Chap. 13). This makes it desirable at some point to make the additional investment of moving to a new waterhole.

According to Lee, the !Kung lifestyle, under normal circumstances, generates a quite positive energy return on investment (i.e., generates a large surplus) from their desert environment, perhaps an average of some 10 Kcal returned per one of their own Kcal invested in hunting and gathering. New studies indicate that hunting may have an even higher EROI than gathering [2]. In normal times, these cultures had plenty to eat, and the people tended to use the surplus time made available from their relatively high EROI lifestyles in socializing, childcare, and storytelling. The downside was that there were periodic tough times, such as droughts, during which starvation was a possibility. It is probable that our ancestors had a fairly positive EROI for much of the time, although periodic droughts, diseases, and wars must have occasionally, or perhaps routinely, taken a large toll. Thus even though the !Kung, and by implication other hunter-gatherers, had a relatively high EROI, perhaps 10:1, human populations tended to be relatively stable over a very long time, barely growing year to year from millions of years ago until about 1900. Thus, even this relatively high energy return was not enough to generate much in the way of net population growth over time.

It is increasingly clear that our Stone Age hunter-gatherer ancestors, as hunter-gatherers today, tended to be quite good hunters. This hunting prowess resulted in an enormous environmental impact on the large birds and mammals of the earlier world. As humans spread about the world, they encountered in each new place large, naive herbivorous animals of the sort we do not see anywhere on Earth today. For example, the new arrivals in North America found giant beavers, rhinoceros, two species of elephants, camels, and so on. Human arrivals in Australia found giant flightless birds, while the first humans into Italy found large turtles no longer extant and so on. None of these large animals are there today, and except in Africa, there are few animal species larger than 100–200 kg left. These large animals were abundant prior to human arrival (Table 6.1). (Of course, bison, bears, moose, and elk are large and still with us, although in greatly reduced ranges.)

What caused their extinction? There are two competing hypotheses. First, since the climate was warming rapidly 10,000 years ago, it is possible they succumbed to some effect of climate change. The second hypothesis is that humans hunted these animals to extinction. These large animals had no previous reason to be afraid of anything as small and puny as a human being, or that humans could simply walk up to these animals and stick a spear into their side. Africa still has many, many very large herbivorous species, probably because the animals coevolved with humans as they slowly became more proficient hunters with better weapons. All around the world where humans came later, most or all the larger animals disappeared within 2000 years of human arrival. This certainly supports the idea that it was humans who did them in [3]. The fact

Table 6.1 Megafaunal extinctions

	Extinct	Living	Total	% Extinct	Landmass (km²)
Africa	7	42	49	14.3	30.2×10^6
Europe	15	9	24	60	10.4×10^6
North America	33	12	45	73.3	23.7×10^6
South America	46	12	58	79.6	17.8×10^6
Australia	19	3	22	86.4	7.7×10^6

Late Quaternary (last 100,000 years) extinct and living genera of terrestrial megafauna >44 kg adult body weight) of five continents. Adapted after [3]. Data for extinct and living European megafauna from Martin (1984). For Australia, it may be that as many as eight genera were already extinct before human arrival (Roberts et al. 2001). If so, this reduces both the number and percentage of megafaunal extinctions that could conceivably be attributed to human activity
Source: Wroe et al. [32]

that these same animal species had survived many previous climate changes lends support to the human-caused extinction idea. Thus, significant environmental impact from humans is hardly new.

6.2.1 African Origin and Human Migrations

All available evidence suggests that humans and their predecessors evolved in Africa, which is the only place we have found human fossils or evidence dating to roughly 1.7–1.8 million years ago [4]. Take a mental time trip to East Africa about 2 or 2.5 million years ago. You would be in the cradle of the evolution and development of all that makes us human. Remarkably you would find not one but perhaps half a dozen types of early humans (or hominids), each group as distinct from one another as chimpanzees from gorillas. Most of these protohominids were found in small migratory bands at the transition of forests to drier savannas. We continue to learn more about our ancestors. The finding in the 1990s of the fossils of what appears to be the ancestor of humans that lived some 4–6 million years ago is cause for great excitement among those who are determining our lineage. This creature, named *Ardipithecus ramidus* (Ardi for short), walked upright but still spent much, perhaps the majority, of its time in trees (Fig. 6.4). There was strong natural selection for developing hands with opposable thumbs that could grasp branches more firmly than with all digits on one side, preadapting humans for our present hands, very useful for the coming agricultural and industrial environments as well as such amenities as musical instruments.

Recent research has found that a human uses only about one quarter as much energy to walk 100 m as a same-sized chimpanzee, so there obviously has been a tradeoff favoring more energy-efficient walking over the ability to both walk and climb trees well, as Chimpanzees can. Probably most of the Ardis made, or at least used, tools of some sort, for we understand now that even chimpanzees have a rather astonishing ability to make many different types of tools, including stone anvils. Most of their tools were made from organic materials and hence are not well preserved, so we know little about the past of tool making of either chimps or protohominids. By about 2.5 million years ago, our ancestors had developed quite sophisticated methods for making stone knives and spear points by striking or stroking one rock on another in repeated and often sophisticated patterns. There are even a number of ancient "industrial complexes" in, for example, Kenya's Olduvai Gorge, a rich hunting ground for information about our ancestors (Fig. 6.5). Spear points and knife blades are actually energy technologies—energy (force)-concentrating devices that allow the strength of a human arm to be multiplied many times when concentrated on a line or point (Fig. 6.6). This allowed humans to exploit many new animal resources and eventually

6

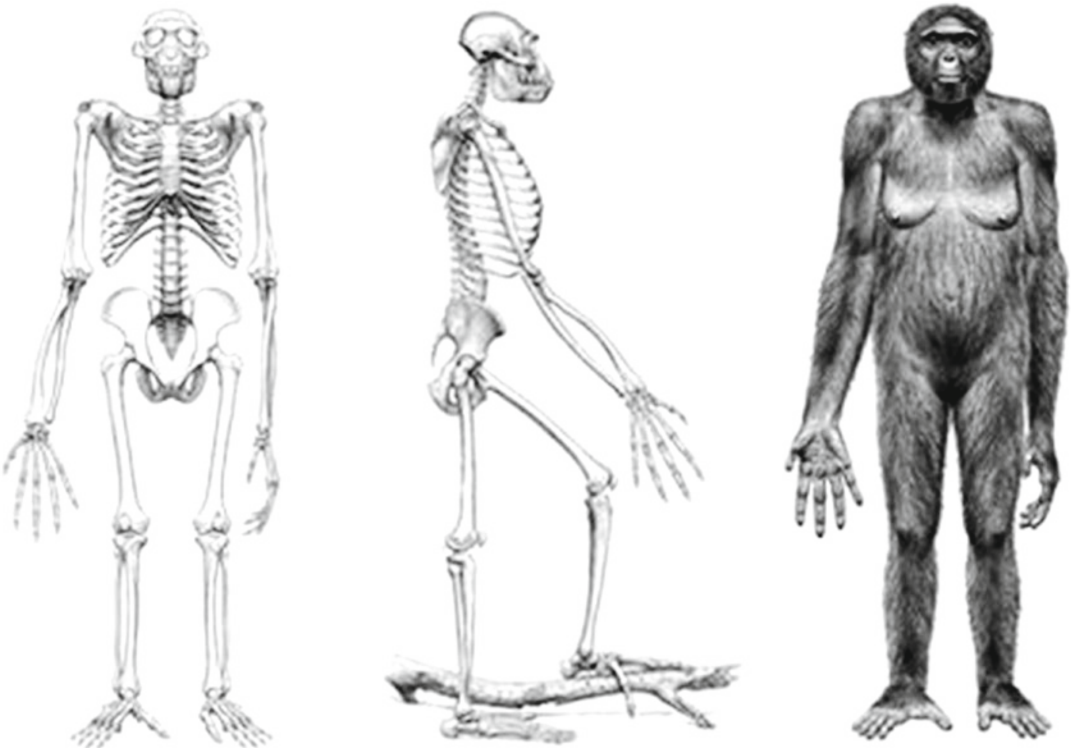


Fig. 6.4 Ardi, *Ardipithecus ramidus*, is a new found fossil that is neither man nor ape but probably represents our human ancestors some 4 million years ago (Source: Science Magazine, Jay Matternes)

Fig. 6.5 Olduvai Gorge (from Shunya website). Many very early human remains have been found here as well as early "industrial" sites, where stone tools were manufactured

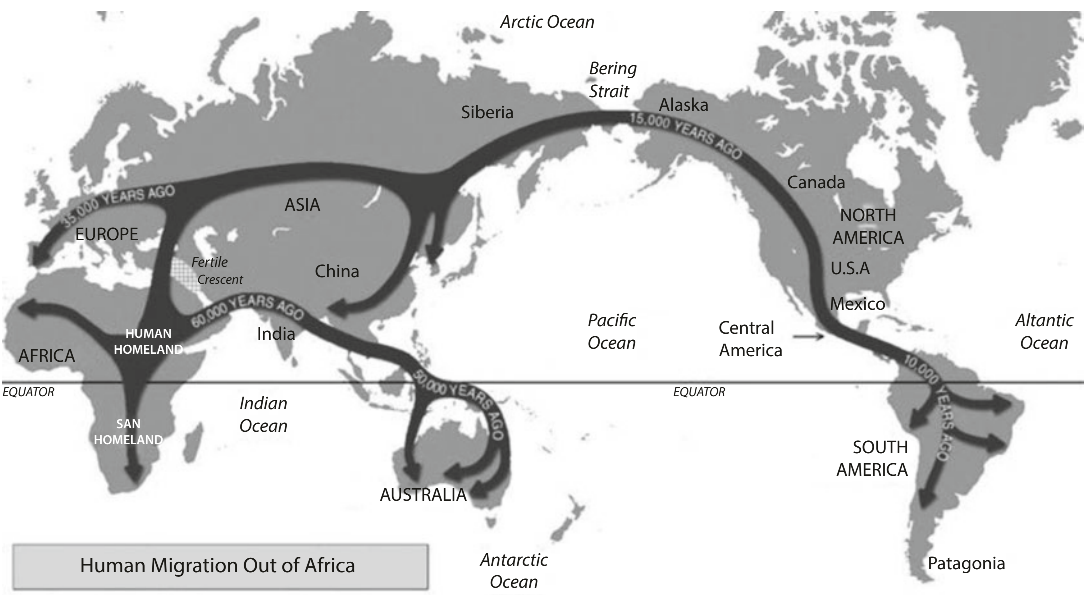
the colonization of cooler lands. Our ancestors were using stone tools for roughly two and a half million years, which is equivalent to about 100,000 human generations.

These stone spear points and knife blades were more or less the first in a long series of technological advances that helped increase the flow of energy to humans, thus greatly expanding the ability of humans to exploit the energy available in the various plant and animal resources in their environment. It also greatly increased the climates in which they could live because of their ability to kill large animals and use their skins for clothes (◘ Fig. 6.7). Another important new energy technology was fire,

◘ **Fig. 6.6** Spear heads

which allowed people to stay warm in cooler climates but more importantly increased the variability and utility of plant foods, as cooking broke down the tough cell walls that plant (but not animal) cells have. Many humans left the relatively benign climate of Africa probably a little less than 2 million years ago. The remains of both humans and their tools of that era have been found in present-day Middle East, Georgia, and Indonesia [5]. By a million years ago, human remains were common all through Asia, but curiously humans did not appear to colonize Europe until roughly 500–800 thousand years ago. The first humanoid colonists of Europe do not appear to be our direct ancestors, for morphologically modern humans (popularly known as "Cro-Magnons" as distinct from the earlier "Neanderthal" stocks) appear to have left Africa in a separate migration only about 100,000 years ago. There are very strong debates in the anthropological literature as to whether all of these groups of people are our ancestors or just the "Cro-Magnon" variety of a large suite of early humans. Modern DNA analysis seems to favor the separate stock concept with some mixing that ended 35,000–40,000 years ago, leaving, it seems as of 2015, a few of their genes mixed with those of Cro-Magnon stock.

One of the many changes that took place as humans moved out of Africa was that humans tended to lose their melanin, a protective pigment



◘ **Fig. 6.7** Human migration patterns. All humans originated in Africa but then took various routes to establish new groups of people that helped people living in Africa avoid various skin diseases such as skin cancer. When humans were exposed to much less sun for long winter periods, while in the meantime covering their skins with animal hides, they did not get the benefit of the sun producing vitamin D within human skin. This made humans much more susceptible to rickets, a debilitating vitamin deficiency disease that results in easily broken bones, obviously a great problem for hunter-gatherers. Since the dark pigment melanin protects the skin, but also decreases its ability to make vitamin D, darker skin is less advantageous in areas with less year-round intense sun. Hence skin color, something of often egregiously misplaced cultural importance, is simply a reasonable evolutionary response to humans leaving or not leaving the tropics.

6.2.2 The Dawn of Agriculture: Increasing the Displacement of Natural Flows of Energy

Sometime about 10,000 years ago, in the vicinity of the Tigris and Euphrates river valleys of present-day Iraq, a momentous thing happened [6, 7]. Humans previously had been completely constrained by their limited ability to exploit entirely natural food chains, due to the low abundance of edible plants there. They found that they could increase the flow of food energy to themselves and their families enormously by investing some seeds into more food for the future. How this happened is lost to antiquity, but as described by Jared Diamond in "Guns, Germs and Steel," [8] it probably happened as people observed that their own kitchen middens (garbage areas) produced new crop plants from the seeds that had been deliberately or inadvertently discarded. This caused hunters and gatherers to experiment agriculturally, and as the climate warmed, more of the experiments were successful.

The implications for humans were staggering. The first, seemingly counterintuitive, is that human nutrition, on the average, declined. One of the best studies to document this was by Larry Angel, who studied the bones of people buried over the past 10 thousand or so years in Anatolia, roughly the border region of modern-day Turkey and Greece [9]. Angel dated the bones he found in ancient burial grounds and could learn many things about the people who once lived there from the bones themselves. For example, their height and general physical condition, as well as functions of the quality of nutrition, could be determined by the length and strength of the bones. Bones could also show the number of children a woman had by the scars on the pubis, whether that person had malaria by the appearance of the bone marrow-producing regions of the bone, and so on. The data indicate that the people became shorter and smaller with the advent of agriculture, indicating a *decrease* in nutritional quality. In fact, the people of that region did not regain the stature of their hunter-gatherer ancestors until about the 1950s. Although agriculture may have given the first agronomists an advantage in terms of their own energy budgets, that surplus energy was translated relatively quickly into more people with only an adequate level of nutrition as human populations expanded. Or perhaps, as outlined below, more of the farmers' net yield was diverted to artisans, priests, political leaders, and war, leaving less for the farmers themselves. One of the clear consequences of agriculture was that people could settle in one place, so that the previous normal pattern of human nomadism was no longer the norm. As humans occupied the same place for longer periods of time, it began to make sense to invest their own energy into relatively permanent dwellings, often made of stone and wood, in which to store the surplus. This left more durable artifacts for today's archeologists.

A second major consequence of agriculture was an enormous increase in social stratification as economic specialization became more and more important. For example, if one individual was particularly skilled at making agricultural implements or understood the logic and mathematics (i.e., best planting dates) of successful farming, it made sense for the farmers of the village to trade some of their grain for his implements or knowledge, initiating, or at least formalizing, the existence of markets. From an energy perspective, relatively low-quality (because so many people had the necessary skills) agricultural labor was being traded for the high-quality labor of the specialist. The work of the specialist can be considered of higher quality in terms of its ability to generate greater agricultural yield per hour of labor. Considerable energy had to be invested in training that individual through schooling and apprenticeship. The apprentice had to be fed while he or she was relatively unproductive, anticipating greater returns in the future. Thus, we can say that the energy return on investment (EROI) of the artisan was higher

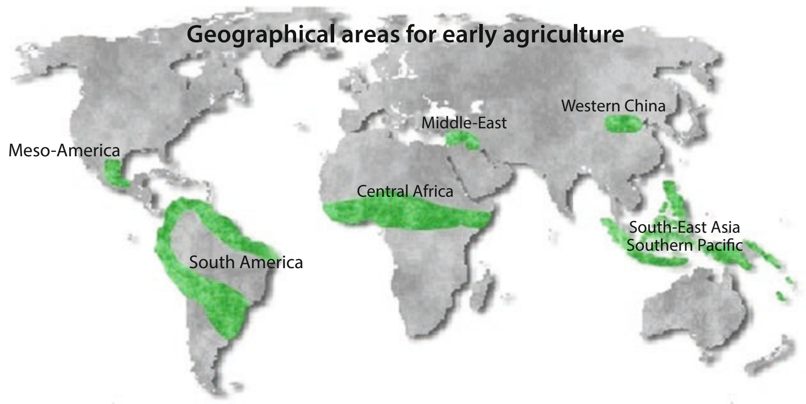


Fig. 6.8 Origins of early agriculture (Source: Wroe et al. [32])

than that of the farmer, even if less direct, and often his pay and status as well. High agricultural yield led to surpluses that could be stored, leading to the concentration of political power by those who built and controlled the granaries.

Eventually, the concept of agriculture spread around Eurasia and Africa (Fig. 6.8). A new phenomenon appeared with the development of agriculture, the large net surpluses from the farmers and the permanent settlement of certain regions: cities and other manifestations of urbanization. The first place this occurred appears to be in the Tigris and Euphrates valleys, and one of the first cities was known as Ur, from which we derive the word urban. Today we call that ancient civilization Sumer and the people Sumerians. There were many great cities of that time (roughly 4700 years ago) and region, including Girsu, Lagash, Larsa, Mari, Terqa, Ur, and Uruk. These cities grew up in what had been at first a heavily forested region, as can be understood from the massive timbers in remaining ruins, although today there are essentially no trees and no cities in that region. In fact, the forests were gone by 2400 B.C., the harbors and irrigation systems silted in or required increasing amounts of work to maintain, the soil became depleted and salinized, barley yield dropped from about 2.5 tons per hectare to less than one, and by 2000 B.C., the Sumerian civilization was no longer extant. The world's first great urban civilization, in fact its first great civilization, used up and destroyed its resource base and just disappeared over a span of 1300 years.

The interaction of people with cultivars (plants that humans cultivate) also changed greatly the plants themselves. All plants are in constant danger of being consumed by herbivores, from bacteria to insects to large grazing or browsing mammals or, formerly, herbivorous dinosaurs. The evolutionary response of plants to this grazing pressure was to derive various defenses, including physical protection (such as spines, especially abundant in desert plants) and more commonly chemical protection in the form of alkaloids, turpenes, tannins, and so on. These compounds, usually derived at an energy cost to the plant, place a heavy burden on herbivores or potential herbivores by discouraging consumption or by extracting a high energy cost on those specialized herbivores that can eat them, for the energy cost of detoxifying poisonous compounds is very high [13, 14]. Humans do not like these frequently bitter, poisonous compounds either and for thousands of years have been saving and planting the seeds from plants that taste better or have other characteristics that humans like. Partial exceptions are, e.g., mustards, coffee, tea, cannabis, and other plants whose bitter alkaloids are poisonous if that was all we ate but an interesting dietary supplement in small doses. Consequently, our cultivars are, in general, quite poorly defended against insects and have led to the invention and use of external pesticides, with complex consequences. Many of our cultivars would not survive in the wild now and have coevolved with humans into systems of mutual dependency. A visitor from outer space might conclude that the humans have been captured by the corn plants who use us for their slaves to make their lives as comfortable and productive as possible! Meanwhile all kinds of pests were themselves adapting to the concentration of humans and their growing and stored food, often with disastrous impacts on humanity [15]. For

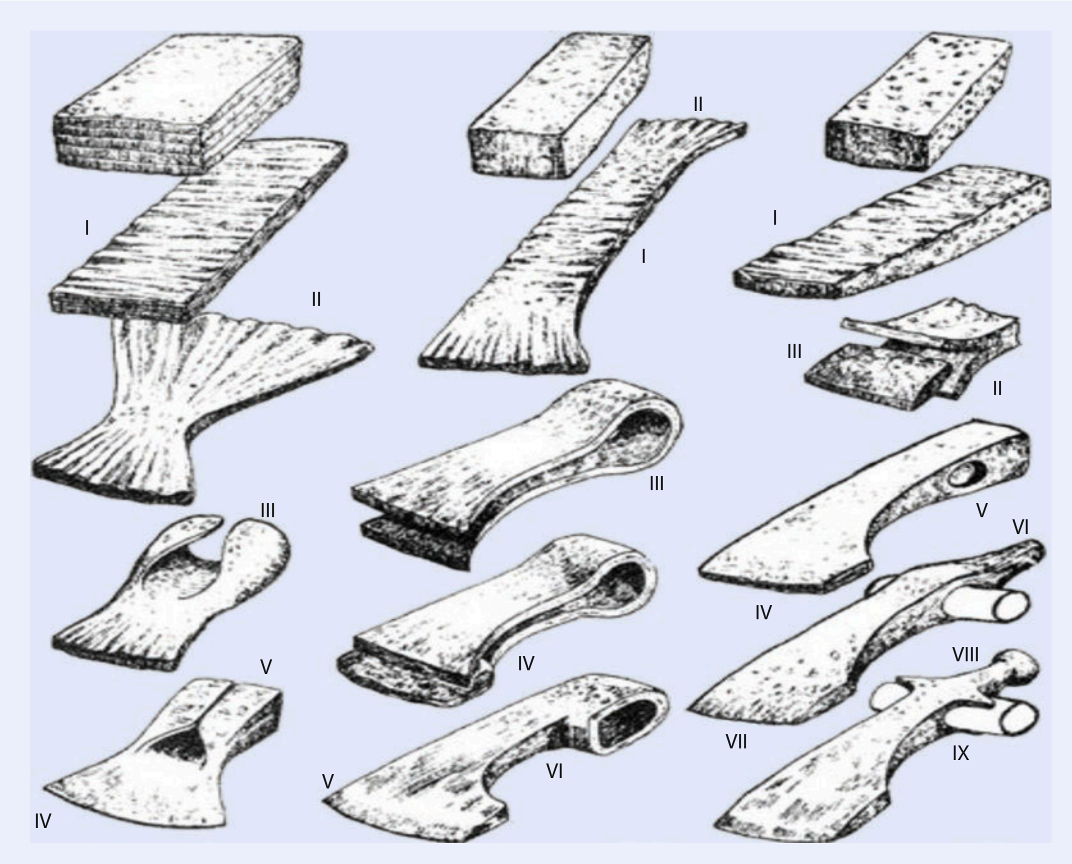


Fig. 6.9 Early metallurgy (Source: National Geographic)

example, crowding is a major factor in the transmission of diseases with epidemic potential such as acute respiratory infections, meningitis, typhus, cholera, and scabies.

At roughly the same time that agriculture was spreading around the world, humans made another extremely important discovery: metallurgy. Prior to the advent of metallurgy, essentially all tools used by humans were derived directly from nature: stone, going back perhaps 50,000 years (Fig. 6.6) fashioned with increasing sophistication, wood, bones, antlers, and so on. According to Ponting [16], the first evidence of the smelting of copper is found in Anatolia in about 6000 B.C.E., although the near contemporaneous existence of residuals of smelting from all continents at only slightly later in time implies that probably many groups of people had roughly the same idea by about 5000 B.C.E. (Fig. 6.9). Eventually very specialized furnaces were developed, as is indicated by archeological digs from 5 to even 10 thousand years ago in Africa, Europe, South America, and Asia. Early copper and bronze tools were replaced over time with iron as people learned to make hotter fires using charcoal. We have been using metal tools for roughly 8000 years, or about 400 generations. So, most of our history as a species is without metal tools. An important component of the transition is that the stone tools could be made with only a very small energy investment, essentially all as human muscle power, whereas the metal tools required a much larger investment in terms of cutting trees, making charcoal, and of course the energy of the wood itself. Early smelting was probably technically inefficient but had the advantage, at least initially, of the availability of very high grades of ore.

Smelted metals had many advantages compared to materials derived directly from nature: metals were harder and could take a sharper edge, increasing the cutting work that could be done by human muscles, and the sharper knife blades and spear points concentrated energy onto a smaller surface and enhanced the process of humans exploiting nature, for example, by accelerating the rate that people could cut trees (and of course each other) with bronze vs. stone axes. Perlin [10] has chronicled the tremendous increase in human cutting of forests in a wonderful book "A Forest Journey." He makes the point in this book that massive deforestation is an old phenomenon and that India, China, and most of the Mediterranean were pretty thoroughly deforested by the time of Christ. In most cases, the most severe deforestation was to get fuel for metallurgy.

The scenario often went something like this (with Crete as a good example). A group of people would find and develop a rich ore deposit of, for example, copper. This metal would be very valuable in trade and the people would become prosperous. Cutting of trees for smelting also cleared land for agriculture, and the wealth and well-being of the people increased not only from the trade in metals but also from the substantial increase in the area under agriculture in the rich forest soils where the trees had been cut. Things would tend to go very well for roughly a century. But once rich forest soils were exposed to agriculture and rain, they would tend to erode, and the agricultural yields would decline. That civilization would decline as ore deposits and soils wore out, until they collapsed: meaning that the number of people being supported decreased dramatically. According to Perlin (and many others [10, 12, 16, 17]) this process has occurred again and again and again throughout history. India and Greece have had three separate major deforestations, with the forests growing back each time human populations became lower. The great works of literature, for example, Thucydides *The Peloponneisan Wars*, were written about events enormously impacted by large resource and environmental events (i.e., the exhaustion of sufficient forests for Athenians to smelt silver or make ships) although such resource issues were rarely considered by historians until recently [10].

Table 6.2 Evolution of power outputs of machines available to humans

Machine Horsepower	
Man pushing a lever	0.05
Ox pulling a load	0.5
Water wheels	0.5–5
Versailles water works (1600)	75
Newcomen steam engine	5.5
Watt's steam engine	40
Marine steam engine (1850)	1000
Marine steam engine (1900)	8000
Steam turbine (1940s)	300,000
Coal or nuclear power plant (1970s)	1,500,000

*Cook, E. 1976. Man, Energy, Society, W. H. Freeman

Other important energy-related events were occurring in these prehistorical times. Perhaps most important was the domestication of useful animals, some of which predated agriculture, while some occurred simultaneously. The domestication of animals and the increased sophistication of animal husbandry were important in increasing energy resources for humans in at least two ways. First, since these animals ate plant material that humans did not, this greatly increased the amount of energy that humans could harvest from nature, especially in grasslands. Second, oxen and especially horses as draft animals greatly increased the power output of a human (Table 6.2). This power was useful for transport, for agricultural preparation (which came later), and for war. A horse, however, did not necessarily increase the speed of communications because over a day, a fit human can outrun a horse!

The story of how the use of animal technology was passed throughout Eurasia has been developed elegantly by Diamond [8]. Most of the important domestic animals came from Eurasia and could thrive more easily at the same latitude. Our most important animals, the sheep, cow, horse, pig, and chicken, were "corralled" in Eurasia and developed into today's domestic animals. The increasing familiarity with beasts of burden and the development of roads and caravan technology in turn allowed for the development of long-distance trade [11]. Meanwhile, sailing and navigational skills were developed and passed on, and Cottrell writes well about the importance of using wind power in ships to greatly enhance the amount of work (carrying goods) that one person could do. Trade between cultures enriched the knowledge and the biotic resources of many human groups.

As agriculture, settlement, and commerce expanded, there became a greater need for maintaining records, and some time about 3000 B.C., formal writing was developed, apparently simultaneously in Egypt, Mesopotamia, and India (and perhaps other places). Writing allowed for technologies to be maintained from one generation to another and transferred among cultures. Cumulatively, these new technologies increased the energy flow to the human population, which slowly but relentlessly increased. These old records have allowed us to estimate some earlier patterns of human population changes (Fig. 6.10). They suggest that the human population record is hardly one of the continuous regular growths but rather one of the periodic growth and decline. Sometimes this is manifested as catastrophic decline and the virtual or absolute cessation of that population or, more commonly, the political structure that once held them together. Edward Deevy [18] has suggested that there were three main increases in human populations associated with first the corralling of animals, then the development of agriculture, and then the industrial revolution. We are still experiencing the latter as global human population growth continues strongly, although at a somewhat lower rate than a few decades ago.

6.2.3 Human Cultural Evolution as Energy Evolution

As we keep pointing out, most of the major changes that occurred in the ability of humans to exploit more and more of the resources around them were either directly about or clearly associated with increased use of energy. Spear points and knives are energy-concentrating devices, fire allows greater availability of plant energy to humans, agriculture greatly increases the productivity of land for human food, and so on. These evolutions of the ability of humans to control more and more energy, for example, the evolution of wind and water power, are probably best told in Fred Cottrell's wonderful book *Energy and Society*, published more than half a century ago [19]. Cottrell was a railroad man for most of his life and then a college professor near the end. Always impressed, like us, with the energy that undergirds all that humans do, Cottrell's focus was on the development of what he called "converters," which are specific technologies for exploiting new energy resources.

Cottrell's early chapters focus on herding and agriculture as a means of exploiting biotic energy and then on water and wind power. He shows the historical importance of a city being located relatively downstream on a river, so that the natural flow of the water would allow that city to exploit easily all upstream resources, such as timber, agricultural products, game, and ores. Of course, there was always a problem with this: barges had a one-way trip so it had to build a new barge at the top of the watershed for each new trip. Also, crews had to walk or otherwise get themselves back upstream. Nevertheless, a barge could carry a much larger load compared to a single individual, who can carry only about 25 kg at a maximum for any serious distance, or a pack animal such as a horse that can carry about 100 kg. Thus, the use of a barge carrying, say, 10 tons of goods and with a crew of four increased the efficacy of each person by a factor of 25–100. This process continued well into the nineteenth century on the Mississippi River until the age of steamboats. Raftsmen simply broke up the rafts and barges for lumber at the end of the journey (Taylor [33]).

The development of a sailing ship likewise increased the energy that subsidized a human porter enormously. According to Cottrell's calculations, an early sailing ship such as used by the Phoenicians (more or less the equivalent of modern-day Lebanon) increased the load that a human could carry by some factors of 10 and by late Roman times as much as a factor of 100. The Romans needed to import large quantities of grain from Egypt because, in part, they had depleted their own soil. But, according to Cottrell, the Romans were not the only ones who had an eye on this grain, and initially the Romans lost a lot of grain to pirates. This required the Romans to transport the grain in heavily guarded narrow warships, and a significant part of the grain was consumed by the soldiers on board. Thus, one further energy investment had to be made by the Romans—clearing the Mediterranean

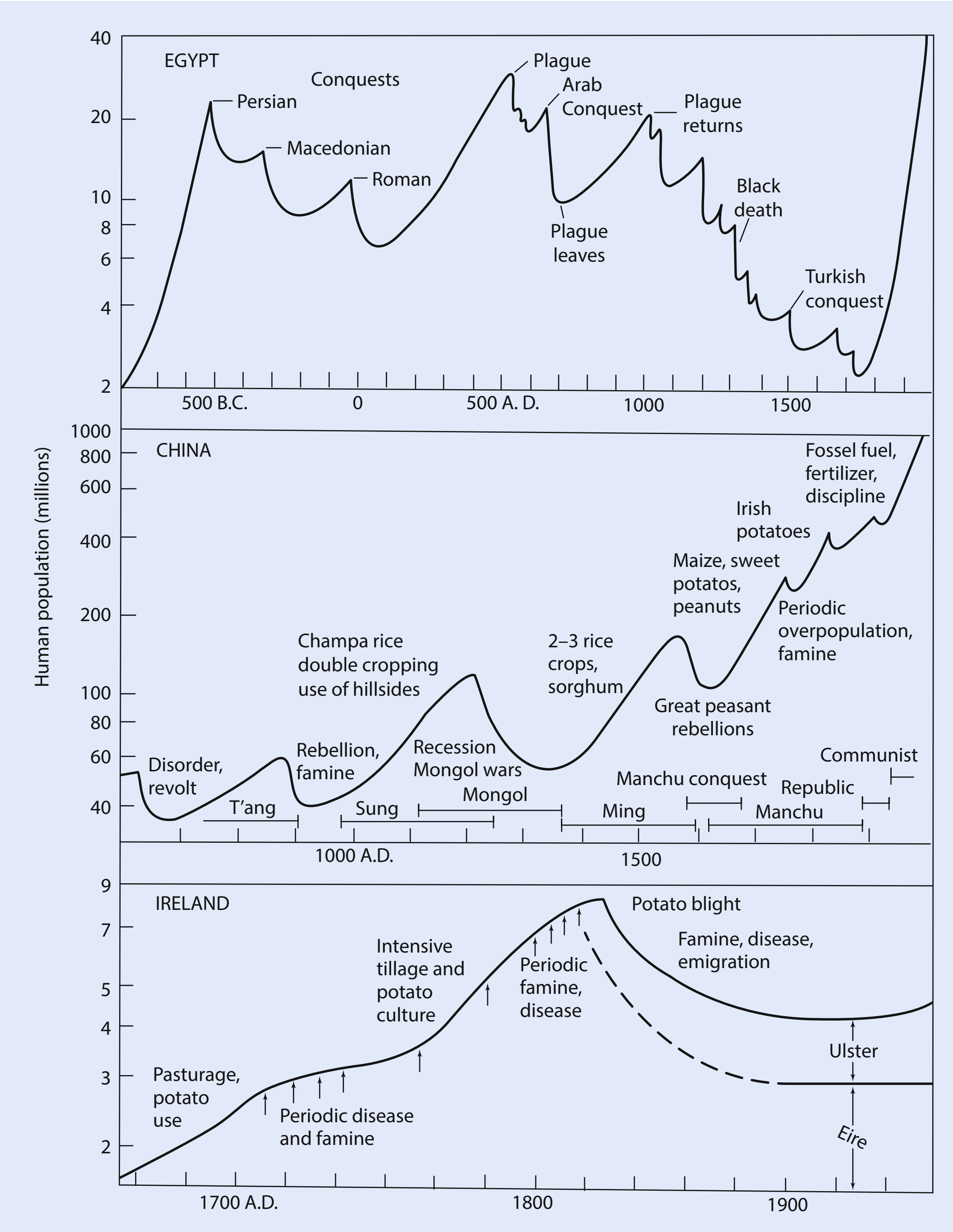


Fig. 6.10 Human population changes in Egypt, China, and Ireland, regions that had relatively well-developed bureaucracies and hence good data (Source: [34])

of pirates. Once this was done, proper wide-beamed sailing merchant vessels could be used, and Egypt finally became a large net energy source for the Romans. Cottrell gives many other examples of the increasing use of energy by humans over time, including very interesting chapters on the growth of railroads in England, steam power, and industrialized agriculture.

Fig. 6.11 Ruins of ancient city of Ur

6.2.4 The Possibility, Development, and Destruction of Empire

Agriculture and its greatly increased yield brought with it the possibility of the concentration and storage of food, specialization, and, through greater populations, military-political power. These concepts are again ably reviewed in Diamond, Tainter, Ponting, and others. From our energy perspective, agriculture allowed for huge energy surpluses as a result of high return ratios (EROIs) from large energy investments. Thus, agriculture allowed a massive increase in the ability of people to generate culture and cultural artifacts. We have bare glimpses of these in the remaining artifacts of ancient cultures such as the main building at Ur (Fig. 6.11), temple complexes, and the great wall of China. What we see of these ancient civilizations today are beautifully shaped and carefully put together stones, and, as we dig more carefully, more sophisticated ornamentation, pottery, and metal household implements. By digging a little deeper, we can find other impressive artifacts of past civilizations: irrigation systems to bring water over large distances and large pyramids of stacked stones. These artifacts imply huge energy surpluses relative to hunter-gatherers, probably much of its vast public works programs to keep farmers occupied during non-planting or harvesting seasons.

In hunter-gatherer cultures, there was normally relatively little differentiation in what different people did, except for divisions by sex and age. Agricultural surpluses allowed a greater division of labor and with it a greater difference in wages, status, and social power. This division of labor led in time to extreme differences in political power. This power was enhanced as professional military men became increasingly common, exemplified in the ancient Assyrian cultures. Most people had very little status or wealth and tilled the soil or took care of domestic matters. Only a very small proportion, large land owners, merchants, technocrats, and political leaders, lived lives of increasing affluence and luxury. Over time, the difference between rich and poor increased drastically.

As the concentration of wealth and power increased, as central granaries became more important and as military power and war became increasingly institutionalized, there were increasing opportunities for the development of empire. An *empire* is defined as large geographic areas under the rule of a central place and chief and maintained through what we might call civil servants or bureaucrats (although "lieutenants" are probably more accurate). Tainter and others [12] have developed the concept of a pattern that they believe has occurred again and again through history. One city or local culture becomes very successful through effective agriculture, mining, or trade and the resultant growth in population and economy. Often it becomes increasingly wealthy, allowing it the surplus energy to support soldiers and expropriate larger and larger areas of land around its periphery while exploiting the subjugated people's energy surplus. Since war is expensive, it becomes increasingly important for the central city to impress others with their wealth, a sign of surplus energy available to be used, potentially, against others. Therefore, huge public investments are made in public structures, temples, administrative centers, markets, roads, food storage facilities, and so on. If they are successful, outsiders decide it makes sense to become aligned with this most powerful culture, even at the expense of tribute in the form of agricultural products, precious metals, or other materials. Thus, the culture expands, often many times over.

At some point, the culture, through its growth, begins to exhaust the initial resources that made it rich. Another problem is that as cultures increased in linear dimensions, the energy cost of moving resources (e.g., taxation grains) to the central city became greater and greater. If the provinces sensed difficulties in the central city, they might become a bit more restless, requiring increasing investments in military forces or status symbols in the central city. According to Tainter, eventually the citizens of both the central city and the provinces become tired of paying the high taxes for what is mostly "maintenance metabolism," that is, the food, roads, and armies needed to maintain the central city. Due to diminished revenues, the physical and social infrastructure is not maintained, leading to the collapse of the empire. Tainter, an archeologist, ecologist, and historian, says that this has occurred repeatedly (he gives more than 20 examples in his first chapter) through prehistory and history. Ponting develops a similar scenario in many detailed examples and with a bit more emphasis on resource depletion, as does Charles Redman.

6.3 Mediterranean Cultures

There are some quite detailed assessments of the rise and collapse of earlier civilizations from the perspective of the energy and other resources required for development and maintenance. Mediterranean cultures are a good place to start thinking about these questions for a number of reasons. First, many of the most important ideas for the contemporary world, including democracy as a form of government, mathematics as we know it, and concepts in art and culture, originated in this region. Second, the Mediterranean world offers a well-documented, well-studied suite of examples for us to explore and to understand the importance of energy and other resources in helping to shape the events that many of us recognize from traditional historical accounts. Third, this region remains today a vibrant and sometimes contentious region with many issues going way back in time. Many of the readers of this book will have been educated on the history of the region, which allows us an opportunity to examine familiar territory through our different lens of energy-based analysis.

6.3.1 Greece

Contemporary Western democracies usually trace their ancestry back to ancient Athens and neighboring cities in what is now Greece. Twenty-five hundred years ago, these were vibrant, dynamic, frequently wealthy cities with some truly remarkable accomplishments, including defeating significantly larger Persian forces and producing some of humanities' still greatest architecture, sculpture, literature, and ideas about government. Athens and its sister city-states were also venal, domineering, and frequently squabbling cultures, squandering remarkable opportunities for "the good life" on pointless wars. The most important city-states were Athens and Sparta. Today we remember Athens as an incredible caldron of art, ideas, and famous men and Sparta as a culture completely dominated by preparing its young men for war (hence, "Spartan conditions" is a term used today for harsh, uncomfortable, and arduous conditions). Athens too was a militaristic and imperialist culture and excelled in maritime combat. Athens and Sparta lived for many years in an uneasy truce which eventually ended in distrust and shifting alliances. From 431 to 404 B.C.E., these states and their allies initiated more than 25 years of intense combat that has been elegantly told by Thucydides [20]. Thucydides was once one of Athens' generals, but the price of losing even one battle in Athens, which had happened to Thucydides, was dismissal from the army. This gave him the time to write a comprehensive history (*The Peloponnesian Wars*, a classic of history) of what ensued during this war, which was a stalemate for decades.

One interesting energy-related analysis of the Peloponnesian Wars, from which the following is borrowed, is found in Perlin's book *A Forest Journey*. Perlin surveys the Peloponnesian Wars from the perspective of the forests and forest-derived energy required for the military activity and generation of the wealth required to finance the war. Anyone visiting Greece today is impressed by the nearly total absence of extensive and robust forests, so that it is quite curious to think of Greece and its southern part, the Peloponnesian peninsula, as heavily forested. Plato, as late as the sixth century B.C., remarked that not long before his own time, the hills surrounding Athens provided the huge building timbers he could still see in the buildings of Athens and that these hills even contained forest-dwelling wolves that were a threat to livestock. Perlin believes that these abundant forests probably saved Greece from Persian domination as they provided timber to construct the Athenian fleet that defeated the Persian monarch Xerxes at Salamis. This was followed by the construction of an even larger 200 ship navy so that ambitious Athens could become the mightiest marine force in Greece. The Athenians were running into timber shortages, however, because of intense demand for fuel and construction wood in the city (including for immense wooden cranes to build the Parthenon) and because of an immense vein of galena ore that had been discovered in the nearby town of Laurion. The ore could be smelted using charcoal as an energy source to produce silver, which was then spent on the new fleet, public works such as the Parthenon, and personal luxuries. While this immense ore deposit made Athens extremely wealthy and powerful, it was at the expense of many of the forests of the region. This became a large problem because the Persians still controlled the timber supply regions to the east and north, including especially the Strymon valley. Ten thousand Athenians sent to colonize the mouth of that river, to insure a timber supply for Athens, were slaughtered by the locals. A second invasion was somewhat more successful resulting in the capture, at least for several years, of the port city of Amphipolis. When that city was lost later, (this was the battle Thucydides lost) Athens struggled for timber throughout the ensuing decades and centuries.

The Peloponnesian War that followed was principally between Athens and Sparta but also included other Greek city-states. It was ruinous to all the participants. Due to wood being used for all instruments and means of war, it depleted the remaining forests of southern Greece and then consequently soil eroded. The war spread even to Sicily, which the Athenians attacked unsuccessfully in a vain attempt to seize the forests to build a giant armada. In the meantime, Sparta had seized the forest reserves on the peninsula that belonged to Athens and other states. Sparta then turned to Macedonia, made a new alliance, and built a new fleet. Meanwhile, plague had entered Athens, greatly decreasing their number of soldiers. The Spartans made an alliance with their former Persian enemies and constructed a new fleet from Persian forests. They caught the Athenian fleet on shore with their crews foraging for dinner, and Athens was finally and permanently defeated, leaving the city destitute and without fuel or too much in the way of food. Thus, although we learn of the war in terms of battles, generals, and so on, much of the background was about energy (to smelt silver to pay the armies and for obtaining timber and metal for weapons and armor) and other resources (e.g., wood for ships), the depletion of which contributed to the eventual outcome. The golden age of Athens was over, as was the city's contribution to our present culture.

Fig. 6.12 Maximum extent of Roman Empire (Source: Tylecote [35])

6.3.2 Rome

Rome, founded in about 750 B.C.E (according to myth by the abandoned twins Romulus and Remus, who were supposedly nurtured by a female wolf), was initially a group of neighboring hill towns that increasingly became incorporated into a city. Rome kept expanding through trade and military conquest until it comprised much of the world known to Romans. The Romans learned early on that wealth could be gained much more easily through conquest and subsequent taxation than through other means and thus kept expanding its boundaries. Subjugation and taxation were of course not especially popular among those subjugated, but the *Pax Romana* (Roman peace) imposed by the strong Roman military force actually decreased local conflict for many. The city was ruled by a series of kings until about 400 B.C.E, when it was changed to a Republic ruled principally by a senate of patricians.

The Roman Empire lasted 500 years from roughly 44 B.C.E, when Julius Caesar appointed himself emperor, to 476 C.E., although the eastern portion at Constantinople lasted for 1000 years more. The Empire reached its maximum extent about 117 C.E., when it encompassed essentially all areas around the Mediterranean, including all or most of the present countries of Italy, France, Spain, England, Greece, and Egypt, as well as the North African coast, Syria, the Middle East, and the regions around the Black Sea (Fig. 6.12). Rome had at its height about 1,000,000 people (of which only about 10% were citizens), and the entire Empire contained as many as 70,000,000 people. This Empire was carved out, maintained, and governed essentially by human energy—by citizen soldiers on foot who traveled on campaigns each year, utilizing wonderfully engineered stone roads that spread throughout the Empire (hence "all roads lead to Rome"), although ships were used over the Mediterranean itself. Imperial Rome was probably the most populous city in the world until the eighteenth century. The task of feeding roughly 1,000,000 people was a great undertaking, especially following the passage of a law that guaranteed free grain to Roman citizens. The Roman invasion and subjugation of Egypt were not simply about Caesar's lust for Cleopatra but also about shoring up Roman food supplies after the soils of Italy had been depleted by Roman farmers. Fortunately for the Egyptians, and for the Romans, the annual flooding of the Nile replenished the Egyptian soils. This occurred every year until the closing of the Aswan Dam in the 1960s. The concentration of artisans in Rome, and the *Pax Romana* that existed within the Empire, brought unprecedented economic prosperity to many, many people, while Roman engineering and architecture (borrowed heavily from the Greeks and others) generated massive and often wonderful public works throughout the Empire. Swamps were drained, creating new agricultural land and ending malaria. While Rome is mostly thought about as a militaristic imperial force, and it certainly was, day-to-day life and influence were probably more a function of extensive and very effective trade, engineering, and agriculture.

Although Roman emperors were often venal, cruel, and corrupt, the best of them espoused very noble ideas about civilization and citizenship. There were a succession of good and bad emperors and other leaders, often representing different classes of people. For example, Julius Caesar although an aristocrat by birth, represented especially the interests of the common citizen class, although those who killed him also claimed to represent more the interests of the general Roman citizen. Either way, like the period when Athens was at its height, this was a remarkable period for civilization. Some of the leaders, including Marcus Aurelius, appear in history's lens as quite enlightened. Edward Gibbon, the eighteenth-century historian who wrote *Decline and Fall of the Roman Empire*, described the period best or at least most eloquently [21]. Gibbon believed that Rome in the second century might have been the greatest time of all for humanity.

» In the second century of the Christian Era, the empire of Rome comprehended the fairest part of the earth, and the most civilized portion of mankind. The frontiers of that extensive monarchy were guarded by ancient renown and disciplined valor. The gentle but powerful influence of laws and manners had gradually cemented the union of the provinces. Their peaceful inhabitants enjoyed and abused the advantages of wealth and luxury. The image of a free constitution was preserved with decent reverence: The Roman senate appeared to possess the sovereign authority, and devolved on the emperors all the executive powers of government. During a happy period of more than fourscore years, the public administration was conducted by the virtue and abilities of Nerva, Trajan, Hadrian, and the two Antonines.

» If a man were called to fix the period in the history of the world, during which the condition of the human race was most happy and prosperous, he would, without hesitation, name that which elapsed from the death of Domitian to the accession of Commodus. The vast extent of the Roman empire was governed by absolute power, under the guidance of virtue and wisdom. The armies were restrained by the firm but gentle hand of four successive emperors, whose characters and authority commanded involuntary respect. The forms of the civil administration were carefully preserved by Nerva, Trajan, Hadrian, and the Antonines, who delighted in the image of liberty, and were pleased with considering themselves as the accountable ministers of the laws. Such princes deserved the honor of restoring the republic, had the Romans of their days been capable of enjoying a rational freedom.

Nevertheless, there were always economic troubles, generally related to natural resources, including grain and wood, and the failure to maintain the solar-based systems that generated them. The general consumption of the Romans always exceeded the revenues. Common and necessary raw materials, such as wood, became more and more difficult to obtain as forests increasingly far from Rome were cut and turned to agricultural land, whose productivity tended to decrease over time. To meet its expenses, the government increasingly debased the gold and silver currency, causing extreme inflation, a fascinating story told in detail by Walker [22]. The Roman denarius was adulterated from being 98% silver in 63 C.E. to zero percent (i.e., all copper or other such metals) by 270 C.E. as the main silver mines at, for example, Rio Tinto, were depleted. As the denarius was adulterated, its purchasing power decreased proportionally.

Lead may have had an impact too, as the bones of ancient Romans have very high levels of lead, probably reflecting its use in pipes and in wine making. Nevertheless, it is quite remarkable what humans can do based on essentially solar energy plus their own (or slave) muscle power alone. Perhaps it is better to conclude that the energy that built and maintained Rome was hardly the muscle power of Romans and the agriculture of Italy but rather that of the millions of subjugated people in the provinces (and their land) who grew the necessary grain and cut the necessary wood to maintain the level of concentrated wealth in Rome. Perlin calculates that to run the baths at Caracalla for 1 year, 114 million tons of wood was required, a truly prodigious quantity that had to be transported from tens to hundreds of miles by human or horse power.

Over time, the Romans "became soft," hiring or forcing others to do their military service and grow their food. Vast expenditures went into public buildings and sports (if that word can be used) complexes, the most important of which is the Coliseum, where thousands of exotic animals were brought in and put into combat with slaves. They even staged naval battles in the Coliseum by flooding the interior with water. Clearly Hollywood has had its precedents. But by 200 C.E., the Empire began to be nibbled away by soil erosion, plagues, crop failures, and the Germans and Asians who desired the wealth that was within. Ultimately the city itself was successfully stormed by the Goths, Visigoths, and Vandals, with the full fall generally agreed to be in 476 C.E. Of course, Rome, the city, is still there, with many artifacts from earlier times, although it is hardly the center of an empire.

The most interesting and, from our perspective, insightful analysis of the decline and fall of Rome (other than Gibbon's monumental books) is that of Joseph Tainter [12], who examined the entire process from the perspectives of the energy cost and gain of each activity. The main way that the ancients gained wealth was through conquest. Whatever wealth had accumulated in a region was the result of the slow accumulation of solar energy. This included mineral wealth, for the metals had to be mined by solar-powered human activity and then smelted using wood for fuel. Obviously, this was hard work, and many preferred the much easier (although possibly fatal) path of conquest. As the Roman Empire became larger and more powerful, it also became more complex to maintain and defend the provinces and eventually Rome itself. According to Tainter, increasing complexity is usually how problems are solved. But there is a high energy cost to complexity that makes its use eventually counterproductive. Tainter develops in a very compelling narrative how complexity, for example, through the maintenance of distant governmental administration and bureaucracies, garrisons, communications, and so on and the import of grain from ever more distant provinces, imposed an ever-increasing energy drain on the Empire and how this led eventually to its susceptibility to decay and invasions. Basically, the necessary investments into maintaining centralized administrative and military control become increasingly expensive and counterproductive, especially as the limits of an empire are pushed further and further from the centralized control, necessitating increasing energy costs for transport and to maintain the compliance of other people. Combining the language of Tainter and that of economists, we might consider this decreasing marginal return to complexity, which Tainter shows us occurred again and again and eventually led to the collapse of most empires.

We know less about the next 500 years in what had been the Roman Empire, partly because few historians have given us as comprehensive assessment of the subsequent events as we have for the years of the Roman Empire. These years are often called the "Dark Ages" or "the Middle Ages" and are left at that. It is important to remember that life went on, Romans or Italians or whatever we wish to call them continued to live in Italy (as French did in Gaul and so on), solar energy was used through agriculture and forestry to maintain people as they had been for millennia, and people lived, loved, fought, and died, while populations grew and sometimes declined from plague. Sometimes they left stone or occasionally literary artifacts but more usually leaving behind only more depleted soils and forests. What was left of knowledge and culture and civilization tended to be kept alive in monasteries and in civilizations further to the East.

6.3.3 The Rise of Islam [23]

The prophet Mohammed, originally a merchant but eventually a political and religious leader, united the Arabian Peninsula in the seventh century AD. His followers expanded the empire under his influence so that within a hundred years after his death they controlled a large area stretching from Central Asia through the Middle East and along North Africa to Spain. The empire expanded again in about 1200 to become what was probably the largest land empire ever. Although the empire, including the political administration, was quite ethnically diverse and far from centralized, the people were united in their devotion (or subjugation) to Islam and in their use of the Arabic language, in which the Muslim holy book, the *Koran*, was written. Known in the West for their fierceness, once subjugation occurred, the Muslim leaders tended to be relatively tolerant and left others within their administrative units (including Christians and Jews) to their own devices as long as they paid their taxes. At that time, most of the economies of the Muslim empire were either agricultural or grazing animal based. Likewise, conquest was generally through foot soldiers or cavalry, so that we can assume that both the economy and expansion was nearly completely based on a solar and biomass base for energy.

The Muslim world was increasingly focused in Cairo following the Arabic conquest of Egypt in the seventh century AD. Originally, Muslims eschewed naval warfare and even sea-based trade, focusing on land-based expansion by trade or voluntary conversion or sometimes conquest. Day to day, the main events were much more likely to be about trade than conquest. For example, Muslims had regular overland trade to China along the very lengthy "silk road." Eventually, they became seafarers, focusing initially on the Sea of Arabia and then the coasts of India and Africa. Their long presence in Africa is reflected in, for example, the name Swahili, which means coast in Arabic, which remains as the principle language in Kenya. Arab traders brought coffee, originating in Ethiopia, to the rest of the world, and this is reflected in the scientific name for the best coffee, *Coffea arabica*.

Increasingly the Byzantines, as the residuals of the Eastern Roman Empire were called by the early Middle Ages, attacked Egypt and other Arabic possessions using ships and caused great destruction. Again, the use of solar energy to make timbers for ships and wind energy to move large quantities of people and goods by ships gave enormous power to those who were able to exploit it. In response, the great Arab leader Caliph Abd al-Malik in the late seventh century initiated a great program of ship building. This program was based in Egypt, but Egypt had few trees and none of a size to allow the construction of strong ships. Large cedars, many 170 feet in length, were imported from Lebanon, although that was very expensive. Consequently, the shipbuilding had to be moved to what is now Tunisia, which at that time was heavily forested. A very strong fleet was constructed which captured Sicily (with its huge forests) and established a beachhead in Spain. In time, the Mediterranean became essentially an Arab Lake, as it had been previously a Roman one. The only ones to challenge this were the Venetians, who had access to the forests of the Po and Adige river basins. Thus, the exploitation of wind energy allowed the Muslims to conquer and hold on to huge new land holdings and to generate great wealth through trade. They were the masters of the Mediterranean world for nearly a thousand years (or more, considering that today most of North Africa is Muslim).

Among the many who accepted Islam as their religion were the Turkic peoples of Central Eurasia, who established a very strong empire beginning near present-day Constantinople and eventually spreading under the Ottoman group influence through much of the Islamic world. They also spread into the West and were finally stopped at Vienna in 1683. According to Rondo Cameron, although this was not a tightly integrated empire, it persisted and spread for a very long time because it did not subjugate those it conquered but only asked for taxes which were not excessive. This approach to Empire seems to be a relatively successful one compared to brutal repression. Arabic influence spread through European culture, leaving, for example, its imprint in the English language with words such as "arsenal" (construction house, originally), "algebra," and "algorithm," both reflecting the great advance made in mathematics within the Muslim world during what we now call "the Dark Ages" in Europe. More recently, an important GIS tool, IDRISI, was named after the great twelfth-century Arabic-Sicilian geographer.

The Muslim world, Ottomans, often found themselves in direct competition with the Christian world. Several specific events stand out. The Christian invasions of Muslim-controlled Jerusalem, known as the great Crusades (1095–1099, 1147–1149, 1188–1192, 1202–1204, 1217–1221, 1228–1229, 1248–1250), reflected the growing wealth, power, and some would say arrogance of Europe and represented not only a chance for the faithful and adventurous to attempt to wrest the "Holy Land" from the "infidels" but also opportunities for plunder, rape, trade, and extension of commercial influence. The first Crusade caught the inhabitants of Jerusalem by surprise, and an enormous blood bath of mostly Muslims (but also Christians) by Christians followed as the city was wrested from "infidels." None of the subsequent Crusades were as successful militarily. Some of the related events were especially pernicious. On the fourth Crusade, the European knights and their camp followers (tinkers, blacksmiths, prostitutes, and so on), tired of walking and riding horses, stopped in Venice to attempt to purchase passage by ship to the Holy Land. The Venetians, crafty businessmen and politicians, took their gold for passage, loaded the heavily armed men onto ships, and set off for what they said was the Holy Land. The Venetians had some old scores to settle with the inhabitants of Constantinople, then a Christian remnant of the old Holy Roman Empire. On the journey, a detour was taken, and the unsuspecting knights were deposited before the city of Constantinople which the Venetians claimed was Jerusalem. When they asked their Venetian ship captains why the city was adorned with crosses, they were told that this was a Muslim trick. So, they attacked the city, eventually subduing the inhabitants, and looted, raped, and pillaged for several months. The Venetians received not only payment for ship passage but insured that Constantinople would no longer be a threat to their commercial interests in the Aegean and Adriatic Seas, for example, for wood in the region, at least for a while. In the long term, the plan perhaps backfired as the weakened Christian City of Constantinople fell later to Islamic invaders from the East in 1453, and the importance of the Venetian Empire and Christianity in that region faded. Those who wish might say that indeed God works in mysterious ways.

Thus, the enmity of much of Islam today toward the West for the exploitation of the region's oil resources is hardly new and lives on today as a great distrust by many Muslim cultures for the motives of the West. As the West has become so dependent upon oil from the Muslim world, it is hardly surprising that many view the relation with great suspicion.

6.3.4 The Lasting Legacies of Ferdinand and Isabella

Another place that Muslims and Christians clashed was in Spain. Muslims came to Spain from the south across the Mediterranean and from the ninth to the thirteenth century controlled most of the Iberian Peninsula. While there they developed very sophisticated agricultural and horticultural systems and, essentially, tolerated diverse other cultures. Christian influence filtered in from the north beginning about the tenth century, culminating in the expulsion of both Moors and Jews by King Ferdinand and Queen Isabella, whose names are familiar to most Americans because they also supported Columbus, both in 1492. The result was disastrous for the Spanish economy because the Moors were much more sophisticated agriculturists than the Christians, at least for the southern part of Iberia, and because many skilled Jewish people were forced to leave. Many of these Moors and Jews probably went as Colonists into the Americas, feeling no longer welcome in Spain. The wealth of Spain, originally based on sophisticated agriculture and trade, was partially restored only by the brutal exploitation by Spain of the inhabitants of the New World as they extracted gold, silver, and other minerals with the aid of slaves, wood fuel, and wind power for their sailing ships. The food production system exported to the New World by the Spanish was one based on cattle raising, as this was the system favored by the Christian Spanish. The often sophisticated agricultural systems (e.g., extensive terracing) in place in Central and South America were displaced, even destroyed, by the very crude cattle-based *latifundia* system brought from Spain. In both southern Spain and Central America, the cattle were turned out to graze in the much more productive Native

American gardens that were often highly terraced, representing generations of careful investments of human energy. Since cattle return much less food per hectare per year than crops, the overall productivity of these systems for food energy was greatly lowered. Thus, in a sense, the actions of Ferdinand and Isabella destroyed two great agronomical systems and replaced them with unsophisticated grazing systems with perhaps one-tenth or one-twentieth the capacity to produce usable food energy for humans but with a greater capacity to produce money income for *haciendas*.

Entire forests, such as in southern Bolivia, were cut to supply timbers for mines and provide energy for smelting. Much of the Tarija region of southern Bolivia, for example, was deforested to support the silver mines in Potosi, and the timbers were transferred nearly a thousand kilometers horizontally and thousands of meters vertically on the backs of mules and slaves [24]. The deforestation resulted in some of the most extensive erosion found on the face of this Earth, which covers nearly 5 million hectares. Spain grew rich on the imported gold, but a curious phenomenon happened. The Spanish efforts in the New World doubled the quantity of gold in the old, but it decreased its value to less than half! What had happened was no different from what happens when a modern country prints too much money: inflation. Gold has little utilitarian value but is rather a medium of exchange. The real wealth of Europe came from the fields, forests, fisheries, and artisans, that is, the investments of solar and human (and occasionally wind and water) energy into the process of turning raw materials into real wealth: food, clothes, shelter, tools, utensils, and so on. Much of that gold ended up eventually in the great cathedrals of Europe.

6.3.5 Other Regions of the Earth

While Europe was living in the "Dark Ages," independent and often very sophisticated cultures were developing in China, India, and the Americas, each of which had much greater and often more sophisticated human populations than did Europe. Again, these were solar-powered agrarian cultures for the most part and depended year after year on intensive human labor and of course the sun as a source of energy. Several grass-based nomadic civilizations, including that in Mongolia led by Genghis Khan, also established very extensive empires that in his case reached nearly to Europe. In the Americas, very extensive city-states developed, flourished, and eventually collapsed. For example, the Olmecs and Maya of present-day Mexico and the Inca in Peru followed such fates long before the arrival of Europeans but more commonly after that. But, as we said at the beginning of the Mediterranean section, these cultures are not our focus here.

6.4 The Energetics of Preindustrial "Modern" Societies: Sweden and the Netherlands

There have been several especially comprehensive analyses of preindustrial solar-powered economies in the Netherlands and Sweden by De Zeeuw [25] and Ulf Sundberg [26]. These analyses indicate that it was possible to generate a very significant energy-based economic machine on plant material alone. The longer view, however, is that eventually these "renewable" systems tend to become depleted, and they require relatively low population densities (compared to the present) to be successful.

In the period 1640–1740, the Dutch had created a very profitable ceramics industry in the vicinity of the city of Delft, near Rotterdam. Even today, it is possible to purchase very fine China by the name of Delft. Making pottery is energy intensive, as the raw material (basically clay with metal decorations—in the case of Delft characteristically blue) must be heated to high temperatures. The fuel for this in the Netherlands was originally peat, partially decomposed Sphagnum moss, which was abundant in the low-lying areas of the Netherlands. To this day, large rectangular holes, called polders, remain where the peat was extracted four centuries ago.

A particularly thorough energy analysis of the economy of that time has been undertaken for Sweden by Sundberg. In 1550, Sweden was overwhelmingly rural and very poor. Most of Sweden is too cold for much agriculture, which was concentrated in the south of the country. Most of the citizens lived scattered throughout the vast forests where they cut trees for charcoal, which was used for a variety of purposes, most importantly for smelting the abundant silver, copper, and especially iron ore. Thus, Sweden had at that time two particular assets in terms of natural resources: vast forest areas and rich iron ore. In order to make iron, high temperatures (above 1000 °C) must be used. This is not possible from timber alone but can be done with charcoal, which is basically wood heated in the absence of oxygen so that it is nearly pure carbon. Charcoal is made by taking trees and piling them into a large earth-covered structure containing from dozens to hundreds of trees. Then the pile is fired and allowed to smolder for days.

In 1600, approximately 15–20% of all Swedes lived in small family groups scattered throughout the forest. Their houses were quite small and simple, and most men worked making charcoal. The resulting charcoal was taken to regional metal processing centers, and the iron and copper ore turned into metallic implements. The principal products of Swedish iron factories in the period 1600–1800 were very good cannons. The Dutch were the first to take full advantage of these cannons and mounted them in warships that made them rulers of the European seas for about 100 years, until the English became better at the game. The Dutch invested in cannons because they allowed them, essentially, to steal whatever they wished from other nations. This was considered fair game, at least by the rules of the newly emerging mercantile capitalist economy (although not by the conquered and colonized).

Over time, more and more of the Swedish forest was cut and burned, and since trees grow slowly in the cold climate, eventually the vast Swedish forests were destroyed in almost their entirety. The Swedes faced an extreme energy crisis, and many froze in the winter because they had insufficient fuel and insufficient food. Starting in about 1850, vast numbers of Swedes moved to the United States, especially to the Northern Midwest, where they felt right at home among the snow and the pine forests.

Crosby [27] has commented upon the particular aggression and greed of Europeans compared to others about the world. By 1641, the Dutch trade and military empire extended as far away as Malaysia, where a Dutch fort and windmill can still be found in Malacca. If other nations wanted to trade in waters where the Dutch ruled, they had to either pay tribute to the Dutch or suffer the loss of some of their ships and ports. Consequently, the Dutch got very rich. Thus, the energy of photosynthesis of Swedish forests was translated into dominance of the seas by the Dutch using wind-driven ships to carry far more Swedish cannons than land armies could muster. These energies also generated a very high level of comfort for Dutch burghers and the leisure to generate some of the world's greatest art. Then as now, affluence had a source somewhere in extensive use of energy. But that affluence for the Dutch did not last either, for it was the British defeat of the Dutch at the Straits of Malacca in 1647 that catapulted the British into prominence as a mercantile power. Then the bulk of the eighteenth century was spent in British conflict with the French. Finally, at the end of the Seven Years' War in 1763 and the great British naval victory at Trafalgar in 1805, British hegemony was established over the world's seas, and the long period of Pax Britannica began.

Throughout world history, however, most people remained very poor. Societies often adjusted to these mean circumstances by generating limited social expectations and mechanisms that allowed people to be comfortable with only these very limited economic circumstances and opportunities. One's rewards would be found (they said) after death, or in serving God modestly, or in leisure (in many societies, men hardly worked but spent much of the day in cafés or smoking cigarettes or hashish, while the women tended the fields or shops as well as the children). Fortunately, death rates were high, and the population did not expand greatly beyond the means of the land to support the people who were there. People may have been as happy as, or even happier than, today. We don't know, but the economic circumstances for most were barely above what it took to remain alive and to have and raise children. Some very few adventurous souls would join armies going to faraway places to exploit new resources and peoples (i.e., the rampant European colonialism of one, two, three, and four centuries ago and the crusades long before that). When the Americas opened, massive numbers of Europeans were ready to move to the new "empty" continents to try to better their fortunes, sometimes paying little respect to the fact that the continents were already heavily peopled with Native Americans. In other words, once material opportunities opened, there were plenty of Europeans ready to give it a try to improve their own personal financial situation. Even so for almost all individuals, it was extremely hard to make a living. This was normally accomplished through hard physical labor to chop down trees or to farm or work a mine or in a factory. Records of colonial Americans, for example, show that people spent almost all their time and money just surviving, although they may have done that in reasonable comfort. The concept of spending money for recreation simply did not exist for most, as there was relatively little surplus wealth or surplus energy in these solar-based societies.

Throughout history in many societies, it was deemed just fine to attack another city or nation and simply steal whatever wealth they had accumulated. While this may sound offensive to us in fact, it was highly regarded by many in antiquity. Great writers of past times chronicled approvingly again and again the stories of a leader of one state who plunders another state, bringing glory and treasure to himself and his own state. Vikings, living in Northern lands of very low productivity, sought wealth in raiding parties that terrorized much of Europe for 1000 years. Wooden Viking ships with charcoal-derived iron nails and weapons and woolen sails were constructed and equipped entirely using solar energy. Europeans stole entire continents from Native Americans on solar power (again winds and charcoal plus genocide and settlement), with God as well as gunpowder and European germs on their side [8, 27]. Today this process continues through the economic principle of "globalization," which is viewed by many principally as a means by which the more developed world legitimizes its extraction of resources and cheap labor from the less developed world. Others believe that trade benefits all.

We stop our history here for we treat the history of industrial society in more detail in ► Chaps. 8–10. Meanwhile we provide some additional references for those who want to think more deeply about energy and the "progress" of civilization [29–31].

6.5 A Somewhat Cynical View of Human History

It is very impressive to examine from today's perspective the views of the ancients with relation to war. Plutarch's Lives [28] is a book about famous ancient Greeks and Romans, written several thousand years ago by a distinguished Roman historian. One of your authors (Charles Hall) tackled this book with vigor, wanting to better himself since his classical education, once the signature of a well-educated person, was limited to two undistinguished high school years of Latin under the fierce eye of Miss Meservey. He was also interested in what might be the characteristics of leaders whose reputation had lasted thousands of years. He was quite surprised by what he found: the largest group of the people singled out for praise by Plutarch made their mark by plundering other culture's cities. Plutarch recounted with favor and apparently without irony how these people brought fame and riches to their own cities or regions. These great leaders of the past appeared to be simply robbers and plunderers of accumulated solar-based profits. Human history has been in large part about mustering armies to rape and plunder and about the efforts of others to counter these robbers. Modern Italy, Scotland, and many other European landscapes are full of ancient stone fortifications that must have taken an enormous portion of the time and energy reserves of the ancient citizens to construct. The evolution of more powerful cannons reduced the effectiveness of these fortifications until they were reconstructed to stronger specifications.

America too is constructed on conquest and plunder, from the obvious example of early English and Spaniards stealing the lands of Native Americans to a US military expedition taking what is now California and the rest of the US southwest from the Mexicans in the time period from 1820s to the 1850s by force. When the United States somehow "forgot" to claim the low passes through the Southern Rocky Mountains, they bought them from Mexico in the Gadsden Purchase of 1853. Classic empires seem to have receded during the twentieth century as a new form of imperialism called globalization has advanced, spawning nationalism and ethnic conflict among the world's hungry and disposed.

Occasionally, we can get a quantitative glimpse of the enormous inputs required to fuel the expansion of empires and also the misery suffered by the common person during both the times of the expansion and the collapse of empires. Little was known about energy during most of history, but we can get some glimpses and make some rough calculations. Napoleon was famous for his "cannon park" of 366 cannons, each capable of hurling a 6- to 12-pound iron ball. He took this formidable machine with him to Russia, an incredible and ultimately disastrous campaign that resulted in the death of most of his army. The Russian army under Kuznetsov chose not to stand up to Napoleon's well-oiled military machine but instead retreated before him, stopping only briefly at Borodino to give some serious resistance before melting away, leaving Napoleon to be defeated later by "General Winter." Military historian John Keegan has calculated the energy requirements to feed that cannon park. The 300 plus cannons required 5000 horses to pull them along plus soldiers and teamsters to handle the horses and man the cannons. The men required about 12 tons of food a day and the horses 50 tons of hay, so many additional horses were required to bring along the fuel for men and horses pulling the cannon. One of Keegan's main points is that Nelson's fleet at Trafalgar carried six times the fire power at one-fifth the logistic cost by exploiting wind energy. This indicates the importance of being able to exploit a relatively large energy resource, in this case the wind.

In three successive summers, one of your authors (Charles Hall) happened to read three historical books on European history and some important military invasions in search of empire: the first Peter Massey's on Peter the Great and his attack south into the Crimea in 1696, the second Phillipe De Segur's (a nobleman in Napoleon's army) record of Napoleon's Russian Campaign in 1804, and the third Antony Beevor's Stalingrad, the story of the furthest point that Nazi Germany had penetrated into Russia in 1942–1943. Each of these books is a masterful summary of enormous military campaigns. But it came as a shock when the first map in the third book turned out to be essentially the same map, with nearly the same national borders, in each of the two previous books, centered on the regions of the Baltic, the Black Sea, Moscow, and the Caspian. Each of the books tells of initial tremendous success and enthusiasm for the "glory" of conquest by the invading armies, but in each of them, the invading armies were humbled eventually by the peasant armies, climate, and lack of enough fuel within the devastated invaded regions to support horses, tanks, and soldiers. The suffering of the soldiers, officers, and the commoners caught in the middle in each was immense, and in each, the tales of massacre and barbarous behavior on all sides were appalling. No additional territory was gained by any of these campaigns, despite the enormous expenditure of resources. Educated German officers in 1942 knew well of Napoleon's appalling retreat in Russia and watched day by day as General Winter imposed the same horrible fate on their own army. At the end, it all seemed so stupid. Except for the massacre and displacement of Native Americans (and other aborigines) by Europeans, it seems that since 1800 (and probably long before) most land has remained in the hands of those who were there first. But that certainly has not stopped many invasions as ambitious generals and leaders attempt to conquer other's land.

Thus, much of history can be seen as times of very limited abilities to do much more than survive on one's own resources and that the main path to personal or national wealth was through exploiting others through warfare. Much of history can be viewed as a series of attempts by one group to exploit or dominate others, either by directly stealing their wealth (represented as the long-term gradual accumulation of net solar energy in precious metals, jewels, and edifices) or by gaining access to their resources. We end our brief historical review at a point before the fossil fuel era gave a tremendous boost to our ability to both generate wealth at home and to inflict carnage and misery upon each other [23]. We do note an optimistic pattern: the long age of arrogant European colonization, empire by conquest, and continuous international conflict appears to be behind us following the end of Second World War. With the rise of industrialization and the enormous ability to increase wealth that fossil fuels and their technologies allowed, plus a growing appreciation of the cultures of others and the costs of war, the concept of empires and subjugation of others seems to have largely stopped. But war and its misery continue for all kinds of other reasons, and exploitation of others continues through economic means.

6.5.1 The Repeated Collapse of Empires

There are several dictums of history that are important here. The first is that "history is written by the winners," and the second is that most human endeavors of the past are barely or not at all recorded. The scholars who think the most about this are archeologists, and the archeologist (and anthropologist, historian, and energy analyst) who has the most to say about this issue is Joseph Tainter. Tainter's magnum opus is *The Collapse of Complex Societies* (although we have found his 1992 paper "Evolutionary Consequences of War" to be equally cogent). Both are incredibly good reading. Tainter lists a minimum of 36 once-great civilizations that exist today only as a series of rocks and other hard materials, often under desert sands. The list goes on and on. One has only to visit the great museums of anthropology in, for example, Mexico City or Jalapa, to get a perspective on what incredible civilizations there were in the past and how so many have crumbled.

Why do most military invasions fail, and how did it come to pass that so many once proud and powerful civilizations fell apart so completely and, often, so quickly? There are probably many reasons, but we believe that the energy-based mechanisms put forth by Tainter and summarized above offer the best clue.

6.6 Summary

All of life, including human life in all of its manifestations, runs principally on contemporary sunlight that enters the top of our atmosphere at approximately 1.4 kW/m^2 (5.04 $MJ/m^2/h$). Roughly half that amount reaches the Earth's surface. This sunlight does the enormous amount of work that is necessary for all life including all economic activity. The principal work that this sunlight does on the Earth's surface is to evaporate water from that surface (evaporation) or from plant tissues (transpiration) which in turn generates elevated and purified water that falls back on the Earth's surface as rain, especially at higher elevations. The rain in turn generates rivers, lakes, and estuaries and provides water that nurtures plants, animals, and civilizations. Differential heating of the Earth's surface generates winds that cycle the evaporated water around the world, and sunlight of course maintains habitable temperatures and is the basis for photosynthesis in both natural and human-dominated ecosystems. These basic resources have barely changed since the evolution of humans (except for the impacts of the ice ages) so that preindustrial humans were essentially dependent upon a constant although limited resource base. Over time, humans increased their ability to exploit larger parts of that natural solar energy flow through technology, initially with spear points, knives, and axes which could concentrate human muscular energy, and then with agriculture, metals, dams, and now with fossil fuels.

The development of agriculture allowed the redirection of the photosynthetic energy captured on the land from the many diverse species in a natural ecosystem to the few species of plants (called cultivars) that humans can and wish to eat or to the grazing animals that humans controlled. Curiously, the massive increase in food production per unit of land brought on by agriculture did not, over the long run, increase average human nutrition but mostly just increased the numbers of people. Of course, it also allowed the development of cities, bureaucracies, hierarchies, the arts, and more potent warfare. For most of humanity's existence, most of the energy used was animate—people or draft animals—and derived from recent solar energy. Often humans themselves did most of the work, often as slaves but more generally as physical laborers which, in one way or another, most humans were. For thousands of years, from the period of the beginning of empires 5000 or more years ago until the widespread use of coal for steam power in about 1850, the principal source of energy for any large-scale agriculture or public works was masses of human power, principally but not always as slaves or near slaves (i.e., serfs). By one account, the Cheops pyramid represents essentially the entire energy surplus of the Nile civilization of about 3 million people at that time and required the labor of 100,000 people over 20 years. A second very important source of solar energy was from wood, which has been recounted in fascinating detail in books by Perlin, Ponting, and Smil. Massive areas of the Earth's surface—Peloponnese, India, parts of England, and many other locations—have been deforested three or more times as civilizations have cut down the trees for fuel or materials, prospered from the newly cleared agricultural land, and then collapsed as fuel and soil become depleted. Archeologist Joseph Tainter recounts the general

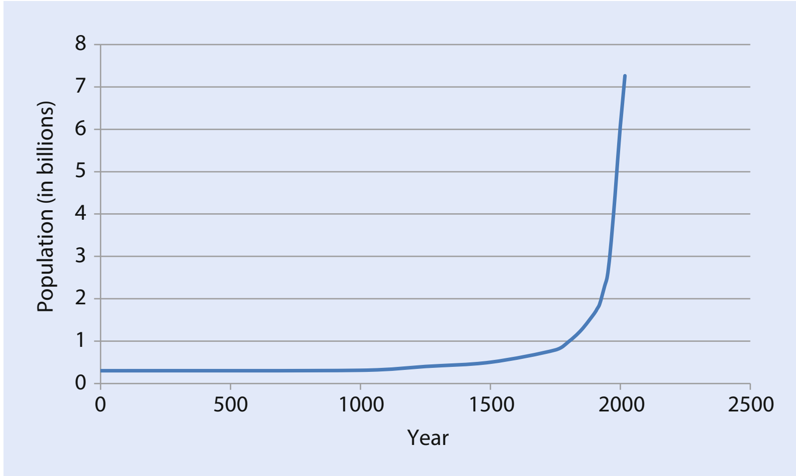


Fig. 6.13 Global human population (Source: United Nations)

tendency of humans to build up civilizations of increasing reach and infrastructure that eventually exceeded the energy available to that society.

Both the natural biological systems subject to natural selection and the preindustrial civilizations that preceded our own were highly dependent upon maintaining not just a bare energy surplus from organic sources but rather a substantial energy surplus, or large net energy, that allowed for the support of the entire system in question—whether of an evolving natural population or a civilization. Most of the earlier civilizations that left artifacts that we now visit and marvel at—pyramids, ancient cities, monuments, and so on—had to have had a huge energy surplus for this to happen, although we can hardly calculate what that was. An important question for today is to what degree does the past critical importance of surplus energy apply to contemporary civilization with its massive although possibly threatened energy surpluses.

6.6.1 Surplus Energy and Contemporary Industrial Society

Contemporary industrial civilizations are dependent on fossil fuels in addition to solar energy. Today fossil fuels are mined around the world, refined, and sent to centers of consumption thousands of miles away. These fuels have allowed for accelerated exploitation of solar energy and for the huge increase in food production, water transport, and sanitation that has allowed the human population to grow enormously over the past 100–200 years. For many industrial countries, the original sources of fossil fuels were from their own domestic resources. The United States, United Kingdom, Mexico, and Canada are good examples. Since many of these initial industrial nations, however, have been in the energy extraction business for a long time, they tend to have both the most sophisticated technology of both production and use and the most depleted fuel resources, at least relative to many countries with more recently developed fuel resources. For example, as of 2017, the United States, originally endowed with one of the world's largest oil provinces, was producing only about half of the oil that it used, Canada had begun a serious decline in the production of conventional oil, and Mexico recently was startled to find that its giant Cantarell Field, once the world's second largest, had begun a steep decline in production at least a decade ahead of schedule. Although new sources of oil are being developed (see ▶ Chap. 13), these are from relatively low-yielding and expensive wells. Meanwhile, the global human population continues its upward course, although at a decreasing rate (◘ Fig. 6.13). The next chapter examines the role of oil in our society in much greater detail.

Questions

1. Discuss several examples of how preagricultural humans exploited solar energy and the relation of the energy they obtained to their own personal energy investments.

2. How are spear points related to energy?
3. How does agriculture concentrate energy for humans? How does this process support a larger population?
4. The human use of fire assisted in opening a huge new food resources of agriculture for humans. Can you explain what the connection might be?
5. What was the relation of agricultural surplus to human specialization?
6. Many former dominant human cultures have collapsed. Can you give an example and the reasons thought likely for that happening?
7. Name at least two important legacies of the reign of Ferdinand and Isabella.
8. What does surplus energy mean to civilizations?

References[1]

1. Lee, R. 1969. !Kung Bushmen subsistence: An input-output analysis. In *Environment and cultural behavior*, ed. A.P. Vayda, 47–79. New York: Natural History Press.
2. Shostak, Marjorie. 2000. *Nisa: The life and words of a !Kung Woman*. Cambridge, MA: Harvard University Press. Glaub, M. and C.A.S. Hall. 2017. Evolutionary implications of persistence hunting: An examination of energy return on investment for !Kung Hunting. *Human Ecology* 45:393–401.
3. Martin, P.S. 1973. The Discovery of America. *Science* 179: 969–974.
4. This analysis is based mostly on a series of papers in Science magazine of 2 March 2001.
5. Culotta, E., A. Sugden, and B. Hanson. 2001. Humans on the move. *Science* 291: 1721–1753.
6. De Candolle, A. 1959. *Origin of cultivated plants*. London: Hafner Publishing Co..
7. Sauer, C.O. 1952. *Agricultural origins and dispersal*. New York: American Geographical Society.
8. Diamond, Jared. 1999. *Guns, germs and steel: The fates of human societies*. Norton: W.W & Company, Inc..
9. Angel, J.L. 1975. Paleoecology, paleodemography and health. In *Population ecology and social evolution*, ed. S. Polgar, 667–679. Paris: Mouton Press.
10. Perlin, John. 1989. *A forest journey: The role of wood in the development of civilization*. New York: W.W. Norton & Company, Inc..
11. Michener, James A. 1963. *Caravans*. New York: Ballantine Books.
12. Tainter, J. 1988. *The collapse of complex societies*. Cambridge, UK: Cambridge University Press.
13. Abrahamson, W.G., and T.N. Taylor. 1990. *Plant-animal interactions*. New York: McGraw-Hill Encyclopedia of Science & Technology.
14. Feeny, Paul. 1970. Seasonal changes in oak leaf tannins and nutrients as a cause of spring feeding by winter moth caterpillars. *Ecology* 51: 565–581.
15. McNeill, W.H. 1976. *Plagues and peoples*. Garden City: Anchor Press/Doubleday.
16. Ponting, C. 1992. *A green history of the world: The environment and the collapse of great civilizations*. Saint Martens: Penguin Books.
17. Carter, V.G., and T. Dale. 1974. *Topsoil and civilization*. Norman: University of Oklahoma Press.
18. Deevey, Edward S., Jr. 1960. The human population. *Scientific American* 203: 194–204.
19. Cottrell, F. 1955. *Energy and society*. New York: McGraw-Hill.
20. Thucydides T. 1954. History of the Peloponnesian war. Trans. R. Warner. New York: Penguin Books.
21. Gibbon, E. 1776. *et Seq. The decline and fall of the Roman empire*. Vol. I. New York: Random House.
22. Walker, D.R. 1974. *et Seq. The metrology of the Roman silver coinage. Part I: From Augustus to Domitian. Part II. From Verva to Commodus. Part III. From Pertinax to Uranius Antonius*. Oxford: British Archaeology Report. Supplementary Series 5, 22, 40.
23. Adas, M. 1993. *Islamic and European expansion*. Philadelphia: Temple University Press.
24. Hall, C.A.S. 2006. Integrating concepts and models from development economics with land use change in the tropics. *Environment, Development and Sustainability* 8: 19–53.
25. Zeeuw, J.W. 1978. Peat and the Dutch golden age. The historical meaning of energy-attainability. *A.A.G. Bijdragen* 25:3–31.
26. Sundberg, U. 1992. Ecological economics of the Swedish Baltic Empire: An essay on energy and power, 1560–1720. *Ecological Economics* 5: 51–72.
27. Crosby, Alfred. 1986. *Ecological imperialism: The biological expansion of Europe, 900-1900*. London: Cambridge University Press.
28. McFareland, John W. 1972. *Lives from Plutarch*. New York: Random House.
29. Thomas, D.W., J. Blondel, P. Perret, M.M. Lambrechts, and J.R. Speakman. 2001. Energetic and fitness costs of mismatching resource supply and demand in seasonally breeding birds. *Science* 291: 2598–2600.
30. Odum, H.T., and R. Pinkerton. 1955. Time's speed regulator: the optimum efficiency for maximum output in physical and biological systems. *American Naturalist* 43: 331–343.
31. Hall, C.A.S. 2004. The continuing importance of maximum power. *Ecological Modeling* 178: 107–113. Brown, M., and C.A.S. Hall. The H. T. Odum Primer: an annotated introduction to the publications of Howard Odum.

1 Dedicated to chemists Frederick Soddy and William Ostwald, anthropologist Leslie White, archeologist and historian Joseph Tainter, sociologist Fred Cottrell, historian John Perlin, systems ecologist Howard T. Odum, economist Nicolas Georgescu-Roegen, energy scientist Vaclav Smil and a number of others in these and other disciplines.

32. Wroe, S., J. Field, R. Fullagar, and L.S. Jermin. 2004. Megafaunal extinction in the late Quaternary and the global overkill hypothesis. *Alcheringa* 28: 291–331.
33. Taylor, G.F. 1951. *The Transportation Revolution, 1815–1860*. New York: Holt, Rheinhart and Winston.
34. Lieth H., Whittaker R.H. (eds) 1975. Primary Productivity of the Biosphere. Ecological Studies (Analysis and Synthesis), vol 14. Berlin, Heidelberg: Springer.
35. Tylcote, R. 1992. *A History of Metallurgy*. Institute of Materials

Energy, Wealth, and the American Dream

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_7

The Americas were the last major liveable continents to be settled by humans. When first Asians ("Native Americans" or "Indians") and then Europeans (and those they brought from Africa and Asia as slaves or laborers) settled in the Americas, they found enormous lands empty (or empty-able) of other humans with incredibly rich resources. Having emigrated from relatively densely populated and socially and economically stratified Asia and Europe, the Americas represented enormous resources, per capita power and the freedom to exploit those resources and hence generate wealth. This of course is a well-known story told to most American children with a focus on the various heroic activities of our ancestors. But it is also a story about energy.

7.1 Waves of Colonists to America: First Asians and then Europeans

Most scientific analysis supports the idea that people first came to the Americas during the low ocean levels that occurred 10,000–20,000 years ago when huge amounts of water were tied up in glaciers during the most recent ice age. (One should respect, however, the view of many Native Americans, including some Native American scientists, that "they have been here indefinitely"). When Native Americans arrived on this continent, they found few other humans, amazing natural ecosystems, and enormous wildlife resources (their principle resource base). Since these people were skilled hunters and had very effective tools (spears and bows and arrows, as well as highly evolved social systems for hunting and, subsequently, agriculture), they had a tremendous economic boom, increasing in numbers to perhaps 50 million people in the Americas. But there was a cost to this tremendous economic growth: the extinction of many of the species that had originally been very important in their diet. For example, we know that 10,000 years ago, there were two species of elephants, 10-foot-tall beavers, and giant sloths in what is today the United States. These and many other large species (known collectively as megafauna, meaning simply "large animals") disappeared soon after humans came. While scientists debate the degree to which climate change vs. human hunting did in these animals, there is no question that everywhere that humans went on the planet, the large animals disappeared soon after [1, 2]. Meanwhile other humans in the Americas were overexploiting soils in many regions, leading to collapse, that is, a radical and sudden decrease in the magnitude and degree of complexity of entire societies. This happened, for example, to the Mayas of the Yucatan and present-day Guatemala [2, 3]. Whether such a collapse will occur with present-day European Americans has been discussed by these and many other authors, most of whom consider it a distinct possibility.

The second wave of humans that entered the Americas came from Europe starting in 1492. They brought with them a whole new suite of plants, animals, and technologies [4]. From our present perspective, the basic result of this was that the overwhelming majority of the people that were in the Americas in 1492 were killed directly by Europeans or by the diseases they brought, as described in *Guns, Germs, and Steel* [5]. It is not a pretty story and would be called genocide today [6]. Thus, the total population again was maintained at a very low level as the new people arriving from Europe were more or less no more than compensating for the net reduction of the original human inhabitants. From the perspective of the next three centuries of economics, this meant that there were still tremendous resources on a per capita basis for each European immigrant and for their children. America was a land of opportunity indeed, for there were enormous untapped resources and not too many people with whom to share them. From roughly 1700 to about 1890, there was always an "empty frontier" to the west with land open for the taking and many opportunities for the ambitious and industrious. Of course European Americans rarely considered that these "empty" lands were already heavily populated with Native Americans, whose sometimes settled but frequently nomadic, non-industrial lifestyle was in fact very well equipped for a sustainable existence based on mostly renewable resources. The economy of the entire continent went from one relatively sustainable to one clearly not. The greatest cause for the war of independence was basically resource scarcity—the cutting off of the trans-Allegheny frontier, first in 1763 by means of the Proclamation Act and then later in 1775, with greater enforcement, with the Quebec Act. Open rebellion soon followed [7].

The technologies that the Europeans brought (such as mining, metallurgy, the moldboard plow, deepwater fishing, and so on, plus the development of a series of self-serving myths (e.g., "Rain Follows the Plow," "Manifest Destiny")) led to a massive exploitation of both renewable (e.g., soil, trees, fish, bison) and nonrenewable (i.e., gold, silver, coal) resources and seemingly limitless wealth for many, although hardly all, people in a way very consistent with Polanyi's definition of economics (▶ Chap. 1). The inventiveness of Americans certainly added to their ability to utilize resources, and important new products and processes were developed which included, among others, the light bulb, intercontinental railroads, steamboats, mass-produced automobiles, and the telegraph. As we said earlier, most people who thought about why the United States had become so wealthy attributed the wealth to the particular industry of the existing or immigrant European populations or to the blessing of God. Probably far fewer thought about the fact that the United States had such a huge, largely untapped resource base and very low population density compared to, for example, Europe. In other words, the United States enjoyed large resources per capita. Alfred Crosby [4] believes that Europeans were especially good at colonizing the rest of the world and exploiting the resources where they colonized because of their unique and self-serving aggressiveness. As evidence he cites that essentially all people in the world today are where they were in 1000 AD except for Europeans (who have colonized North and South America, South Africa, Australia, and New Zealand—i.e., all regions with temperate climates) or who have been moved by Europeans (African slaves and their descendants, Chinese workers to the Western United States). This view of the essential aggressiveness of Europeans and their ability to successfully exploit others and their resources is the essence of Jared Diamond's highly successful book *Guns, Germs, and Steel* which also focuses on certain geographical advantages that Europeans had. Europeans were not necessarily good inventors, but they were extremely good adapters—of gunpowder, agriculture, animal husbandry, metallurgy, communication through the written word, and so on. All of this was transferred to the United States, where it was applied with great gusto to a continent rich in unexploited resources, including, as we have said, timber and grass for fuel, good soils with summer rains, rich mineral deposits, and so on. Thus, immigrants from Europe found that they could own what was, by ordinary European standards, a massive amount of fertile land whose fertility depended basically upon their own initiative and energy. This was the beginning of the "American dream"—the ability to exploit large quantities of solar energy by massive numbers of ordinary individuals.

Transforming nature is a hard work. In the past when this work was done mostly with one's own muscles, the amount of transformation an individual could do was physically difficult and limited in magnitude. Wealthy people of the past often did this through the hard work of others by means of social conventions such as low-wage labor, serfdom, and slavery. Think of the lovely houses and lives of ease of southern US plantation owners 150 years ago, an affluent lifestyle generated on the backs of dozens to hundreds of laborers working to clear forests and plant and harvest crops. In fact slavery has been a common situation mentioned frequently in the bible and in many ancient historical accounts. It was not a nice life (to put it mildly), and the concept became increasingly repugnant even to many of the owners of slaves. The Civil War ended slavery in the United States, but de facto slavery continued as former slaves continue to work the lands and as many poor immigrants were brought into the United States from Ireland, Italy, China, and elsewhere to do hard physical work at "slave wages" or as indentured servants. People were helped in this work by the physical power of horses and by the physical work obtained from burning wood and the power of falling water. Wind was exploited by sailing ships and an occasional windmill, and increasingly coal was used for railroads and in factories. But overall most work continued to be done by human labor assisted by animals through the turn of the century. This is not to say that most people were not happy, often they were. But the production of wealth was a difficult, sweat-generating process, and most people were very poor materially by today's standards.

7.2 Industrialization and Isolationism

By the late eighteenth century, new sources of energy were being developed in, especially, New England where the abundant water power potential allowed enormous, by the standards of the time, new factories to be built making textiles, shoes, chemicals, and all manner of iron tools and equipment. This allowed the development of great concentrations of workers in such towns as Manchester, New Hampshire, Lowell, Massachusetts, Boston, and New York City. Water-powered machines greatly increased the amount of goods a laborer could generate in an hour (i.e., labor productivity) and the subsequent wealth of at least some in New England. Meanwhile as forests were cleared for agricultural land and for homesteads in New England, Europeans spread to the Southeast and then westward to virtually the entire Midwest, where enormous amounts of wood fuel were available for all manner of local industries [8]. Fish and other marine life were abundant too, and the world's vast numbers of whales were greatly decreased by Massachusetts seafarers in order to get whale oil, the principle source of lighting. At the start of the nineteenth century, England and Germany had begun their great industrial transformation, using the concentrated solar energy found in coal to generate enormous new amounts of high temperature heat that allowed far more work to be done than was the case with water power, wood, or charcoal. This technology was transferred to the United States which had very rich coal reserves. In 1859 Colonel Edwin Drake drilled the nation's first oil well, and kerosene began to replace whale oil as the lighting source of choice. The enormous wealth generated by the new industrialization allowed the "captains of industry" to become enormously rich by world standards. This, along with the great disparity in wealth between them and their workers, generated the phrase "the Gilded Age" for the 1890s. But it was not a smooth pattern of growth as periodic depressions caused a serious loss of wealth for many people, rich and poor. Most people continued to be poor, or at least far from affluent, making barely enough to survive and have a family. Still, in America, despite the disparities in income, the wealth distribution was quite equitable compared to Europe and most of the rest of the world, in part due to the ability of many to have access to land and its solar energy (once the Native Americans were displaced) through farming or with an axe.

7.3 Spindletop and the Beginning of the Affluent Society

Then in 1901 something special happened. The generation of wealth for entire societies (especially in the United States and also much of Europe) suddenly changed and the proportion of people with at least moderate wealth took a great upswing, as did the total quantity of wealth in the world and even the wealth per capita (Fig. 7.1). Perhaps the single most important event in a series of similar events was the development of the Spindletop oil field in Beaumont, Texas, in 1901, which gave a new realization that serious wealth could be generated for the many by finding, selling, and using oil (Fig. 7.2). Before Spindletop, oil certainly had been found and developed, but individual oil fields were relatively rare, small, and difficult to develop, with production of hundreds or thousands of barrels of oil per year. Spindletop alone changed all that, by producing up to 500,000 barrels *per day*, essentially doubling the nation's petroleum production. It was then understood that a great deal of wealth could be had for many with relatively little investment from the oil business, and soon other areas were found to be nearly as productive as Spindletop. Other people looked at how relatively small investments could produce a great deal of money and by using the ideas and technologies, developed at Spindletop, oil production increased rapidly. Large additional finds were made not only in Texas and Louisiana but also in Indonesia, Persia, Romania, and many other areas. As the production of oil increased more and more every year so did the nation's wealth, far more rapidly than ever before. Oil's original use was for kerosene but soon a waste product, gasoline, found an important new use as automobile fuel. Oil and oil-driven vehicles began to be applied to all economic areas, such as growing food and transporting it to long distances, catching fish, cutting, moving and milling lumber, running all kinds of factories, and most other economic processes. While the few continued to get most of the direct wealth, its use spread afflu-

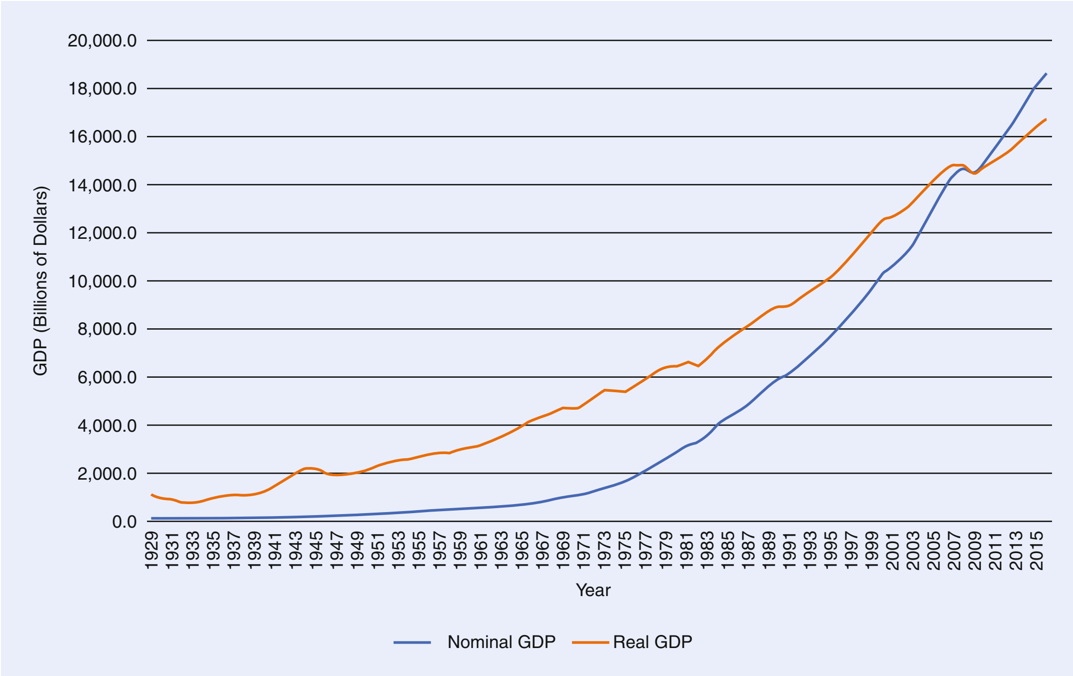


Fig. 7.1 US nominal and real GDP (2005 dollars) from 1900 to 2016 (Source: US Department of Commerce)

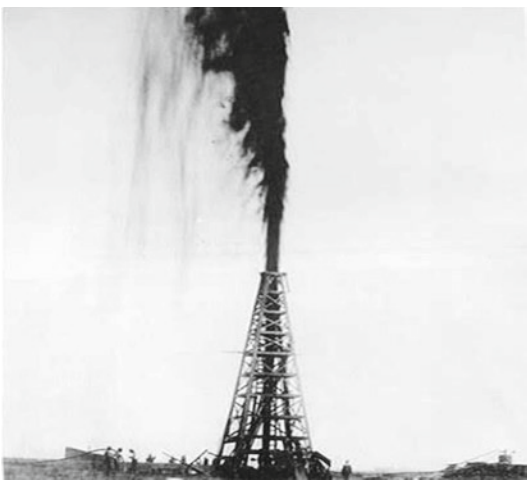


Fig. 7.2 Spindletop, Beaumont Texas, 1901 (Source: Texas Energy Museum)

ence to the many. The gross domestic product (GDP), an index of the total production of wealth by the country, began to grow exponentially (i.e., as compound interest), decade after decade, (although interrupted by periodic depressions in 1921 and of course 1929) something almost unheard of before. Thus began the age of affluence for the many or what can be called mega affluence. This led many economists and politicians to consider 3 or even 5% per year economic growth as "normal" when in fact it was something very new. That oil-based growth spread increasingly around the world and has continued for many until now.

This pattern of exponential growth of oil (and energy more generally) use and wealth for the United States continued at least up to the "oil crises" of the 1970s, (Fig. 7.3) fits in well with our more general energy perspective, for it focuses on the raw materials needed and the energy required to do any process, including economic production. Quite simply, it is energy that has allowed our economy to undertake the transformations that extract and process those materials into the economic products and services we desire. Other things are needed of course, such as the technology to get and use energy and a supportive political and economic environment, but the driver of wealth production is the energy to do the work of economic production. To make this clearer we will examine in more detail what has probably been the largest generation of wealth to have ever occurred—the production of the vast amount of wealth represented by "the American dream."

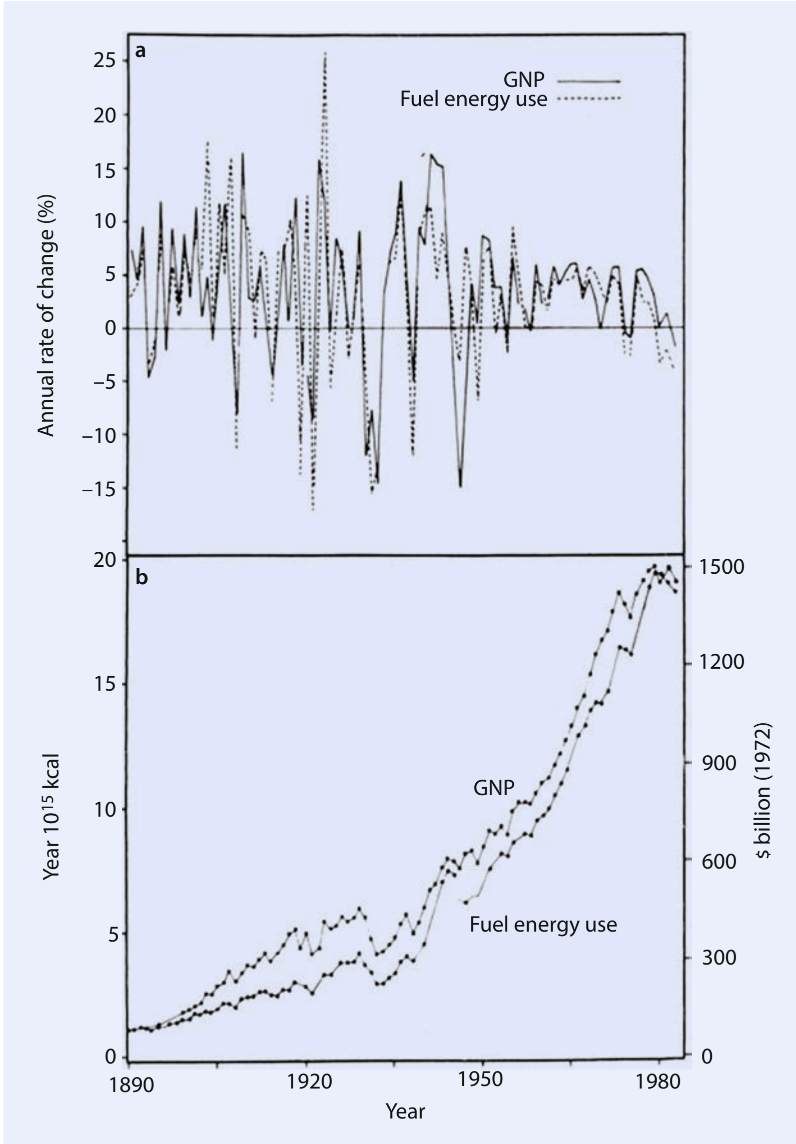


Fig. 7.3 Total income produced plotted along with total energy used for the US economy 1905–1984 (Source: Cleveland et al. and Hall et al. [21]). **a** Annual rates of change; **b** raw data from US Energy Information Agency

7.4 The Creation and Spread of the "American Dream"

At the beginning of the twentieth century, the United States, dominated by European Americans, was becoming the world's emerging agricultural and industrial giant. In 1900 the United States ran principally on coal, wood, and animal power, but oil became increasingly important with the new oil wells and the development of automobiles, trucks, and tractors that ran on gasoline. Large dams, built with the help of massive oil and coal-powered machines, brought irrigation water and electricity to many in rural areas, resulting in huge additions to the availability of biological and physical energy for each American. For the first time, a very large proportion of the population of an entire country was becoming fairly affluent, and some were becoming extraordinarily so. This enormous affluence was associated with, and clearly dependent upon, an increasing use of energy that expanded at almost exactly the same rate as the increase in wealth and that made each worker much more productive (Fig. 7.4 and Table 7.1).

Curiously, even as the United States became more and more dependent upon fossil fuels for basic transportation (meaning mostly railroads), people became even more dependent upon horses for transportation of people and goods at either end of the journey [9]. Because coal-fired railroads

Fig. 7.4 Mean US labor productivity per worker hour (in constant dollars) and energy used per worker hour, 1905–1984, when data acquisition was stopped (Source: Cleveland et al. and Hall et al. [21])

generated a great deal of noise and were very smelly, and especially because they threw out sparks that often set houses on fire, they tended to be banned from city centers. Thus, until the dominance of the internal combustion engine after about 1920, freight and passengers tended to be delivered from the railheads to the center of the city by solar (i.e., grass)-powered horse drawn vehicles!

For the past century and a half, the United States has been considered the world's richest nation and, perhaps of greater importance, as a place where someone with a lot of skill and effort can make a great deal of money if they work hard enough, regardless of the circumstances of their birth. For many working class Americans, the dream was not affluence, but stability—a steady job, a house of one's own, the ability to pay one's bills, take a vacation, have a cushion for old age, or have a better life for one's children. This economic success is usually attributed to the characteristics of the people who live within its borders, to their genes, to their hard work, the beneficence of God, or some other such factor. Education is usually considered important, and the United States has traditionally led the world in the quantity and quality of its higher education, especially at the postgraduate level. Many of the readers of this book in the United States may be taking economics or business classes in order to learn the skills necessary to become more affluent. The idea that possessing more money makes you better off is central to the economic theory of consumer behavior, which in turn is an underpinning of modern economic thought. We will explain this idea in more detail in ▶ Chap. 14, including the fact that there is no clearly convincing evidence that it is true. The idea that a better education will lead to more affluence is also deeply engrained in the American psyche, and from the first days of the Republic land was to be set aside for schools (as per the Northwest Ordinance) largely for this reason.

In fact the ability to achieve wealth in the United States is largely a consequence of the incredible resource base once found on the North American continent. These include initial endowments of huge forests, immense energy and other geological resources, fish, grass, and, perhaps of greatest importance, rich deep soils where rain falls during the growing season. While many other regions of the world also have, or had, a similarly huge resource base, the United States has several other somewhat unique important attributes. The fact that these resources have been exploited intensely for only a few hundred years (vs. many thousand as in Europe or Asia), the presence of large oceans separating Americans from others who might want our resources, and an extremely low human population density in the past and even now, results in resources per capita that was very large and is still relatively high (Table 7.2). There

7

Table 7.1 Getting a feel for energy units and their conversions

Useful conversions:	
1 calorie	4.1868 joule
1 kilocalorie (Cal or kcal)	4187 joule
1 BTU	1.055 KJ
1 kWh	3.6 MJ
1 therm	105.5 MJ
1 liter of gasoline	35 MJ
1 gallon of gasoline	132 MJ (million joules)
1 gallon of diesel	140 MJ (million joules)
1 gallon of ethanol	84 MJ (million joules)
1 cord dried good hardwood	26 GJ
1 barrel of oil	6.118 GJ
1 ton of oil	41.868 GJ (= 6.84 Barrels)
Some basic energy costs:	
1 metric ton of glass	5.3 GJ
1 metric ton of steel	21.3 GJ
1 metric ton of aluminum	64.9 GJ
1 metric ton of cement	5.3 GJ
1 MT of nitrogen fertilizer	78.2 GJ
1 MT of phosphorus fertilizer	17.5 GJ
1 MT of potassium fertilizer	13.8 GJ
1 joule	Picking up a newspaper
1 million joules (1 MJ)	A person working hard for 3 h
3 million joules (3 MJ)	A person working hard for 1 day
11 million joules (11 MJ)	Food energy requirement for one person for 1 day
1 billion joules (1 GJ)	Energy in 7 gallons of gasoline
1 trillion joules (1 TJ)	Rocket launch
100 10^{18} joules (100 exajoule)	Energy used by the United States in 1 year (2009)
607 10^{18} joules (607 exajoule)	Energy used by World in 1 year (2015)

Thanks in part to R. L. Jaffe and W. Taylor Energy info card, Physics of energy 8.21, Massachusetts Institute of Technology

Table 7.2 Population numbers and density of the Unites States and other countries in 2009–2010

	Total population (thousands)	Density (people/km²)
World (land)	6,828,134	46
Bangladesh	162,221	1127
Palestinian Territories	4013	667
South Korea	46,456	487
Puerto Rico	3982	449
Netherlands	16,618	400
Haiti	10,033	362
India	1,182,328	360
United Kingdom	62,041	255
Jamaica	2719	247
Germany	82,689	229
Pakistan	169,792	211
China	1,338,153	139
Nigeria	154,729	168
France	62,793	113
United States	309,535	32
Argentina	40,134	14
Russia	141,927	8
Greenland	57,000	0.026

*Source: Wikipedia

are a number of reasons that the population density is low. Probably time is most important. To the best of our knowledge, humans have been in North America, at least on any substantial scale, for only 10 or 15 thousand years vs. 50,000 for Europe and much longer for Africa and Asia. Second is the vast depopulating of the original native population that occurred after 1492. The third is the slowing of population growth rates commonly observed as humans become more affluent – which occurred in the United States (◘ Table 7.2).

7.5 Two World Wars Separated by the Great Depression

By the early 1900s, the spirit of isolationism was strong among the citizens of the United States who were deeply suspicious of Europe and its entrenched rivalries and frequent wars. The United States had mostly isolated itself by choice from Europe and indeed most of the rest of the world. After a long delay, the United States entered the First World War, greatly accelerating our involvement with the rest of the world even as antiwar sentiment at home was especially strong. Indeed, incumbent president Woodrow Wilson based his reelection campaign on the slogan, "He kept us out of war." The military value of oil and petroleum-based transportation was first realized by Winston Churchill who had begun the transformation of the British fleet from coal to oil just before the war. England, however, had no oil. Parliament passed the final piece of the conversion, a guaranteed contract for Anglo-Persian Oil Company (now BP) 9 days after the outbreak of hostilities, and thus began the long and often contentious association of the increasingly oil-dependent Western world and the oil-rich Middle East. The value of oil was shown clearly when the French, faced with a potential large military defeat during the battle of the Marne in 1914, rushed 6 thousand French soldiers from Paris to the battlefield in taxicabs where they helped to achieve a great victory. Petroleum was also used for the first time for airplanes and primitive tanks. The war, begun with coal-powered ships and railroads and millions of horses, ended as an increasingly petroleum-based conflict. Thus, the ability of petroleum to enhance all things, including mass murder of and by armies, was tremendously enhanced.

After the war, the United States had a decade-long period of peace and greatly increasing affluence, except for the depression of 1921, fueled in large part by ever-increasing production of oil. In retrospect it is clear that much of that affluence, however, was wealth only on paper or speculation. In contemporary terms the increase in oil prices became an asset bubble. *Speculation* refers to people purchasing land or other resources not for their own use but in anticipation of being able to sell it later to someone else at a higher price. To do this, banks in the 1920s loaned out far more money than they actually had as assets (i.e., "money in the vault" or ownership of houses) to cover the loans. Simplistically one can think of banks as the place that people put their excess money, saving "for a rainy day," while other people can borrow that money to buy a home, for example. Since most homeowners want to keep their home and will try hard to make their payments, this is normally considered a fairly safe way to loan money, at least if the bankers have done their homework and determined that the borrowers have the means to do so. Since the early days of capitalism (which many attribute to the rise of the Medici family in Florence, Italy), banks have also loaned out some portion of this money for others to use as investment capital, that is, money to start or expand a business, to buy equipment, to build buildings, and so on in anticipation of using them to make additional money. Banks increasingly loaned out more money than they had in the vaults or even on paper. Nobel Laureate Paul Samuelson wrote that this process, called *fractional reserve banking*, probably had its origins with ancient goldsmiths who gave receipts or notes for the storage of gold. Eventually, the notes began to circulate as money when the smiths realized not all depositors were likely to return for their gold at the same time. Both processes have allowed banks to pay interest to those who put their money in the bank. Traditionally the prudence of the bank owners and directors, or sometimes government regulators, led bankers to keep a significant portion of the bank's money in the actual bank vaults, so that the people who own the money can withdraw it if they want. All banks, however, live in fear of a "run on the bank," that is, a time when too many people want to get their money out of the bank at the same time. Some speculation has always been with us, but it became much larger toward the end of the 1920s. This was because in the expanding economy the price of land and securities had been pushed up to be far higher than their real worth by people paying higher and higher prices in anticipation of even higher prices in the future. Reality caught up with the speculators on October 20, 1929, a day called "Black Tuesday" because of the enormous loss of wealth and remembered today as a time when, at least according to legend, a number of investors committed suicide by jumping off their Wall Street buildings. On that and ensuing days, speculators and other investors lost 100 billion dollars, a huge sum at the time. Although no more than 2% of Americans owned stock at that time, the impacts of the Wall Street collapse filtered downward to local banks, who then loaned out far less money to protect themselves and thus to local economies. Speculators had borrowed money from their stockbrokers who, in turn, borrowed from banks. The spectacular losses in asset values left investors unable to repay their brokers, who then defaulted upon their own loans. Runs on the banks ensued and insolvencies rose to more than 5000 by 1931. Before long nearly 20% of Americans had lost their jobs.

This began the period we now know as "the Great Depression" when the country slipped into a long period of little or negative economic growth, high unemployment, and the general financial difficulties of the 1930s. President Herbert Hoover, who had previously shown great skill in combating postwar starvation in Europe, attributed the primary cause of the Great Depression to the "war of 1914–1918" and the economic consequences of the peace treaty that ended the war. This attitude encouraged American isolation and individualism, which was made even stronger by the press, especially in the Midwestern states. The publisher of the influential *Chicago Tribune* carried on an enthusiastic campaign to stop the country from any international entanglements, such as aiding Britain in the days before the United States joined the First World War. He even considered Hoover's mild reforms to try to deal with the early days of the depression and his tepid contact with international leaders to be dangerous, going as far as to call Hoover "the greatest state socialist in history." This is pretty ironic as today Hoover is usually considered as one of our most conservative presidents. Hoover believed that the economy would correct itself given time and used an unemployed man selling apples on a street corner as an example of someone working individually toward a recovery for all.

In fact the economy got worse, and in the next election, the country rejected Hoover and turned to Franklin Roosevelt. Roosevelt ran as a fiscal conservative and believed in a balanced budget. This belief led him to raise taxes to pay for social programs. Consequently his "New Deal" did not provide a great fiscal stimulus. Yet Roosevelt had also long believed in the idea that the government should strive to improve the life of its people, especially in hard times, and he became increasingly convinced that the government should spend more and more money, even if it were borrowed, to try to "prime the pump" of the economy. This belief (essentially, but not explicitly Keynesianism; see pp. 55) took many forms, which ranged from job creation programs such as the Civilian Conservation Corps and Works Progress Administration to Social Security and the reform of labor relations. Most economists, including liberal and Keynesian economists, agree that this approach actually did not generate enough deficit spending to add a great deal to economic recovery. That took the huge increase in public spending associated with Second World War, during which time the economy had tremendous growth fueled by massive increases in government spending and government debt. The commitment to a balanced budget disappeared during the war, and the use of deficit spending to stimulate the economy, along with the social structure that the war helped create, led to a long period of rapid economic growth. What was not so well understood was that all of this economic expansion required cheap oil, which established our long-term structural dependence upon petroleum. The combination of increased government spending and the rekindling of the moribund industrial power of the nation had been a primary factor that clearly worked for winning the war and maintaining an ever-increasing standard of living and thus the American dream.

Nevertheless there are many to whom Roosevelt's (and later presidents') intervention in the economy was anathema, for they believed that government should stay out of what they consider people's own private business. But their voices were few and far between at the time. The era of the "New Economists," who based their principles on the work of John Maynard Keynes but emphasized economic growth over all other goals, was about to begin. It was the era in which economists believed they had "conquered the business cycle."

The New Economists believed that with wise application of prudent policies regarding taxing, spending, money, and interest, they could enhance the efficiency of markets and relegate depressions to the past. This confidence could not last beyond the 1970s, however, a period characterized by high unemployment and inflation. The questions of the effectiveness of government regulation are with us again in the 2010s, at a level of venomous discourse that few economists of the 1950s and 1960s could have possibly imagined.

What is especially interesting from our energy perspective is that the depression was a time of tremendous energy availability in the United States. The East Texas field, the nation's largest ever except for Prudhoe in Alaska, was discovered in 1930, the first full year of the depression. Oil was cheap, but there was virtually no market for it. But when the US economy finally began to recover, especially in the 1940s, there was a great deal of energy to power that expansion. Thus, it is clear that one needs not only available cheap energy but additional economic conditions to generate economic growth, an issue we addressed in ► Chap. 5.

Meanwhile Japan, a relatively small country without a large resource base and which had formerly looked inward for centuries, increasingly became industrialized and of necessity looked outward for the resources it needed. Buoyed by their success against a giant Russian fleet at the battle of the Tsushima Straits in 1905, the Japanese built a huge, modern fleet. As much as half of the gross national product of Japan went to building up their military machine, and this expansion took up such a large portion of the resources available to them that, for example, Japanese families were encouraged to feed their rice to make their boys, the future soldiers, strong, while the girls got to eat only the water the rice was boiled in. Japan invaded China and Korea for coal and iron and began to expand outward into the Pacific Ocean, for example, into Okinawa. The United States had worked to contain the imperial ambitions of the Japanese in the 1930s by both negotiated treaties and a limited military buildup in the Pacific. The Japanese realized that the expansion of their economy depended upon reliable access to oil. That oil was to be found in the Dutch East Indies (now called Indonesia). The United States, in a largely overlooked overt act of war, blockaded Japan's access to that oil using warships in 1941. The most militant voices in the Japanese military were convinced that the only way to protect their oil resources was to deliver a knockout blow to the United States' Pacific Fleet. Thus, the desired and partly successful isolation of the United States from the rest of the world came to a screaming halt on December 7, 1941, when the Japanese attacked US Naval bases on the island of Oahu in the Hawaiian Islands. On the day after the attack, President Franklin Delano Roosevelt asked Congress for a declaration of war. Germany and Italy subsequently declared war on the United States. The Second World War that began in Europe in 1939 had begun for the United States. In many ways it was the world's first war based upon oil, and in many ways it greatly accelerated the industrialization of the world. The role of oil in the Second World War has been especially well told by Daniel Yergin [10] in *The Prize*, his very comprehensive book about oil.

Our entry into the shooting war began with the Japanese bombing of the United States fleet in 1941, although as noted this was not the first act of war in the Pacific. The war ended in Europe with the military defeat of the Italian and German militaries and the surrender of the Fascist and Nazi governments. Again, the availability or lack, thereof, of fossil fuels played a key role. Toward the end of the war, Germany, having lost access to the petroleum supplies of Africa and the Middle East, produced limited amounts of gasoline from coal, pioneering the same technologies (called Fischer Tropsch) currently being considered for making liquid fuels from coal. Their production facilities, however, were destroyed by allied bombing once the allies gained air superiority. Air superiority was itself enabled by the fact that US companies invented and then produced 100-octane aviation fuel that allowed the use of higher compression, more powerful engines, which helped the British to win the battle of Britain and the allies to eventually gain general air superiority. The Germans were so depleted of liquid petroleum by late in the war that they had to bring the first ballistic missiles (the V-II rocket) to the launching pad with mules. In the Pacific theater, the Japanese too had run so short of oil that they initially had to leave the world's largest battleship in port for lack of fuel and then sent it out to a last battle with only a one-way supply of oil. They used turpentine as fuel to fly some of the kamikaze (suicide) airplanes that were attempting to sink the ship that the father of one of this book's authors (Hall) was on in Okinawa. Hall's friend and colleague Tsutomu Nakgatsugowa remembers clearly as a child that all of the pine trees in his Japanese village were uprooted to make turpentine for fuel. The war ended in 1945 after the first use of atomic weapons during wartime, representing again an enormous increase in the human use of energy, both in the nuclear explosions themselves but also in the huge amount of fossil and hydroelectric energy that had been used to separate the isotopes of uranium. It was only a matter of time and technology until America's vast industrial strength prevailed. Perhaps it was more accurately put by Pulitzer Prize-winning historian David Kennedy who said that the war was won with Russian lives and American machines. And, we add the petroleum to run them.

7.6 The Rise of Affluence for Many

On the home front, something unique occurred. The standard of living *rose* for a people engaged in war, as the war effort rekindled the US economy that had been devastated by the Great Depression. Unemployment, which stood at more than 17% of the labor force in 1939, fell to less than 1.2% in 1944. The value of economic output more than doubled in a mere 6 years. Large social changes occurred during the war years too. Women entered the paid labor force in unprecedented numbers, often earning high wages in both clerical and production jobs. There was little to spend one's money on, and savings as a percent of income rose to the highest levels in history, providing massive investment monies. People patched their clothes, recycled their metals, and, encouraged by gasoline rationing, stopped driving to aid the war effort. African-Americans found relatively high-paying jobs in the labor-scarce factories and began the slow and painful process of integrating into White society. The conflict between labor and management that so characterized the depression era declined as the major industrial unions signed a no strike pledge for the duration of the war while seeing both corporate profits and their wages and benefits increase.

Even larger changes were to come with the end of the war, changes that dramatically impacted the drive toward affluence. A new social contract between workers, employers, and the government was in the process of creation, and this social contract provided a newly powerful nation with the "pillars of postwar prosperity" [11]. These vehicles to maintain prosperity and social stability were based on domestic economic growth and enormous international power (military and economic) internationally, specifically:

1. Basic accord between capital and labor, at least, after a period of intense strike activity following the war, especially between the largest multinational corporations and the largest manufacturing unions, facilitated by giving labor a share of productivity gains in the form of higher wages.
2. *Pax Americana*. The United States became the dominant military and economic power after the Second World War, holding most of the world's nuclear weapons and gold as well as being the largest exporter of oil. Additionally the international monetary system was reworked with the US dollar as the key currency, and the fractional reserve banking system was internationalized to allow the expansion of the money supply to accommodate growth globally.
3. Accord between capital and citizens. Large-scale oligopolies, the government, and the average citizen united around three basic premises: economic growth would replace redistribution as the means of improving well-being; government policy should be focused on the availability of cheap nuclear and other energy and anti-communism.
4. The containment of intercapitalist rivalry. The tight oligopolies constructed from the 1890s onward controlled destructive price competition and allowed large corporations to control their rivalries by means of mechanisms such as price leadership, market division, and use of advertising. Initially the United States was the dominant producer worldwide, having the only viable industrial economy at the end of the war. Stable oligopolies competed on the basis of market share, not price.

A critical component of these patterns was the large increase in labor productivity during that time. This allowed both industry owners and labor, especially of the largest corporations, to do better and better. What was less emphasized but enormously clear in retrospect was that to allow these four pillars to operate and expand, it was both possible and necessary to massively increase the production from oil, gas, and coal fields, some new, and some old but barely tapped previously, so that once the economic engine was started, there was a great deal of high-quality energy available even though the war itself had consumed some 7 billion barrels of oil (about the same as recent annual consumption by the United States). The United States began using many times as much energy per person as had been the case relatively few decades before.

In addition, the nation was left with the enormous munitions facilities built at taxpayer expense at, for example, Muscle Shoals, Alabama. These facilities used the Haber-Bosch process, invented in Germany just before the First World War, to make ammonia [12]. This chemical process for the first time allowed humans to access directly the enormous amount of nitrogen in the atmosphere, which was extremely valuable for the munitions, agricultural, and chemical industries. Before Haber and Bosch perfected their chemical synthesis, the primary sources of nitrates were manure, the large deposits of bird guano found off the South American coast and the sodium nitrate deposits in the Atacama Desert. Peru and Chile had fought the guano wars over access to the bird droppings. But eventually the mining of the guano exceeded the replenishment and the resource vanished. Another source needed to be found. Seventy-eight percent of the atmosphere is nitrogen (N_2), but this nitrogen is very difficult to access because of the triple bonds in the di-nitrogen molecule (i.e., N_2). Until 1909 only the tremendous energy of lightning or some very selected algae and bacteria could break these bonds. Gunpowder and fertilizer depended upon the exploitation of rare deposits of nitrates concentrated by birds over millennia. Fritz Haber, in one of the most important scientific discoveries ever made, found that by heating and compressing air mixed with natural gas, that is, by adding hydrogen and large amounts of energy to the nitrogen in the air, and with the right catalyst, the N_2 molecule could be split and turned into ammonia (NH_3). This in turn could be combined with nitrate (itself created by oxidizing ammonia) to generate ammonium nitrate which is the basis for both gunpowder and the most important fertilizer. When in 1946 there was no further need for massive amounts of explosives, the US Federal Government asked whether there might be any other use for these factories. The answer came back from the agricultural colleges: yes, we can use it to greatly increase agricultural yield, and this is what happened. This "industrialization of agriculture" freed food production from its former dependence upon manure, and, encouraged by the concurrent development of machinery, far fewer Americans were needed to grow our food. This increased the exodus to the growing number of urban industrial jobs, the increased use of oil, gas, and coal, and the massive generation of wealth. Over the course of the twentieth century, America continued to change from a relatively poor, largely agricultural, rural country into an increasingly industrialized and urban country while becoming vastly more wealthy, by most accounts, in the process. Meanwhile the energy required to do all this economic work was increasing exponentially (Fig. 7.3). New economic theories were launched to explain the enormous increase in wealth with, however, essentially no mention of the energy enabling and facilitating the expansion by those chronicling the process.

European and Japanese industry had been destroyed by the fighting. Every warring nation except the United States saw their industry and infrastructure in ruins, and the allies, especially Britain, were deeply in debt to the United States. The new peace was to be an American-dominated peace, with the terms dictated by Americans. The American-led Marshall Plan helped rebuild the war-devastated economies of Europe. After some 15 years of depression and war, the international monetary system was in need of serious rebuilding. The gold standard, which had served as the foundation of international trade since the mercantile days of the 1600s, was a casualty of the depression. In 1944 an International Monetary Conference was convened at a ski resort in New Hampshire called *Bretton Woods*. Under the auspices of the new system, known as the Bretton Woods Accords, the US dollar replaced gold as the basis for international trade and investment. Only the dollar was stated in terms of gold; the value of all other currencies was expressed in dollar terms. In essence, the rest of the world was willing to give the United States interest-free loans in their own currencies just to hold our dollars. The United States reaped several benefits from the new configuration on the world level. The value of US investments abroad grew at nearly 9% per year from 1948 to 1966. The terms of trade or the ratio of export prices to import prices grew by 24% over the same period. People of this country bought in a buyer's market (i.e., in conditions favorable to the buyer), and US corporations sold in a seller's market. Finally, US business gained access to crucial raw materials and additional cheap energy, despite the fact that the United States was the world's leading exporter of oil at the time. America's industrial might and monetary control formed an important foundation for growing affluence. America became extremely powerful both economically and in terms of energy use.

The depression era had witnessed a considerable amount of strife between labor and capital. By the late 1930s, strong industrial unions organized to win recognition, higher wages, and better working conditions. But the large spending of the war brought jobs, relative prosperity, and relative peace between capital and labor. After a flurry of strike activity immediately following the war, relations between large businesses and their employees stabilized. In 1948 an epoch-making contract was signed between General Motors and the United Auto Workers. In this contract the UAW gave up their claim to joint management of the company and control over the trajectory of technology. In return they received a larger share of the company's profits in the form of higher wages and benefits. The contract linked increases in wages to increases in productivity or output per worker. In this climate of American peace, labor stability and productivity (i.e., value added per hour of worker input) grew at a brisk pace, and the after-tax earnings of American manufacturing workers grew by more than 50% from 1948 to 1979. This was responsible for spreading some of the wealth earned by business to the pockets of the American worker. More than most factors, this new social contract, based on shared gains from increased productivity, helped establish the American dream. One specific example of the link between energy and economic prosperity rarely understood by most economists is that of the roll of energy in the dollar value of the products generated by a worker working for 1 hour. Increased labor productivity allowed the employer to pay his or her worker more even while making a larger profit. This increased productivity is normally assigned to technological progress. What is less understood is that labor productivity increased in direct proportion to the amount of energy used per worker hour (◘ Fig. 7.4). At that time labor productivity in the United States was two or three times that of a European worker, not because the worker worked harder or was more clever, as commonly assumed, but because he or she had big machines using two or three times more energy helping him do the job! Again what is often attributed exclusively to technology was in fact equally based on increasing the availability and the use of cheap energy, which was much cheaper in the United States than in most other nations.

7.7 The Increasing Role of Government

The idea that government participation in the economy should be minimal, which had been around at least since the time of the Physiocrats and Adam Smith, went by the wayside starting with the Great Depression and continuing into the postwar years. The strategy for ending the depression, the New Deal, created not only an alphabet soup of government agencies but also an attempt to involve the Federal government in economic planning. This planning was augmented and extended in the Second World War, undoubtedly the greatest public works program in the history of the United States. After the war Congress passed a law entitled The Employment Act of 1946. This law mandated the government to pursue taxing and spending policies that would result in reasonably full employment, stable prices, and economic growth. In this era of "New (Keynesian) Economics," budget deficits were sometimes purposefully created. They became an important tool of economic policy rather than a dangerous aberration that must be avoided at all costs. The increased spending, which was often financed by debt rather than taxes, injected increased purchasing power into the economy to help maintain postwar affluence. Government created new programs to subsidize home mortgages and home ownership, an important component of the expanding realization of the American dream. Spending on social programs also increased. In 1968 a state-supported health initiative for the elderly called *Medicare* was passed into law to supplement the retirement insurance program (Social Security) created during the Great Depression. For the first time, being old no longer meant being poor for the majority of American workers. This act represented the culmination of a whole series of social spending programs during the 1960s. Spending for income maintenance programs and education increased during the presidency of Lyndon

Johnson, who envisioned a "Great Society." But spending also rose for military purposes, as the United States became more deeply involved in a prolonged war in Vietnam. While this expansion of spending eventually helped to initiate the end of the American dream, more than two decades of prosperity and increasing affluence for a growing number of Americans ensued. The United States was affluent enough to spend more on health care and education and create more opportunities for those formerly left out of the general economic expansion. General affluence increased even while waging war—at least initially. Wages and profits continue rising, at least for a large proportion of the population.

The engine that held this increasing prosperity together was economic growth, that is, the increase in the material economy expressed in the dollar value of the goods and services we produced in a year (this is called gross domestic product or GDP). The fuels for that were a social structure that prompted growth, expanding international markets and the exponentially increasing use of oil and coal and gas, as all through this period energy use increased in almost direct proportion to the economy—the fossil energy was there to do the actual work of an expanding economy. GDP more than doubled from 1945 to 1973, increasing from about $1.8 trillion to over $4.3 trillion in inflation-corrected (e.g., year 2000) dollars. Energy was readily available and very cheap, and the incentives to use it abounded as "the good life" was increasingly sold using advertising.

7.8 The "Oil Crises" of the 1970s: Hints at Limits to Economic Growth

As the 1970s approached, all four pillars of the American success story began to fracture. Europe and Japan caught up and surpassed the United States in terms of technology and economic growth. New technologies, a more restrictive regulatory climate, and a new type of mergers (international conglomerates) destabilized the tight oligopoly control of manufacturing. This would further destabilize corporate structure in the 1980s and 1990s. The rise of state-owned oil companies in the Middle East and elsewhere presented another threat to the control of intercapitalist rivalry. Bretton Woods was abandoned, and US oil production peaked in 1970. In 1973 the United States experienced the first of several "oil supply shocks" that seemed, for the first time, to inject a harsh note of vulnerability into the united chorus of the American dream for all. Before the 1970s nearly all segments of American society—including labor, capital, government, and civil rights groups—were united behind the agenda of continuous economic growth. The idea that growth could be limited by resource or environmental constraints or, more specifically, that we could run short of energy-providing fossil fuels was simply not part of the understanding or dialog of most of this country's citizens. But this was to change during the 1970s.

In the popular phrase of economists, the economy began to "overheat." Consumer spending had more than doubled from $1.1 trillion in 1945 to nearly $2.5 trillion in 1970 (in 2000 dollars) as workers spent the dividends from the social contract from 25 years earlier on the many goods they had been deprived of in the depression and the war and as general affluence increased. As the US economy retooled in the postwar era, investment spending likewise rose from about $230 billion to $427 billion in 2000 dollars, aided by steadily increasing numbers of people, consumer credit, and corporate profits. Government spending, driven by the expansion of social programs during the time of President John Kennedy's "New Frontier" and President Lyndon Johnson's "Great Society," while the costs of fighting the Vietnam War, in constant 2005 dollars, increased from $405 billion in 1950 to more than $1 trillion during the same 20-year period. Unemployment fell at a relatively steady pace, dropping from about 6.5% of the labor force in 1958 to only 4% in 1969. Hourly earnings of manufacturing workers after taxes rose from about $2.75 per hour in 1948 to about $4.50 in 1970 when both were expressed in 1977 dollars. As spending increased faster than the ability to produce goods (given the relatively modest levels of unemployment), prices began to rise. The specter of "creeping inflation" began to enter the lexicon of economists and citizens alike.

In 1973, the United States (and much of the world) experienced the first "energy crisis." Crude oil, selling for $2.90 per barrel in September (a price that had been nearly constant for decades), soared to $11.65 by December. The price of gasoline shot up suddenly from 30 to 65 cents a gallon in a few weeks, while the available supplies declined. Americans became subject to gasoline lines, large increases in the prices of other energy sources, and double-digit inflation. Home heating oil became much more expensive, as did electricity, food, and even coal! Few people understood that the production of oil in the United States had reached a peak in 1970 and had begun to decline. While the specific initiation of the price increase began with a bulldozer that in 1970 ruptured a pipeline carrying oil from the Persian Gulf to the Mediterranean, the peak of oil production in the United States, the United States resupplying the Israeli military in their war with Egypt, the long history of Western arrogance in the Middle East, and the exponential increase in the use of oil set up the circumstances in which a minor event could generate an enormous impact. In 1979 the world experienced another oil shock. According to the Energy Information Agency, the current dollar price of domestic crude oil rose from $14.95 in 1978 to $34 per barrel in 1980. This would amount to nearly $90 per barrel in 2017 prices. Consequently, the 1980 price of gasoline increased again to an average of $1.36 per gallon, equal to $3.70 in 2017 prices. The increases were directly in response to the withdrawal of supply by the new Islamic Republic in Iran, after the collapse of the US-backed government of Reza Pahlavi, but again the inability of the United States to supply its own consumption underlay all. Many of the economic ills of 1974, such as the highest rates of unemployment since the Great Depression, and rising prices were repeated in the late 1970s and early 1980s when oil once again became less available and more expensive due to restrictions in supply brought about by the Organization of Petroleum Exporting Countries (OPEC—including many oil-rich countries in the Persian Gulf, Venezuela, and Indonesia).

Americans became used to energy as a topic that was in the newspaper every day, and especially in the colder Northern Tier of the United States, conversation was often about wood as a fuel to heat one's house or the fuel efficiency of the then-new Japanese imported cars vs. the familiar Fords and Chevrolets. The American economy, used to being overwhelmingly the strongest in the world, suffered as businesses in the countries American aid helped restore after the Second World War now became effective competitors. This was partly due to energy prices, which were once much cheaper in the United States, became effectively the same around the world, with the result that higher-priced American labor was no longer compensated for by cheaper American energy. On the contrary, real wages began to fall in the United States. By the end of the 1970s, Japanese autoworkers were earning more per hour than their American counterparts. The unemployment rate increased to nearly 10% in 1982, a number unheard of since the Great Depression of the 1930s, while prices of everything increased at nearly 10% per year. But unemployment and inflation were supposed to be inverse to each other according to the economist's well-established Phillips curve! Here they were simultaneously increasing, something called *stagflation*. Labor productivity ceased to increase, also something formerly unheard of (Fig. 7.4). The news was so bad that the Reagan administration stopped gathering data on this important economic parameter. For many, it seemed like the world was falling apart.

Stagflation, which was difficult to explain by means of standard Keynesian theory, is easy to explain from an energy perspective: as energy prices increased and supplies declined the dollars circulating in the US economy were increasing more rapidly than new energy was added to do economic work. As a result each dollar bought fewer goods and services, which was perceived as inflation. In addition the relatively monopolized corporate structure allowed business to pass on increased costs of production in the form of higher prices. As more of society's output was required to get the energy necessary to run the economy, costs of everything from food to packaging were pressured upward. This resulted in an increase in joblessness as there was less money available for purchases. In fact adding the energy and historical perspective provides a ready explanation for stagflation: as energy use was increasingly restricted (by supply and higher prices), the economy contracted. And, as we said, since the energy supply contracted more than the dollar supply, there was also inflation. This explanation shows the power of energy analysis and the inadequacy of pure economic models that exclude the fundamental role of energy. In systems language, the economic models focused almost entirely on the internal dynamics of the system but were insensitive to changes in forcing functions because they had not been included in the model structure.

7.9 The Limits to Growth

As we developed in ► Chap. 4 at about this time, a series of quite pessimistic reports about the future came out, most importantly the "Club of Rome's" *Limits to Growth* [13], *The Population Bomb* by Paul Ehrlich [14], and analyses of the likely future of oil production by Hubbert [15]. These reports implied in various ways that the human population appeared to be becoming very large relative to the resource base needed to support them—especially at a relatively high level of affluence—and that it appeared that some rather severe "crashes" of populations and civilizations might be in store.

7.10 Crumbling Pillars of Prosperity

In retrospect, we can now say that the pillars of postwar prosperity began to erode in the 1970s and early 1980s and that changes in the social sphere also began to complicate and add to the biophysical changes derived from the decline in the availability of cheap oil. Even though the oil market had stabilized and cheap energy returned to the United States in the late 1980s, the changes in the structure of the economy were long lasting. The economy ceased growing exponentially, although it continued to grow linearly but at a decreasing rate, from 4.4% per year in the 1960s to 3.3, 3.0, 3.2 to 2.4 to about 1% in the following decades. Many formerly "American" companies became international and moved production facilities overseas where labor was cheaper, and oil, no longer cheaper in the United States compared to elsewhere, was the same price, although cheap enough to pay for the additional transport required. The decrease in labor costs when production facilities were moved to other countries outweighed the costs and the process of globalization accelerated. Productivity growth (formerly strongly linked to increasing energy used per worker hour) in manufacturing industries began to slow, falling from 3.3% per year in the 1966–1973 period to 1.5% from 1973 to 1979 to essentially zero in the early 1980s. Reductions in the rate of growth in the energy-intensive sectors of utilities and transportation were even greater, while construction and mining showed actual declines in output per worker hour. As productivity growth slowed so did the growth in workers' hourly income, from a substantial 2.2% per year from 1948 to 1966 (which would lead to a doubling of incomes in 32 years) to 1.5% in 1973 to 0.1% in 1979. Corporate profits also decreased from nearly 10% in the mid-1960s to a little more than 4% by 1974. Things seemed bad for both capital and labor [16, 17].

Mainstream economists seemed at a loss to explain this phenomenon. Their statistical models, which relied on the amount of equipment per worker, education levels, and workforce experience, left more factors unexplained than explained. Even the profession's productivity guru, Edward Denison, had to admit that the 17 best models explained only a fraction of the problem. Fortunately two other approaches yielded far better explanations. Economists associated with the "Social Structure of Accumulation" (SSA) approach (Bowles et al. [18]) developed a statistical model that explained 89% of the decline and attributed most (84%) of the slowdown in productivity growth to decreases in work intensity. Under the social contract of the postwar era, unions were able to limit speedup by a series of work rules that limited how hard workers could be driven. Despite increases in the numbers of supervisors, businesses (especially manufacturing firms) could not increase the amount of output per worker at will, especially without increasing wages. The biophysical approach also yielded promising results. Howard Odum had been writing about the importance of energy in the economy for a decade, as had others [19, 20]. In a 1984 article in the prestigious journal *Science*, Cutler Cleveland, Charles Hall, Robert Costanza, and Robert Kaufmann [21] found that they could explain 98% of the decline in output growth by the decline in fuel energy after the oil crises of the 1970s. They also explained many basic attributes of economics in energy terms, an approach that introduced the concept of biophysical economics. The two concepts (biophysical economics and the social structure of accumulation) are linked because the increase in fuel-intensive machinery is one factor in how intensive work can be made [22, 23].

Things looked increasingly difficult for the United States in the international arena as well. The United States had rebuilt Europe and Japan with the latest technology soon after Second World War, and by the 1970s these former "second rate trade partners" turned into fierce competitors. The commitment to energy efficiency in

Europe and Japan far surpassed that of the United States, and they had much newer capital equipment because of needing to rebuild after the war. Moreover, labor relations in the countries were far less contentious than they were at home. Terms of trade or the ratio of export prices to import prices fell from about 1.35 in the early 1960s to only 1.15 by 1979. Adding to the difficulties faced by the United States, the world monetary system came unglued by the early 1970s. The system, developed in Bretton Woods, New Hampshire, depended upon the United States being the world's most productive economy and upon its willingness to let other countries redeem their dollar holdings in gold. However, when declines in productivity and the terms of trade, and the mounting costs of the Vietnam War, came home to roost the value of the dollar relative to other currencies plummeted. President Richard Nixon suspended the convertibility of dollars to gold. The international trade system was now a free-for-all, and the new and more chaotic system contributed to a fall in corporate profits. Something had to give, but Presidents Nixon, Ford, and Carter were unable to break the political stalemate of rising labor costs caused by union power and a commitment to low rates of unemployment in spite of their best efforts. Something had to give.

In 1979 the editors of *Business Week* opined that to restore the nation's affluence, labor would have to learn to accept less. The *Wall Street Journal* was calling for "supply-side economics," an approach associated with increasing the rate of exploitation of natural resources by decreasing government environmental and other regulations. In the same year, on the steps of the Statehouse in Concord, New Hampshire former actor and California Governor Ronald Reagan, then a presidential candidate, declared that "for the country to get richer, the rich have to get richer." Reagan won the 1980 presidential election and instituted what the Social Structure of Accumulationists termed "a program for business ascendancy" or what the *Wall Street Journal* praised as "supply-side economics." This constituted a sharp turn to the right in American politics. The Reagan administration focused far more on inflation than on the restriction in growth, immediately confronted unions, and further disciplined workers by moving to create a sharp recession by means of policies that raised interest rates and hence severely restricted the amount of money in the economy and, consequently, jobs. By the mid-1980s, home mortgages carried 20% interest rates, and business loans were nearly as expensive. In order to increase America's power in the world, they instituted an aggressive program of military buildup and returned to what former President Theodore Roosevelt termed "big stick diplomacy." Inflation rates subsided and corporate profits rose, but these victories came at a cost. Unemployment rose to almost 10%, inequality increased as the percent of Americans living in poverty jumped from about 11% to a little more than 13%, while the number of rich households (who earned more than nine times the poverty level) went from less than 4% in 1979 to nearly 7% in 1989. Compared to earlier times, most Americans thought that the economy was a mess. Few blamed it on energy, but in retrospect we can say that the pillars of postwar prosperity were eroded in the 1970s and early 1980s because there was no longer unlimited supplies of cheap energy, which caused changes in the economic and social sphere that had begun to impact prosperity.

7.11 The 20-Year Energy Breather

By the mid-1980s, the price of gasoline had dropped again as the inflation-adjusted (2010) price of crude oil fell from $98.52 per barrel in 1980 to $15.84 in 1998. The new Prudhoe Bay field in Alaska, the largest ever found in America, added to our oil production and helped mitigate, to some degree, the decrease in production of other domestic oil. Around the world many earlier discoveries had become worth developing in the 1970s, and cheap foreign oil flooded the market. As a result, energy as a topic faded away from the media and so in the perception of most people. For most people who thought about it at all, the reason that the energy crisis was "solved" was that the market was allowed to operate by generating incentives from the higher prices. In fact this was largely true, for although domestic production continued to fall year by year (Fig. 7.5), foreign-derived oil was increasingly imported to the United States from other countries, and we shifted the production of electricity away from oil to coal (a generally dirtier but more abundant form of energy) to natural gas (generally a cleaner form) and to nuclear energy. So it indeed did look

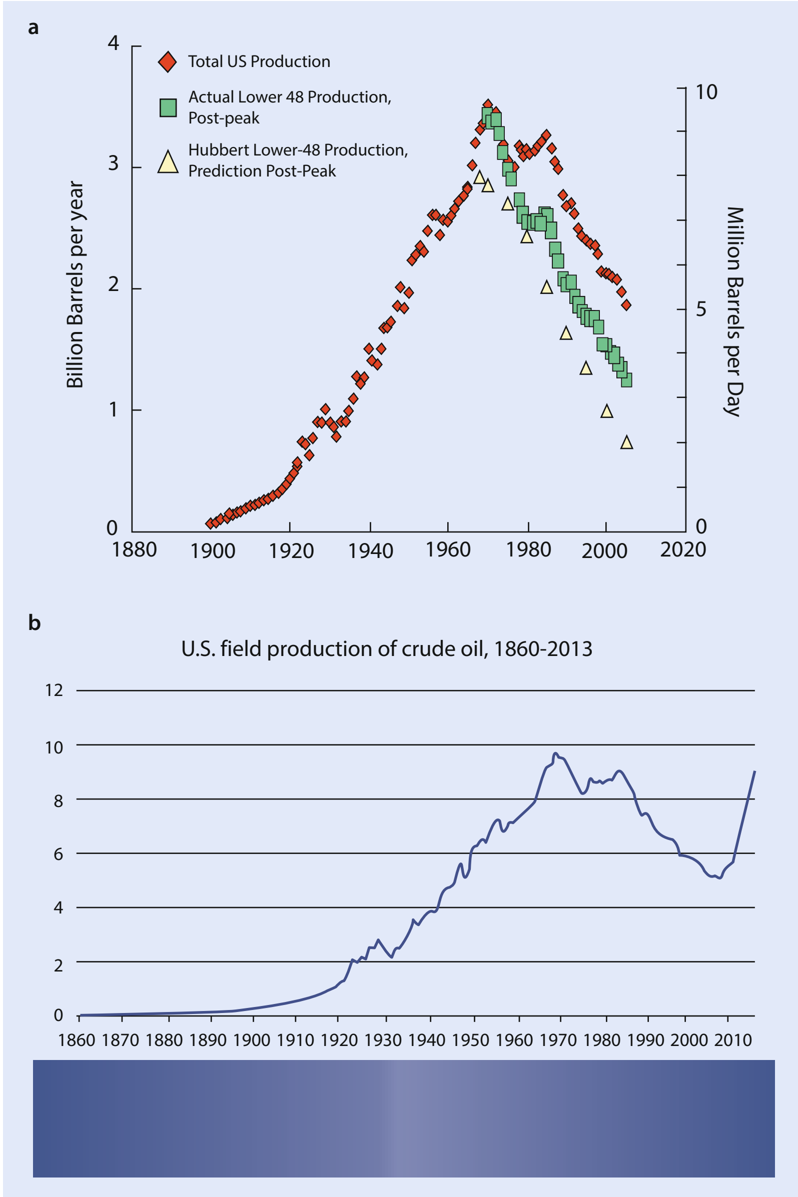


Fig. 7.5 **a** Production of conventional oil in the United States (with and without Alaska) compared to Hubbert's 1969 prediction for the lower 48 (Source: 2006 Cambridge Energy Research Associates) **b** Oil production extended to 2017 (in millions of barrels per day) and including unconventional oil. Even with the increase in domestic production the United States as of 2017 imports nearly half the oil it uses

like the economy, through price signals and substitutions, had in fact responded to the "invisible hand" of market forces. Conservative economists felt vindicated, and the resource pessimists beat a retreat, although the economic stagnation of the 1970s, as indicated by declining rates of GDP growth, continued until the present day in the world's mature economies.

By the early 1990s, inflation had subsided and the world economy grew at about 3% a year. Inflation-corrected gasoline prices, the most important barometer of energy scarcity for most people, stabilized and even decreased substantially from $3.41 per gallon in March 1980 to $1.25 in December 1998, in response to an influx of the foreign oil (Fig. 7.6). Much of this new wealth was generated not through working for wages but by owning stocks. Wages fell while assets surged, but, as in earlier times in history, stock ownership was not spread evenly throughout the economy. The majority of stock market gains accrued to the top 1% of the income distribution. Increasingly many landscapes were filled with very large houses that were far larger than

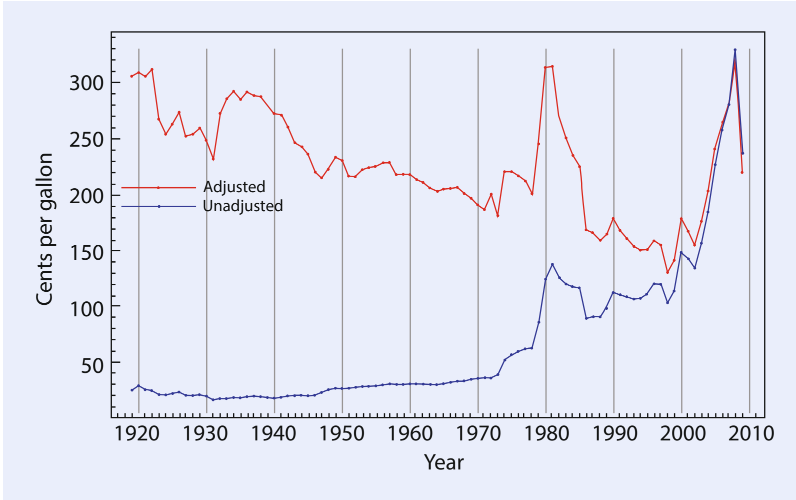


Fig. 7.6 Gasoline price corrected and not corrected for inflation (2005 dollars) in the United States (Source: USDOE)

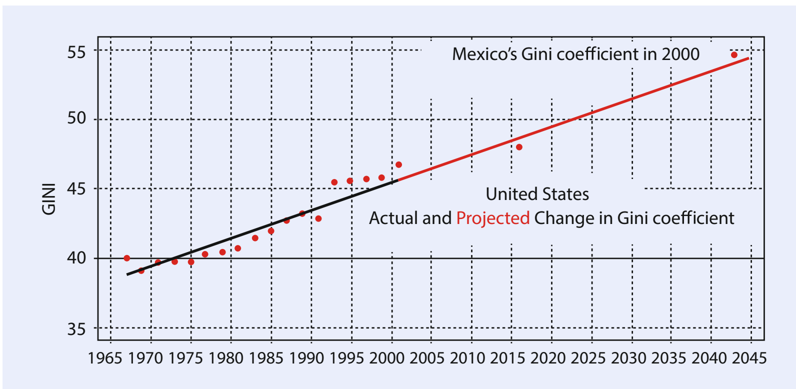


Fig. 7.7 Gini coefficient for the United States, which is the ratio of the income earned for by the top 20% compared to the income earned by the bottom 20%. This graph shows that since about 1970, there has been increasing inequality of wealth in the Unites States, with the wealthiest 20% gaining an increasingly large proportion of the economic pie, while the poorest 20% get a smaller and smaller portion (Source: ► SustainableMiddleClass.com)

the basic needs of a family and purchased primarily as luxury items, for the perceived status or on speculation—that is, to sell for a higher price a few years in the future. This process was driven by market forces, as housing represents both investment and shelter for most Americans. One's house is generally a person's greatest asset or repository of wealth. However, large houses, especially those filled with myriad electronic appliances, are also extravagant energy users. So declining real energy prices combined with market forces produced a growing stock of larger houses that used more energy even though many appliances had become much more efficient. Discussions of energy or resource scarcity largely disappeared from public discourse or were displaced by new concerns and courses about environmental impacts on tropical forests and biodiversity. Income inequality between the rich and poor, as measured by the Gini index, increased greatly both absolutely and in comparison with other industrialized nations (Fig. 7.7; Table 7.3). Indeed it seemed that some 100 years after the first "Gilded Age," America had entered a new one.

In the United States conservatives led by President Ronald Reagan were successful in convincing many formerly apolitical or even labor union people that their own personal conservatism in issues such as family, society, religion, and gun ownership could be best met through conservative economic and political groups whose agendas were historically opposed to the interests of the working

Table 7.3 Recent Gini indexes for a select group of nations. The lower the number, the more equitable the distribution of wealth

Japan	24.9
Sweden	25.0
Germany	28.3
France	32.7
Pakistan	33.0
Canada	33.1
Switzerland	33.1
United Kingdom	36.0
Iran	43.0
United States	**46.6**
Argentina	52.2
Mexico	54.6
South Africa	57.8
Namibia	70.7

Source: Sustainable ► middleclass.com

people. These groups and their representatives in government were very much opposed to the government in general and any interference with individual "freedom," especially intervention in the market. Thus, they opposed, for example, government programs to generate energy alternatives (such as solar power or synthetic substitutes for oil), believing that market forces were superior for guiding investments into energy and everything else. They also tended to be opposed to restrictions on economic activity based on environmental considerations and even mounted campaigns to discredit scientific investigation into environmental issues such as global warming. (However, it is important to point out that many conservative people are extremely interested in conservation of nature.) One specific thing that President Regan did was to remove the solar collectors installed by President Carter on the roof of the White House even though they were working fine.

These new conservative forces tended to be opposed to government policies that restricted such freedoms (i.e., gas mileage standards and speed limits). Both liberals and conservatives tended to support free trade and hence contributed to the movement of many American companies or their production facilities overseas where labor was cheaper and pollution standards often less strict. One effect was probably a substantial contribution to the improved efficiency of the economy (GDP per unit of energy used) as polluting and expensive heavy industries were moved overseas. For example, strong federal programs to improve solar collectors and the like were often eliminated as government interferences. By 2000 the country seemingly had recovered from the stagnant 1970s and the recessions of the 1980s and early 1990s, although the prosperity was based on a growing level of debt, just as it was in the mid-1980s. Stock values began to increase steadily, and the general economic well-being of many Americans led to a general sense of satisfaction in market mechanisms. The collapse of the Soviet Union and the end of its influence in Eastern Europe effectively brought the Cold War to an end, and the free market approach to economics came to dominate the economics profession. The ideas of John Maynard Keynes, emphasizing government intervention and considered to be orthodox in the golden age of postwar prosperity, fell into disrepute in many of the nation's leading graduate schools. The apparent success of England in the late 1980s under conservative Margaret Thatcher led to additional impetus that the conservative free market approach to economics worked. The presidential administrations of George Herbert Walker Bush and Bill Clinton alike pressed a free trade agenda and reduced spending on social programs. As markets became "liberalized," prices of basic commodities from coffee to cotton to oil declined by more than 100%. The terms of trade greatly improved for the United States, but poverty rates and debt soared in Africa and other developing regions where coffee growers, for example, had to compete with each other for the limited markets in the rich countries.

Our energy perspective has a different view, of course. First of all, much of the economic expansion of Presidents Reagan and George H. W. Bush was paid for with debt, so that the administrations of these supposedly fiscally conservative presidents (and Congresses at the time) actually generated far more debt, even when corrected for inflation and increased GDP, than even the supposedly "free spending liberal" Franklin Roosevelt did for domestic programs in earlier times. It is important to understand that while the United States and Great Britain, for example, appeared to be doing much better economically under "conservative" administrations, both countries happened to enjoy low prices for oil and for energy in general while there were conservative leaders. (Less conservative Bill Clinton and Tony Blair benefited too.) In the United States, there was a decline in the proportion of GDP needed to pay for energy, from a maximum of 14% in 1981 to about 6% in 2000. This effectively gave the US people some 6–8% additional discretionary income (that not required for basic food, shelter, and clothing), which could be spent on big houses and stocks. In addition, declining oil prices, an input to most basic commodities, reduced general inflation. In England Margaret Thatcher received a great deal of credit for her nation's economic recovery, but few attributed her success to the simple fact that the vast North Sea oil field came on line during her administration, greatly reducing former costs for imports. Since a large part of that oil was sold abroad, very large revenues for the government were generated, allowing the reduction of other taxes. Clearly conservatism alone could not explain fully England's success, as nominally socialist Netherlands was also doing very well economically at that time fueled by the vast Groningen gas field, whose profits allowed for social benefits to be extended to everyone. Energy analyst Doug Reynolds [24] generates a strong case that the collapse of the Soviet Union, often attributed to strong actions by the United States, was actually mostly a consequence of the partial collapse of Soviet oil production over the previous 3 years, greatly reducing the revenues that went to the central government and leading to many problems such as the inability to pay military pensions. So, again, these historical data about energy help explain what is normally attributed solely to political or economic leadership. A more difficult question, of course, is how to govern well when the abundant resource "rug" is pulled out from under the economy, a question of great importance as we write this book.

7.12 Political and Economic Response to Oil Price Increases Since 2000

A rather comfortable economic situation in America became subject to some disquiet as oil prices once again increased in the early 2000s and the overpriced stock market fell by nearly 20% as the "technology bubble" burst. Those who particularly benefited from the extensive surplus wealth of the 1990s often shifted their money into the housing market, as real estate was perceived to be a safer investment than technology. Government programs initiated by the Clinton Administration, and encouraged through the Bush campaign, were designed to put more people into their own homes for political and social reasons. Oil prices relaxed a bit through about 2006 but then increased rapidly in 2007 and enormously more in the first half of 2008. For those who read widely, there was a new set of economic predictions emanating from various oil industry analysts. Followers of M. King Hubbert, including Colin Campbell and Jean Laherrere [25], warned that the "peak" in oil production was soon upon us and that the end of cheap oil would almost certainly follow-and with it significant economic consequences. The new Bush Administration, apparently with its own inside information on declining oil production prospects, called for the drilling of oil in the Alaskan Wildlife Preserve and enhanced oil and gas development. Something that was barely noticed was that global oil production stopped growing in 2004. Colin Campbell had predicted at the Association for the Study of Peak Oil meeting in Lisbon that we were likely to see an *undulating plateau*, rather than a steep peak, for global oil. He reasoned that initial shortfalls in oil would lead to price increases, which would lead to economic recession, a reduction in demand, economic recovery, and a new cycle. This basic pattern seems to have been exactly what has happened from 2004 to at least mid-2017.

The stock market continued to be sensitive to oil price changes, and the value of the Dow Jones kept struggling to increase beyond its inflation-adjusted peak in 1998 when corrected for overall price increases (Fig. 7.8). While in some senses (high employment and increasing wealth of the more affluent) the economy was doing quite well in the first 7 years of the 2000s, many questioned how much of the apparent affluence was real and how much was based on debt, as both real estate speculation and debt soared. From 1997 to 2005, the financial sector debt grew from 66 to more than 100% of GDP. Household debt rose accordingly, from 67% to 92% of GDP. Many private and public pension systems were based on the assumption that stocks would continue to grow at historical rates of 8% or

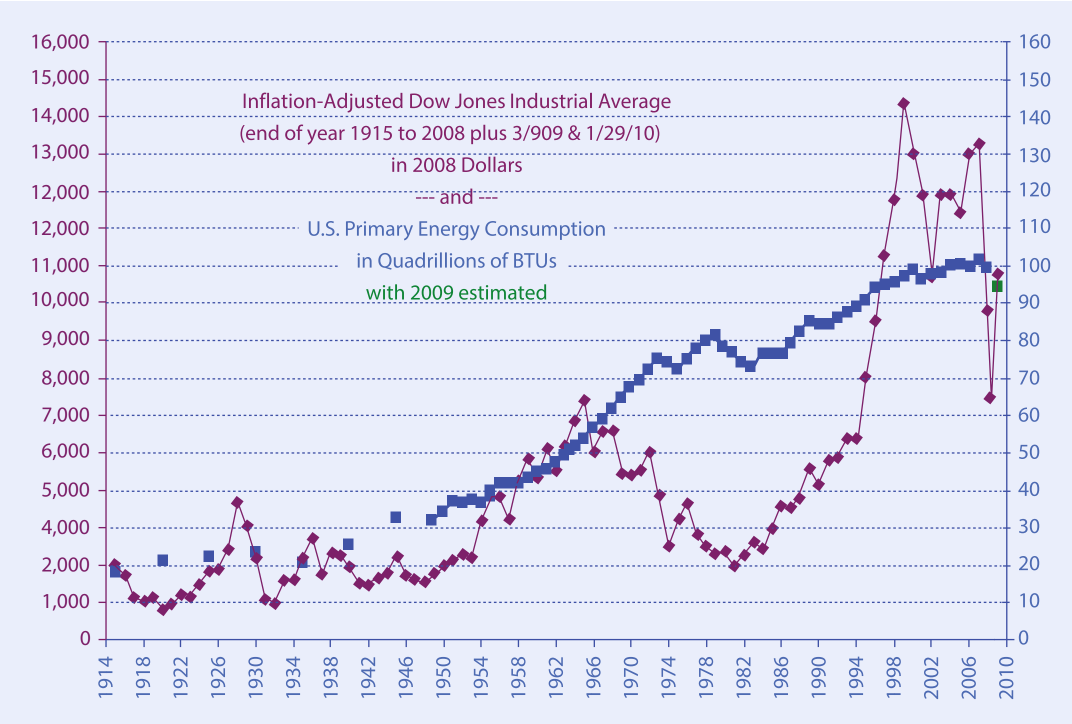


Fig. 7.8 Inflation-corrected Dow Jones industrial average (2008 dollars) scaled to plot in the same graph space as total energy used by the US economy. Over long periods of time, the slopes are very similar, but the Dow Jones snakes around the total US energy use, with the deviations from that line presumably reflecting psychological aspects (Figure courtesy of William Tamblyn)

more, as had been seen in the "good times" (but speculative) period during the late 1990s. When the stock bubble disappeared in 2000, many large companies were found to have placed not nearly enough money into their pension funds. Many workers who had worked hard all their lives with the expectation that they would have a good, solid pension found they had little or nothing. Some were fortunate in having the Federal government bail them out, but there is not enough money in that fund to cover even a fraction of the people who will have lost their pension. Public entities, which are required by law to meet their pension obligations, fell about 500 billion dollars in the hole. All forms of debt, including that of the Federal government, increased faster than did the economy as a whole, as measured by the growth of gross domestic product. The Federal Government took in about $55 billion dollars more in taxes than it spent in the last year of President Clinton's administration. By 2003, powered by tax cuts at the top end of the income distribution and increased military spending, the debt soared to an annual deficit of more than $500 billion by 2006. The Federal government, attempting to avoid inflation, did not "print" more money but became increasingly dependent on loans from Asia, especially China, to pay its bills. These loans, and those of the 1980s, will be a tremendous financial burden on young people who are reading this as undergraduate or graduate students, yet our government is unwilling to raise taxes or reduce total spending. As of 2006 wars in Iraq and Afghanistan cost more than $8 billion *per month*! The situation only got worse following the financial meltdown of 2008, and Presidents Bush, Obama and now Trump are setting new deficit records (Fig. 7.9). Various health-care initiatives imply enormous future federal spending requirements. In addition, individuals had been living far beyond their means by borrowing heavily on credit cards. There is another unseen debt as well, that of delayed maintenance of society's infrastructure such as bridges, roads, levees, schools, and so on, not to mention degradation of the natural infrastructure of clean water, soil, and biodiversity.

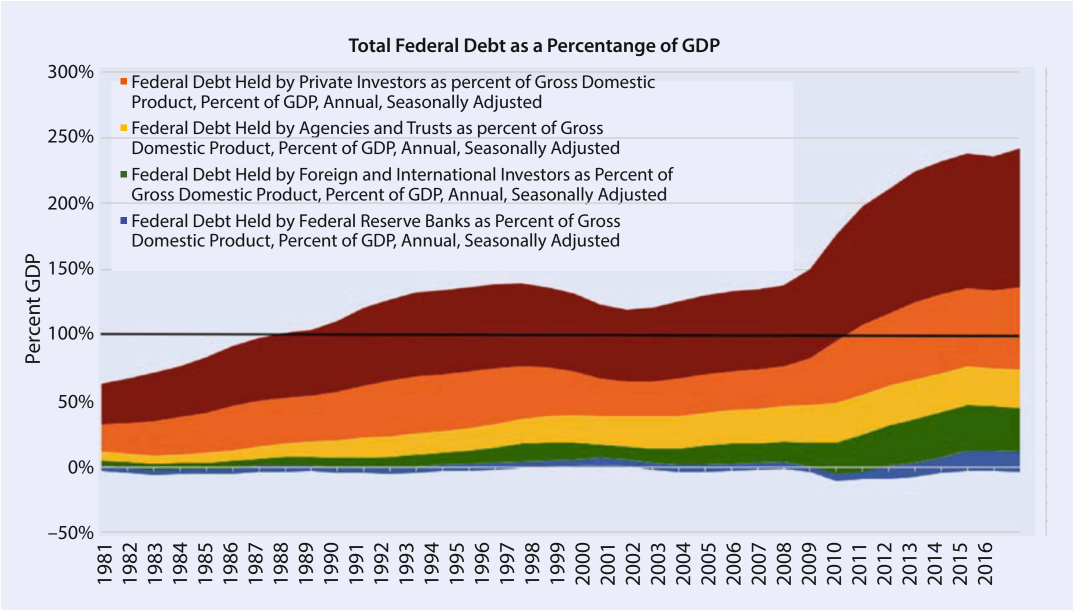


Fig. 7.9 Cumulative debts for the United States in trillions of dollars (Source: Wikimedia Commons)

What does this debt mean in energy terms? In 2008 the United States owed various entities such as banks and pension funds in Japan and China and other countries some 8000 billion dollars. If we were to pay off this debt all at once and those who received these dollars (say retired Japanese Toyota workers) chose to spend it on beef, fish, rice, or Ford automobiles from America, it would take at the average energy use per unit of economic activity of the US economy (about 8 MJ/dollar then) an estimated 64 exajoules worth of energy to make those goods, equal to 10 billion barrels of oil or half of US known oil reserves remaining then to make the products those foreign people purchased. Meeting the interest payments transfers income to the holders of debt and has nearly the same effect. In other words with our massive foreign debts, people in other countries have a huge lien (i.e., obligation to repay a debt) on our remaining energy reserves or whatever their replacements might be. If this debt becomes too burdensome, one way out of this problem is hyperinflation. After the Treaty of Versailles in 1918, Germany was obligated to pay some $30 billion in "reparations" to France and England. They paid the international debt in "hard currency," mostly borrowed from US banks, while they paid off their own domestic debts in deflated marks. The impact was to enormously devalue their currency (Fig. 7.10). Prices rose by nearly 21% per day and doubled every 3.7 days. In 1918 it took 1 mark to purchase about 0.4 gram of gold. By November 1923 it took 100 billion Reichsmarks to buy the same quantity. This greatly undermined the entire financial system and helped lead to the rise of the Nazi party. Moreover, when the crises of the early 1930s crippled the US banking system, American banks were either unwilling or unable to continue the loans to Germany, the Germans defaulted on their reparations, and Britain and France suspended the repayment of their war debts to the United States. The ensuing collapse of international trade and the gold standard was a primary factor in the depth and length of the Great Depression. The East African nation of Zimbabwe recently experienced just such hyperinflation with an inflation rate so that prices doubled each day. Thus, while the United States continues to have enormous wealth and potential for creating wealth, it may be increasingly constrained by the new a shortage of cheap energy. Our debt to other nations and to nature also makes our financial future potentially precarious. We need a careful systems analysis using both conventional and biophysical accounting to determine what is real wealth production and what is not, whether there is, as in the past, much potential for future growth to pay off this debt, and what might be the effects of future increases in energy costs.

Fig. 7.10 Elderly Germans buying a loaf of bread with a very large number of hard-earned marks, which had completely lost their value

The financial "crises" that occurred in the second half of 2008 add another dimension to our analyses. Many financial firms, highly respected for decades, collapsed or were accused of excessively risky and even quite shady financial undertakings. The government was asked to bail out all kinds of financial entities, and many people lost from one-third to one-half of their savings as housing and Wall Street prices collapsed. As of this writing, it is far too early to tell whether the stock market expansion of 2017 is excessive speculation and "irrational exuberance" or a genuine new direction for Wall Street. We suspect that if Wall Street is to grow again in the future sustainably beyond where it was in 2007, huge new energy supplies or an unprecedented and unlikely increase in efficiency will need to be found. Barring that, many Americans will have to readjust greatly and permanently their perspective on wealth production through the stock market and probably in regard to economic growth in general. That this transition would be difficult, financially, intellectually, and emotionally, for many, is an understatement.

We leave an examination of whether the vast increase in oil prices in the first half of 2008 was directly responsible for the economic meltdown of the second half of 2008 for a more comprehensive analysis in ▶ Chap. 18. In the meantime note that the total increase in energy use in the United States began to flatten out considerably starting in about 2000 (Fig. 7.8). Thus, if the production of real wealth is as dependent upon the use of energy, as we believe, then we have left a long period of increasing energy and wealth and entered a period where it may no longer be possible to produce much more, or perhaps even as much, of either.

As we prepare the second edition in 2017 the United States economy is doing well or poorly—depending on who you are. A radically different president and suite of advisors promise to "make America great again" without quite telling us what that means. He also promises to bring back economic growth of 4% a year. Environmental regulations and protections are relaxed wherever one looks which is supposed to assist economic growth but is likely to have adverse impacts on many people. Oil and especially natural gas prices are low which according to our analysis should help economic growth—but so far the results are lukewarm at best. Many petroleum companies are not making profits and investments are decreasing, which is likely to constrain future production. Wealth is increasingly concentrated in the hands of the most wealthy, so that about half the nation's wealth is owned by 1% of the people. The president says we are becoming a large energy exporter, but we still import half our oil. Renewable energy is growing but still represents only a few percent of our energy supply. The increasing use of robots is likely to have adverse effects on the need for labor. Some sectors of the American economy have recovered, more or less, from the 2008 crash, but others have not. The stock market valuation is high but potentially unstable, and employment is higher. But wages have not recovered and middle class incomes have stagnated at best. Oil prices are moderate at this writing so we might expect some growth from that perspective. This issue is considered further in ► Chaps. 10 and 13. There are still no economists of national influence attempting to understand the relation of the economy to resources.

7.13 Why Does the Energy Issue Keep Emerging?

Given the very large jumps in the price of crude oil and gasoline, both down and up, that have occurred in the first 17 years of the new millennium, many people have started to think about energy again. Why do "oil crises" keep reoccurring? Despite conservative claims that market processes and technology make considerations of any "limits to growth" and physical restrictions on energy resource supplies obsolete, world shortages and price fluctuations continue. Why have middle class incomes been failing so frequently to increase in value as it had in the past?

In the long term, markets and technologies have been a means of enabling humans to increase their wealth and material well-being. But wealth does not come out of thin air but only from the use of energy and the exploitation of physical resources. Thus, an associated and necessary aspect of this increase in wealth is that the same factors, markets and technologies that have enabled and encouraged us to become wealthy have also enabled and encouraged us to run through the world's resources more rapidly. It is quite possible that we are beginning to reach the limits of the Earth's ability to provide cheaply and easily the resources we have taken for granted. The periodic oil price increases are small reminders that, eventually, the piper must be paid, for as we like to say, Mother Nature holds the high cards. Humans are indeed industrious and ingenious, but that industriousness and ingenuity still require the Earth itself to provide the raw materials and fuels that are the basis for most wealth production and the capacity to absorb our wastes. Humans appear to be increasing energy and economic costs to the economy through indirect impacts of industrialization, for example, increased damage from hurricanes from a warmer ocean, increases in sea level, possible crop production declines, tropical soil drying, increased rates of flooding and tornadoes, and so on. The *Stern Report* [26] says that the price for mitigating future environmental impacts from global warming might be 20 times more than the cost of acting now to reduce our impact on the planet. This is one of many reasons why we must include more natural science in our economics, which we do throughout this book.

7.14 Debt, Inequality, and Who Gets What

Whatever the future of the total production of wealth in the United States, there are several clear and unsettling trends that will affect the future of energy supplies. The first is the enormous increase in debt in recent years (Fig. 7.9). Thus, much of the apparent prosperity of the recent past was based on debt and the ability to run an economy when debt expands more rapidly than income and wealth is highly suspect. The limits to debt constitute a limit to growth. The US generated huge debts (at the time) as the administration of Franklin Roosevelt, especially during World War II, spent far more money than it took in. Since then the debt economy has escalated even further. The standard answer of mainstream economists is that economic growth allows us to carry the debt without peril to the rest of the economy. The crucial question for the world economy is whether there is the energy available today to facilitate or even allow the growth that might make the debts of today and the future fundable? If the traditional internal limits to growth, such as demand and productivity, coincide with the biophysical limits, every economic problem will be rendered more difficult. The era of peak oil is likely to be the era of degrowth.

The long history of the United States is based upon the strength of the middle class. Since about the 1960s, however, capital and wealth more generally have been concentrated increasingly in the hands of the wealthy (Fig. 7.7). The postwar history of the United States was based upon the spreading of income among workers and the poor. This provided the income to purchase the tremendous increase in the world's output after the Second World War. But increasingly in recent years, the wealth of the United States has been concentrated in the hands of the rich, mostly as a consequence of expanding financial markets and tax policies that increasingly favor them, at least relative to earlier past tax policies that would (in the Second World War) tax up to 94% of a wealthy person's income. As we will describe in greater detail in ► Chap. 10, progressive taxes were curtailed in the 1980s, and the corporate tax burden has been falling since the early 1950s. Historically, increases in inequality have faced limits as well. When income becomes too concentrated at the top, as it did in the 1870s, 1890s, and 1920s, a depression followed, as citizens lacked the purchasing power to buy the products of industry. Excess capacity increased, investment declined, and unemployment soared. By many measures this is the situation the United States faces today.

7.15 Are We Seeing the End of the American Dream?

A central question that we will continue to explore in this book is whether this American dream (or European or Chinese or whoever else's dream) is sustainable, and what we might do to maintain it over the long term. Sustainability is a relatively new issue in economics but one that is increasingly on many economists' agenda. What sustainability is, of course, is highly dependent upon the perspective of who is asking the question. To an anthropologist or developmental economist, a sustainable economy might mean one that persists in time in the face of competition or aggression from other cultures or entities; to a conservation biologist, sustainable economy might mean one that does not degrade biodiversity, and to a resource-oriented person (like ourselves), it might mean not "living beyond the planet's biophysical means of supporting one's culture." We prefer a concise biophysical definition of sustainability. To be sustainable, an economy must live indefinitely within nature's limits. In other words an economy must persist over the long haul without excessive depletion or degradation of the energy and material flows—and the physical milieu—of the biophysical system that contains and supports economic activity. A sustainable economy must be able to provide not only jobs but, ideally, also meaningful work and meaningful lives for those human beings who make up "the economy." By this definition we are very, very far from sustainable. To us it is dishonest and unethical to declare as sustainable so many "green" entities, as we see daily in the media, that in fact require the use of fossil fuels and nonrenewable fuels or other depletable resources. The fact that a product or process is marginally better or greener in these respects than their competition (or can be made to look so) does not make them sustainable.

Historically, especially in the post Second World War era, the vehicles to maintain prosperity and social stability were economic growth domestically and enormous power (military, monetary, productivity) internationally. The productivity increases and cost containment that was the basis for these pillars were dependent upon cheap oil while ignoring many environmental issues such as CO_2 release. As we can no longer do this, the inherent tendency toward stagnation that characterizes mature market economies is exacerbated by biophysical limits. As these pillars of prosperity have weakened, the prospects for a dream, instead of a nightmare, decline as well [27, 28]. While such a decline may seem far away for those in comfortable circumstances reading this, the reality for people in Egypt, Syria, Nigeria, Venezuela, and other once-wealthy oil-producing regions who have indeed reached and passed peak oil is that economic and social turmoil is now a daily fact of life. This is especially well documented in the book by Nafeez Ahmed *Failing States, Collapsing Systems: BioPhysical Triggers of Political Violence* [29]. We think it extremely important to understand this book and adjust our economic aspirations to what is likely to be a new biophysical reality.

What is the basis for our perspective? What does it mean to live within nature's limits? Ultimately it comes down to maintaining the per capita resource stocks and flows required for human existence (and at what level of material well-being?) and the degree to which the atmosphere and oceans can handle the wastes of the human economy. The number of people, in the United States and elsewhere, continues to increase greatly (◘ Fig. 7.11). For example, when Hall was born (1943), there were about 137 million Americans and a little more than 2 billion people in the world. There are now more than 330 million people living in the United States and 7 billion people in the world. So the resources that form the basic inputs into our national and global economies have to be divided by roughly three times more people, and this is in only one person's (incomplete) lifespan. Global populations may well double or at least increase by another 50% in the reader's lifetime. Our most important mined resource is oil, and while it is not clear yet whether global oil production has peaked for all time, it is clear that per capita oil use (or oil use per person)

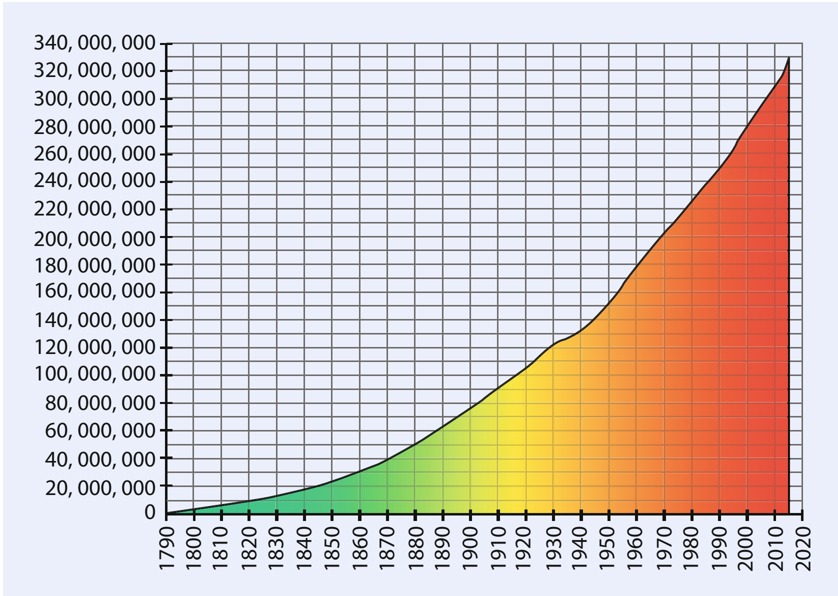


Fig. 7.11 US population (excluding many native Americans)

peaked in about 1978 (Fig. 7.12). In other words, a growing amount of oil (until recently) has been used by an even more rapidly increasing world population. The traditional economist argues that this is not crucially important as various technologies have allowed humans to generate more resources or more wealth from the resources that we do use. While we do not argue with the idea that technological innovation is important, we also will show in later chapters how this concept is extremely misleading and cannot form the only solution for our future and that of our children.

We can begin by considering petroleum, perhaps our most important resource beyond sunshine, clean water and soil. Most everything we do is based on cheap oil [19–21]. Where we live relative to where we work, what we do for our work, how much leisure time we have and how we spend it, the price of our food, most of our purchases and how much education we can afford, to name but a few, are largely dependent on adequate supplies of cheap oil. For example, it takes the energy of about a gallon of oil a day to feed each of us, about 80 barrels of oil to provide an undergraduate education at one of our colleges, and the energy equivalent of about 10 gallons per day to keep us supplied with all the goods and services that we demand through our economic activity. In earlier days this level of affluence was available to only a tiny elite of society and was usually provided by slave labor or indentured servitude. The net effect is that each of us today has some 60 to 80 "energy slaves" doing our bidding, effectively "hewing our wood and hauling our water."

The incredible thing about oil and gas is the almost complete absence of an understanding of its importance to the average American and their failure to understand how critical it is to our economy. At a meeting of ASPO-USA (the Association for the Study of Peak Oil) in 2006, Denver Mayor Hickenlooper, who understood the importance of oil and its restrictions, said "This land was originally settled by the Arapahoe and Cheyenne tribes. Everything that the natives depended upon, their food, clothing, shelter, implements and so on, came from the bison. They had many ceremonies giving thanks and appreciation to the Bison. We today are as dependent upon oil as the Sioux were on the bison, but not only do we not acknowledge that, but most people do not have a clue." The second critical thing about oil is "peak oil" and that as of 2017, global conventional oil production is no longer clearly increasing and may indeed be decreasing (Fig. 7.13). Almost certainly it will decrease substantially in the near future as we enter, in the words of geologist Colin Campbell, "the second half of the age of oil."

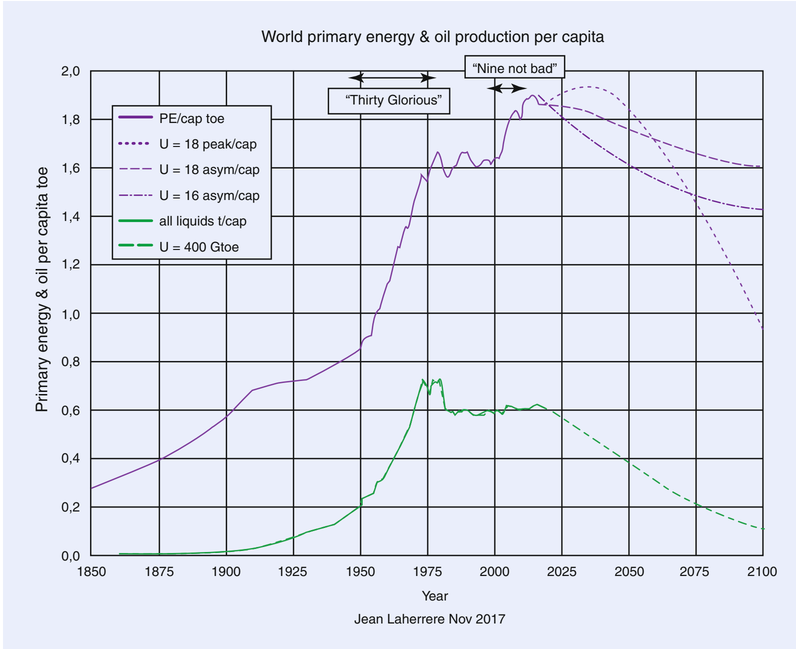


Fig. 7.12 Total and per capita oil production for the world, with projections. *TOE* tons oil equivalent

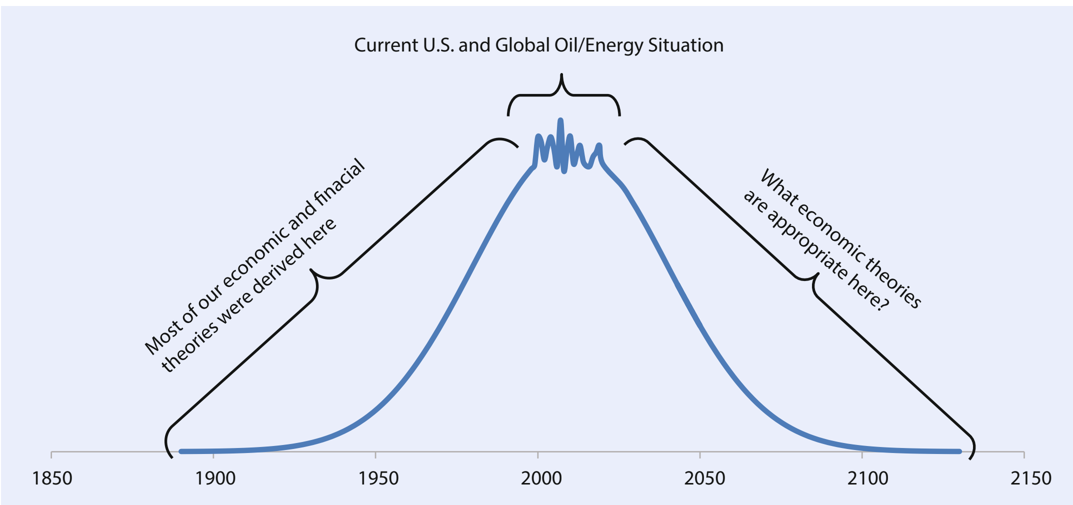


Fig. 7.13 Conceptual view of relation of our economic concepts and the Hubbert curve for global total oil use. Most of our economic concepts were derived during a period of increasing energy use. They may be having trouble explaining economic events during the present period of peak oil. How will they do during the decline in energy availability?

We conclude by saying yes, the American dream was the product of industrious and clever people working hard within a relatively benign political system that encouraged business in various ways but that all of these things also required a large resource base relative to the number of people using it. A key issue was the abundance of oil and gas in the United States, which was the world's largest producer in 1970. But in 1970 (and 1973 for gas, although there may be a second peak), there was a clear peak in US oil production, and while the continued increase in oil production worldwide buffered the United States (and other countries) from the local peak, it seems clear by 2010 that global oil production has reached its own peak, while demand from around the world continues to grow. This mismatch between supply and demand resulted in a sharp increase in the price of oil and many economic problems that we believe it caused, at least in part, including the stock market declines, the subprime real estate bust, the failure of many financial corporations, and the fact that some 40 odd of 50 states are officially broke and that there is a substantial decrease in discretionary income for many average Americans. As developed later in ► Chap. 19, we believe that all of these economic problems are a direct consequence of the beginning of real shortages of petroleum in a petroleum-dependent society.

7.16 The Future of the American Dream

"Macondo (The Gulf of Mexico oil spill site) eventually gripped the media and political eye. It is time for sober reflection on the global energy predicament and not for knee jerk reactions. How important is primary energy production and consumption for the OECD way of life? It links to economic growth, tax receipts and all that these pay for, pensions, manufacturing, food production, defense, leisure, comfort and security." These words, presented by energy analyst Euan Mearns on The Oil Drum (Europe) June 16, 2010, sum up our dilemma. As the country wrestles with the terrible environmental and economic consequences of the oil spill, does the disaster signal the end of finding new oil to support the fishing and recreation industries that are bemoaning the impact? Will this single event be the turning point in our hope of maintaining our national affluence [27, 28]. Have we finally caught up with living beyond our means through debt while seeing the beginning of the end of the continual hope that technology, such as deep sea drilling, will extend the American dream forever? A decade later it turns out that this oil spill has made little difference. American economic activity continues unabated but with little growth and increasing disparity between rich and poor, so that the American dream seems increasingly beyond the grasp of more and more people. How much of this is related to the essential cessation of growth in oil production and availability?

It is not just the United States but Europe and Japan that are experiencing much lower economic growth than in the past. In Europe growth has essentially stopped for a decade, something called "secular stagnation". This is something that is quite unprecedented in the economics of the past century but which now seems to be quite established in many places. It has meant that it is very difficult to meet pension plans (because investments do not return as they once did) and that money in the bank does not earn interest. Those who still assume that growth should be taking place have ended up in very bad financial situations. This is correlated over time with a discontinuance of the growth of oil and fossil fuels more generally. It is behind the difficulties that many states have with meeting pension obligations, assisting students, and so on and perhaps the disaffection that many have with governments.

Franklin Roosevelt ran up huge debts to reconstruct the economy after the Depression and during the Second World War. Except for Bill Clinton, all presidents since (and including Ronald Reagan) increased debt (both corrected for inflation and the size of the economy) more than did Franklin Roosevelt during the new deal (Fig. 7.9). Will the economy continue to grow so that we can pay off those debts if there is no longer cheap energy? What kind of jobs will be available if Americans have less and less discretionary income? Do we still want more labor productivity, which has usually meant subsidizing each laborer with more and more energy? Or do we want less labor productivity, that is, more energy productivity, by subsidizing each increasingly valuable unit of energy with more and more labor, to keep our people employed? How will we maintain and enhance the value of pension funds if the stock market no longer grows in real terms? What about our inner cities? Can we find ways to employ those desperately in need of a job? Do population increases enhance our economic well-being or simply divide up our remaining untapped resources among more and more people? Do we need a completely new approach to economics during times when energy is declining? What indeed will the American dream mean in the future? Can we generate a new American dream with fewer material goods and more leisure? Will these issues limit our ability to support and educate our children? These are a new set of economic questions that require a new way of thinking about economics. Much of the rest of this book will try to provide some of the information to answer these questions.

Questions

1. What does the word biophysical mean? How does it relate to economics? With what word(s) would you contrast it with?
2. What factors are likely to influence your own economic success in life?
3. Although energy is barely discussed in physiocratic, classical, or neoclassical economics, explain how each of these schools of economic thought focuses on the dominant energy flows of their time.
4. Why was Spindletop an event of great economic importance for the United States?

References

1. Martin, P.S. 1973. The discovery of America. *Science* 179: 969–974.
2. Tainter, J. 1988. *The collapse of complex societies*. Cambridge: Cambridge University Press.
3. Diamond, J. 2004. *Collapse: How societies choose to fail or survive*. New York: Viking Press.
4. Crosby, Alfred. 1986. *Ecological imperialism: The biological expansion of Europe, 900–1900*. Cambridge, UK: Cambridge University Press.
5. Diamond, J. 1997. *Guns, germs and steel: The fates of human societies*. New York: Norton.
6. Zinn, Howard. 1980. *A people's history of the United States*. New York: Harper and Row.
7. The two greatest causes for war among the perhaps one third of the colonial population who supported independence was basically resource scarcity – the cutting off of the trans-Allegheny frontier, first in 1763 by means of the Proclamation Act, and then later in 1775, with greater enforcement, with the Quebec act. Open rebellion soon followed.
8. Perlin, J. 1989. *A forest journey: The role of wood and civilization*. New York: Norton.
9. Greene, A.N. 2008. *Horses at work: harnessing power in industrial America*. Cambridge: Harvard University Press.
10. Yergin, Daniel. 1991. *The prize: The epic quest for oil, money, and power*. New York: Simon & Schuster.
11. Bowles, S., D. Gordon, and T. Weisskop. 1990. *After the wasteland*. Armonk: ME Sharpe.
12. Smil, V. 2001. *Enriching the earth. Fritz Haber, Carl Bosch and the transformation of western food production*. Cambridge: MIT Press.
13. Meadows, D., D. Meadows, and J. Randers. 2004. *Limits to growth: The 30-year update*. White River: Chelsea Green Publishers.
14. Ehrlich, P. 1968. *The population bomb*. New York: Ballantine Books.
15. Hubbert, M.K. 1969. Energy resources. In *The National Academy of Sciences–National Research Council, Committee on Resources and Man: A Study and Recommendations*. San Francisco: W. H. Freeman.
16. Hardesty, J., Norris C. Clement, and Clinton E. Jencks. 1971. The political economy of environmental destruction. *Review of Radical Political Economics* 3 (4): 82–102.
17. England, R., and B. Bluestone. 1971. Ecology and class conflict. *Review of Radical Political Economics* 3 (4): 31–55.
18. Bowles, S., D. Gordon, and T. Weisskopf. 1990. *After the wasteland*. Armonk, New York: M.E. Sharpe.
19. Odum, H.T. 1973. *Environment, power and society*. New York: Wiley Interscience.
20. Daly, Herman. 1973. *Towards a steady-state economy*. London: W.H. Freeman and Company, Ltd. Tables.
21. Cleveland, C., R. Costanza, C. Hall, and R. Kaufmann. 1984. Energy and the U.S. economy. A biophysical perspective. *Science* 225: 890–897; Hall, C.A.S., C.J. Cleveland and R. Kaufmann. 1986. *Energy and Resource Quality: The ecology of the economic process*. New York: Wiley Interscience. 577 pp.
22. Jorgenson, D.W., and Z. Grilliches. 1967. The explanation of productivity change. *Review of Economic Studies*: 249–283.
23. Maddala, G.S. 1965. Productivity and technical change in the bituminous coal industry. *Journal of Political Economy*: 352–265.
24. Reynolds, D. 2000. Soviet economic decline: Did an oil crisis cause the transition in the Soviet Union? *Journal of Energy and Development* 24: 65–82.
25. Campbell, C., and J. Laherrere. 1998. The end of cheap oil. *Scientific American* (March): 78–83.
26. Stern, N. 2007. *The economics of climate change. The stern review*. Cambridge: Cambridge University Press.
27. Luce, E. 2010 Goodbye, American Dream. The crisis of middle-class America. *The Financial Times*. 30 July 2010.
28. Brinkbaumer, K., M. Hujer, P. Muller, and T. Schulz. 2010. Is the American dream over? *Das Spiegel*: 2010.
29. Ahmed, N. 2017. *Failing states, collapsing systems: BioPhysical triggers of political violence*. New York: Springer.

The Petroleum Revolution and the First Half of the Age of Oil

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_8

8

8.1 The First Half of the Age of Oil [1]

This chapter will focus on the importance of fossil fuels (coal, gas, and oil) and especially petroleum (meaning natural gas and oil, or sometimes just oil). First we want to ask why petroleum, and especially oil? Why has petroleum been so important, and why is it so hard to unhook ourselves from it? To do that we need to look more broadly for a moment at the energy situation that has faced, and that faces, humanity. Solar energy, either directly or as captured by plants, was and is the principal energy available to run the world or the human economy. It is enormous in quantity but diffuse in quality. As we have developed in the previous chapter, the history of human culture can be viewed as the progressive development of new ways to exploit that solar energy using various conversion technologies, from spear points to fire to agriculture to, now, the concentrated ancient energy of fossil fuels. Until the past few hundred years, human activity was greatly limited by the diffuse nature of sunlight and its immediate products and because that energy was hard to capture and hard to store. Now fossil fuels are cheap and abundant, and they have increased the comfort, longevity, and affluence of most humans, as well as their population numbers.

But there is a downside, for fossil fuels are made principally of carbon. The use of carbon-based fuels generates a gaseous by-product, carbon dioxide (CO_2) that appears quite undesirable. Now we are constantly bombarded with recommendations of our need to "decarbonize" our economy because of the environmental impacts, such as climate change and ocean acidification that the increases in carbon dioxide appear to be causing. These impacts are likely to become much more important into the future. Consequently there have been considerable efforts to come up with fuels or energy sources not based on carbon. To date that effort has failed completely, for, according to the data compiled by the US Energy Information Agency, the amount of CO_2 produced most years continues to increase (unless there is a recession). With so many apparent options how come we cannot unhook ourselves from carbon? Why is it that most of our energy technologies continue to rely on the chemical bonds of carbon (most usually combined with hydrogen as hydrocarbons)?

The answer lies in basic chemistry: the only effective and large-scale technology that has ever been "invented" for capturing and storing that solar energy is photosynthesis. Humans use the products of photosynthesis for all or most all of our fuels simply because there is no alternative on the scale we need. This is because nature, the source of our fuels, has favored the storage of solar energy in the hydrocarbon bonds of plants and animals. The reasons are that these elements are abundant and "cheap" to an organism, and, most importantly, capable of forming *reduced* or energy-containing chemical compounds. Hydrogen and carbon, which essentially do not exist in elemental form at the Earth's surface, are so important that plants have evolved the technology to split water and atmospheric carbon dioxide to get hydrogen and carbon, which they combine to form energy-rich *hydrocarbons* and, with a little oxygen, *carbohydrates*. There simply are not other elements in the periodic table that are sufficiently abundant and capable of such ready reduction. Nitrogen, for example, is abundant as N_2 but much more expensive energetically to split, and sulfur is less available. In addition carbon has four valence electrons, capable of forming four bonds with other atoms and hence the very complex structures of biology. Bonds with hydrogen greatly increase the capacity to store energy in a molecule. Thus plants and animals are carbon and hydrogen based because nature had no choice. Human cultural evolution has exploited this hydrocarbon energy profitably mostly because they had no choice but to use the products of photosynthesis. Now we are stuck with the carbon dioxide while we try to figure out if there possibly can be an alternative that is energetically feasible.

8.2 The Industrial Revolution

Beginning on a small scale about 1750 but then increasingly rapidly about 1850, there was a rather remarkable change in the hydrocarbons that humans used, from the recently captured solar energy of wood, water and muscle power to the enormously more powerful fossil fuels. This was the beginning of the "industrial revolution," although perhaps a more proper name would

Table 8.1 Energy density of oil and other fossil fuels (may vary somewhat with specific fuels)

Fuel type[a]	MJ/l[a]	MJ/kg[a]	kBTU/Imp Gal	kBTU/US Gal
Regular gasoline/petrol	34.8	~47	150	125
Premium gasoline/petrol		~46		
Autogas (LPG) (60% propane and 40% butane)	25.5–28.7	~51		
Ethanol	23.5	31.1	102	85
Methanol	17.9	19.9	78	65
Gasohol (10% ethanol and 90% gasoline)	33.7	~45	145	121
E85 (85% ethanol and 15% gasoline)	33.1	44	143	119
Diesel	38.6	~48	167	139
Biodiesel	35.1	39.9	152	126
Vegetable oil (using 9.00 kcal/g)	34.3	37.7	148	123
Aviation gasoline	33.5	46.8	144	120
Jet fuel, naphtha	35.5	46.6	153	128
Jet fuel, kerosene	37.6	~47	162	135
Liquefied natural gas	25.3	~55	109	91
Liquid hydrogen	9.3	~130	40	34

Coal 29, *Biomass* 15–28 MJ/GJ
[a]Mj/l = MegaJoules per liter. Neither the gross heat of combustion nor the net heat of combustion gives the theoretical amount of mechanical energy (work) that can be obtained from the reaction. (This is given by the change in Gibbs free energy and is around 45.7 MJ/kg for gasoline.) The actual amount of mechanical work obtained from fuel (the inverse of the specific fuel consumption) depends on the engine. A figure of 17.6 MJ/kg is possible with a gasoline engine and 19.1 MJ/kg for a diesel engine. See brake-specific fuel consumption for more information

be the "hydrocarbon revolution." Humans had begun to understand how to use the much more concentrated energy found in fossil (meaning old) fuels. Why did they do this? The answer is simple. People wanted to do more work because to do so is profitable. They want more of some raw material transformed into something useful that they can eat, trade, or sell. Fossil hydrocarbons have greater energy density than the carbohydrates such as food and wood, and as a consequence they can do much more work—heat things faster and, to a higher temperature, operate machines that are faster and more powerful and so on (Table 8.1). The first fossil hydrocarbon used at any significant scale was coal, first used at a large scale in the nineteenth century, then oil in the twentieth century, and now increasingly natural gas. The global use of hydrocarbons for fuel increased nearly 800-fold since 1750 and about 12-fold in the twentieth century alone, and this has enabled our enormous economic growth (Fig. 8.1).

Economists usually call rapid increases in economic activity development. Hydrocarbon-based energy is important for three main areas of human development: economic, social, and environmental [2]. Most importantly, hydrocarbons have generated an enormous increase in the ability of humans to do all kinds of economic work, greatly enhancing what they might be able to do with their own muscles or with those of work animals by using fossil-fueled machines such as

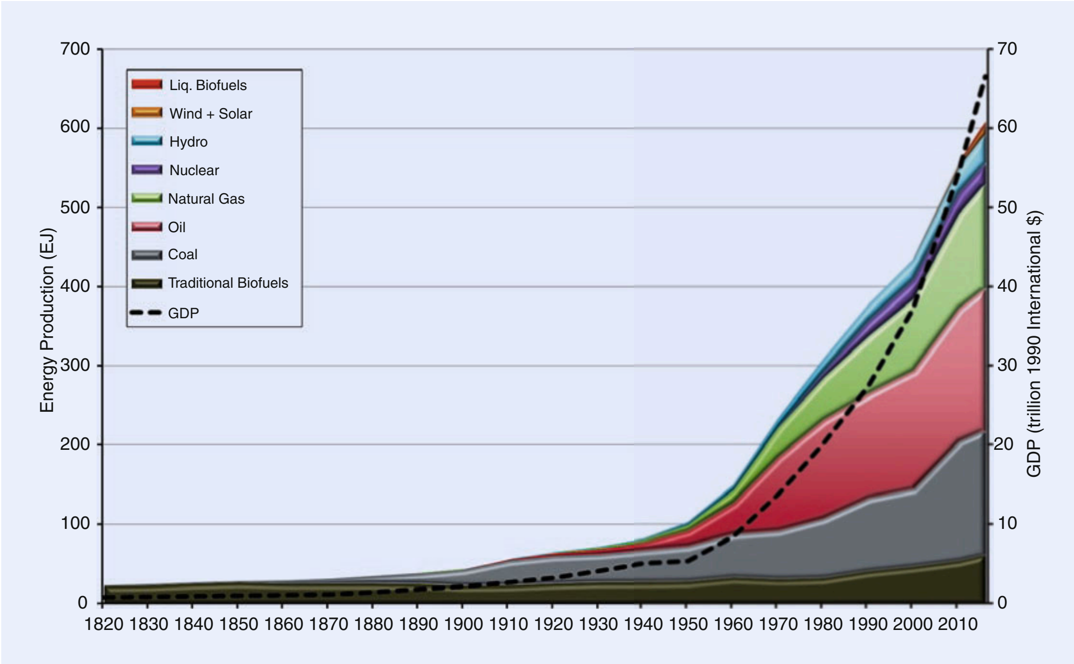


Fig. 8.1 The global use of hydrocarbons for fuel by humans has increased nearly 800-fold since 1750 and about 12-fold in the twentieth century. The most general result has been an enormous increase in the ability of humans to do all kinds of economic work, greatly enhancing what they might be able to do by their own muscles or with those of draft animals (Source: Authors)

trucks and tractors (Table 8.1). Perhaps most importantly this work includes an enormous increase in the production of food.

The industrial revolution started in England with coal in roughly 1750, but by about 1960 the world was using more petroleum than coal, and oil continues to be our most important energy source [3]. Now we live in, overwhelmingly, the age of oil. Some have said that we now live in an information age or a post-industrial age. Both are only partly true. Overwhelmingly we live in a petroleum age. Just look around. All transportation, all food production, all plastics, most of our jobs and leisure, much of our electricity, and all of our electronic devices are dependent upon gaseous and especially liquid petroleum. This has been, and continues to be, the age of oil and of hydrocarbons more generally. Perhaps the industrial revolution should be renamed the "hydrocarbon revolution" because that is what happened—humans moved from using various carbohydrates as their principle means of doing economic work to using hydrocarbons.

One reason that this is the age of oil, and hydrocarbons more generally, is that there continues to be a strong connection between energy use and economic activity for most industrialized [4] and developing economies [5] (Fig. 8.1). Some have argued that through technology and markets, we are becoming more efficient in our use of energy. But the evidence for that is ambiguous at best. As yet unpublished top-down macroeconomic analysis (i.e., simply dividing inflation-corrected GDP by total energy used) undertaken by Ajay Gupta indicates that for most countries of the world, there remains a very strong link between energy use and economic activity, as measured by inflation-corrected GDP and that there is no general trend of countries becoming more or less efficient in turning energy into GDP. One apparent exception is the United States, where there is an apparent decline in the ratio of energy used per unit of gross domestic product. Energy analyst Robert Kaufmann suggests that while there has been some real improvements in fuel efficiency (driven by higher fossil fuel prices), the increases in efficiency are due principally to a shift to higher-quality fuels and especially to structural changes in national economies as richer nations move their heavy industries overseas to reduce pollution or find cheaper labor [6]. There may be another reason as well that the United

States, but few other nations, appears to be becoming more efficient in our use of energy. According to the organization Shadowstatistics, the United States has been engaged in a systematic "cooking of the books" on the official measure of inflation, that is, a deliberate official underestimate of inflation since 1985 to make governments look good. Correcting for any or all of these actions would greatly decrease the perceived improvements of efficiency in the US economy. In addition it is clear from Gupta's data that the main way that countries develop (i.e., get richer) is through using more energy to do more economic work [7].

Energy prices have an important effect on almost every major aspect of macroeconomic performance because energy is used directly and indirectly in the production of all goods and services. Both theoretical models and empirical analyses of economic growth suggest that a decrease in the rate of energy availability will have serious impacts on the economy [8]. For example, most US recessions after the Second World War were preceded by rising oil prices, and there tends to be a negative correlation between oil price changes and both stock prices and their returns in countries that are net importers of oil and gas [9]. Energy prices have also been key determinants of inflation and unemployment. There is a strong correlation between per capita energy use and social indicators such as the UN's Human Development Index, although that relation is much more important at low incomes than high—in other words increasing energy use is far more important at improving quality of life for poor than for rich [10]. By contrast, the use of hydrocarbons to meet economic and social needs is a major driver of our most important environmental changes, including global climate change, acid deposition, urban smog, and the release of many toxic materials. Increased access to energy provided the means to deplete or destroy once-rich resource bases, from megafaunal extinctions associated with each new invasion of spear-equipped humans, to the destruction of natural ecosystems and soils through, for example, overfishing and intensive agriculture and other types of development. Harvard biologist E.O. Wilson has attributed the current mass extinction to what he calls HIPPO effects: Habitat destruction, Invasive species, Pollution, Population (human), and Overgrazing. All these activities are energy-intensive. Such problems are exacerbated by the increase in human populations that each new technology has allowed, as well as the overdependence of societies on previously abundant resources. Energy is a double-edged sword.

8.3 Peak Oil: How Long Can We Depend on Oil?

The critical issue with oil is not when do we run out, but when can we no longer increase or even maintain its production and use. We believe that "peak oil," the time when humans can no longer count on increasing oil production no matter what their effort, is more or less now and that this will become the most important issue facing humanity. This critical issue can be understood at two levels: first as a simple fact, less, not more, oil over time, and second by a more thorough understanding of the properties and attributes of oil, which we do next. While the exact timing of peak oil for the world remains somewhat debatable, it is clear that it must be soon because each year, we use two to four times more oil than we find. What is even more obvious is that our old rate of increase of 3 or 4% a year has declined since 2004 to from 0 to 1% and that oil availability per capita is declining.

At present, oil supplies about 32% (and natural gas about 20%) of the world's non-direct solar energy, and most future assessments indicate that the demand for oil will increase substantially if that is geologically, economically, and politically possible. While the use of nonfossil energy resources (e.g., photovoltaic and wind) is increasing rapidly, they still provide only about 2% of global energy use. While the percentage of solar is anticipated to increase, the absolute amount of fossil fuels is predicted to increase for the indefinite future for as long as that is possible. What do we know about the future availability of oil? Predictions of impending oil shortages are as old as the industry itself, and the literature is full of arguments between "optimists" and "pessimists" about how much oil there is and what other resources might be available. There are four principal issues that we need to understand in order to assess the availability of oil and, by extension other hydrocarbons, for the future. We need to know the quality of the reserves, the quantity of the reserve, the likely patterns of exploitation of the resource over time, and who gets and who benefits from the oil. All of these factors ultimately affect the economics of oil production and use.

8.4 Quality of Petroleum

Oil is a fantastic fuel, relatively easy to transport and use for many applications, very energy dense, and extractable with relatively low energy cost and (usually) low environmental impact compared to most other energy sources (Table 8.1). What we call oil is actually a large family of diverse hydrocarbons whose physical and chemical qualities reflect the different origins and, especially, different degrees of natural processing of these hydrocarbons. Basically oil is phytoplankton kept from oxidation in deep anaerobic marine or freshwater basins, covered by sediments, and then pressure-cooked for 100 million years [11]. In general, humans have exploited the large reservoirs of shorter-chain "light" oil resources first because larger reservoirs are easier to find and exploit and lighter oils require less energy to extract and refine [12]. The depletion of this "easy oil" has required the exploitation of increasingly small, deep, offshore, and heavy resources. Oil must first be found, then the field developed, and then the oil extracted carefully over a cycle that typically takes decades. Oil in the ground is rarely like what we are familiar with in an oil can. It is more like an oil-soaked brick, where the oil must be pushed slowly by pressure to a collecting well. The rate at which oil can flow through these "aquifers" depends principally upon the physical properties of the oil itself and of the geological substrate, but also upon the pressure behind the oil that is provided initially by the gas and water in the well. Progressive depletion also means that oil in older fields that once came to the surface through natural drive mechanisms, such as gas and water pressure, must now be extracted using energy-intensive secondary and enhanced technologies. As the field matures, the pressure necessary to force the oil through the substrate to the collecting wells is supplied increasingly by pumping more gas or water into the structure. EOR or enhanced oil recovery is a series of processes by which detergents, CO_2, and steam have been used—since the 1920s—to increase yields. Too-rapid extraction can cause compaction of the "aquifer" or fragmentation of flows which reduce yields. So our physical capacity to produce oil depends upon our ability to keep finding large oil fields in regions that we can reasonably access, our willingness to invest in exploration and development, and our willingness to not produce too quickly. Thus, technological progress is in a race with the depletion of higher-quality resources.

Another aspect of the quality of an oil resource is that oil reserves are normally defined by their degree of certainty and their ease of extraction, classed as "proven," "probable," "possible," or "speculative." In addition, there are unconventional resources such as heavy oil, deepwater oil, oil sands, and shale oils that are very energy-intensive to exploit. Thus while there are large quantities of oil left in the world, the quality of the actual fields is decreasing as we find and deplete the best ones. Now it takes more and more energy to find the next field and, as they tend to be of poorer quality, more and more energy to extract and refine the oil to something we can use.

8.5 Quantity of Petroleum

Most estimates of the quantity of conventional oil resources remaining are based on "expert opinion," which is the carefully considered opinion of geologists and others familiar with a particular region (Table 8.2). The ultimate recoverable resource (URR, often written as EUR) is the total quantity of oil that will ever be produced from a field, nation, or the world, including the 1.3 trillion barrels extracted to date. URR will determine the shape of the future oil production curve. Recent estimates of URR for the world have tended to fall into two camps. There is a great deal of controversy—or rather range of opinion—about how much oil remains (Fig. 8.2). Lower estimates come from several high-profile analysts, many of them retired petroleum geologists, with long histories in the oil industry who suggest that the URR is no greater than about 2.3 trillion barrels (in other words the 1.3 we have used and another 1.0 we will extract in the future), and may be even less [12]. The USGS (United States Geological Survey) "low" estimate is that this number may be about 2.4 trillion barrels, half from new discoveries and half from reserve growth, which is increased estimates of oil available from existing fields. A "middle" estimate is three trillion barrels and the highest credible estimate is four trillion barrels (Table 8.3). These latter three values are from a very comprehensive study by the

Table 8.2 How reliable are official energy statistics? (All values in gigabarrels) (From Lewis L. Smith)

OPEC	Cum prod end 2003	% Depleted	Indicated total	Remaining reserves Gb (as of 2004)				BP estimates interpreted
				PFC	ASPO	Salameh	BP	
Iraq	28	22%	127	99	62	62	115	Total discovered
UAE	19	31%	61	42	49	37	98	Total discovered
Kuwait	32	35%	91	59	60	71	97	Total discovered
Libya	23	39%	59	36	29	26	36	
Saudi	97	42%	231	134	144	182	263	Total discovered
Algeria	13	50%	26	13	14	11	11	
Nigeria	23	50%	46	23	25	20	34	High estimate
Iran	56	51%	110	54	60	64	131	Total discovered
Venezuela	47	58%	81	34	35	31	78	Total discovered
Qatar	6.8	62%	11	4.2	4.1	4.6	15	Total discovered
Indonesia	20	75%	27	6.7	9.4	12	4.4	
Total	**365**		**870**	**506**	**492**	**520**	**882**	

Statistics for the oil industry are not as bad as those for the wine industry, but still, they are pretty bad! This is especially true for reserves; the amounts of oil which engineers and geologists estimate could be extracted in the future from active reservoirs or promising geological formations, given present prices and technology. The three most important compilers of statistics for the oil industry are the BP, *Oil and Gas Journal*, and the US DOE's Energy Information Administration. And that is all they are, compilers. They do not audit, check, or question the information supplied to them by their diverse sources, and they use different definitions of e.g. reserves. One reason is rumored to be that they are afraid of being "cutoff" by any source to which they pose embarrassing questions! Just out of curiosity, I (LLS) checked the table, "Worldwide look at reserves and production," in the December 21 issue of the *Oil and Gas Journal*, pp. 20–21. Of the 200 or so political jurisdictions which merit statistical recognition by the UN, 107 got a line in the table, because they have "proven" oil reserves, gas reserves, or both. There are five good reasons why an estimate of reserves for a nation should change [up or down] every year. Indeed it is almost impossible for them to remain unchanged, if the engineers and geologists have done their work correctly. These five reasons include new findings, revisions in old estimates, and, clearly, production. However, I note that in the referenced table, only 29 countries [27% of the total] report no oil reserves or changed their estimate from last year. The other 78 [73%] reported exactly the same figure for this year as last year. This includes one country which is widely believed to be exaggerating its "official" estimate by more than 100%! Some of the "no changers" include Indonesia, Iraq, Kuwait, Norway, Russia, and Venezuela. Ironically Norway is one of the few countries that publishes good production data by oil field. You may draw your own conclusions! I gather that the situation for natural gas is a little better, but not enough to trust the data for all important producers USGS 2000
Source: Iran's reserves less than half. OPEC's reserves overstated by 80%. From ► mushalik@tpg.com.au and ► http://www.energiekrise.de/e/aspo_news/aspo/newsletter046.pdf

1.1 Trillion	Depleted	Statistical reliability	Production outlook	Technical basis
Actual reserves: 0.9 trillion	Proven >90%	Proven oil in place—high confidence Developed—clear recovery factor Undeveloped—good. recov. est.	Growth thru actual reservoir mgmt. and performance	Improved oil recovery thru existing technology
	Probable >50%	Probable oil in place—confident Developed—prelim. recovery factor Undeveloped—est. fair recovery	Growth thru delineation, testing, and development	Clear opportunity with existing technology
	Potential >5%	Potential oil in place—low confiden. Drilled—v. low recovery factor Undrilled—recovery likely poor	Growth thru pricing, delineation, or IOR/EOR technology	Indicative data and potential opportunity
Contingent resources: 1.1 trillion	Resource: Uneconomic volume and commerciality	Likely presence but undelineated OIP or GIP	Profitability or technology currently inadequate	Available access but lacks good reservoir and fluids data
Prospective and speculative resources: 2.0 trillion	Oil, gas, shales, EHC, and to be discovered resources (speculative outlook)	Technically present but physically inaccessible hydrocarbons	Future resolution thru exploration and relevant technology	General geological, seismic and/or physical indications
		Conceptually possible hydrocarbons, incl. EHCs		

Fig. 8.2 How much oil remains in the world is highly uncertain. For example, "Reserves" are inflated with >300 B bbls of "resources" Source: From ▶ mushalik@tpg.com.au

Table 8.3 Published estimates of world oil ultimate recovery

Source	volume (trillions of barrels)
USGS, 2000 (high)	3.9
USGS, 2000 (mean)	3.0
USGS, 2000 (low)	2.25
Campbell, 1995	1.85
Masters, 1994	2.3
Campbell, 1992	1.7
Bookout, 1989	2.0
Masters, 1987	1.8
Martin, 1984	1.7
Nehring, 1982	2.9
Halbouty, 1981	2.25
Meyerhoff, 1979	2.2
Nehring, 1978	2.0
Nelson, 1977	2.0
Folinsbee, 1976	1.85
Adam and Kirby, 1975	2.0
Linden, 1973	2.9
Moody, 1972	1.9
Moody, 1970	1.85
Shell, 1968	1.85
Weeks, 1959	2.0
MacNaughton, 1953	1.0
Weeks, 1948	0.6
Pratt, 1942	0.6

From Hall et al. [1]

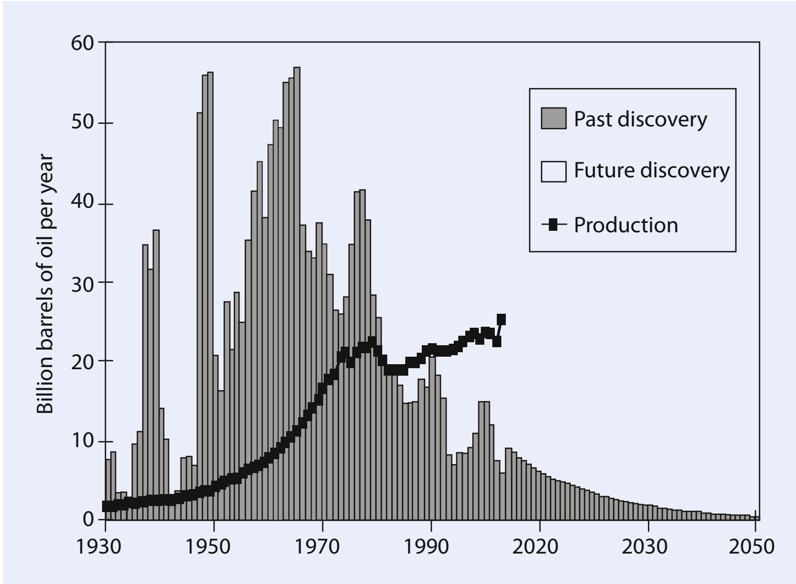


Fig. 8.3 Rates of finding and rate of production for conventional oil globally where field updates have been updated to the year that the initial strike was found (Source: Colin Campbell). Note: There is another way of graphing this data by attributing "revisions and extensions" to the year of revision, not the year of initial strike. This exaggerates the finding rate of more recent years

US Geological Survey in 2000, which if nothing else tend to cover the range of other estimates (USGS [13–15]). Even in that study, the lower values tend to be from their staff of geologists, and the larger ones reflect increasingly the opinion of USGS economists who believe that price signals will allow lower grades of oil to be exploited through technical improvements, and there will be corrections of earlier conservative estimates. This relatively new addition to the USGS methodology is based on experience in the United States and a few other well-documented regions. The new totals assume, essentially, that petroleum reserves everywhere in the world will be developed with the same level of technology, economic incentives, and efficacy as in the United States. Although time will tell the extent to which these assumptions are realized, the last 10 years of data have shown that the majority of countries are experiencing patterns of production that are far more consistent with the low rather than medium or higher estimates of ultimately recoverable reserves (URR) [16, 17]. Increasingly other estimates by, e.g., US and European energy agencies (AIE and IEA), are coming in on the low side. An assessment by oil expert (the best in our opinion) Colin Campbell shows that we are now producing and consuming 2–4 barrels for each barrel we find (Fig. 8.3) One would think that the best way to find and produce more oil would be to drill more, but in fact the finding of oil and gas is almost independent of drilling rate, at least at the levels we have been used to undertaking, because time is needed to determine where the next good place to drill is (Fig. 8.4). The impact of new drilling technology (horizontal drilling and "fracking") is considered in ► Chap. 13.

8.6 Pattern of Use Over Time

The best-known model of oil production was derived by Marion King Hubbert, who proposed that the discovery and production of petroleum over time would follow a single-peaked, more or less symmetric, bell-shaped curve (Fig. 8.5). A peak in production would occur when 50% of the URR had been extracted (he later opined that there may be more than one peak). This hypothesis seems to have been based principally on Hubbert's intuition and his tremendous experience examining the patterns of many, many oil fields. It was not a bad guess, as he famously predicted in 1956 that US oil production would peak in 1970, which in fact it did [15]. Hubbert also predicted that the US production of natural gas would peak in about 1975, which it did, although it has since shown signs of recovery and there is a second peak following 2010 based on "unconventional" and "shale" gas. He also predicted that world conventional oil production would peak in about 2000. In fact, conventional oil production continued to increase until 2005, after which it appears to have entered an oscillation or "undulating plateau," as predicted earlier by geologist Colin Campbell.

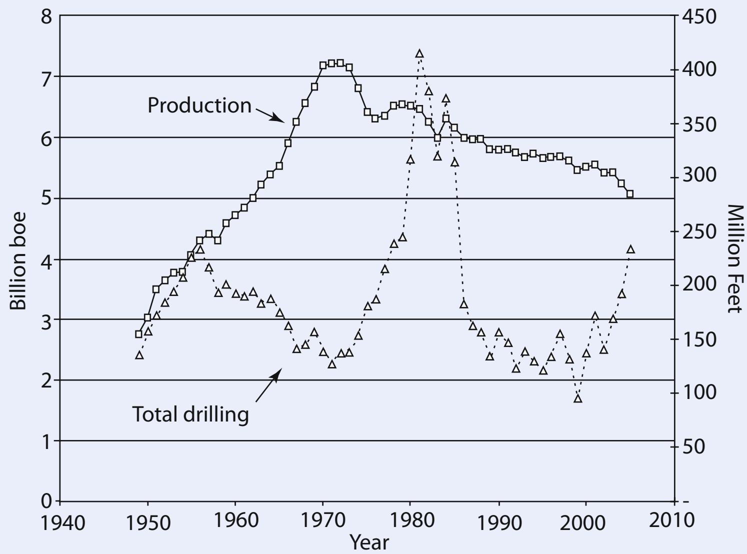


Fig. 8.4 Oil and gas production appears to be independent of drilling effort. There has been essentially no correlation between drilling effort and the rate of finding and (here given) production of oil and gas in the United States except in the first years plotted. Production increased until 1970, then peaked, and then declined steadily despite enormously increased, and subsequent decreased, drilling effort. Drilling rates tend to increase when prices are high and the converse (Source: U.S. EIA; note there has been no data made available since 2007)

In the past decade, a number of "neohubbertarians" have made predictions about the timing of peak global production ("peak oil") using several variations of Hubbert's approach [15–24]. These forecasts of the timing of global peak have ranged from one predicted for 1989 (made in 1989) to many predicted for 2005–2015 to one as late as 2030 [18]. Most of these studies assumed world URR volumes of roughly two trillion barrels and that oil production would peak when 50% of the ultimate resource had been extracted. The predictions of a later peak begin with an assumption of a large volume of ultimately recoverable oil. How much oil will we actually recover? The USGS study quoted above gives a low estimate (which they state has a 95% probability of being exceeded) of 2.3 trillion barrels and a "best" estimate of 3 trillion barrels. One analysis fitted the left-hand side of Hubbert-type curves to data on actual production while constraining the total quantity under the curve to two, three, and four trillion barrels for world URR. The resultant peaks were predicted to occur from 2004 to 2030 [19]. Brandt [20] shows that the Hubbert curve is a good prediction for most post-peak nations, which includes the great majority of all oil-producing nations. Other recent and sophisticated Hubbert-type analyses by Kaufmann and Shiers [21] and Nashawi and colleagues [22] suggest peaks in conventional oil about 2013–2014, consistent with the low URR estimates of, e.g., Campbell and Laherrere, at least as long as there is not much more recoverable oil than seems likely at this time [12]. If that is the case, the peak may be displaced for one or two decades. An important issue that most of these studies do not consider is that most of the oil left in the ground will take an increasing amount of energy to extract.

Most recent results of curve-fitting methods showed a consistent tendency to predict a peak within a few years, and then a decline, no matter when the predictions were made. This is consistent with the fact that we are using at least twice as much oil as we are finding. Other forecasts for world oil production do not rely on either assumptions about URR or the use of "curve-fitting" or "extrapolating" techniques but simply draw straight lines into the future based on past increases. According to one forecast by the US Energy Information Agency (EIA) (2003), world oil supply in 2025 will exceed the 2001 level by 53% [13]. The EIA reviewed five other world oil models and found that all of them predict that production will increase in the next two decades to around 100 million barrels per day, substantially more than the 77 million barrels per day produced in 2001. Several of these models rely on the 2000 USGS higher estimates of URR for oil. Clouding the empirical assessment is that official estimates are newly including "nonconventional" resources (notably natural gas liquids, but also "heavy" and "ultra-deepwater" petroleum, and

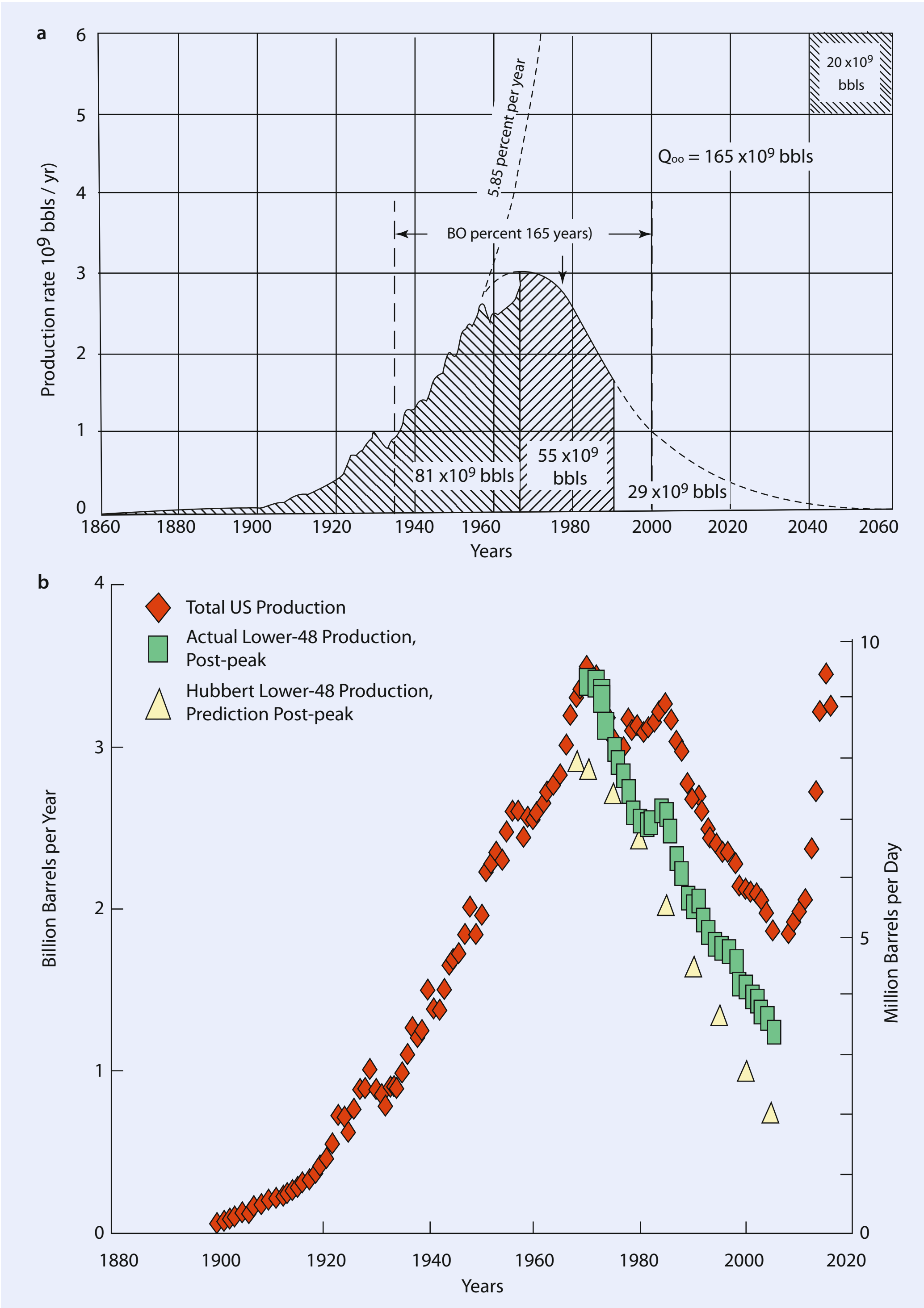


Fig. 8.5 Hubbert curves. **a** Original for United States (Source: Hubbert 1969). **b** Present for United States (Source: 2006 Cambridge Energy Research Associates) and author updates. **c** Original for world (Source: [15]). **d** Present world data (Source: Jean Laherrere ; XH = Extra Heavy (tar sands); LTO = Light, tight (e.g. "fracked" oil. Brown line is what is usually considered "conventional" oil. Thin lines are LaHerrere's predictions). **e** American whaling industry. Depletion was enormous: from 90 to 99% of many whale species were killed (Source: Ugo Bardi)

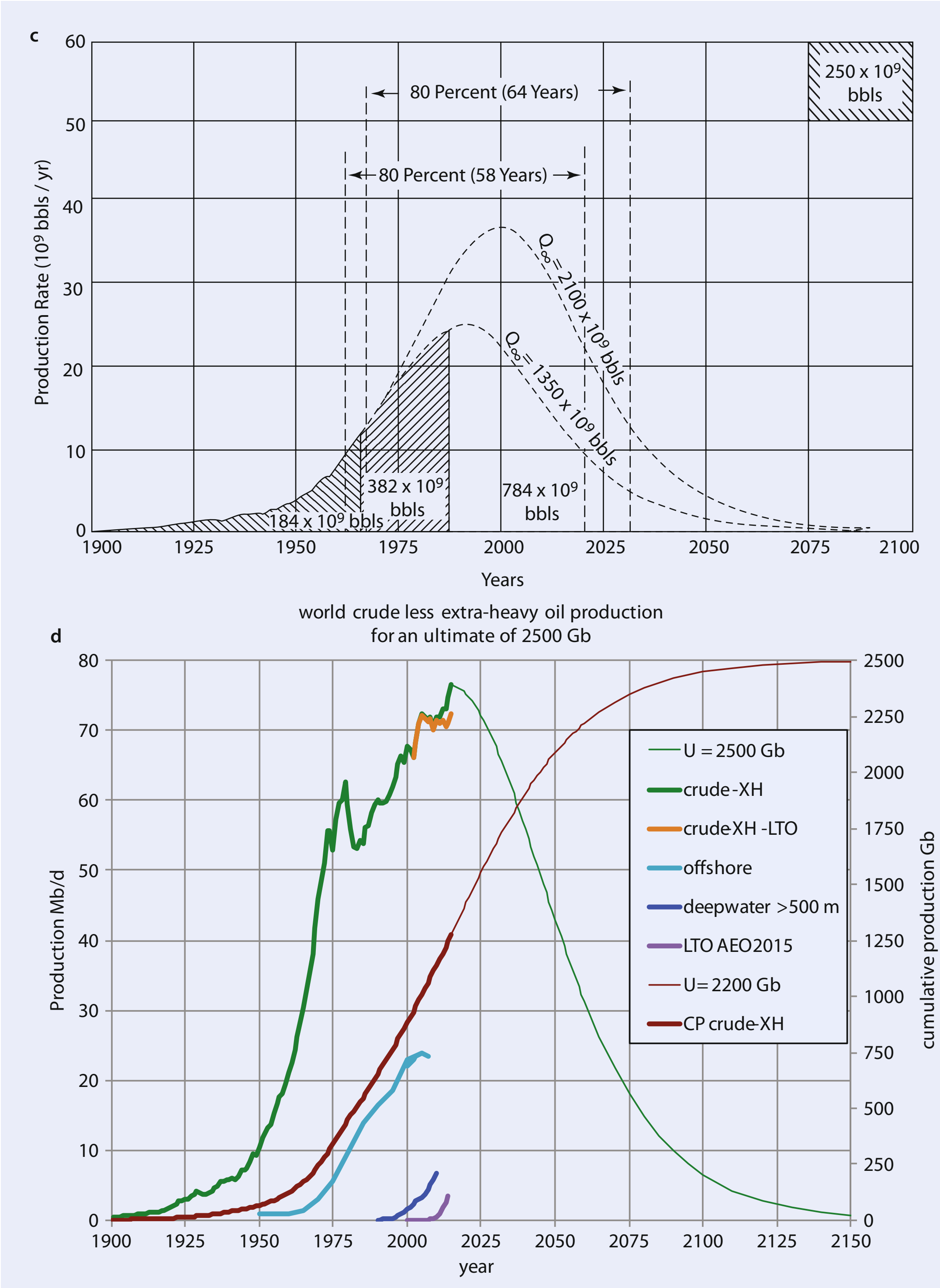


Fig. 8.5 (continued)

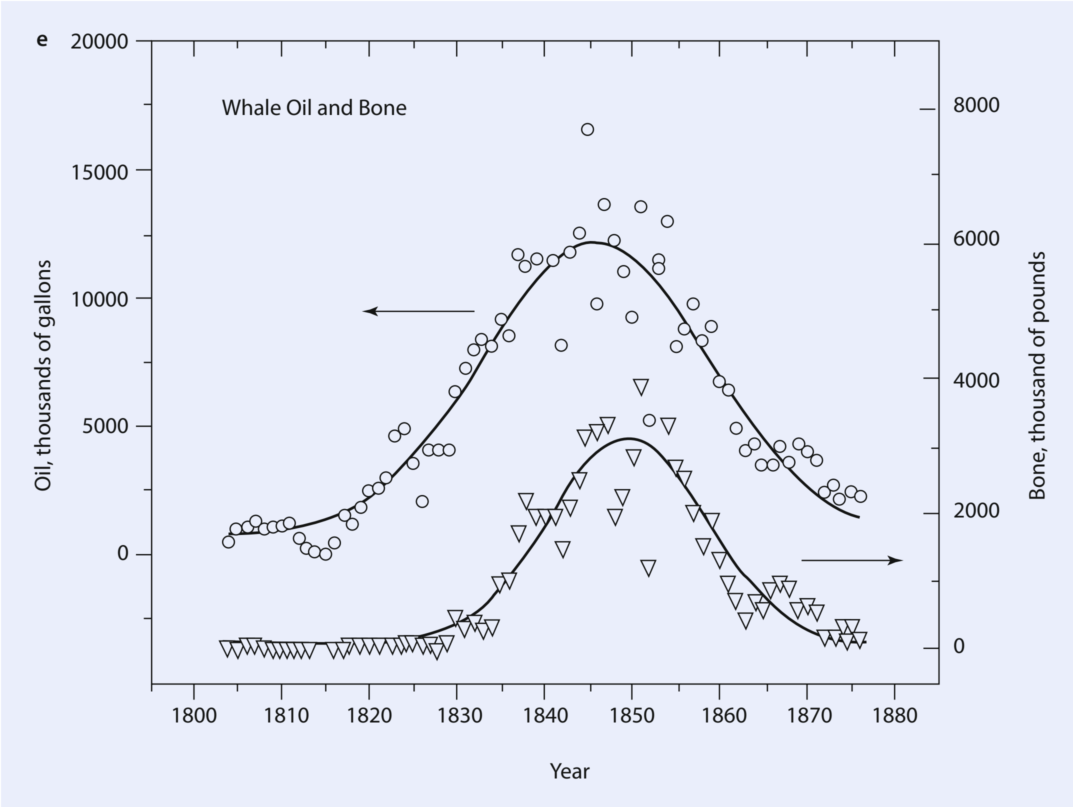


Fig. 8.5 (continued)

biologically derived ethanol) into the empirical estimates of "oil." If these are not included, then the production of conventional oil seems to be essentially flat since 2005 Fig. 8.5d.

It should be noted that the majority of oil-supply and oil price forecasts which we examined (such as that undertaken by Cambridge Energy Research Laboratory and with the possible exception of "post-peak Hubbert" analyses) had a poor track record, regardless of method. It is now a well-established fact that economic and institutional factors, as well as geology, were responsible for the US peak in production in 1970 [23, 24], forces that are explicitly excluded from the curve-fitting models. Thus, the ability (or the luck) of Hubbert's model (and its variants) to forecast production in the 48 lower states accurately should not necessarily be extrapolated to other regions.

On the other hand, one excellent study (in our opinion, but careful, Hall was an author!) by Hallock et al. made predictions for all major oil-producing countries assuming Hubbert curves and using low, medium- and high URR estimates from USGS [13]. They then returned 10 years later and examined the actual behavior of oil production vs. their predictions [16, 17]. They found that the vast majority of oil-producing nations followed a Hubbert curve; most had peaked by 2012 and most followed a path consistent with the USGS low (vs. medium or high) estimates of available oil (Fig. 8.8). Exceptions are several very large oil producers (e.g., Iraq, Iran] whose trajectory is still uncertain due to political events or for whom it is too early to tell. The actual data on global conventional oil production certainly shows at least an undulating plateau from 2005 to 2015 or so at the time of this writing and perhaps even a production peak (Fig. 8.5d). Certainly the old growth rate of 3–4% per year has slowed way down. This is astonishing given the previous continuous growth in production year after year from 1940 through the early 2000s, and this slowdown occurred during times of greatly increasing oil prices. Clearly Hubbert-type peaks have occurred for oil for many nations [17] and for other resources, such as whale oil and perhaps phosphorus (Fig. 8.5e).

So, why is global oil production decreasing or at least no longer increasing? The principle reason is that most oil production comes from very large oil fields (called "elephants"), and we have found very few elephants since the 1960s. Now these large oil fields are aging, and the production in many of these fields is declining by 2–10% a year. Thus while it is true that we are finding additional new oil supplies, these new fields are equal in volume to only about one-fifth of the existing fields, and hence a decline is expected (Fig. 8.6). According to Chris Skrebowski, editor of Petroleum Review, at least one-quarter of the 400 largest oil fields in the world are in decline, and it appears impossible that new oil discoveries, most of which are not large, can possibly make up for the decline in the elephants.

Economic forecasts have not fared well in explaining US oil production. In the period after the Second World War, oil production often increased as oil prices decreased, and vice versa (Fig. 8.4), a behavior that is exactly the opposite of predictions of conventional economic theory. Economic theory also assumes that oil prices will follow an "optimal" path toward the choke price—the price which is sufficiently high to cause the quantity of oil demanded to begin to fall to zero. Thereafter, at least in theory, the market signals a seamless transition to substitutes. In fact, even if such a path exists, prices may not increase smoothly because empirical evidence

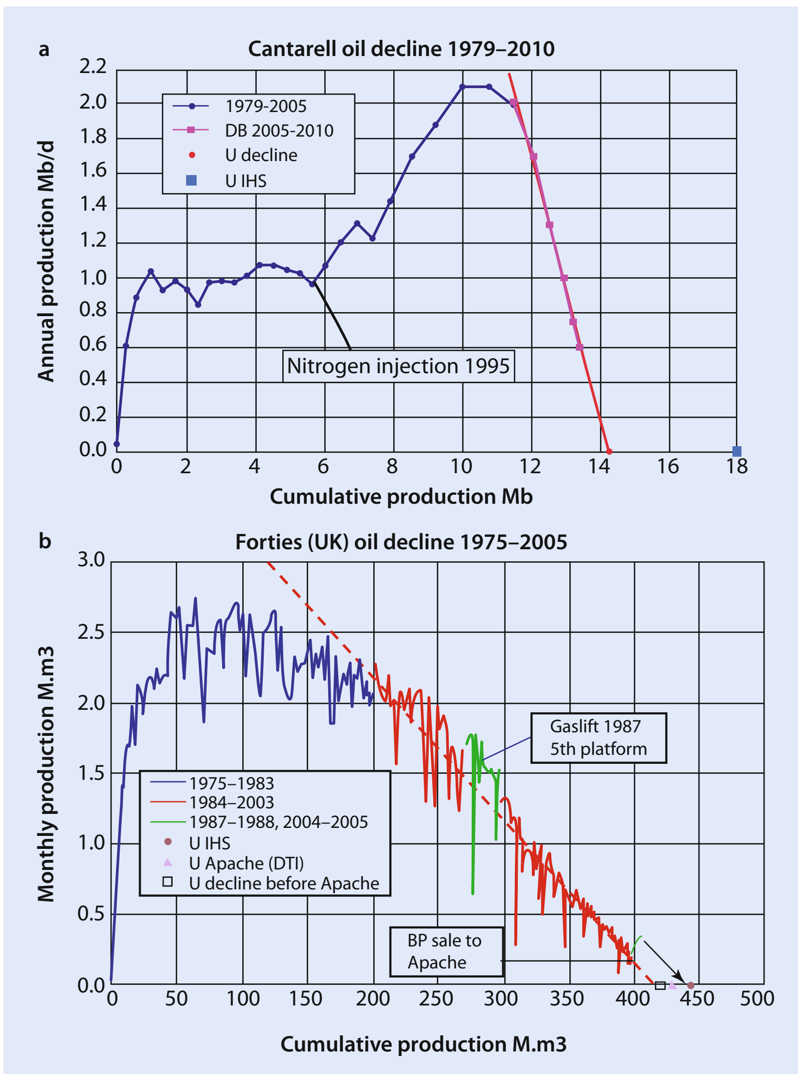
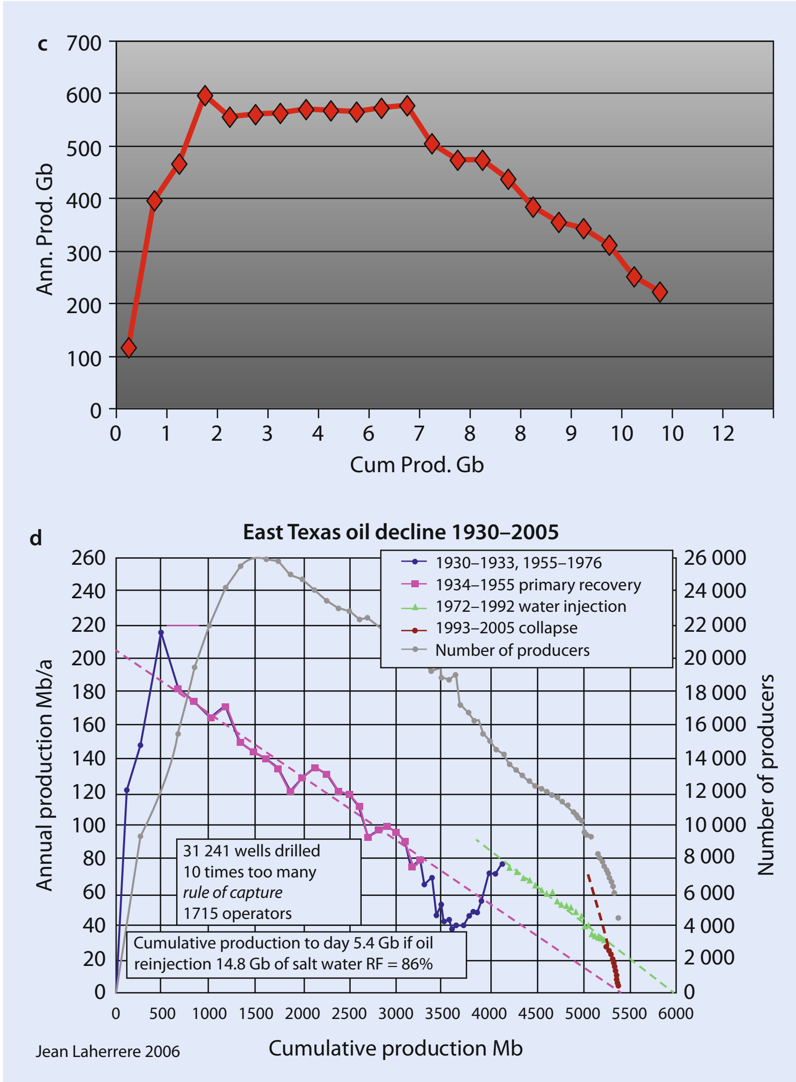


Fig. 8.6 Decline in the production of a number of important "elephants" (Source: Jean Laherrere). **a** Canterell, Mexico, once the world's second largest field. **b** The forties field in the North Sea. **c** Prudhoe field, the largest in the United States. **d** East Texas, the second largest field in the United States

Fig. 8.6 (continued)



indicates that producers respond differently to price increases than they do to price decreases [24]. In the presidential campaign of 2008, one often heard in response to the increased price of oil "drill, drill, drill!." In fact there is little evidence that there is any relation between drilling rate and the production of conventional oil and gas, with the exception of the early 1950s (Fig. 8.4). One way to think about this is that "Mother Nature holds the high cards." In other words oil production will be determined much more by what is geologically possible than by human efforts or economics [12]. Significant deviation from basic economic theory undermines the de facto policy for managing the depletion of conventional oil supplies—a belief that the competitive market will generate a smooth transition from oil. We see little evidence of this happening thus far.

Whatever the exact details or the dates of peak oil, it is clear that we are, in the words of Colin Campbell, in transition from the first half of the age of oil to the second half of the age of oil [25]. Each half is and will be equally oil dependent, but the difference will be that between an increasing quantity being used each year to a flat and then decreasing quantity.

8.7 Net Energy from Oil

Our view is that the question is not how much recoverable oil is left in the Earth. We agree that there is a great deal, possibly (but probably not) near the high end of the estimates. But what is missing from the debate is how much of that oil can be recovered with a significant, or perhaps any, net energy gain. These are old arguments about peak oil, but most assessments are made in the absence of net energy costs [26]. If we extrapolate essentially any time series analysis of the net energy returned from oil, all of them show (if present trends continue) a break-even point within decades. Thus we think we will reach the energy break-even point long before we are able to exploit the larger estimates of reserves given by, e.g., the USGS [13] (Fig. 8.7). In other words the total amount of oil in the ground is

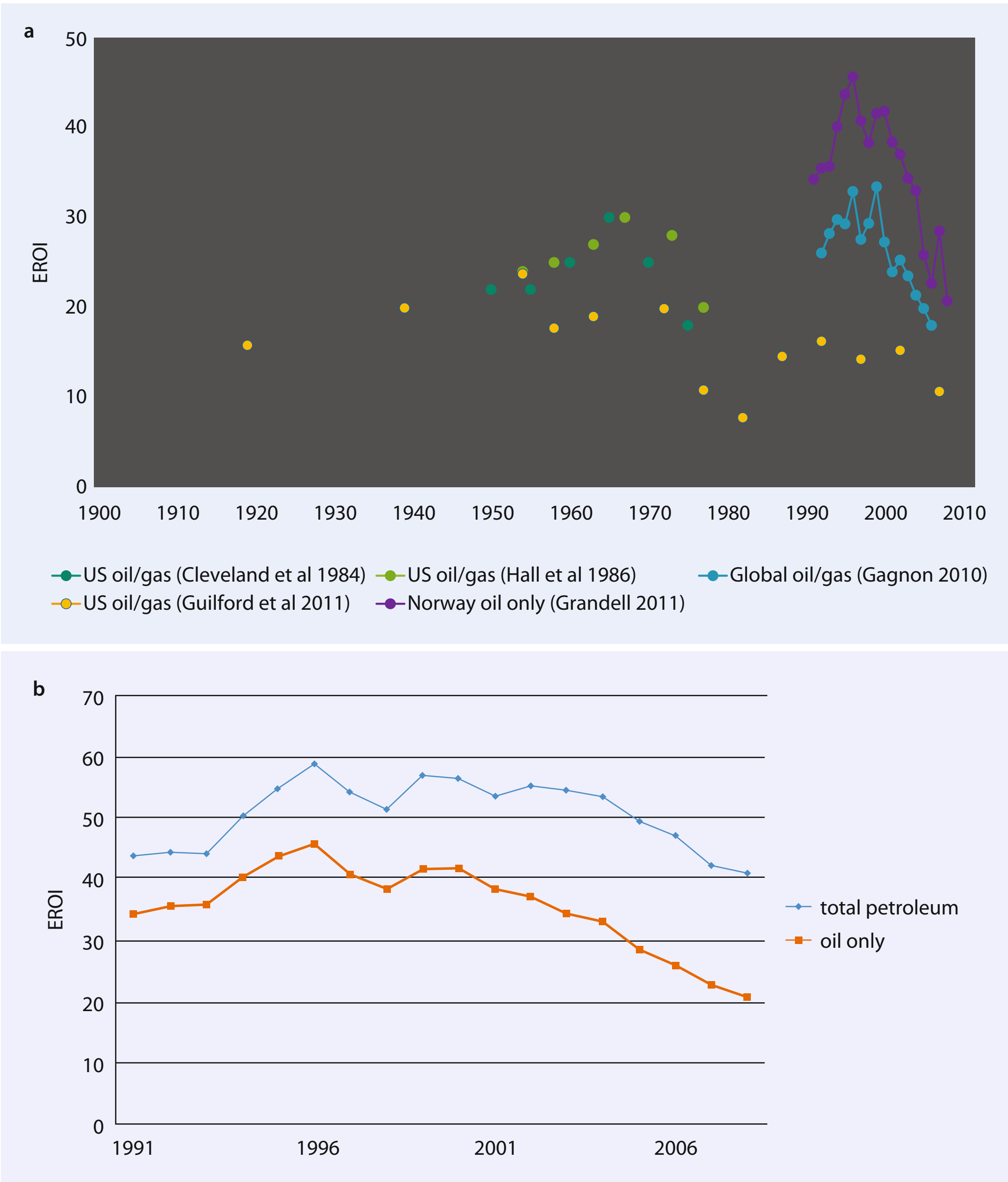


Fig. 8.7 **a** Three estimates for EROI for producing oil in the U.S., along with Estimates for Norway and for all independent oil companies (Source: Hall et al. 2014 and references therein). **b** Example of EROI for oil, and oil plus gas for a country, in this case Norway. One can see the effect of the development, and then gradual depletion, of the important North Sea oil fields. **c** EROI for all fuels for England, including only direct and also direct and indirect energy costs. (From Brand-Correa L.I., Brockway P.E., Copeland C.L., Foxon T.J., Owen A., Taylor P.G., (2017) Developing an Input-Output Based Method to Estimate a National-Level Energy Return on Investment (EROI). Energies 2017, 10(534)

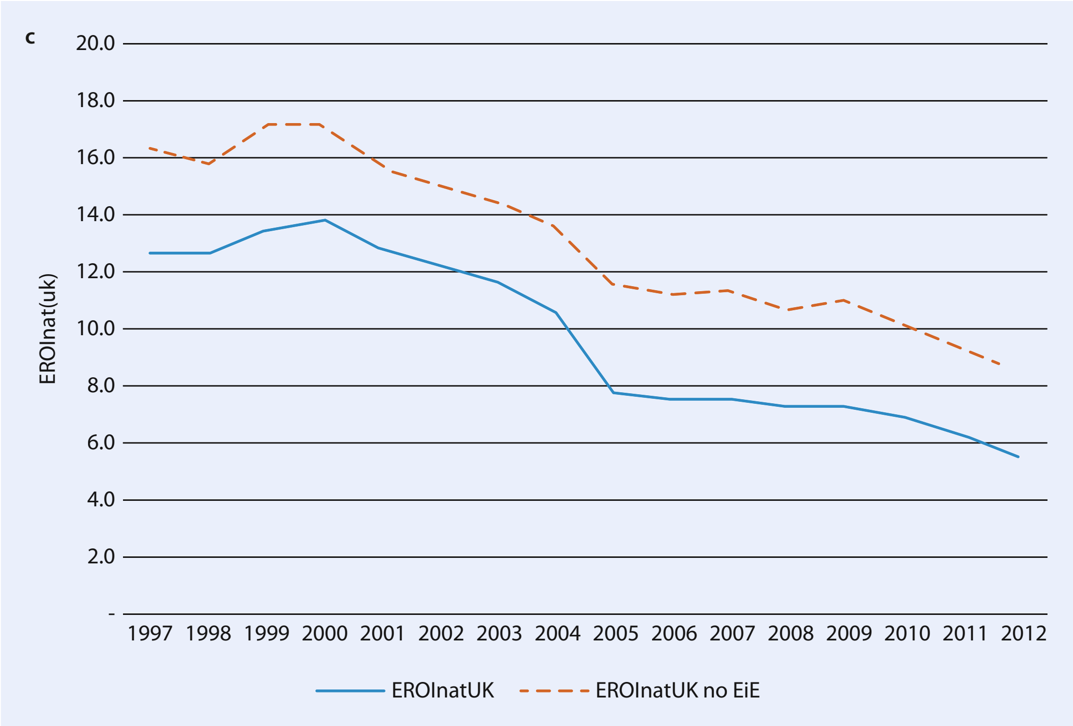


Fig. 8.7 (continued)

not a relevant number. Rather we need to know how much of that can be extracted with a significant net energy profit. This important issue of the energy cost of getting additional quantities of oil, and how that might influence URR, is given in ▶ Chap. 18.

Meanwhile most realistic projections of the availability of all fossil fuels indicate a peak in availability within one or two decades, even including unconventional fuels (e.g. Mohr et al. in 32; Fig. 8.9). Additionally, there may be additional limits to use imposed by efforts to limit climate impacts. However these issues are resolved, it is clear that plans for future economic growth cannot assume that, as in the past, the energy to allow this to happen will be available.

8.8 Geography of Oil

Oil is used by all of the nearly 200 nations of the world, but significant amounts are produced by only about 42 countries, 38 of which export important amounts. This number is declining because of the depletion of the once-vast resources of North and South America, the North Sea, Indonesia, and many other regions and owing to the increasing domestic use of oil by many of the exporters. The number of exporters outside the Middle East and the former Soviet Union will drop in the coming decades, perhaps sharply, which in turn will greatly reduce the supply diversity to the 160 or so importing nations [27]. Such an increase in reliance on West African, former Soviet Union, and especially Persian Gulf oil has many strategic, economic, and political implications. Much of the world's reserves are found in nations not known to be especially friendly to the United States or the West more generally, in part because of the West's long history of "boots on the ground" expropriation of oil or interference with the governments of producing countries. The EROI for discovering oil and gas in the United States has decreased from a value of more than 200:1 in the 1930s to less than 5:1 in the 2010's, and for production from about 30:1 in the 1970s to less than 10:1 today (Fig. 8.7). The enormously increasing demand for oil from China and their large reserves of money are also likely to have a large impact because the Chinese should have little trouble paying for their oil even as prices raise. On the other hand, the efficiency of using oil may be improving, so that in many OECD countries, there is little or no growth in the use of oil, or even a decline.

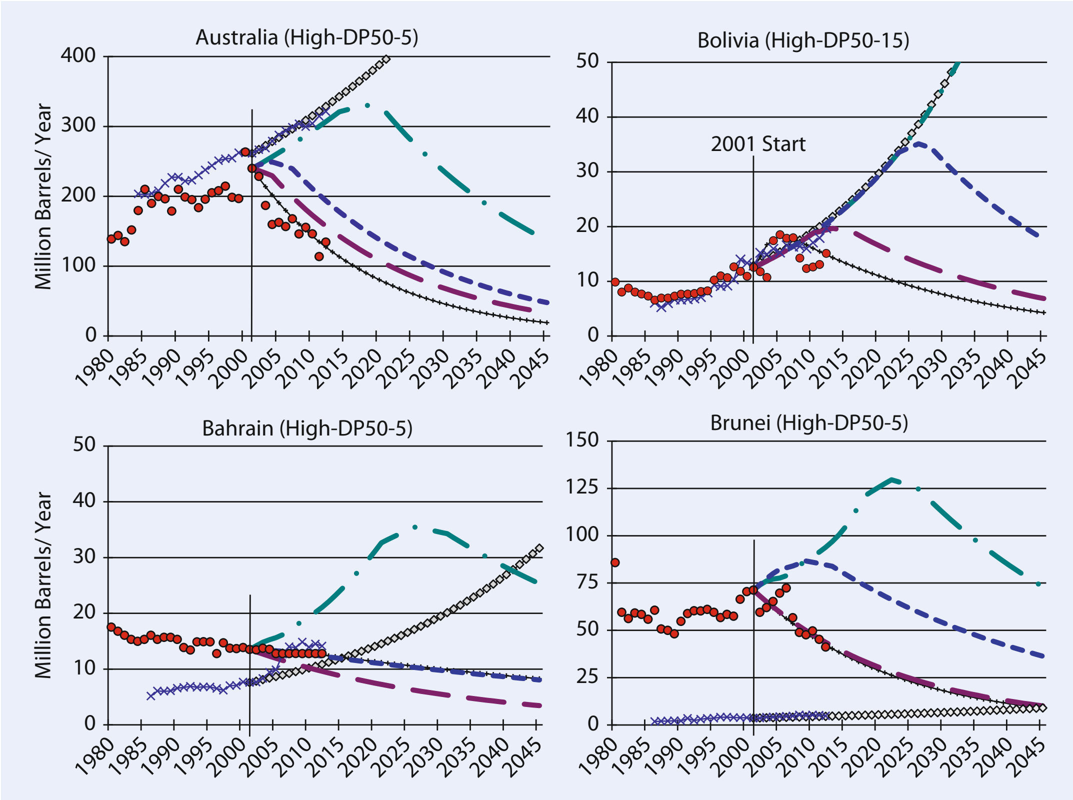


Fig. 8.8 Typical patterns of oil production for most oil-producing countries showing common patterns of Hubbert Curves for conventional petroleum. *Red dots* represent data. *Blue* and *purple lines* indicate predictions made in 2004 (*vertical line*) for low, medium and high estimates of oil reserves. Crosses and diamonds equal consumption. (Source Hallock et al. 2014; see for all other major producing countries)

8.9 Energy and Political Costs of Getting Oil

The future of oil supplies is normally analyzed in economic terms, but economic costs are likely to be dependent on other costs. In our earlier work [e.g., 26, 27], we summarized the energy costs of obtaining US oil and other energy resources and found, in general, that the energy returned on energy invested (EROI; see ▶ Chap. 18) tended to decline over time for oil and most other energy resources examined. This includes the energy cost of obtaining oil by trading (energy-requiring) goods and services for energy itself [26]. Likewise the EROI for the production of oil and gas globally has declined from about 36:1 in the 1990s to about 19:1 in 2006 [27]. In other words, with all of our super technology we can continue to get oil and gas, but the energy cost per barrel continues to increase as we deplete the best resources. This is also true for such estimates for other parts of the world, and we do know that both heavy oil in Venezuela and oil sands in Alberta require a very large part of the energy produced, as well as substantial supplies of hydrogen from natural gas, to make the oil fluid [28]. The very low economic cost of finding or producing new oil supplies in the Arabian Peninsula implies that it has a very high EROI value, which in turn supports the probability that production will be concentrated there in future decades. Alternative liquid fuels, such as ethanol from corn, have a very low EROI, perhaps not even a positive gain over the fossil fuels invested to plant and distill the alcohol [29]. An EROI of much greater than 1:1 is needed to run a society, because energy is also required to make the machines that use the energy, feed, house, train, and provide health care for necessary workers and so on (▶ Chap. 19).

No one who watches the news can fail to be aware of the importance of cultural and political differences between those nations that have the most oil and those that import it. How these factors will play out over the next few decades is extremely important, but also impossible to predict. Most of the remaining oil reserves are in Southern Russia, the Middle East, and North and West Africa, countries or regions with either Muslim governments or significant Muslim populations. For a long period, frustration and resentment have been building up among many Muslim populations, not least because of their perception that the main Western powers have failed to generate even-handed policies to address the conflicts in the Middle East over the past half century. Iranians still have vivid memories of the role the Central Intelligence Agency played in the overthrow of their democratically elected prime minister, Dr. Mohammed Mossadeq, on behalf of the Anglo-Iranian Oil Company (now BP). Another factor is that the huge revenues earned by the oil-exporting nations have been very unevenly distributed among their respective populations, adding to internal and external pressure to adopt a more equitable approach to human development. The "Arab spring" of 2011, where new pressures for governmental reform have greatly increased the instability of many Middle Eastern oil-producing nations—and oil prices for at least a few years. Much of the unrest stems in part from the failure, and some would say impossibility, of these economies to produce sufficient jobs and even food for their growing populations. Suffice it to say that there will continue to be high risks of international and national terrorism, overthrow of existing governments, and deliberate supply disruption in the years ahead [Ahmed in 35]. In addition, exporting nations may wish to keep their oil in the ground to maintain their target price range. Thus, there are considerable political and social uncertainties that could result in less oil being available than existing models predict.

8.10 Deep Water and Extreme Environments

Although considerable uncertainty remains about how much oil we will extract, eventually one thing is clear: oil is getting harder and harder to find [25–33]. This can be seen by the increasing dollar, energy, and environmental cost of getting oil and by the fact that we are undertaking major exploration and development in areas (such as very deep ocean) that were thought too difficult and expensive just a decade ago, so that half of new US drilling effort now takes place far offshore. There have been amazing developments in technology that have allowed this new exploration: drilling ships unanchored to the bottom kept in place by GPS systems and huge thrusters, drill strings that go down through 2000 meters of ocean and then 5000 meters or more of rock, and so on. The Deepwater Horizon oil spill of 2010 in the Gulf of Mexico has brought all these operations to the attention of the public, and one of the first questions asked was: why are we working in such a difficult and potentially dangerous environment? The answer is that while the oil fields that have been discovered at these depths appear to be the only large fields left that have not yet been exploited—in other words we went after the easy stuff first and left the most difficult until later. So if we are to continue to have oil, we need to undertake these expensive and risky operations. The most interesting analysis of this issue is by Tainter and Patzek [34], where the authors ask whether we have expanded the complexity of our American "empire" to the point that the energy cost of getting energy itself to the center of the "empire" exceeds the gain from that energy. They point out that this may be analogous to other ancient empires (such as Rome) which expanded until they reached the limits of managing the complexity necessary for maintaining the society [35]. A similar analysis might be made of our large efforts in militarization in support of maintaining oil flows.

A final important issue relating to the development of new oil or its possible substitutes has been put forth by Robert Hirsch and his colleagues in several extremely insightful papers [36, 37]. Their basic point is that a critical element in finding a substitute for petroleum (if indeed a substitute exists) is time—that is, even if a workable substitute can be found (and they examine, e.g., shale oil, biomass fuels and even greatly increasing the gas mileage of our vehicles) and assuming that government (or private) programs can be developed and money is no object that it would take decades simply to scale up the approach. In other words if we could maintain liquid-fuel use at the level of the peak of oil (perhaps about what we have in 2005–2010), it would take decades to construct the needed infrastructure. The importance of trucks to our present way of life and the implications of not having enough petroleum to run them have been wonderfully assessed by Alice Friedemann [38]. It is a very sobering perspective.

8.11 How About Natural Gas?

Petroleum usually means liquid and gaseous hydrocarbons and includes oil, natural gas liquids, and natural gas. Thus a chapter on oil is incomplete without some consideration of natural gas. Natural gas is often found associated with oil, although it has other possible sources, including coal beds and organic-rich shales. Oil is a natural hydrocarbon where the original plant material, often composed of hundreds to thousands of carbons linked together, has been cracked or broken by geological energies to a length of (ideally) eight carbons (octane). If the cracking continues to the extreme, the carbon bonds are broken completely to a length of one carbon, usually surrounded by four hydrogen molecules, a gas called methane. This makes gas an ideal fuel because oxidizing hydrogen releases more energy and releases less carbon dioxide than oxidizing carbon. Methane is much more easily obtained, stored, and moved than is hydrogen, partly because the much smaller hydrogen molecule leaks more easily. When natural gas is held in a tank, some heavier fractions fall out as natural gas liquids, and these materials can be used, essentially, either directly or as inputs to refineries. Natural gas was once considered an undesirable and dangerous by-product of oil production, and it was flared into the atmosphere. With time its commercial value was recognized and a complex pipeline system evolved. Now natural gas is more or less tied with coal as the second most important fuel in the United States and the world. An important question is: if oil falters can natural gas take over its role? It can even be used to propel vehicles with minimal changes to the engine, and it has essentially displaced the role of oil in electricity production. It is not as energy dense or transportable as oil, but it comes close, and because it is clean it has many special uses such as for baking and as a feedstock for plastics and nitrogen fertilizer.

Beginning in 2010 there was a great deal of excitement and debate about whether "unconventional" natural gas from, e.g., the Marcellus shale, can provide an energy renaissance for the United States. While it has been known that considerable gas exists in association with certain shales, it was too difficult to get out because the shale formations were too thin and a conventional vertical well simply passed through the formation without intercepting much gas. New technologies, including horizontal drilling and fracturing or "fracking" the rocks with very high-pressure water, have allowed considerable amounts of gas to be produced. But since the environmental impacts are barely known and possibly large, and tens of thousands of wells are needed to get a significant amount of gas, there is large controversy about the degree these wells should be drilled. Something less well known is that most of the gas in those areas we know best (e.g., the Barnett shale in Texas) comes from a relatively few "sweet spots" and that the total regional production may go through most of a full Hubbert cycle in only 15 years. Meanwhile conventional gas production has peaked and dropped off to less than half the peak, so that so far the unconventional gas of all kinds is simply compensating for the drop-off of conventional gas (◘ Fig. 8.10). Thus natural gas is likely to be very important as oil production and availability decline, and it will extend the petroleum age by a few decades. But then that too will be gone, and the United States will be left with little domestic production of domestic oil or gas. The younger people reading this book will have to deal with the decline and even termination of the petroleum age (◘ Fig. 8.9).

8.12 The Future: Other Technologies

The world is not about to run out of hydrocarbons, and perhaps it is not going to run out of oil from unconventional sources any time soon. What will be scarce is cheap petroleum, the kind that allowed industrial and economic growth. What is left is an enormous amount of low-grade hydrocarbons, which are likely to be much more expensive financially, energetically, politically, and especially environmentally. As conventional oil becomes less available, society has a great opportunity to make investments in different sources of energy, perhaps freeing us for the first time from our dependence on hydrocarbons. There are a wide range of options, and an equally wide range of opinions, on the feasibility and desirability of each. Nuclear power faces formidable obstacles. Experience of the past several decades has shown that electricity from nuclear power plants can be a reliable and mostly safe source of electricity, although expensive form of power, when all public and private costs are

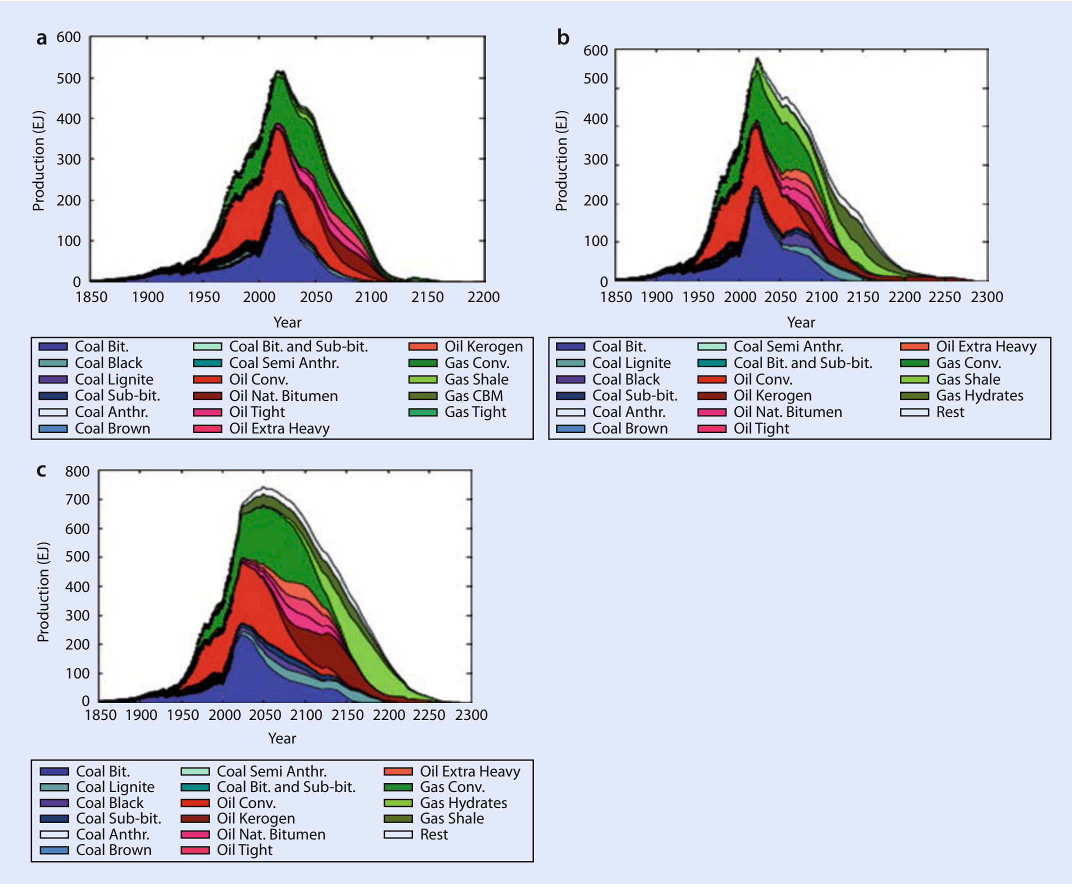


Fig. 8.9 Data on, and prediction of future usage of all fossil fuels using low (**a**), "best" (**b**) and high (**c**) estimates of supplies (From Mohr et al. 2015). Similar patterns have been found by other investigators

considered. The earthquake-tsunami-induced accident at Fukishima may make continued expansion unlikely in many nations. Other unresolved issues include that nuclear power generates high-level radioactive wastes that remain hazardous for thousands of years, possible nuclear weapons proliferation, and whether there is enough uranium to allow a significant contribution to global energy supplies. These are high costs to impose on future generations. Even with improved reactor design, the safety of nuclear plants remains an important concern. Can commercial nuclear power be divorced fully from nuclear weapons proliferation? Can these technological, economic, environmental, and public safety problems be overcome? Can new reactors using thorium fuel be created that decrease the problem of dangerous by-products generated from uranium while expanding the fuel supplies? These questions remain unanswered while we increase our use of fossil fuels essentially every year.

Renewable energies present a mixed bag of opportunities. Some argue that they have clear advantages over hydrocarbons in terms of economic viability, reliability, equitable access, and especially environmental benefits. But nearly all suffer from very low energy return on investments compared to conventional fossil fuels. In favorable locations, wind power has a high EROI (18:1 or more). The cost of photovoltaic (solar electric) power has come down sharply, making it a viable alternative in areas without access to electricity grids, but the EROI remains relatively low, perhaps only 4:1 or less, when considered on a systems level although some argue that because the input tends to be fossil fuel and the output electricity the realized EROI is considerably higher [39]. Both of these solar energies require very expensive backups or transmission systems to compensate for intermittent production, as they are available only 20–30% of the time. With proper attention to

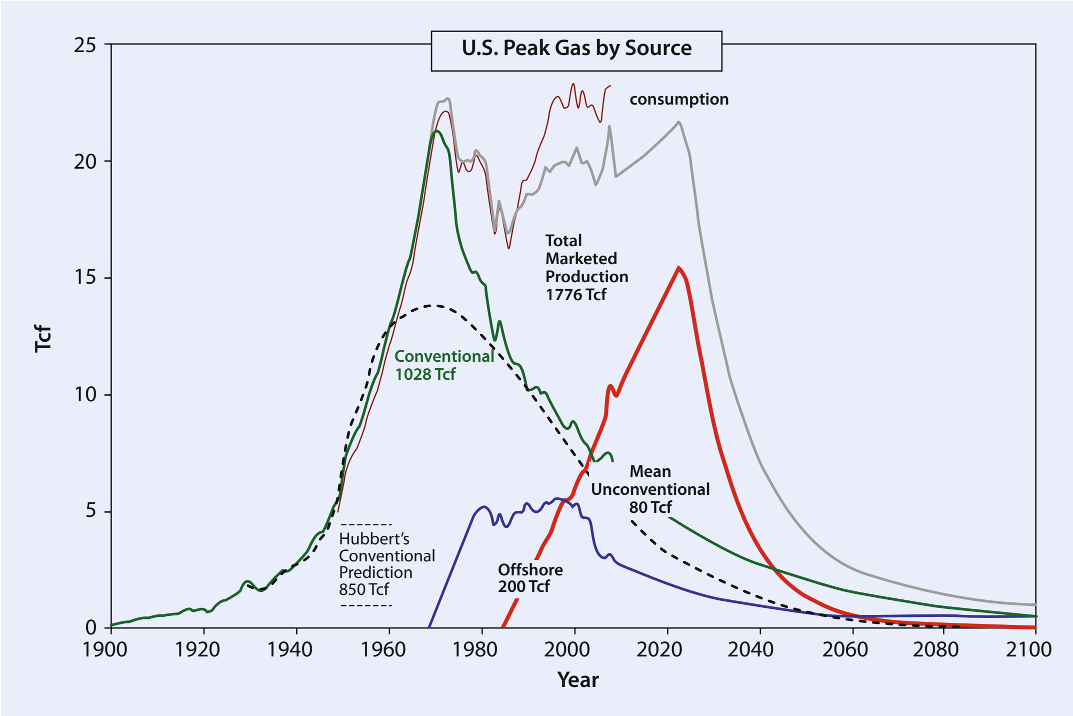


Fig. 8.10 Patterns of past and projected production of conventional, deepwater, and unconventional (e.g., shale gas) US natural gas (Source: Bryan Sell, Sustainability, in 2011)

environmental concerns, biomass-based energy generation is competitive in some cases relative to conventional hydrocarbon-based energy generation. At present liquid-fuel production from grain has a relatively low EROI [34, 35]. Hydrogen, advocated by many, is an energy carrier, not an energy source, and thus requires some kind of fuel to be used to split water or run some other process to generate the hydrogen. Additionally there are many problems to overcome because the small molecules leak easily and are hard to store. Hydrogen generated from renewable energy sources or electricity-driven hydrolysis is currently expensive for most applications, but it merits further research and development.

A disquieting aspect of all these alternatives, however, is that as energy delivery systems (i.e., including backups, transmission, etc.), they all have a much lower EROI than the fossil fuels we would like them to replace, and this is a major reason for their relatively low economic feasibility in most applications [32]. This may be changing rapidly now, we shall see. But going to half renewables would be an extremely tough nut to crack. Subsidies and externalities, social as well as environmental, add difficulties to this evaluation but are poorly understood or summarized. This presents a clear case for public policy intervention that would encourage a better understanding of the strengths and weaknesses of renewable (and traditional) forms of energy. Policy intervention, in concert with ongoing private investment and also markets, may be necessary to speed up the process of sorting the wheat from the chaff in the portfolio of renewable energy technologies, necessary if for no other reason than to protect our atmosphere.

8.13 The Social Importance of These Supply Uncertainties

Many once-proud ancient cultures have collapsed, in part, because of their inability to maintain energy resources and societal complexity [35]. Our own civilization has become heavily dependent on enormous flows of cheap hydrocarbons, partly to compensate for other depleted resources (e.g., through fertilizers and long-range fishing boats), so it seems important to assess our main energy alternatives. Oil is quantitatively and qualitatively most important. Investments in oil have continued to increase, but supply remains flat and is likely to decrease. Some of the most promising new oil fields have turned out to be very disappointing [32, 33]. Global findings in 2016 were only about 10% of global use. If indeed we are approaching the oil scarcity that some predict, it is barely reflected in oil prices, and few investments in alternatives are being made at anything like the scale required to replace fossil hydrocarbons—if indeed that is possible. Unfortunately the majority of decision-makers hold on to the fantasy that the market has resolved this issue before and will do so again. Further, an increasing number of US citizens believe that government programs are too ineffective to resolve any problem, including energy problems. We view this is a recipe for disaster. It is enhanced by the failure of science to be used as fully, effectively, or objectively as it should be. Failures in proper government funding for good energy analysis have led to the dominance of "science" in the media and decision-making whose role is basically to support the predetermined position of those who support it. In 2017 the state of official oil-supply modeling is in some ways no different than it was in Hubbert's time: a wide range of opinion exists and there is little or no objective and reliable overview.

This issue is critical at this point in time because if civilization is to survive the next 50 years, enormous new investments are necessary in whatever we will need to replace existing flows of conventional oil and gas and even coal. Energy costs now are only for the costs to extract fuels from existing reserves, not to come up with replacements once those fuels are gone. As energy prices increase, citizens are probably not going to be too excited to pay even more for a program to develop the research and infrastructure to generate replacement fuels, even if we knew what they should be. According to one of our best energy analysts, Vaclav Smil, at this time there seems to be few really good options except to decrease our appetites for energy [40].

What can science do to help resolve this uncertainty? Our principal conclusion is that these critical issues could be and should be the province of open scientific analysis in visible meetings where "all sides" attend and argue and where financial resources are provided to objective analysts to reduce uncertainty and understand different assumptions. This analysis should be informed by professionalism, the peer review process, statistical analysis, hypothesis generation and testing, and so on, rather than by simply the opinions of the experts one chooses or the quips on the blogosphere. These issues should be the basis of open competitive government grant programs, graduate seminars, and even undergraduate courses in universities, and our courses in economics should become at least as much about real biophysical resources, such as hydrocarbon reserves, as about market mechanisms. Also, we need to think much harder about the alternatives, including their energy cost of implementation and also the need to develop a lower energy-using society. None of this appears to be part of the plans of any existing government or governmental agency in the United States or most other industrial countries.

Questions

1. What is meant by the phrase "the first half of the age of oil"?
2. What are energy-dense materials?
3. We have been told that we live in an "information age." Argue for or against that statement.
4. Is peak oil a fact or a concept? Defend your view.
5. What does EUR mean? How is it related to peak oil?
6. What were Hubbert's basic ideas?
7. If there are huge amounts of oil left in the Earth, does this imply adequate supplies for the foreseeable future? Why or why not?
8. How is natural gas related to oil?
9. What does "cheap oil" mean relative to the remaining oil we might be able to extract from the Earth?
10. With many alternatives, why do you think that we have continued to rely so much on oil and other hydrocarbons?

Acknowledgments We thank S. Ulgiati, R. Kaufmann, Jean Laherrère and C. Levitan for discussions.

References

1. Derived, with substantial modifications and permission from Hall, C., P. Tharakan, J. Hallock, C. Cleveland, and M. Jefferson. 2003. Hydrocarbons and the evolution of human culture. *Nature* 426: 318–322. Updates on EROI are available in a special issue of the Journal Sustainability (2011) and Hall, C.A.S., J.G. Lambert, S.B. Balogh. 2014. EROI of different fuels and the implications for society. *Energy Policy* 64: 141–152.

2. Munasinghe, M. 2002. The sustainomics transdisciplinary meta-framework for making development more sustainable: Applications to energy issues. *International Journal of Sustainable Development* 5: 125–182.
3. Interlaboratory Working Group. 2000. Scenarios for a clean energy future. Lawrence Berkeley National Laboratory LBNL-44029, Berkeley. http://www.ornl.gov/ORNL/Energy_Eff/CEF.htm.
4. Hall, C.A.S., D. Lindenberger, R. Kummel, T. Kroeger, and W. Eichhorn. 2001. The need to reintegrate the natural sciences with economics. *BioScience* 51: 663–673.
5. Tharakan, P.J., T. Kroeger, and C.A.S. Hall. 2001. Twenty-five years of industrial development: a study of resource use rates and macro-efficiency indicators for five Asian countries. *Environmental Science & Policy* 4: 319–332.
6. Kaufmann, R.K. 2004. The mechanisms for autonomous increases in energy efficiency: a cointegration analysis of the US energy/GDP ratio. *The Energy Journal* 25: 63–86; Wiedmann, T. O., Schandl, H., Lenzen, M., Moranc, D., Suh, S., West, J. and Kanemotoc, K. (2012). The material footprint of nations. Proceedings of the National Academy of Sciences of the United States of America. 112, 10, 6271–6276.
7. Gupta, A. 2015. Energy and GDP for various countries and the world.
8. Smulders, S., and M. de Nooij. 2003. The impact of energy conservation on technology and economic growth. *Resource Energy Economics* 25: 59–79.
9. Sadorsky, P. 1999. Oil price shocks and stock market activity. *Energy Economics* 21: 449–469. See also Hall, C.A.S., Groat, A. 2010. Energy price increases and the 2008 financial crash: a practice run for what's to come? The Corporate Examiner. 37: No. 4-5:19–26.
10. Lambert, J., C.A.S. Hall, S. Balogh, A. Gupta, and M. Arnold. 2014. Energy, EROI and quality of life. *Energy Policy* 64: 153–167.
11. Tissot, B.P., and D.H. Welt. 1978. *Petroleum formation and occurrence*. New York: Springer-Verlag.
12. Campbell, C.J., and J.H. Laherrère. 1998. The end of cheap oil. *Scientific American* 278: 78–83; see also Jean Laherrère's discussion of uncertain definitions of oil reserves at ASPO France and his (and other's) papers at http://theoilage.org/the-oil-age-journal/.
13. United States Geological Survey (USGS). 2003. The world petroleum assessment 2000. www.usgs.gov; Energy Information Administration, US Department of Energy. 2003. International outlook 2003. Report no. DOE/EIA-0484(2003), Table 16 at http://www.eia.doe.gov/oiaf/ieo/oil.html.
14. ———. 2000. United States Department of Energy long term world oil supply. http://www.eia.doe.gov/pub/oil_gas/petroleum/presentations/2000/long_term_supply/index.htm.
15. Hubbert, M.K. 1969. *Energy resources (Report to the Committee on Natural Resources)*. San Francisco: W.H. Feeeman.
16. Hallock, J., P. Tharakan, C. Hall, M. Jefferson, and W. Wu. 2004. Forecasting the availability and diversity of the geography of oil supplies. *Energy* 30: 207–201.
17. Hallock, J.L., Jr., W. Wu, C.A.S. Hall, and M. Jefferson. 2014. Forecasting the limits to the availability and diversity of global conventional oil supply: Validation. *Energy* 64: 130–153.
18. Lynch, M.C. 2002. Forecasting oil supply: Theory and practice. *The Quarterly Review of Economics and Finance* 42: 373–389.
19. Bartlett, A. 2000. An analysis of U.S. and world oil production patterns using Hubbert-Style curves. *Mathematical Geology* 32: 1–17.
20. Brandt, A.R. 2007. Testing Hubbert. *Energy Policy* 35: 3074–3088.
21. Kaufmann, R.K., and L.D. Shiers. 2008. Alternatives to conventional crude oil: When, how quickly, and market driven? *Ecological Economics* 67: 405–411.
22. Nashawi, I.S., A. Malallah, and M. Al-Bisharah. 2010. Forecasting world crude oil production using Multicyclic Hubbert model. *Energy & Fuels* 24: 1788–1800.
23. Kaufmann, R.K., and C.J. Cleveland. 2001. Oil Production in the lower 48 states: Economic, geological and institutional determinants. *The Energy Journal* 22: 27–49.
24. Kaufmann, R.K. 1991. Oil production in the lower 48 states: Reconciling curve fitting and econometric models. *Resources and Energy* 13: 111–127.
25. Hall, C.A.S., and Ramirez-Pasualli, C. 2013. *The first half of the age of oil. An exploration of the works of Colin Campbell and Jean Laherrere*. New York: Springer.
26. Cleveland, C.J., R. Costanza, C.A.S. Hall, and R. Kaufmann. 1984. Energy and the United States economy: A biophysical perspective. *Science* 225: 890–897.
27. Hall, C.A.S., J.G. Lambert, and S.B. Balogh. 2014. EROI of different fuels and the implications for society. *Energy Policy* 64: 141–152.
28. Hall, C.A.S. (ed). 2011. Special issue of journal "Sustainability" on EROI Sustainability: 2011 (3): 1773–2499. Includes: Guilford, M., C.A.S., Hall, P. O'Conner, and C.J., Cleveland. 2011. A new long term assessment of EROI for U.S. oil and gas: Sustainability: Special Issue on EROI. Pages 1866–1887; and Sell, B., C.A.S, Hall, and D., Murphy. 2011. EROI for traditional natural gas in Western Pennsylvania. Sustainabilities: Special Issue on EROI. 2011. Pages 1986–2008; Hall, C.A.S. 2017. Energy Return on Investment: A unifying principle for Biology, Economics and sustainability. Springer Nature, N.Y.
29. Gagnon, N., C.A.S. Hall, and L. Brinker. 2009. A preliminary investigation of energy return on energy investment for global oil and gas production. *Energies* 2 (3): 490–503.
30. Poisson, A., and C.A.S. Hall. 2013. Time series EROI for Canadian oil and gas. *Energies* 6 (11): 5940–5959.
31. Murphy, D.J., C.A.S. Hall, and R. Powers. 2011. New perspectives on the energy return on investment of corn based ethanol. *Environment, Development and Sustainability* 13 (1): 179–202. Giampietro, M. and K. Mayumi. 2009. The biofuel delusion. Earthscan, London.
32. Masnadi M.S. and Brandt, A.R. (2017). Energetic productivity dynamics of global super-giant oilfields, Energy & Environmental Science, 10, 1493–1504; Mohr, S.H., Wang, J., Ellem, G., Ward, J. and Giurco, D. (2015). Projection of world fossil fuels by country. Fuel 1(141), 120–135.
33. Hakes J. 2000. Long term world oil supply: A presentation made to the American Association of Petroleum Geochemists, New Orleans, Louisiana. http://www.eia.doe.gov/pub/oil_gas/petroleum/presentations/2000/long_term_supply/index.htm.

34. Tainter, J., and T. Patzek. 2011. *Drilling down: The Gulf oil debacle and our energy dilemma*. New York: Springer.
35. Tainter, J. 1988. *The collapse of complex systems*. Cambridge: Cambridge University Press; Ahmed, Nafeez. 2016. Failing States, Collapsing Systems: Biophysical Triggers of Political Violence. SpringerNature.
36. Hirsch, R., R. Bezdec, and R. Wending. 2005. Peaking of world oil production: impacts, mitigation and risk management. U.S. Department of Energy. National Energy Technology Laboratory. Unpublished Report.
37. Hirsch, R. 2008. Mitigation of maximum world oil production: Shortage scenarios. *Energy Policy* 36: 881–889.
38. Freidemann, A. 2016. *When the trucks stop running. Energy and the future of transportation*. New York: Springer.
39. Prieto, P., and C.A.S. Hall. 2012. *Spain's photovoltaic revolution: The energy return on investment*. New York: Springer. The issue of what is the proper EROI for solar fuels is quite contentious: see e.g. Hall, Charles A.S. 2017. Will EROI be the primary determinant of our economic future? The view of the natural scientist vs the economist. Joule 1(2):3–4.
40. Smil, Vaclav. 2011. Global energy: The latest infatuations. *American Scientist* 99: 212–219.

Energy, Economics, and the Structure of Society

The development of a fossil fuel–intensive society had many obvious impacts on the ability of the economy to produce wealth and to provide humans with an increased material standard of living. What is perhaps less obvious is that this development also enabled and forced many other changes to society by giving economic and political power to those who controlled access to this energy. This section examines some of those changes as they occurred in the United States (and many other areas of the world).

Contents

The Petroleum Revolution II: Concentrated Power and Concentrated Industries

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_9

9.1 Introduction

In earlier chapters, we developed the link between the historical development of energy sources and the development of human society. More energy has allowed humans to do more work, including that of producing more wealth and more humans. We use the joule, for those not steeped in physical science, as the standard measure of energy. One joule is the amount of energy needed to accelerate a mass of 1 kg by a constant force of 1 Newton for a distance of 1 meter along a horizontal frictionless surface. A joule is equal to about one-quarter of a calorie. Our more familiar unit is the kilocalorie (often written as Calorie) and is found, for example, on the back of food packages. One kilocalorie is 1000 calories, equal to about 4 kJ. Thus, if you consume a drink that says it has 100 kilocalories, you will have consumed 418 kJ. Later, in ▶ Chap. 15, we explore the relation between energy and power from a scientific perspective. Power is the rate of doing work and is commonly measured in watts. From the standpoint of physics, *power* is energy used or expended per unit of time or the work that power causes or allows to be done. The most common unit of power is the watt, where 1 W = 1 J/s.

9

But power means something else in a political and economic context, and here we want to extend the definition to ways in which power is used in the social sciences and day-to-day life more generally. English can be a difficult language for many people to learn because the same word, in this case power, can mean very different things. According to *The Oxford English Dictionary*, power means, in addition to the sense used in physics, the "possession of control or authority over others, or a movement to enhance the status or influence of a specified group, lifestyle, etc." This definition seems equally appropriate to social realms, and this chapter reflects both perspectives—physical and social—on power. In most cases there is physical power behind any economic or social power. The latter cannot be measured as clearly and explicitly as can physical power, but all are clearly related. When the physical power to run an economy was solar, the economic and political power tended to be widely distributed in the hunting and gathering era. Then, after the Neolithic transition, land ownership became very concentrated in the hands of an aristocracy. People who owned land intercepted lots of solar power and tended to have a lot of political power. The increased use of fossil fuels, which are concentrated energy, tends to concentrate both economic and political power in less area. Hence in the nineteenth and early part of the twentieth century, political power tended to pass from the landed gentry to those who owned factories in cities and then increasingly to those who owned the energy sources.

9.2 Petroleum and Economic Concentration

In ▶ Chap. 6 we developed the concept that control over energy and power, in the scientific sense, and led to increased output and an increase in status, wealth, and power in the social sense. The development of petroleum fuels allowed a previously unimaginable increase in the ability to do physical work as well as unheard of concentrations of economic power. This is true both for nations (in the United States throughout the last century, Britain and Germany during the previous century) and for corporations or individuals. There has never been, and probably never will be, an energy source as concentrated as petroleum, with the exception of fissionable elements such as uranium and plutonium. At the same time, there have been few, if any, industries as concentrated in the economic sense as "the old house" of Standard Oil. Concentrated economic and physical power emerged together in the United States and elsewhere. During the past century, many hundreds of small oil companies coalesced into the "seven sisters" that essentially controlled global exploration and production. The revolution in industrial structure and large monopolized firm occurred during the same historical time period and not merely by coincidence.

Economic concentration is synonymous with the process of monopolization. We use the term *monopoly* not in the narrow context of an industry that consists of a single seller, but in the broader meaning of an industry being dominated by a few very large companies. (The technical term for this is *oligopoly*.) In most of the developed world, monopolized or concentrated industry is neither rare nor an anomaly. This is true despite the textbook model of businesses favored by mainstream economists: competitive industries of many powerless firms operating in impersonal markets that allocate resources with maximum efficiency. Rather economic concentration is an explicit strategy on the part of firms themselves to control their economic environments and protect their opportunities to achieve profits in the long run [1]. The economic power controlled by a firm is regularly threatened by a host of internal and external forces: new products and markets, technological change, government regulation, and most importantly the rise of excess capacity and ruinous price competition. If a firm expands its productive capacity and then fails to sell the products, or can sell the products only at lower prices, its profits evaporate. The history of big business is largely the story of coping with excess capacity and avoiding price competition, often by getting favorable consideration by government. Perhaps no person stated the desperation business felt for a strategy to protect profits from price cutting better than nineteenth-century steel magnate Andrew Carnegie:

> Political economy says goods will not be produced at less than cost. This was true when Adam Smith wrote, but it is not quite true today… As manufacturing is carried on today, in enormous establishments with five or ten millions of dollars invested and with thousands of workers, it costs the manufacturer much less to run at a loss…than to check his production. Stoppage would be serious indeed. While continuing to produce may be costly, the manufacturer knows too well that stoppage would be ruin… Manufacturers have balanced their books year after year only to find their capital reduced at each successive balance… It is in soil thus prepared that anything promising relief is gladly welcomed. The manufacturers are in the position of patients that have tried in vain every doctor of the regular school for years, and are now liable to become the victim of any quack that appears. Combinations, syndicates, trusts—they are willing to try anything [2].

Initially Carnegie was of the mind that combinations of firms to control prices by controlling markets (i.e., monopolies) were folly. Carnegie Steel was a technologically dynamic company that could benefit from price cutting because it could outproduce all its rivals at a lower cost. Initially the company sought competitive advantage by cutting its prices, and buying up its weakened competitors, not through monopolies. Yet Carnegie Steel would eventually become the core of "the steel trust" monopoly as US Steel (itself absorbed by the interests of banker J.P. Morgan). As we shall see, the same phenomenon of concentration by means of price cutting would characterize the largest trust of the era and the champion of the petroleum revolution—Standard Oil.

9.3 Why Study Monopoly

We believe that a new set of abstractions and economic theory must be developed for the second half of the age of oil. All theories of how the economy works commonly used today were developed in the age of rising oil availability and high energy returns on investment. Will these theories work in an age of declining availability of oil? To build a new theory, we need not abandon everything from the past. Rather we need to refine prior approaches and adapt them to a new era of biophysical constraints and limits to growth. But, more than anything, we need to begin this theoretical development from the perspective of understanding the economy as it actually exists, which is not simply a collection of small powerless companies who accept passively the impersonal forces of the market and forego large economic profit in the interests of low consumer prices and a stable general equilibrium. Rather the economy as it actually exists is dominated by giant corporations, operating on a national and international basis. These companies want to control market forces that threaten not only short-term profits but also their long-term growth in profits. These forces include ruinous price competition, rising costs of production, periodic recessions, excess capacity, unwelcome taxation and regulation, and the destabilizing effects of rapid technological change.

The study of the concentrated economy is important beyond the microeconomic level of the individual producer or consumer; the effects of monopolization are equally, if not more, important for the overall, or aggregate, macroeconomy. Some argue that a monopolized economy tends to stagnate rather than grow because of the internal dynamics of capital formation, as well as pricing and output decisions on the part of the large-scale firm in a concentrated industry. In simpler terms a concentrated economy cannot always create the growth needed to provide other laudable social goals such as full employment and poverty reduction. The solutions to depression-borne problems of the nineteenth century, which often favored corporate concentration, have become the cause of different economic and social problems in the twentieth and twenty-first centuries. Market economies suffer from essentially two sets of limits. The first are the familiar set of internal limits revolving around the process of capital formation and investment, business cycles, and the uncertainties of competition with other firms. Strategies of industrial concentration first developed to transcend this set of limits. The second set of limits are generating an increased impact in the second half of the age of a suite of external, or biophysical, limits to growth as the raw materials necessary for the earlier strategy of growth become increasingly limited. Economics in the second half of the age of oil will require an understanding of the interaction of both the internal and biophysical limits to economic activity. In this new era continued high growth is highly unlikely because of biophysical constraints such as peak oil, declining EROI, climate change, and degradation of our oceans and soil fertility. Let us begin with the study of the petroleum revolution in the context of the ongoing decline in energy prices, concentration of economic power and the development of large-scale industry (◘ Fig. 9.1).

9.4 Petroleum and the Social Revolution

In 1850 the "civilized" world was illuminated at night principally by whale oil, which was undergoing its own peak and decline as species after species of whales were hunted to near extinction (◘ Fig. 8.4e). By the late 1850s, kerosene was being refined regularly in Europe from crude oil obtained from hand-dug pits. The invention of a lamp with a glass chimney that would reduce the smoke and brighten the flame contributed greatly to the demand for kerosene. But kerosene could become "the new light" only if adequate and cheap supplies could be located. The limiting factor was the cost of hand-digging pits; the solution was to be found in well boring, soon to be known as drilling. The first commercially viable oil well in the United States was drilled by a promoter named Edwin Drake, given the appellation of "Colonel" by his supporting bankers to impress the rural population. Drake and his drillers struck oil in August 1859 near Titusville, Pennsylvania. Within a year and a half of Drake's successful well, another 75 wells were producing oil. Early successes created new boom towns such as Pithole and Oil City. Production in the oil regions of Pennsylvania soared from 300 barrels per day in 1859 to 3 million in 1861. As a result of the surge in supply, prices fell from $10 per barrel in January of 1861 to 10 cents per barrel in June of 1861. Within a year demand expanded and oil prices rose again to over $7 per barrel.

As we enter the second half of the age of oil, an age characterized by declining growth and declining energy returns on investment (EROI) for oil, and rising prices, one should not forget that the history of the first half of the age of oil was quite the opposite: increasing production, high EROIs, periodically plummeting prices, and overproduction. During the 1860s and 1870s, many small producers began to merge. This increasing monopoly concentration appears to have been a strategy to cope with the falling profits, prices, and bankruptcy caused by overproduction of an easily obtainable resource. Moreover, the legal basis of the new industry stemmed from the old English common law principle of the "rule of capture." The petroleum beneath the ground belonged to the owner of the land above. But since the oil beneath was part of a common pool that could be depleted by a few, the incentive was to extract as much as possible as soon as possible in a process known as "flush production." No place in the oil regions serves as a better example of the excesses of flush production and speculation than the town of Pithole, Pennsylvania. With the discovery of oil, property values soared, especially as oil production increased to over 6000 barrels per day. Derricks were erected on myriad tiny lots. Rapid extraction damaged the underlying strata, leaving a large share of the petroleum unextractable, due largely to the collapse of underground pressure. Property values and the town too collapsed.

Despite the demise of Pithole, production in the Pennsylvania oil region as a whole continued to increase, reaching 3.6 million barrels a year by the end of the Civil War. Given this much production, producers struggled to find adequate markets for the output, another problem that characterized the industry in its prepeak years. Crude pipelines were constructed to avoid the bottlenecks imposed by poor roads

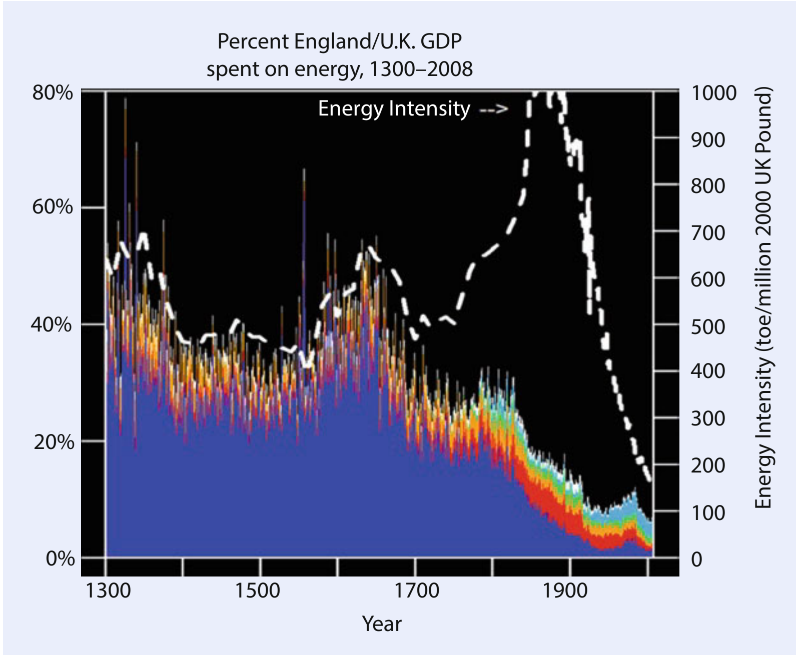


Fig. 9.1 The price of energy as a percent of GDP for England 1300–2010. (Source: Carey King)

and recalcitrant teamsters, and the Titusville Oil Exchange opened in 1871 in an attempt to shorten the link between supply and demand. It was on this exchange that present structure of long-term contract prices, short-term "spot market" prices, and very long-term futures markets was established [3].

Once the chimney lantern became common in the United States, the expansion of demand for kerosene was largely a function of the general economic expansion and political stability that emerged at the end of the Civil War. This rise in economic activity affected many of the country's primary industries and would be accompanied by an increase in the scale and scope of both manufacturing and transportation. The postbellum period saw the creation of the national corporation, the expansion of long-lived fixed capital, and the replacement of the craftsperson operating on a local scale with semiskilled operative labor and centralized management. It was also the beginning of the nation's dependency upon fossil fuels. The energy density of concentrated fuels combined with the new organization of labor produced dramatic increases in productivity and output. The new large-scale industry opened opportunities for large-scale businesses to control factors often left to chance in the older competitive economy.

9.5 The Rise of Standard Oil

No company is as closely associated with the concentration of economic power as Standard Oil. Standard Oil began modestly as a trading partnership in post-civil War Cleveland, Ohio, and rose to become the largest and most powerful company in the nation and the world's first multinational corporation by the end of the nineteenth century. By the middle of the twentieth century, it was the largest corporation in the world. Standard Oil originally rose to power in the first stage of the petroleum revolution—the provision of kerosene for illumination.

The construction of a new rail line into Cleveland, Ohio, which had access to the Great Lakes and proximity to the Pennsylvania oil regions made Cleveland an ideal center for petroleum refining. By 1865 a young general merchant by the name of John David Rockefeller had become the largest refiner in the city. The refining industry was still competitive and the techniques of refining simple enough to preclude advanced technology as a barrier to entry. The result was a large number of small producers and intense price competition. As Rockefeller's refining capacity grew, he realized he needed to find markets to absorb the output. To assure profitability Rockefeller developed a multipronged strategy, centered upon the production of a high-quality product at a lower cost than his competitors. The very name Standard Oil stems from the quality of the company's kerosene. Standard Oil was able to control quality so that Standard's kerosene contained a negligible quantity of the dangerous by-product gasoline. Cost control was accomplished by a combination of large-scale production, reduction of transportation costs, and *vertical integration*, that is, amassing all stages of an industry from refining, to marketing, to transportation on an in-house basis. It was only later, when new oil fields were discovered, that Standard integrated backward into oil extraction.

The primary method Standard used to reduce transportation cost was the use of the railroad rebate, which was enabled by Standard's scale of production. Business historian Alfred Chandler reports that the first railroad, the Lake Shore running from Cleveland to New York City, willingly reduced transportation costs per barrel in 1872 from $2.00 to $1.35 in return for a guarantee that Standard would supply 60 carloads of oil per day to be transported. The increased output benefitted the railroad as well as Standard by allowing a more consistent use of the railroad's capacity [4]. Standard then extended the policy of extracting rebates on the shipment of their oil to receiving a rebate, or drawback of 25 cents per barrel, on the shipment of their *competitor's* oil. According to energy analyst Daniel Yergin, "For what this practice really meant was that its competitors were, unknowingly, subsidizing Standard Oil. Few of its other business practices did as much to rouse public antipathy toward Standard Oil as these drawbacks—when eventually they became known" [3].

The problems of price instability, cost control, and capacity utilization, a regular feature of the industry since its inception, were exacerbated by the decline in overall economic activity following a financial panic in 1871 and the subsequent depression that lasted from 1873 to 1879. Chandler reports that the index of wholesale commodity prices, which stood at 151 in 1869, fell to 82 in 1886 [4]. Standard's production of refined kerosene continued to rise over the course of the 1870s, but its ability to market its product at a profitable price did not. Standard's strategy to address the threat of ruinous price competition was consolidation. In today's technical terms, this is called *horizontal integration* or the absorption of potential competitors for the purpose of controlling market price. Thus Standard undertook both vertical and horizontal integration and became increasingly the only game in town.

Merger was Standard's favored means of consolidation, and price cutting was its tactical method. Lower costs of production, made possible by economies of scale and cheaper transportation costs, allowed the company to undersell potential rivals. When faced with an independent producer that would not sell willingly, Standard subjected them to "a good sweating." They would increase output until the market price dropped below the rival's cost of production. Standard would then purchase the nearly bankrupt company at a favorable price and then restrict output so the price would once again climb. In the process, they brought the most able executives into Standard's management. By 1881 Standard controlled 90% of the kerosene market and sold 70% of its output in Europe. By the mid-1880s Standard's controlled 80% of marketing as well [4]. Despite the greatest degree of monopoly control that the nation has ever seen, the Standard alliance remained vulnerable to outside forces and reacted in a number of different ways to dissipate those threats and bring stability and control to the market for petroleum.

Price competition was not the only threat that faced Standard. Others included new sources of supply and new modes of transportation, as well as legal challenges. One threat was the attempt of independent producers to break Standard's hold of railroad transportation by building their own pipeline from the oil regions to the markets in the eastern United States. Standard then quietly acquired an interest in the Tidewater pipeline in 1879 and gained effective control of pipeline transportation within 2 years. Another problem was the discovery of fields outside of the Pennsylvania oil regions, first in Lima, Ohio. The additional production flooded the market and resulted in a price decline. After much debate, the Standard interests became directly involved in production, circumventing the oil exchanges. By 1891 Standard controlled approximately 25% of oil production. Standard had succeeded in building a truly integrated company, from extraction to refining and transportation to marketing [3].

By the mid-1890s Standard had become a fully consolidated and vertically integrated company. This form of business organization allowed Standard to withstand legal challenges to its strategy of price control by means of merging with competitors and fixing prices. Control over prices and ruinous competition was codified beyond a mere association of producers in 1882 when Standard formed the perfectly legal Standard Oil Trust. Stock shares of the various operating companies were ceded to Standard Oil of Ohio in return for trust certificates. Decisions about the direction of the company were made by a set of directors acting on behalf of the shareholders of the Standard Oil Trust rather than in the interests of the separate operating companies. While popular lore focuses upon price fixing, the first actions of the trust were to control costs. They reduced the number of refineries and concentrated production. Forty percent of output was produced by three refineries, and the average cost per gallon of refined oil fell from 1.5 cents to 0.5 cents. Standard expanded their marketing apparatus to assure adequate outlets for the newly expanded production, establishing wholly owned subsidiaries Continental Oil and Standard Oil of Kentucky as marketing companies [4]. Popular opinion and outrage led to the passage of the Sherman Antitrust Act in 1890, which banned conspiracies in restraint of trade. However, the Sherman Act was not intended to address the benefits of cost reduction by means of vertical integration, only price fixing due to horizontal integration. The cost cutting by expanding scale and controlling market allowed Standard to survive three significant challenges to become, by the mid-twentieth century, the largest and most profitable corporation in the world.

Beginning in the 1890s, several states filed suit against the Rockefeller interests, as well as against John D. Rockefeller himself. In 1907 the Federal government filed suit in the circuit court alleging that Standard Oil was in violation of the Sherman Antitrust Act. The circuit court found in favor of the government and Standard Oil appealed the case to the Supreme Court. In 1911 the Supreme Court validated the decision of the Federal Circuit Court: Standard Oil had conspired to restrain trade. The Standard Oil trust was dissolved into 34 separate operating companies, the most prominent being Standard Oil of New Jersey, Standard Oil of New York (Socony-Vacuum), and Standard Oil of California. Despite the breakup Standard of New Jersey (later Exxon) remained the second largest industrial corporation in the country [4]. Jersey Standard is of particular interest. In an attempt to circumvent state-level legal challenges, popular opposition to the trust, and lackluster acceptance of the Certificates of Trust by financial markets, the company took advantage of holding company legislation, recently passed in New Jersey in 1889. The holding company legislation allowed manufacturing companies to purchase the stock of other corporations and issue its own securities for the acquisitions. The holding company replaced the trust as the legal vehicle for consolidation and merger and provided for even tighter control over the pricing and output decisions than did the trust. More effective and consolidated management was able to exert control over all phases of an operating company [5]. The Standard Oil Trust reincorporated in 1899 as a holding company: Standard Oil of New Jersey. Its capitalization increased from $10 million to $110 million, and it controlled the stock of 41 other companies [3].

9.6 Further Challenges to the Standard Empire

A new legal form, vertical integration, and virtual control of the world market for kerosene did not insulate Standard Oil entirely from external threats to their control and profitability. They were to face new challenges at the twilight of the nineteenth century. These challenges came from new and substantial sources of supply, both foreign and domestic, new rivals to production and marketing. The new domestic sources of supply were discovered in Texas, Oklahoma, and California. Along with these discoveries came large and powerful new companies that are today as recognizable as is the name Standard: Texaco, Gulf, and Unocal. Other abundant sources of supply came into production in Russia, Romania, Indonesia, and, by the early 1900s, Persia. New international companies such as Royal Dutch Shell and BP were born of these discoveries. Another fundamental transformation of the petroleum industry occurred in this same period: the eclipse of kerosene by the electric light. Next another new innovation, the gasoline-powered automobile, would give vast new sources of growth and profit to the petroleum industry.

9.7 New Sources and New Rivals

Standard Oil initially satisfied domestic and world demand from its Pennsylvania oil fields. That was to change in the latter decades of the nineteenth century. The existence of oil on the shores of the Caspian Sea had been chronicled by Marco Polo. The first wells replaced hand-dug pits by 1872, and by 1873 some 20 small refineries were located in the Russian city of Baku. The industry expanded rapidly, from less than 600,000 barrels in 1874 to 23 million barrels in 1888, aided by the financing of the Nobel family. The Nobel Brothers Petroleum Producing Company was fully integrated, both backward into wells, tankers, and storage facilities and forward into refining and marketing. The demand for kerosene in Russia alone was insufficient to absorb the output of the Baku refineries. The short winter days and need for illumination could not overcome the poverty of the Russian peasantry. The Nobels success brought new competitors in the form of the Rothschilds, who purchased the railroad from Baku to the port of Batum on the Black Sea. Russian kerosene was now able to compete with that of Standard, which had previously controlled European markets. The American company then launched the type of price war that allowed them to consolidate their domestic empire. But the Russian-based companies fought back. The Nobels established a marketing company in the United Kingdom, while the Rothschilds improved technically the Baku-Batum railroad and eventually constructed a pipeline. By 1891 the Russian share of the world's kerosene exports rose to 29%, with a commensurate decline in US exports [3].

The Rothschilds, especially, were plagued with the age-old problem that characterized the industry in the first half of the age of oil: how to market the surplus resulting from the expanded production and refining of the new sources of supply. They turned their sights to East Asia and found an agent by the name of Marcus Samuel to sell their product to a wide network of merchants and traders. In the early 1890s, Samuel had developed the bulk tanker to reduce shipping costs and, by 1893, achieved access to the newly opened Suez Canal, cutting 4000 miles from the traditional route to Asia around the Cape of Good Hope. In the same year, Samuel founded a tank syndicate to reduce ruinous price competition in oil storage. By 1902 more than 90% of the oil transported through the Suez Canal was under the control of Samuel's company, Shell Oil.

Another threat to Standard's control came after the discovery of oil on the Indonesian (then Dutch East Indies) island of Sumatra. In 1885 the first successful wells were completed, and production was concentrated under the auspices of the Royal Dutch Company in 1890. By 1892 Royal Dutch constructed a pipeline from the oil fields to coastal refineries, and by 1897 output had increased by five times from a mere 2 years earlier. Standard had previously marketed kerosene in Indonesia and considered Royal Dutch a threat which they desired to incorporate into the Standard operation. Instead, Standard was spurned, and negotiations commenced to amalgamate the company soon to be known as Royal Dutch Shell. The Asian producers and marketers wanted a greater degree of concentrated power to withstand what they perceived to be the immanent Standard tactic of price cutting [3]. The new company would survive to become one of the world's majors.

In addition to international challengers to its foreign markets, Standard was subject to declines in its domestic reach two decades before the Supreme Court ordered its dissolution in 1911. First, Pennsylvania-independent oil companies, united under the name of Pure Oil, constructed a pipeline to market their output on the east coast of the United States. Second, as early as 1885, it was clear that the output of Pennsylvania Fields had peaked and begun serious decline. The State Geologist of Pennsylvania stated that "the amazing exhibition of oil is only a temporary and vanishing phenomenon – one which young men will see come to its natural end" [6]. The oil boom of the entire Appalachian Basin was already over by 1900. Third, in the early 1890s, large fields were discovered in Southern California. By 1910 California's 73 million barrels represented 22% of the world's output, mostly controlled by the independent company Union Oil (now Unocal). Standard finally commenced operation in the California fields, establishing Standard Oil of California (now Chevron) in 1907. However, the monumental change in the oil business occurred in January 1901 with the discovery of the Spindletop field previously mentioned in ▶ Chap. 6. The original gusher produced 75,000 barrels per day and a new oil boom had begun. Land values skyrocketed and population soared from 10,000 to 50,000. In an experience similar to the one that occurred in the

Pennsylvania oil regions, numerous tiny leases led to more than 400 wells on Spindletop itself. Prices collapsed to 3 cents per barrel. The original promoters needed markets for their oil and found a likely buyer in Marcus Samuel's Shell Oil at a long-term price of 25 cents per barrel. The glut caused by the Spindletop find was augmented by another discovery in Oklahoma. The common problem of overproduction led not only to falling prices, but in this case, as in Pithole, flush production depleted the well. Underground pressure for Spindletop gave way in 1902, the year after discovery.

The stabilization of the Texas industry would fall to the Pittsburgh bankers (the Mellons) who had financed the initial operation. The original promoters were dismissed, the contract with Shell renegotiated, and the Mellons began the development of a vertically integrated company based on the extraction and refining of petroleum. Their first task was to come to terms with the overcapacity that the construction of the new refinery and pipeline network created. The corporation that restructured and further integrated into nationwide marketing became known as Gulf Oil. In addition, another significant corporation, Texaco, was built upon the expansion of transportation, storage, refinery capacity, and the currying of important political connections.

Every discovery would bring a glut of new oil and price declines into the market. This, in turn, created the need for constant expansion into new markets. Standard's control of the industry was clearly in decline. In 1880 Standard controlled 90% of kerosene refining in the country. By 1911, the year of its dissolution, the former monopoly controlled but 65% of domestic kerosene output, while its international markets likewise declined in the face of new discoveries and new competitors [3]. Yet while Standard's control was declining, its profits and output increased. The new century was to bring the end of the kerosene era but the dramatic expansion of oil demand as we entered the age of the internal combustion engine and the automobile.

9.8 Markets Lost and Markets Found

As we have said, the primary use of oil in the first stage of the petroleum revolution was for illumination purposes. The market for kerosene, however, was to all but disappear at the end of the nineteenth century. In 1879 Thomas Edison perfected the incandescent light bulb and began operations of a generating plant in 1882. Edison made sure to price electricity competitively. Electricity overcame many of the drawbacks of kerosene such as smoke, soot, and oxygen use. But the adoption of electricity was not immediate. The original generating plants, located near load centers until the adoption of alternating current, were powered by coal-fired piston engines which were very noisy and dirty. Moreover, electricity was considered dangerous and the cause of myriad great fires that swept the urban centers of the Northeastern United States at the dawn of the twentieth century. While a young man, Klitgaard spent many years as a restoration carpenter and saw the reason. He observed and corrected many situations where electricity entered urban dwellings at 240 volts over bare wires, with only ceramic insulators separating the wires from the dry roof beams upon which they were placed. But once these safety constraints were overcome technically, the use of electricity for light and power caught on quickly. In the time period from 1885 to 1902, demand for lightbulbs soared from 250,000 to 18 million per year. In 1890 only 15% of urban railways and streetcars were powered by electricity. By 1902 94% used electricity as a motive force [4]. Problems with carbon emissions as greenhouse gases had barely been recognized theoretically. The switch to electrical power virtually eliminated the very serious public health problems associated with the use of horses as beasts of burden.

Electricity fundamentally changed the process of production. When factories were powered by a central source, steam, or water, the layout of the factory was dictated by distance from the central source, and power was delivered to places of use by dangerous and inefficient system of pulleys and belts. Factories had to be multistory affairs on a small footprint. Much time was lost to the movement of semifinished goods between floors. The advent of the electric motor allowed sprawling single-story sheds with the power source decentralized to the individual machine. Here again, we see the role of energy in the improvement of productivity. The same process of industrial concentration occurred in the electrical industry itself. In 1892 the New York Banker J.P. Morgan consolidated Edison Electric with Thompson-Houston to form General Electric which shared the market only with Westinghouse. In the type of co-respective behavior common to oligopolies, Westinghouse and GE regularly shared patents [5].

9.9 The Age of Gasoline

In the first phase of the petroleum revolution, gasoline was a dangerous by-product. But gasoline becomes the primary petroleum product with the invention of the automobile powered by the internal combustion gasoline engine. Automobiles gained acceptance in Europe by 1895 and soon after began to sweep personal transportation in the United States. Eight thousand cars were registered in 1900. By 1912 nearly a million vehicles were on the road [3]. One year later Ford took advantage of the possibilities afforded by the electric motor and single-story shed production when he built his first assembly line in Highland Park, Michigan. In the early days of the industry, autos had been assembled by teams of skilled workers, often bicycle mechanics, who built each car from the wheels up. Automobiles were little more than luxury items for the affluent. Ford's Model T, introduced in 1908, sold for $850, then an enormous sum. After the construction of the Highland Park, plant cars were assembled by semiskilled operatives on a continuous line. The price of a Model T fell as the cost of production fell with the expansion of scale and an increase in the throughput of materials and labor. By 1925, the peak of the first automobile boom, a Model T sold for $240. Mass production changed the automobile from a luxury item to one that workers could afford. Ford workers were paid above the industry average. Ford nearly doubled industry standard wages when he commenced his famous "$5 day" in 1915, essentially as a cost-saving measure. Previously assembly line work was seen as so degrading that the Ford plants had a difficult time retaining an adequate workforce. Absenteeism was 10.5% and turnover reached 470% in 1913. Turnover costs in 1913 alone were nearly $2 million. So Ford raised wages to keep his workers. "There was…no charity involved… We wanted to pay these wages so that business would be on a lasting foundation. We were building for the future. A low wage business is always insecure. The payment of $5 for an 8-hour day was one of the finest cost-cutting moves we ever made" (Ford, quoted in Perelman 2006: 135) [7].

As the price declined, and credit was offered, sales and registrations of automobiles increased steadily, reaching 23 million in 1925. Registrations fell during the depression, and new cars were not produced during the Second World War, as auto plants were converted to produce tanks and airplanes. Moreover, gasoline and tires were rationed during the war. The second automobile boom commenced following the war and produced lasting impacts upon the nation. In 1950, 40 million cars were registered in the United States. This figure climbed to over 65 million in 1962 and more than 250 million by 2007.

The automobile qualifies as what economists call an *epoch-making innovation*. Few other such technological changes qualify. An epoch-making innovation must not only absorb large amounts of capital investment, but must create more opportunities for investment in other industries. Baran and Sweezy contend that only three innovations transformed society, absorbed sufficient capital, and created new industries and processes: the steam engine, the railroad, and the automobile. To this Richard Duboff adds electrification and Michael Perelman contends that computerization must be considered [5, 7, 8]. The automobile not only absorbed tremendous amounts of fixed capital, accounting for 6.3% of all value added in manufacturing in 1929, but also created myriad peripheral industries. Repair shops, drive-in movies, motels, gas stations, and the fast-food industry owe their existence to the automobile. The automobile itself is dependent upon petroleum for energy. Indeed all epoch-making innovations have been energy-intensive, indeed among the most energy-intensive products of their day. Moreover, these innovations have been subject to a similar degree of industrial concentration as was the petroleum industry, largely for the same reasons: the need to rationalize production, reduce costs, expand market share, and avoid ruinous price competition.

9.10 Industrial Concentration as a Consequence of Concentrated Energy

Before the massive use of fossil fuels, production was essentially organized on the basis of small shops using skilled labor. Skilled master craftspeople were generally responsible for all or many stages of production and agreed to be responsible for the training of apprentices. Upon completion of their apprenticeships, new craft workers were deemed fit by the society of masters to travel to obtain independent unsupervised work. In fact they were called journeymen. After a long period of learning the technical and business skills of their respective trades journeymen could rise to the rank of master. Societies of masters, which were called guilds, decided collectively upon prices and standards of quality. This world of small business did not display the type of price competition found in microeconomics texts. As an institutional structure, guilds limited the type of competition that could ruin a master's fortune. Instead the guilds brought stability to the preindustrial economy. Thus, the modern concept that competition is necessary for efficient operation of businesses was not the historical norm.

Few examples existed of alternative organizations. Large-scale textile mills appeared along the swiftly flowing rivers of New England at locations such as Lowell and Lawrence, Massachusetts, and Manchester, New Hampshire, by the 1820s. These mills not only employed larger numbers of workers than the typical small shop, but they were not organized around the principle that every entry-level worker would become eventually a master. The labor force of the early textile mills consisted mostly of young women recruited from the hardscrabble New England farms, whose employment, frequently boring and arduous, was expected to be temporary.

In the decades after the Civil War, the US economy went through a process that economic historian Richard DuBoff termed "the grand traverse" and what we call industrialization or the development of the hydrocarbon economy. This transition entailed the transformation of a primarily local and regional economy utilizing local natural sources of energy into an economy that was based on large-scale industry, mass production, and the use of fossil energy, generally derived from far away. The railroads were the nation's first big business. Railroad building commenced in earnest in the late 1840s, following the nation's first depression. There were only about 2300 miles of track when the decade of the 1840s began. Another 5100 miles of track were added in the 1840s and 21,400 in the 1850s. After the Civil War, track building increased significantly. In the 1880s additions to track construction peaked, when another 74,700 miles were built. By the time railroad travel was supplanted by the automobile and freight was hauled primarily by truck, the railroads had established themselves as the nation's first large-scale enterprise. Railroads accounted for 15% of all gross private domestic investment in the 1850s and 18% in the 1870s and 1880s [10]. Moreover railroads augmented the communications networks, as telegraph wires were built along railroad rights-of-way. The construction of a viable transportation and communications infrastructure was vital for the transformation of the economy as a whole. Recall how Standard consolidated its hold on refining by achieving lower-cost transportation by means of an existing railroad network. The ability to manage a nationwide market was greatly enhanced by a functioning transportation and communications infrastructure.

The economy was transformed fundamentally in the years following the Great Depression of the 1870s as industrialization increased more and more. Not only did the scale of production increase, but so did the organization of labor. As in the case of Standard Oil, the control of costs became a fundamental element in the competition between large enterprises. Jobs were subdivided in a way that Adam Smith himself could barely imagine. The essence of competition became based on increasing productivity. Craft workers were supplanted in manufacturing by an immigrant force of unskilled and semiskilled labor who were content with boring repetitive piecework for secure wages. Behind the ability to mechanize, transport, and impose the detail division of labor was the access to cheap energy. Business historian Alfred Chandler states the matter succinctly: "Cheap coal permitted the building of large steam-driven factories close to commercial centers and existing pools of labor. In the heat-using industries the factory quickly replaced the artisan and the craftsman… Coal, then, provided the source of energy that made it possible for the factory to replace the artisan, the small mill owners, and the putting-out system as the basic unit of production in many American industries" [4].

9.11 Threats and Opportunities

Chandler also makes the important point that the revolution in transportation, itself based upon cheap energy, further transformed the distribution of products. The modern corporation was not born with the advent of mass production but rather necessitated the unification of mass production with mass distribution. If a company produces more than it can sell, the incentive to produce even more output or invest in capital equipment declines. This will be a theme that recurs through subsequent years of economic development. Capital accumulation, brought about by investment in capital goods, is the engine of growth in a private enterprise economy. Periods of lagging investment bring about economic downturns, and the low-profit potentials of a sluggish economy further reduce the ability to find profitable outlets for one's investment capital. The percentage of net national product (or gross national product minus depreciation) that went into investments climbed steadily over the course of the nineteenth century from 10% in the 1840s to 18% in the 1870s to 20% in the 1890s [5]. This growing level of investment aggravated the problems that can occur from producing more than can be consumed. As Andrew Carnegie had realized, large-scale companies would attempt any alternative to shutting down; the consequence of walking away from the considerable costs embodied in the capital equipment was unthinkable. The costs sunk into the purchase of capital equipment typically drove most company leaders to concentrate in order to protect their capital investments from price competition.

Various forms of economic concentration, such as vertical integration, horizontal integration, trusts, and holding companies, were responses to a number of chronic problems that plagued American enterprises operating in the new world of expanding markets, rapid technological change, financial uncertainty, and the availability of cheap energy. Concentrated fuels certainly opened up vistas of low-cost production and transportation unheard of before the harnessing of fossil fuels, but cheap energy alone was insufficient to protect producers from a set of internal limits to capital accumulation such as the tendency to produce too much to sell at a profit. Viewed in this light, monopoly is not a minor aberration to an otherwise competitive economy. Rather it is the eventual outcome of a competitive process as companies attempt to control their economic environment and protect profits and potential growth by avoiding the type of competitive behavior that could perhaps ruin them. In essence, the history of the American industrial revolution is the history of both cheap energy and monopoly concentration and is understood best as a combination of these factors.

Thus economic concentration emerged not as a mistake in the competitive process, as today's mainstream microeconomic theory would have us believe, but as an explicit strategy.

Even as neoclassical economists were perfecting the elegant theory of the "perfect competition," industrialists such as Carnegie, Rockefeller, and other captains of the oil industry were decrying the ruinous effects of "cutthroat competition." For the theorist, price competition was necessary for their view of economic perfection. Resources flowed to their most lucrative use, while the market system forced competing firms to produce at the lowest possible cost and pass the savings onto consumers in the form of low prices. In the end the system balanced in a stable equilibrium. The only way to insure a perfectly competitive equilibrium, however, is to ignore the problem of fixed cost. In fact the initial assumption of the economists of no barriers to entry precludes the analysis of the cost of long-lived fixed productive assets. But industrialists operated in the real world where large-scale industry required substantial investment in fixed capital. If, at the same time, the cost of producing one more unit of output (what economists call marginal cost) is low, real-world producers face a dilemma: competition drove the market price down below the level at which industrialists could turn a profit or even recoup their fixed costs.

The standard economic theory of competition asserts that competition will bring prices down to the level of marginal cost. Theoretically entrepreneurs are willing to accept the going rate of normal profit as all else is competed away by rivals lowering their prices in order to capture more customers. Moreover, the system is stable and there is no tendency to change. But in the real world of business, managers who earned no profit and had no prospects for profit growth would quickly be out of a job. If a real-world industrialist borrowed money to purchase large-scale equipment and then finds prices competed down to the level of producing one more unit of output, the company would never be able to generate revenue sufficient to repay their bondholders and bankers. One may think of railroads, where most of the cost is in tracks and locomotives and little of the cost is in cheap fuel or labor, or in the modern world airlines, for such real-world examples. Chandler summarized the position of the railroads when he said:

> "Competition between railroads bore little resemblance to competition between traditional small, independent unit commercial or industrial enterprises. Railroad competition presented an entirely new business phenomenon. Never before had a very small number of large enterprises competed for the same business. And never before had competitors been saddled with such high fixed costs. In the 1880s fixed costs, those costs that did not vary with the amount of traffic carried, average two-thirds of total cost. The relentless pressure of such costs quickly convinced the railroad managers that uncontrolled competition for through traffic would be ruinous… To railroad managers and investors, the logic of such competition appeared to be bankruptcy for all" [4].

9.12 The Loss of Worker Power and the Gain in Financial Power

Labor productivity continued to rise as the result of the prolonged investment boom and the increase in the energy subsidy to each worker [11]. Productivity growth averaged only 1.6% per year from 1889 to 1919. After the 1920–1921 recession until the late 1950s, it averaged 2.3% annually. New processes such as electrification increased industrial efficiency, and the new technologies of the automobile further reduced the costs of transportation. These innovations, of course, depended upon an ample supply of cheap fossil energy, much of it from the newly discovered sources in California, Texas, and Oklahoma. But consumer demand did not increase as rapidly as productivity or organizational innovations such as scientific management, resulting in wage growth that did not keep up with production. The lack of purchasing power combined with the ebbing of the investment boom, created the conditions underlying the Great Depression. Automobile sales peaked in 1925, the year before the peak in investment as a whole. Construction of skyscrapers in major eastern cities ground to a halt. The decline in demand for autos and skyscrapers reduced the demand for steel, and declining demand for steel further reduced the demand for coal. In another blow to investment, a hurricane devastated South Florida, destroyed the railway through South Florida and the Keys promoted by John D. Rockefeller's early partner, Henry Flagler, and brought a speculative boom in suburban housing to a close.

Yet even while the real economy was "softening," the demand for financial securities continued to rise, fueled by margin buying. Investors could purchase a stock by putting up only a fraction of the value of the stock (the margin) and borrowing the remainder from their brokers. (This is called *leverage* today.) The volume of such loans (the broker's call market), according to John Kenneth Galbraith, was the most accurate index of speculation, as it was money borrowed to purchase stocks, and not real assets. In the early 1920s, the volume of these loans was approximately one to one and a half billion dollars. By 1927 the market increased to a volume of three and one half billion. 1928 saw broker's call loans increase to four billion and, by 1929, six billion dollars. With all this debt-fueled buying, stock prices registered impressive increases throughout the summer of 1929, enhancing the optimism of the market and increasing further the demand for call loans. But reports of the underlying weakness in the real economy began to sap the confidence of some knowledgeable investors throughout the fall of 1929. By October the markets were wavering, although the confidence of investment bankers remained high. Charles Mitchell of National City Bank believed that the underlying fundamentals of the economy were sound and that too much attention was being paid to broker's call loans. Nothing, according to Mitchell, could arrest the upward trend [11].

9.13 The Great Crash

On October 29, 1929, the stock market collapsed. Stock values plummeted by $26 billion. In relative terms, the stock market lost approximately one-third of its September value. The economy was soon plunged into depression. GNP declined by 12.6% from 1929 to 1930, and unemployment increased from 3.2% in 1929 to 8.7% in 1930, peaking at 24.9% in 1933. But how did this happen given that less than 2.5% of Americans owned stock? [12].

The answers lie partly in the weakness of America's banking system. Rural banks, in particular, were chronically undercapitalized without adequate funds to repay their depositors in case of an emergency. Most of their reserves had been loaned out already. More than 500 per year failed even in good economic times. However, the crisis of bank failures climbed after the stock market crash to include urban money center banks. After the collapse of the stock market, heavily leveraged investors could not repay their brokers who, in turn, could not repay the banks. An additional 1352 banks (above the normal 500) failed by the end of 1930. Policy decisions exacerbated the failure of the banking system as the Hoover administration tightened credit and raised interest rates, partly to punish speculators and partly to shore up the British Pound. Moreover, the international gold standard was rendered unworkable after the stock market crash and wave of bank failures. According to the dictates of the gold standard at the time, all trade deficits had to be paid in gold at the end of the year. But gold also functioned as the domestic currency. Squaring international accounts under the prevailing institutional arrangements meant reducing a nation's domestic money supply. This exacerbated the deflationary tendencies already touched off by the collapse of banks and financial markets. In addition, the Versailles Treaty ending First World War had imposed $33 billion worth of reparations on Germany. Germans borrowed heavily from US banks to pay their reparations to England and France. England and France used the reparations payments to repay their loans to US banks. The collapse of the US banking system precluded more loans to Germany. Germany thereby defaulted on their reparations payments, and England and France suspended payments upon their war debts. The international trade system simply collapsed, hastening the reemergence of hostilities in a world shaken by long-term depression [12].

The world that emerged from the Great Depression and subsequent world war was a world fundamentally transformed. The ideology that markets would find their own efficient equilibria was dealt a near-fatal blow by the depth of the depression. The New Deal and Keynes' *General Theory of Employment, Interest, and Money* were to establish the role of government intervention into the economy. Commodity money in the form of the gold standard would give way to government-generated fiat money. International oil supplies would remain in the hands of the allied powers, and oil would soon become officially denominated in US dollars, soon to become known as petrodollars. In short, the postwar social and economic order would soon become dominated by the United States as a political power, by the large-scale corporation as an economic power, and by petroleum as a source of energy and power. Power over the control of oil became political power over the rest of the world.

9.14 Conclusion

In the years following the Civil War, the American economy was transformed from a small-scale, regional endeavor based on skilled labor, hand tools, and natural sources of energy such as wood and grass into a large-scale, national economy powered by cheap fossil energy, long-lived fixed capital in the form of machines and factories utilizing deskilled operative labor. Long before the peak of US oil production, the economy experienced myriad periodic downturns, including three great depressions in the 1870s, the 1890s, and the 1930s. During these times, the pressure on the large-scale industries became intense, and many were driven toward bankruptcy by competitive price devaluations. Facing bankruptcies, the favored strategy was the concentration of industry by means of consolidation and merger. By the 1890s two merger movements had produced most of the characteristics of big business we recognize today, from a few firms controlling the majority of an industries output, to the rise of non-price competition, such as competing to reduce price and expand market share. Horizontal mergers were designed to eliminate ruinous price competition, and vertical integration reduced costs by bringing all aspects of production, distribution, and marketing within the control of a central management and creating the economies of scale. By the end of the century, these concentrated industries had devised mechanisms, such as trusts and holding companies, to cope with the chronic problems of overproduction and excess capacity that accompanied price competition [9].

The evolution of the large corporation and the concentrated industry was a fundamental part of the industrial revolution itself and was enabled and encouraged by the fossil fuel revolution. Many economic historians have chronicled the role that

9

the rise of monopoly concentration played in the American economic experience. Few, however, have focused on the role played by cheap energy. Since we believe that economics should be both a social and a biophysical science, it is important to link the development of energy and power as physical entities with the social and economic factors that they allowed and generated. We can achieve a better understanding how the economy works, historically as well as contemporaneously, by viewing the development of economic power in the context of power in the physical sense; national and corporate powers alike still work to consolidate economic strength just as military or political strength is often consolidated.

The economy still experiences a roller coaster of expansion followed by depression or recession despite the existence of dramatic technological change, the availability of cheap energy in the form of coal and then petroleum, economic concentration, and organizational innovation. Even in times of abundant cheap energy, such as the 1930s, the economy experienced downturn due to the internal dynamics of technology, investment, productivity, demand, and excess capacity. Historically this internal tendency is periodically reversed by the introduction of epoch-making innovations such as the steam engine, the railroads, electrification, and the automobile, allowing for the long-term expansion of productivity, investment, and economic growth. All of these innovations were energy-intensive and depended upon the availability of cheap energy. The digital revolution, energy intensive in its collective impact, may or may not qualify as a major epoch-making innovation, but it seems not to have resolved the problems inherent with the others, as the major economic downturns of 2001 and 2008 seem to indicate.

What is the fate of the concentrated economy as the age of cheap energy comes to an end? In other words, will the biophysical constraints combine with the already existing internal limits to bring about the end of the growth economy? What are the chances of another epoch-making innovation will usher in another buoyant era of economic growth? Can some kind of "green" energy do this? Could this take place while nearly every scientific measurement of the human impact upon the planet indicates we are already in overshoot? If we are already exceeding the biophysical limits of the planet, we doubt severely that we can grow our way into sustainability. But economic growth is at the heart of a monopolized economy. How do we reconcile the need for living within our biophysical limits with the need to produce jobs and opportunity for the next generation and reduce poverty, all of which have relied on growth, at least historically? Much of the rest of the book will focus on that question.

Questions

1. How did the emergence of the fossil fuel age result in a concentration of political and economic power?
2. What is an oligopoly?
3. What was the first large-scale use of petroleum? What resource was it replacing? Why?
4. What is vertical integration?
5. What is horizontal integration? How was it accomplished by Standard Oil?
6. We see kerosene replacing whale oil, and electricity replacing petroleum, both fairly rapidly. What do you think will replace electricity, if anything?
7. Why didn't the end of the kerosene age mean the end of Standard Oil?
8. What was Henry Ford's idea about guaranteeing sales for his Ford automobiles?
9. What is an epoch-making innovation? Can you give three examples and tell how each is related to energy, and do you believe there are any happening now?
10. What was the relation of the rise of coal to skilled labor?
11. Can you give several perspectives on the role of competition in the economy?
12. What was the objective of the Sherman Antitrust Act in 1890?
13. Do you think the basic business conditions of the early 1900s were very different from those of today? Why or why not?
14. "The ideology that markets would find their own efficient equilibria was dealt a near fatal blow by the depth of the 1930s depression. The New Deal and the *General Theory of Employment, Interest, and Money* established the role of government intervention in the economy, as well as a focus on the inability of the private sector alone to create sufficient overall demand to maintain full employment." Discuss these two sentences in light of today's economy.

15. A general problem of industrial capitalism is that the economy is usually unable to absorb all that is produced by the very productive fossil-fueled economy. What were some of the approaches used in the 1950s to deal with this problem?
16. How might the end of cheap oil change the way that our industrial economy operates?

References

1. Galbraith, John Kenneth. 1967. *The new industrial state*. Princeton: Princeton University Press.
2. Carnegie, Andrew. 1889. The bugaboo of the trusts. *North American Review* 148: 387.
3. Yergin, Daniel. 2008. *The prize: The epic quest for oil, money, and power*, 6. New York: Free Press.
4. Chandler, Alfred. 1977. *The visible hand: The managerial revolution in American business*, 321. Cambridge, MA: The Belknap Press.
5. Duboff, Richard. 1989. *Accumulation and power*. New York: M.E. Sharpe.
6. Pascualli, R.C., and Hall, C.A.S. 2012. *The first half of the age of oil. An exploration of the works of Colin Campbell and Jean Laherrere*. Springer, NY.
7. Perelman, Michael. 2006. *Railroading economics*. New York: Monthly Review Press.
8. Baran, Paul, and Paul M. Sweezy. 1966. *Monopoly capital*. New York: Monthly Review Press.
9. Piore, Michael, and Charles Sabel. 1984. *The second industrial divide*, 49–72. New York: Basic Books.
10. Hacker, Louis. 1940. *The triumph of American capitalism*. New York: Simon & Schuster.
11. Galbraith, John Kenneth. 1988. *The great crash 1929*, 68–107. Boston: Houghton-Mifflin.
12. Kennedy, David. 1999. *Freedom from fear: The American people in depression and war, 1929–1945*. New York: Oxford University Press.

Twentieth Century: Growth and the Hydrocarbon Economy

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_10

10.1 Introduction

The economy of the early twenty-first century is not just a larger version of the economy of the nineteenth century. It is fundamentally different. This chapter views the development of the American economy from the middle of the twentieth century through the financial crisis and recession of 2008 and beyond. In Barack Obama was elected president of the United States with a great deal of optimism. But 2010 saw a conservative resurgence based on poor economic growth and by 2016 the election of a right-wing "populist" Donald Trump. To answer this crucial question, we need to look carefully at the patterns of history as well as carefully examine the scientific data, which we do with the remainder of this chapter.

The years following the end of the Second World War were a time when the wealth and power of the United States were on the rise. After the stagnation of the depression and the sacrifice of the war years, there was, once again among a large proportion of the American population, a belief in abundance. From the depths of the depression was born the "golden age" of the American economy. The era was characterized by the growing international power of the United States, both economically and militarily. The wealth that flowed in from the rest of the world was shared more broadly, and with a greater segment of the working population, than at any time since the industrial revolution. Home ownership became a reality for a greater share of the population, and it could be achieved upon a single income. Disney's "Tomorrowland" showcased "the house of the future" replete with all-electric appliances, futuristic design, and virtually no attention to insulation or energy conservation. The days of conservation and sacrifice were gone. Spacious automobiles traversed newly constructed freeways to arrive at Disneyland in Anaheim, California, from far-flung suburbs. And they brought kids, lots of them, as the "baby boom" was just gaining headway. The future looked promising. It was a future based on cheap oil and economic growth.

But the year following the opening of Disneyland in 1955 was a year of warning. In 1956 the nationalization of the Suez Canal by Egypt halted briefly the shipments of cheap and abundant oil to Europe and threatened the existing international order. Roger Revelle and Charles Keeling first began to measure carbon dioxide concentrations in the atmosphere, and M. King Hubbert wrote his famous paper predicting the peak of domestic oil production a mere 15 years in the future. But the academics were ignored and the crisis in North Africa was quickly brought under control. It was a time when Americans could seemingly do anything, including building the dream of happiness through material abundance and perpetual growth.

To recap: It was a time of peace, and peace on American terms, *Pax Americana*. The capabilities of other industrialized nations were decimated. But the war rekindled US industry from the depths of the depression. No other nation could match US industrial output. Rather than seeing the European nations as serious competitors, national and international policy sought to shore up their devastated infrastructures and their demands for goods, particularly US goods. In addition, the US dollar replaced the repudiated gold standard. The rest of the world was willing to give the US interest free loans in their own currencies to hold the dollar, stemming from its surging strength. Since the world's resources, including oil, were denominated in dollars, the country could buy in a buyer's market and sell in a seller's market as the terms of trade (or ratio of export price to import price) consistently favored the United States. Most of the revenues of oil-producing companies were recycled back into the US economy as foreign nations used their petrodollars to by bonds from the US Treasury.

US business could look at the world as its oyster. Little foreign competition existed to threaten the nation's large oligopolies, the dollar was the international currency, and US demand was stable and rising. The government pledged to use its economic policies to limit the kind of ruthless cutthroat competition that characterized the earlier industrial era, and the war mobilization itself was highly favorable to business. Antitrust policy seemed to be more directed toward keeping new firms from upsetting the industrial balance than breaking up the older concentrated industries that had just helped win the war. Industry after industry such as automobiles, breakfast cereals, and petroleum refining settled comfortably into "Big Threes" or "Big Fours." In fact, a new merger movement was about to begin, based on the conglomerate merger of concentrating firms from seemingly unrelated industries. Finally, it was the age of cheap oil, and the United States was still the dominant oil producer in the world. The great finds of the 1930s found little use during the depression and indeed allowed the United States to supply 70% of the oil for the allied war effort. Cheap oil, in conjunction with the aforementioned structural changes, helped fuel the mass consumption, economic growth, and military muscle for years to come.

The "good times" were about to end. By 1970 Hubbert's ominous prediction turned out to be accurate as US oil production for the "lower 48" peaked. (This did not include the Alaskan reserves, because Alaska did not become a state until 1959 and therefore did not figure in Hubbert's calculations.) Oil price shocks buffeted the economy in 1973 and 1979 threatening both the mobile lifestyle and economic growth. The US producers no longer had the spare capacity to keep foreign producers from using "the oil weapon." This was the era that saw the rise to power of OPEC, the Organization of the Petroleum Exporting Countries. The 1970s and early 1980s were the time of stagflation or simultaneous inflation and recession. Under mainstream Keynesian theory, inflation was supposed to occur only if demand continued to expand past the level that would support full employment. But prices rose even in the presence of substantial unemployment. Keynesian policies no longer seemed to work. If the government pursued an expansionary policy, inflation worsened. If it cut its spending or raised taxes to reduce budget deficits, or made money harder to come by, unemployment soared to politically unacceptable levels. Moreover, the international monetary accords conceived and born in Bretton Woods, New Hampshire, in 1944 collapsed under their own weight and US policy. The accords had been built upon the willingness of the United States to convert holdings of dollars to gold at $35 per ounce. By the early 1970s, foreign claims exceeded the magnitude of the gold supply. President Richard Nixon closed the gold window in 1971, ushering in a new era in international monetary politics: one that was far less favorable to the growth of the United States.

Part of the expansion of foreign dollar holdings was based on the expansion of American business abroad, and part was attributed to increased military expenditures such as rental payments for bases and wages and profits paid to local workers and firms, plus the expenditures of military personnel in foreign countries. The war in Southeast Asia was not going well. Rand Corporation analyst and respected neoclassical economist (for his work on decision-making under conditions of uncertainty) Daniel Ellsberg expressed dismay after briefing high-level government officials as to the conditions on the ground only to have them turn around and tell a far more optimistic story to the nation. Ellsberg released the "Pentagon Papers" showing the disconnect between the assessment of war planners and public officials to the *New York Times* in 1971, earning a spot on Richard Nixon's "enemies list" and the honor of being called "the most dangerous man in America" [1]. Ellsberg was correct in his assessment, and by May Day 1975, North Vietnamese tanks broke down the gates of the American Embassy heralding the end of America's longest war to date. By 1979 the "friendly" government of the Shah fell in Iran. Hopes for a Democratic government were dashed as the Mullahs seized power and proclaimed the "Islamic Republic." Oil prices soared and *Business Week* lamented "the decline of US power" in their special issue of March 12, 1979 [2]. Terms of trade began to favor the United States less and corporate profits fell [3].

America, along with much of Europe, took a more conservative turn. Ronald Reagan in the United States and Margaret Thatcher in the United Kingdom gained power and began to develop new economic policies, which departed from the typical Keynesian ideology, based on lowering taxes, remilitarization, anti-union campaigns, the reduction of domestic spending and the decreasing of regulations on business and finance, and restrictive monetary policies to reduce inflation. Things were changing in other nations as well. Social Democratic governments in Germany, Sweden, France, and Italy were replaced by conservatives as well. The Soviet Union, crippled by falling oil prices and cold-war military spending, did not achieve the state of advanced socialism called for by the politbureau in the post-Brezhnev days, and the openness (Glasnost) and restructuring (Peristroika) called for by Mikhail Gorbachev led to the dismantling of the USSR. The Chinese Communist Party began to court openly entrepreneurs. The cold war was won, and there were no viable alternatives to multinational capitalism. Yet economic growth did not respond over the long term, despite great new finds of oil in the North Sea and the North Slope of Alaska. Debt swelled as well, from less than $5 trillions in 1980 to approximately $15 trillions in 1990. The United States went from being the world's greatest creditor to the world's greatest debtor in less than a decade. The Clinton Administration completed the work of the "Reagan Revolution," deregulating fully the US financial services industry, trading proposed environmental legislation for a North American Free Trade Agreement, and ending "welfare as we know it." Eight years of the administration of George W. Bush saw two inconclusive wars and the explosion of a debt economy that ended with the financial meltdown and housing crisis of 2008. Oil prices rose to historic highs in the same summer. As the first edition of this book went to press, the financial crisis has turned into the worst economic downturn since the Great Depression.

In 2008, Barack Obama was elected president of the United States with a great deal of optimism. But 2010 saw a conservative resurgence and the 2016 elections resulted in conservatives controlling all three branches of the U.S. government. The decade since the election of Barack Obama witnessed the rise in the techniques of extracting oil and natural gas from tight shale formations by means of hydraulic fracturing, along with a decrease in oil prices. Rates of economic growth rose from a depression-level 1.3% in the decade of the 2000s to an anemic 2.1% in the decade of the 2010s. Unemployment fell from 10% to less than 5% by 2016. Yet wages did not rise as productivity increased. The hallmark health-care legislation, the Affordable Care Act, was subject to constant attack, and was not as affordable as envisioned. This is not surprising to us. Health care remained in the hands of concentrated pharmaceutical and insurance companies, along with monopolizing hospitals. Monopolies restrict output and raise prices. That is their fundamental business model. To answer that question, we need to understand historically how we got to where we are now. To this question the rest of this chapter turns.

10.2 Historical Antecedents: Depression and War

It is our contention that the events of economic history cannot be explained by social and economic forces alone. The role of energy must be included. Neither does a pure analysis of energy availability and use explain our current situation by itself. Rather they should be analyzed in conjunction. Historically the United States economy has experienced three major depressions, in the 1870s, 1890s, and 1930s. All came after the discovery and exploitations of fossil hydrocarbons. The ability to acquire and use energy allowed the dramatic expansion of production, as the concentrated and highly dense new sources of energy could transcend the strength and often the skill limitations of humans. However, the economy is still limited by the capacity to sell the products at a profit, expand markets, and realize the gains of productivity. When this does not occur, the economy slips into depression. "Overinvestment at the end of economic booms has characterized the economy since the middle of the nineteenth century." Another factor came into play in the oil industry. 1930, the first year of the depression, was the peak year of oil discoveries. With a limited market due to depressed conditions the oil was merely stored. Severe economic downturns may result from a lack of crucial resources such as oil, but they may also result from an overabundance of them. The end of the twentieth century, from the 1950s until the present, was characterized as an age of economic growth. The 1950s and 1960s were "golden years," while the 1970s were an age of stagnation. Economic growth revived somewhat in the 1980s, but the burden of debt soared. The long-term consequences of the creation of a casino economy came due in 2008. But what does the future portend? Will we, through social reorganization, transcend our current problems, or will a set of external, biophysical, limits augment the preexisting social ones to produce an age of austerity. To answer this crucial question, we need to look carefully at the patterns of history as well as viewing carefully the scientific data.

The world economy collapsed into depression for the entire decade of the 1930s. In the United States the presidential election of 1932 pitted two candidates with opposite opinions as to the depression's origins. Incumbent President Herbert Hoover believed the cause stemmed from the Great War and subsequent Treaty of Versailles that ended the war. The victorious allied powers redrew the map of the Middle East as they dismembered the Ottoman Empire, which had sided with Germany and the Austro-Hungarian Empire in the war. The new map showed a curious phenomenon. Places with large populations had little oil, and places with abundant oil reserves had very few people. The Austro-Hungarian Empire was also broken up, creating new nations of Austria and Hungary. While Germany did not possess an empire, it was stripped of its African colonies, forced to accept sole responsibility for the war, and pay some $33 billion in reparations to Britain and France. Germany was also deindustrialized, and the area to the west of the Rhine River was demilitarized. Without the industrial wherewithal to pay the reparations, the German economy was essentially crippled. To pay the reparations, they borrowed money from banks in the United States (the Germans had also borrowed heavily from US banks in the decade preceding the First World War). The British and French then used the reparations payments to repay their wartime loans from the United States, who had emerged from the war as an international creditor. The US banks then loaned the money back to Germany. The stock market collapse of 1929 and subsequent banking collapses of the early 1930s disrupted this precarious and unstable system. Unable or unwilling to continue, US banks stopped the loans to Germany, who then defaulted on their reparations payments to England and France. The British and French no longer had the funds to repay their loans to US banks. Without the infusion of funds from the United States, the system collapsed and world trade evaporated. The United States Congress passed high protective tariffs of up to 67% on selected agricultural commodities to protect their own markets. President Hoover reluctantly signed the Hawley-Smoot Tariff despite the opposition of the nation's most prominent economists. The British created an Imperial Preference System, and Germany contemplated a policy of economic self-sufficiency. World trade, which stood at $36 billion in 1929, dropped to $12 billion in 1932 [4].

The tariff and trade situation was exacerbated by the international gold standard. Under its provisions, a nation was obligated to pay off any trade deficit in gold on an annual basis. However, since gold also functioned as a domestic currency, nations had to drain their domestic currencies in order to square their international balances. Theoretically this was supposed to reduce prices and make a nation's exports more attractive to potential importers. In practice, the reduction of money touched off not only falling prices (deflation) but also unemployment, recession, and international speculation of debtor nation's currencies. Panicked investors in the United States withdrew their deposits, precipitating a banking panic in 1930. Faced with just such a gold drain, the British suspended the gold standard in 1931, adding to the predicament of banks with the withdrawal of international deposits. In addition Hoover advanced legislation to increase tax rates in order to enhance revenue and balance the domestic budget. He believed that balancing the nation's budget would provide the banking system with desperately needed liquidity. However, the economy slipped deeper into depression as wealth creation declined, along with tax revenues. The Federal budget slipped into a deficit of $2.7 billion, which was the largest peacetime deficit in American history. Much of this deficit resulted from Hoover-era policies to stimulate the economy by means of injecting funds into the struggling sectors of the economy.

Congress passed the Glass-Steagall Banking Act in 1931 which not only made the banking system safer by separating speculative securities trading (investment banking) from taking deposits and making loans (commercial banking) but also made possible for the nation's central bank (the Federal Reserve) to release large amounts of gold from its holdings thereby expanding the monetary base. In 1932 Congress passed the Federal Home Loan Bank Act which allowed banks to present mortgage paper for rediscounting at the Federal Reserve and allowed banks to use mortgages for collateral in obtaining loans of badly needed capital. Finally Hoover proposed the creation of the Reconstruction Finance Corporation (RFC) which was designed to allow the government to loan taxpayer dollars directly to struggling financial institutions. Congress initially capitalized the RFC at $500 million and authorized it to borrow up to $1.5 billion. The RFC was the progenitor of the Troubled Assets Relief Program (TARP) created in the waning days of the administration of George W. Bush to deal with the financial collapse of 2008. The reaction in 1932 was as mixed and varied as was the reaction in 2008–2009. Progressives called it "socialism for the rich." *Business Week* hailed the RFC as "the most powerful offensive force that governmental and business imagination has, so far, been able to command" [5].

However, given Hoover's position that the depression was of foreign origin, his domestic policies were both tepid and hamstrung by his view of how the international economy functioned. Hoover remained committed to the principle of voluntarism and had to begrudgingly accept institutions such as the RFC. But more importantly he was more strongly committed to two of the most sacred principles of classical economics: the belief in balanced budgets and an unwavering fealty to the gold standard as the lynchpin of the international economy. He raised interest rates and taxes when the system cried out for increased credit and increased spending, largely because of his belief that not doing so would increase the gold drain and jeopardize the position of allies and trading partners such as Great Britain.

Hoover's Democratic rival, New York Governor Franklin Delano Roosevelt, had an entirely different conception of the causes of the depression. He believed its cause was primarily domestic. While a candidate FDR surrounded himself with a number of Columbia academics that was branded "the brains trust" by a *New York Times* reporter. Chief among his economic advisors was Rexford Tugwell, who was an adherent of the "stagnation thesis" advocated by economists such as Alvin Hansen and Paul Sweezy. Roosevelt came to accept Tugwell's perceptions that the mature economy has reached its frontiers and that no great epoch-making innovations would be forthcoming. The problem was one of overproduction of capital and not a shortage of it, along with the flip side of underconsumption. Roosevelt enunciated his belief in underconsumption in two speeches while a candidate one at the Commonwealth Club of San Francisco in September of 1932 and another at Oglethorpe University in Atlanta, Georgia, on May 22. The Commonwealth Club speech is worth quoting at length, as it foreshadowed the tenor of New Deal programs to come. The new Deal was to be about consumption instead of production, and equity instead of growth. Roosevelt's main theme was how to deal with the generalized overproduction he thought was the cause of the depression. This overproduction characterized the oil industry as well:

> » Our industrial plant is built; the problem just now is whether under existing conditions it is not overbuilt. Our last frontier has long since been reached, and there is practically no more free land…We are not able to invite the immigration from Europe to share our endless plenty. We are now providing a drab living for our own people. Clearly this calls for a reappraisal of values. A mere builder of more industrial plants, a creator of more railroad systems, an organizer of more corporations is as likely to be danger as a help. The day of the great promoter or financial Titan, to whom we granted everything if only he would build, or develop, is over. Now our task is not discovery, or exploitation of natural resources, or necessarily producing more goods. It is the sober, less dramatic business of administering resources and plants already in hand, of seeking to reestablish foreign markets for our surplus production, of meeting the problem of under-consumption, of distributing wealth and products more equitably [6].

The New Deal was neither a well-enunciated program nor a manifesto for economic growth. Rather it was a set of sometimes contradictory experiments to pursue the goals of rescue, recovery, reform, and restructuring. FDR's lieutenants, acting on incomplete information and in collaboration with Hoover's financial advisors, declared a national bank holiday, closed insolvent banks, recapitalized them through the RFC, and reopened them for a trusting and newly confident public. FDR's "fireside chats" themselves helped to restore confidence among a battered and beleaguered public. The chief advisor and organizer of the Brains Trust, Raymond Moley held the belief that the efforts essentially saved capitalism in 8 days [7].

Since that time the administration of Franklin Roosevelt has set the standard for presidential performance. He passed 16 major bills in his first 100 days in office, most reflecting his concerns about overproduction and his fiscal orthodoxy which entailed a belief in balanced budgets. In retrospect this fiscal orthodoxy accounts partially for the fact that unemployment remained stubbornly high throughout the course of the depression.

Year	Unemployment rate
1929	3.2
1930	8.7
1931	15.9
1932	23.6
1933	24.9
1934	21.7
1935	20.1
1936	16.9
1937	14.3
1938	19.0
1939	17.2
1940	*14.6*

Source: Historical Statistics of the United States, p. 73

In addition to the banking bill, the first 100 days saw the Beer and Wine Act which was designed to raise revenue in anticipation of the repeal of Prohibition and the Economy Act designed to cut $500 million from the Federal budget. FDR advanced two bills to deal with the stubbornly persistent problem of unemployment. The Civilian Conservation Corps (CCC) put a quarter million youth to work beautifying the nation's countryside and working on flood control and forestry projects. The Federal Emergency Relief Act injected Federal money directly into depleted state coffers for the purpose of unemployment assistance. Concerns over energy were also a crucial component of the legislation of the first 100 days when Congress created the Tennessee Valley Authority (TVA). The Federal government had built a dam at Muscle Shoals, Alabama, to provide power for the production of nitrates, which are the basis of not only explosives but fertilizer. The dam was completed too late to be of use for the war effort, and a cohort of private utilities successfully blocked the efforts of progressive Republican George Norris to have the Federal government operate the dam. The act created not only the authority to operate the dam but also charged the TVA with flood control, the combating of soil erosion and deforestation, and the construction of additional dams to bring electricity to the depressed rural South.

Faced with a 95% decline in home construction since 1929, Congress created the Home Owners Loan Corporation, rather than committing to the large-scale expansion of public housing, as recommended by New York Senator Robert Wagner. The HOLC stopped the surge of defaults (up to 1000 per day) and introduced standard accounting practices into mortgage lending. This was followed by the creation of the Federal Housing Administration in 1934. Traditionally mortgages required a 50% down payment and a short-term, interest-only loan. If the homeowner was diligent with his or her payments, the note would be refinanced for another 5 years. But when the banking system crashed repeatedly from 1929 to 1933, banks were simply not in a position to refinance the loans even if those homeowners who had retained their jobs were able to make the interest payments. The FHA replaced these traditional mortgages with low-down-payment, long-term (up to 30 years), low-interest, amortized loans where both principal and interest were repaid in equal monthly payments. Moreover, the FHA insured these mortgages from default. Despite the insurance, bankers were reluctant to write FHA loans. Some were worried about government intrusion, while others were concerned about holding on to a low-yield asset for some 30 years. To allay the fears of the bankers, Congress subsequently created the Federal National Mortgage Association (FNMA—better known as "Fannie Mae") to bundle the mortgages into securities which could be sold on short-term markets. FNMA functioned successfully as a government corporation until it was privatized in 1968 [8].

Congress passed the Agricultural Adjustment Act on underconsumptionist grounds. The bill was designed to restore the balance between industry and agriculture and raise farm incomes by restricting crop output in order to raise agricultural prices. Increased rural incomes would provide the wherewithal for the purchase of the output of industry. The bill was paid for by increased taxes on agricultural processors. The hallmark of the first 100 days was the passage of the National Industrial Recovery Act. The NIRA, along with the AAA, was aimed not just at recovery but also restructuring of the economy on the basis of rational economic planning to replace the newly-failed market system as the basis for regulation of prices and output. However, the Supreme Court found the NIRA and AAA unconstitutional in 1935. The conservative bloc was joined by liberal anti-monopoly crusader Louis Brandeis, who objected to the suspension of the antitrust provisions.

The National Industrial Recovery Act (NIRA) established the National Recovery Administration (NRA). It provided a series of complex codes by which business would comply with the need to combat overproduction in order to receive funds. The act also allowed for labor unions to bargain collectively, and it established minimum wages and maximum hours. The law virtually suspended antitrust laws. Economic theory holds that monopolies restrict output and raise prices, a strategy tailor-made for remediating falling prices and overproduction. This allowed the Federal government to plan rationally the output and prices for whole industries. The law also established the Public Works Administration (PWA) designed to administer a large-scale and ambitious infrastructure construction agenda. The PWA was charged not only with the construction of energy-related projects, but it also assumed the duties of stabilizing the near-anarchy of the oil fields of the Southern Plains [9]. After the First World War, fears of oil shortages reared their head. These fears were allayed by two large oil discoveries. In 1926, interestingly enough the peak of the 1920s automobile boom, oil was discovered in the Permian Basin in West Texas and Oklahoma. As is common, large new additions to the supply of oil depressed prices. Oil that was selling at $1.85 per barrel in 1926 averaged only about $1 per barrel in 1930. Then, in 1930, another huge discovery was made in East Texas, one that dwarfed the combined output of Pennsylvania, Spindletop, and Signal Hill in California. The East Texas wells added another half a million barrels per day to the oil supply. Consequently prices dropped again to as low as 10 cents per barrel in the glutted market, adding to the already falling price level precipitated by the Depression. The Texas Railroad Commission, established in the Populist era to exert control over railroads, assumed the responsibility (despite dubious legality) of regulating oil production by regulating its transport. The strategy of the Railroad Commission was one of "pro-rationing" or limiting oil shipments to a fraction of oil reserves. Problems arose in Texas and Oklahoma (where the Interstate Commerce Commission employed a similar strategy), when producers exceeded their allotted shares, shipping illegally what came to be known as "hot oil." The problem became so pronounced that Texas Governor, Ross Sterling, declared that East Texas was in a state of insurrection and called upon the Texas Rangers and the National Guard to quell the problem.

The NRA was first called upon to impose its codes to reduce competition and stimulate economic recovery. The problem was severe enough, however, that newly appointed Secretary of the Interior Harold Ickes brought the regulation of the East Texas fields under the aegis of the interior department when he was informed, in August 1933, that oil prices had fallen to three cents per barrel. The Oil Code, established under the NRA, gave Ickes the power to set monthly quotas for each state. The anarchy in the oil fields abated under the auspices of the NRA and Interior Department. However, when the NIRA was declared unconstitutional in 1935, a separate law, the Connally Hot Oil Act, was established to maintain price stability [10]. The Texas Railroad Commission remained effective at reducing cutthroat competition and stabilizing prices until the 1970s. Petroleum Geologist Kenneth Deffeyes, a colleague of Hubbert, realized that the US oil supply had indeed peaked when in 1971 he read in *San Francisco Chronicle* that the Commission instructed oil companies that they could produce at 100% of capacity! [11]. The Roosevelt Administration responded to the Supreme Court's decision that the NIRA and AAA were unconstitutional by launching a broad and progressive agenda of reform, restructuring and redistribution in 1935, often called "The Second New Deal." The year of 1935 saw the passage of the Social Security Act, providing pensions for the elderly. It was ostensibly devised to reduce unemployment by removing the aged from the labor force to reduce unemployment and was constructed on the principle of private insurance rather than as a dole. Once again, FDR's fiscal orthodoxy necessitated that the program be funded by regressive payroll taxation rather than from the Treasury. The increase in taxes precluded any large-scale stimulative effect. The Social Security Act also provided for Aid to Dependent Children, later modified to become Aid to Families with Dependent Children (AFDC) soon to become the backbone of the Great Society welfare programs of the 1960s. The government also became an employer with the creation of the Works Progress Administration (WPA). The WPA created jobs for construction workers who built miles of highways, public buildings, and university campuses. The WPA also employed engineers writers and artists. In the first year of the program, the WPA employed more than 3 million people and 8.5 million over the life of the agency [12].

Further provisions were advanced to reform structurally the nation's financial system. The Federal Reserve was given increased powers to conduct open-market operations which entail the buying and selling of preexisting Treasury securities, needed now that the gold standard was abandoned. Moreover, a tax bill created a strongly progressive income tax in order to achieve the goal of fairness embodied the New Deal philosophy. Roosevelt's program was certainly eclectic, with a blend of progressive and regressive taxes in conjunction with increased spending. It depended upon no clearly enunciated economic theory, such as that of John Maynard Keynes. These rates, up to 79% for the top incomes, were accompanied by high inheritance taxes which were designed to reduce the intergenerational transmission of wealth. Perhaps the most important law of the reform era was the creation of the National Labor Relations Board. The senator Robert Wagner also inserted a provision (Section 7A) into the NIRA. This section established labor unions, formerly seen as "conspiracies in restraint of trade," as the legitimate representatives of workers in the process of collective bargaining would increase wages and serve the goals of redistribution, but it would also bring about labor peace. The new board would replace the organizational strike with a monitored election. It was also the vehicle that enabled the development of the capital-labor accord that would become a crucial pillar of postwar prosperity. The New Deal ostensibly came to an end in 1938 with the passage of the Fair Labor Standards Act. This act established the 40-hour standard work week and further solidified minimum wages [13]. While the New Deal was successful in establishing significant structural reforms and developing a faith in government that has not been seen since, it was never successful in eliminating the stubborn specter of unemployment. Moreover, New Deal policies were not directed toward economic growth. Contrary to public opinion. He would try contradictory policies to see if they would work. He also believed in a balanced budget, so most spending programs were accompanied by tax increases to pay for them. As Keynes would later tell us, this reduced the "multiplier effect" and led to a very tepid recovery, that is, until the Second World War. The focus of government policy would change significantly with the advent of the Second World War.

10.3 The Second World War and the End of the Depression

The United States entered the Second World War on December 8, 1941. However, the country had been providing food, armaments, and much-needed oil to embattled Britain for more than a year, as President Roosevelt officially declared the United States to be the "Arsenal of Democracy" in December of 1940. The country had been supplying war materiel to the allies since 1939. Historian David Kennedy states the matter concisely: the war was won with Russian lives and American machines. "…the greatest single tangible asset the United States brought to the coalition in World War II was the productive capacity of its industry" [14]. While the war ended the depression, the conditions of the depression were also instrumental in mobilizing for the war. At the war's onset, nearly 9 million workers were unemployed, and half of the nation's productive capacity lay idle. By war's end the impressive economic machine produced nearly 300,000 aircraft, 5777 merchant ships, 556 naval vessels, nearly 90,000 tanks, and over 600,000 jeeps. Of the 7.6 billion barrels of oil used during the war, 6 billion came from the United States. Given the tremendous finds of the late 1920s and early 1930s, the United States possessed an enormous surplus of 1 million barrels per day out of a total production of 3.7 million barrels per day. By war's end oil production had risen to 4.7 million barrels per day. Moreover, the technological change making 8-ringed gasoline in a circle (octane) so that higher compression ratios could be used, along with a guaranteed market for this expensive process, allowed petroleum engineers to refine 100 octane aviation gasoline. This allowed American planes to fly farther and maneuver more agilely with up to 30% more speed and power than their German and Japanese rivals.

The United States supplied more than 90% of the 100 octane aviation fuel. The development of long-distance warplanes allowed for escort cover in the all-important trans-Atlantic tanker routes, which previously had been decimated by German U-boat activity. In addition, the new long-distance bombers destroyed the German coal gasification (Fischer-Tropsch) plants. By war's end German commanders were ordered to move troops and equipment with horses and mules, saving precious gasoline only for battle. Aviation victories so crippled the Japanese war machine of fuel that they had to leave the world's largest battleship in port for lack of fuel and they attempted to fly their technologically advanced Mitsubishi fighters (the famed Japanese Zero) on turpentine [15].

The United States was to be much-changed by the war. It was the only belligerent nation in the history of the world to see its standard of living rise during wartime. Economic concentration would increase, labor union militancy would be tamed in support of the war effort, and women and African-Americans would enter the ranks of industrial production and clerical work in unprecedented numbers. In 1939 the unemployment rate stood at more than 17%. By 1944 it fell to 1.2%. Not only did the rate fall, but the economic prowess of the country absorbed an additional 3 million new labor force entrants along with more than 7 million workers who were previously excluded from active labor force participation, mainly women. Perhaps most importantly, from a perspective of economic policy, the agenda of the Roosevelt Administration turned from one of stability and social equity to one of more and more production. The Second World War saw the birth of growth economics.

Industrial concentration increased during the war, abetted by government policy. Two-thirds of all procurement contracts were given to 100 corporations. The thirty-three largest accounted for half of all government contracting. After-tax corporate profits rose from $6.4 billion in 1940 to $11 billion in 1944. At war's end, the government turned over some $17 billion of publicly funded plant and equipment to private industry at bargain-basement prices. Two-thirds of it was purchased by 87 companies. Changes in production techniques accounted for a great deal of the increase in output. Everything from tanks to planes to Liberty Ships was constructed using the mechanized division of labor, which eliminated the need for overall skill that had been used so successfully by the automobile industry in the 1920s. In an attempt to deal with rising prices occasioned by shortages of crucial inputs and a plethora of money, the Office of Price Administration (OPA) would impose comprehensive wage and price controls. Nonetheless the inflation rate during the war was 28% and farm prices rose by 50%. Things had not been so good on the farm since the early days of the Republic. Organized labor would receive a reward for their slowly growing wages and no-strike pledge in the form of "maintenance of membership" provisions, guaranteeing that business accepts the closed shop requiring union membership as a condition of employment. Union dues were collected by firms themselves through payroll deductions. Gasoline was rationed. The owner of the standard "A" coupon would receive somewhere between 1.5 and 4 gallons per week, depending upon their location. The lucky few with an "X" coupon (e.g., doctors, clergy) still received unlimited supplies. Gasoline consumption fell by 30% from 1941 to 1943 [16].

To fund the war the Roosevelt Administration raised taxes. The income exemption at the bottom was lowered bringing some 13 million new taxpayers into the system. They paid at work, as the innovation of withholding tax made its first appearance. The top marginal tax rates were increased to 94%, so that the wealthy paid most of their income to taxes. Despite the tax increases, the new revenues were able to cover only 45% of the war's cost, as the United States devoted fully half of its gross national product to war spending. The rest was borrowed. Working people bought war bonds, sold to them by celebrities such as Hollywood actors (including Ronald Reagan) and popular musicians as a matter of patriotic duty. Commercial banks did their part, increasing their purchases of Treasury bonds from less than $1 billion in 1941 to more than $24 billion in 1945 [17].

The long and destructive war, started mostly because of a quest for oil and land to grow food for a rising German population, ended in August of 1945. The sheer might of American productive capacity was too much for the beleaguered axis powers to withstand. The Red Army had stopped the Nazi advance toward the Caspian oil fields. Rommel's tanks ran short of gasoline, losing North Africa and opening up Italy for allied invasion and victory. Japan's objective of control over Indonesian oil supplies was never realized. Short on fuel to run their war machine and pummeled by incendiary and atomic bombs, Japan surrendered on August 14, 1945, thereby ending the war. Their great Admiral Isoroku Yamamoto, the designer of the very successful Pearl Harbor attack, had studied in the United States and understood that Japan could never win the war because of America's enormous industrial potential.

10.3.1 The Postwar Economic and Social Order

The United States emerged from the war in an unprecedented position of economic, political, and military power. The nation was the only industrial power in the world, as those of its traditional rivals were decimated in the war, and it supplied the majority of the world's oil. European cities lay in ruins. The allies were deeply in debt, while the United States was the world's greatest creditor. At war's end the allies met in Bretton Woods, New Hampshire, to reconfigure the international monetary system. Unlike the aftermath of the last Great War, no pretense was made of returning to the gold standard which had worked so poorly and helped create the conditions of poverty that helped precipitate the next war. The dollar was "as good as gold" and tremendous advantages flowed toward the United States, consolidating its dominant position. Basic commodities were priced in dollars, and the country did not have to contend with international price fluctuations. Sufficient money was available for the expansion of American business into the devastated markets of Europe and Asia, and American exports soared, as did foreign direct investment. The dollar alone was denominated in gold, and the rest of the world's currencies were pegged to the dollar. The US agreed, in return, to redeem foreign currencies in gold at $35 per ounce. To rebuild war-torn Europe, the International Monetary Conference created the International Bank for Reconstruction and Development, better known as the World Bank. They were to make large-scale loans for the rebuilding of infrastructure—roads, bridges, power plants, refineries, office buildings, and factories. To provide adequate liquidity, or readily available money, the International Monetary Fund was created. In addition the Fund was charged with buying and selling currencies in order to keep them in balance with the dollar at the agreed-upon rate. Since the use of protective tariffs and beggar-thy-neighbor policies dried up world trade and helped transmit the depression internationally, the conference also created a General Agreement on Tariffs and Trade (GATT) to encourage free and open trade. The belief was that nations that trade with one another do not go to war. While the Conference proceeded on Keynesian lines, the plan of British delegate John Maynard Keynes, for an international clearing union was not accepted. Keynes' plan provided a framework whereby nations with large trade surplus would redistribute money to nations with large trade deficits in order to keep trade balances within reasonable bounds. The United States was not only the world's most powerful nation; it was also the world's largest creditor. American representatives were in no mood to adopt Keynes' plan, and they had the power to prevent its implementation. The GATT would have to suffice, although those present hoped for a more fully functional World Trade Organization. The WTO was finally created in 1995. However, the United States did supplement the World Bank funds with its own initiative known as the Marshall Plan.

10.3.2 The Marshall Plan

The theoretical ideas behind the Marshall Plan, conceived by General George C. Marshall, were economic and political. Many political parties in Western nations such as Italy, West Germany, France, the Netherlands, and even Britain found socialism and social democracy appealing in the chaos that followed the war. In a sense the Marshall plan was an attempt to save capitalism in the industrialized world.

The United States provided almost $9 billion into the European economies to ward off the growth of indigenous socialist movements by strengthening the financial markets and production capacity of European democracies. Most of that money (up to 80%) was used to purchase US exports. The framers of the Marshall Plan realized that no single market economy could thrive in a sea of economic stagnation. The Marshall Plan brought countless young scholars to be educated in "the American way of life." It also insured that American corporations would gain entry into formerly protected colonial markets. The United States also agreed to sacrifice some of its domestic declining industries to the greater good of free trade. At the time, this was highly favorable to the expansion of American business. US foreign direct investment increased from $11.8 billion in 1950 to $76 billion in 1970. The share of total profits from foreign operations also rose from 7% in the early 1950s to 21% by the early 1970s. At the same time, up to 46% of all deposits in major New York banks were derived from foreign sources [18].

Back at home the economic scene changed on the domestic front. With the international economy serving as a lucrative source of income and profits, large corporations began to share more with labor in order to achieve labor peace and create a domestic source of demand for their products. They could have both rising profits and rising wages. After the "Treaty of Detroit," productivity bargaining became the pattern in large industry. Since wages increased with productivity, labor had a strong incentive to increase productivity. Since a modicum of democracy was written into work rules, and wages were supplemented with retirement pensions and health-care benefits, once militant workers now had a strong stake in maintaining the system they once struggled against. Consequently productivity (or output per worker) grew at 2.9% per year in the 1950s and 2.1% per year in the 1960s. In contrast it would fall to 0.3% in the stagnant 1970s and "recover" to a tepid 1% per year in the supposedly prosperous 1980s. Wages rose by an average of 2.9% per year in the 1950s and 2.1% per year in the 1960s, while gross national product grew at an annual rate of 3.8% and 4.0% in the same time period. Corporate profits remained strong.

From the late 1940s when the Marshall Plan was implemented until the oil boycott of 1973 after-tax profits grew at 7% annually. During the epoch of stagflation of the 1970s, they fell to 5.5% [19]. The American public exited the war with the greatest accumulation of savings relative to income in any time in the country's history. Wages rose and unemployment fell, prices were controlled, and consumption was held in check by tax increases, patriotism, and the fact that so many crucial materials were requisitioned for the war effort. The prominent economist John Maynard Keynes reasoned, in *The General Theory of Employment, Interest, and Money*, that the buildup of excess savings was a primary cause of the Great Depression. But such was not the case in the postwar United States. Deprived of consumption by 10 years of depression and 5 years of war, Americans were once again, like they were in the decade of the 1920s, on the verge of being consumers once again. Economists called this "pent-up demand."

Accrued savings plus the additional worker and business income translated into growing consumption expenditures, especially with regard to gasoline, automobiles, and housing. Capital formation grew as well growing at 3.5% per year from war's end until the mid-1960s and 4.3% per year from the mid-1960s until the beginning of the economic crisis in 1973. Horsepower in manufacturing grew from 49,893,000 in 1939 to 151,498,000 in 1962. Total consumption expenditures increased dramatically from $70.8 billion in 1940 to $191 billion in 1950 to $617.6 billion in 1970. Spending on gasoline and automobiles increased as well. In 1943, the year the last automobile was constructed for the duration of the war, only 100 cars were sold in the United States, but by 1950 more than 6.6 million cars received new tags. The pre-stagflation-era figure peaked in 1965 when more than 9 million cars left the showroom floor. One could tell something ominous was happening for the automobile-crazed population. By 1970 passenger car sales declined to less than the 1950 level. A similar pattern existed in housing. In the depths of the depression, only 221,000 new dwellings (public and private) were started. In 1950 the nation's building contractors and trade workers constructed close to 2 million homes. After that a high level, exceeding 1 million new homes per year, existed in every year whether prosperity or recession. However, by 1970 only 1.5 million new homes were started. The new suburban homeowners motored to their new dwelling, many made possible by Federal Housing Administration mortgages, or the even more attractive mortgages offered by the Veterans Administration (no money down and the mortgages were guaranteed not just insured). Spending on gasoline soared from $332,000 in the war years of rationing to nearly five and a half trillion dollars in 1970 [20]. Gasoline prices remained cheap, as the United States, which at the time still produced 52% of the world's oil, was relatively unaffected by world events and price spikes such as the one caused by the Suez Crisis of 1956. In 1950 the price per barrel of oil was $2.77 or an inflation-adjusted price of $25.10. The real price of oil did not exceed this level until 1974, during the first oil crisis of the 1970s [21]. Thus the general progress of industrialization was accelerated by the incredibly cheap source of its fuel.

10.3.3 Emergence of the Importance of the Middle East

US oil companies strengthened their position in the all-important Arabian Peninsula, soon to become the world's largest source of crude oil. The original concession was given to Standard of California in 1933 for an up-font payment of $175,000 and an additional $500,000 to be given to King Ibn Saud if oil were to be found. Standard of California was soon to bring Texaco into the consortium to form Aramco (the Arabian-American Oil Company). In 1933 Gulf Oil, headed by Hoover's Treasury Secretary Andrew Mellon, received a 50% share of the oil newly found in Kuwait, a concession they would share with Anglo-Iranian Oil Company (soon to become British Petroleum). After the war Aramco found that they had insufficient marketing operations to dispose of all the oil being pumped from the Saudi fields. They entered into a broader consortium with Standard Oil of New Jersey (soon to become Exxon) and the Standard Oil Company of New York (soon to become Mobil). Aramco overcame the stranglehold of Shell and Anglo-Iranian for marketing in Europe, and fears of overproduction were allayed. Gulf Oil, which was long on crude and short on markets, entered into a consortium with Shell, which was long on markets and short on crude. The basic conditions for expansion, increased production, and increased marketing capabilities were in place. The era of economic growth, based on a social structure of accumulation amenable to business ascendency and lots of cheap oil, was in place [22]. In 1991 the testimonial given on the back cover of *The Prize* by Nobel Laureate in Economics Paul Samuelson put the matter succinctly. "Dan Yergin lucidly and with grace explores the dynamics of the global business that has helped shape the modern economy and *fueled the economic growth on which we have come to depend*" (emphasis added).

The immediate postwar period was also the era of decolonization. Throughout Africa and Asia nation after nation gained independence. Oil-producing nations moved to increase the share of Ricardian rents, or return to pure ownership, for their precious resource. The original concessions of the late nineteenth and early twentieth centuries gave the international oil companies ownership rights of the oil for initial payments and an agreed-upon royalty per barrel. Countries that granted concessions were interested in having the oil companies pump as much oil as possible as it enhanced their revenues. The oil companies, however, were ever mindful of the industry's history of gluts and falling prices. The companies, therefore, had an incentive to limit production to what they could market, and the companies were in charge of production. The aforementioned oil deals resulted in a tight oligopoly which Italian industrialist, and head of *Azienda Generali Italiana Petroli* (AGIP), Enrico Mattei dubbed "The Seven Sisters" (Standard of New Jersey, Standard of New York, Standard of California, Gulf, Texaco, British Petroleum, and Royal Dutch Shell). Oligopolies, as you may recall, pursue a strategy of maximizing profits in the long term by means of limiting output, maintaining stable prices, and enhancing control over production, marketing, and distribution. Fearing nationalization of their Venezuelan concession, Standard Oil of New Jersey agreed to split the rents on a "fifty-fifty" basis. The deal was soon transmitted to the Middle Eastern producers, and the potential instability abated, albeit at higher costs to the oil companies. Royalties were to be paid at an official "posted price" that could differ from the market price. At the time of the deal, the posted price generally exceeded the market price, which was kept low by the tremendous surplus capacity of oil. This transmitted an even greater share of the rents to the producing countries. However, US oil companies were aided by their government, as cost increases were softened by a provision in the US tax code that allowed them to count the new rent payments as taxes and deduct them from their US obligations. Essentially the stability of the oil industry was paid for by US citizens. But oil was cheap and plentiful and incomes were rising. There was no tax rebellion in the United States. However, as we saw in ► Chap. 6, new forms of competition can destabilize an oligopoly structure. Independent oil companies wishing to break into Middle Eastern production such as Getty Oil in the United States and Enrico Mattei's AGIP simply offered a greater share of the rents as the price of entry. The era of colonial subservience on the part of producing nations was beginning to end. Yet the acquiescence of oil companies and governments to the new rent sharing plan provided stability for years to come [23].

10.3.4 The Age of Economic Growth

At the end of the war, all the pieces for a renewed era of prosperity were in place. American companies gained vast and profitable international markets. Few if any foreign corporations were in a position to compete effectively. The United States was the most powerful nation in the world, economically and militarily. The world monetary system was based on the dollar. American workers received wages that grew with productivity. As a result, productivity growth, much of it derived from the application of cheap oil [24], fueled increased profitability, and the increased wages, along with historically unprecedented savings, served as the basis for an explosion in consumption. The war showed more than anything that Keynesian economics, based on deficit spending and public funding of infrastructure, worked. In this era American Keynesians, now calling their approach "The New Economics," began to transform the works of Keynes from a theory based on the problems of uncertainty and speculation into a herald call for economic growth.

The year 1945 saw more workers out on strike than any year in American history as labor unions sought to recoup the perceived losses from wage controls and from signing a no-strike pledge during the war. Part of this dilemma was solved by the generalized acceptance of productivity bargaining and limited capital-labor accord following the collective bargaining agreement between the United Auto Workers and General Motors, known as "Treaty of Detroit." Congress also restricted labor union rights by passing the National Labor Management Act (better known as the Taft-Hartley Act) over the veto of President Truman. In addition Congress moved, on the advice of the New Economists, to deal with the fears that large-scale unemployment would emerge once the stimulus of the war ended by passing the Employment Act of 1946. The measure started originally as Senator Robert Wagner's "Full Employment Bill." Wagner's proposal gave every American the statutory right to a job. If they could not find one in the private sector, the government would create one for them, as they had during the depression under the auspices of the Works Progress Administration (WPA). The bill was to be paid for by a tax on employers. Not surprisingly American Business opposed the bill. Not only did they dislike the taxes to be levied on them, but the general belief was that the absence of the power to dismiss workers would make labor discipline and productivity increases impossible. The eventual legislation was the result of political compromise. The act directed the government to pursue policies that would result in "reasonably" full employment, stable prices, and economic growth. Growth would be the mechanism that enabled the other two goals. Economists Samuel Bowles, David Gordon, and Thomas Weisskopf argued that this stalemated the traditional goals of the labor movement, those of full employment and income redistribution, and replaced them with the imperatives of economic growth [25]. The act also obligated the president to give an annual economic report to the Congress, as well as mandating the creation of a Council of Economic Advisers.

The movement toward a strategy of economic growth, which was not at all apparent in the work of Keynes, began in earnest with the work of the Council of Economic Advisers (CEA), especially after Leon Keyserling advanced to the chairship in 1949. The philosophy of secular stagnation was banished to the past as population grew with the baby boom, military technologies began to impact the civilian world, and new frontiers emerged as former farmland was converted to suburban homes. In Keyserling's imagination, growth could achieve two goals beyond the attainment of reasonably full employment. If the economy grew, more could be given to those at the bottom of society's income distribution without raising taxes and taking it from those at the top, which might adversely impact production and profits. The council firmly believed that only growth could "reduce to manageable proportions the ancient conflict between social equity and economic incentives which hung over the progress of enterprise in a dynamic economy" [26]. With a growing economic pie, the benefits could be shared more easily with more sectors of the economy. The other imperative for growth lie in the needs of the cold war. In 1949 the Soviets detonated an atomic bomb, and in 1950 President Truman directed the departments of State and Defense to devise a new set of priorities for the new realities of the world. The resulting document prepared by the National Security Council was NSC-68. Economic growth was at the heart of the strategy. Only through economic growth could the United States meet its domestic priorities of achieving reasonably full employment and stable prices yet, at the same time, fund its new military objectives of arming "friendly" client states that was the heart of the "Truman Doctrine." However, Truman was somewhat tepid in his acceptance of economic growth, and the subsequent president, Dwight Eisenhower, was rather indifferent, preferring a strategy of price stability. The true era of the liberal growth agenda would come during the presidencies of John Kennedy and Lyndon Johnson.

The "New Economists" of the era believed they had conquered the business cycle such that recessions and depressions would be a thing of the past. By means of fiscal policy (taxing and spending) and monetary policy (money supply and interest rates), the New Economists could fine-tune the economy as if it were a well-oiled machine. If the economy performed sluggishly, the government could stimulate the economy and the increased spending would translate into an expansion of output and jobs. If prices rose to uncomfortable levels, inflation would be controlled by subtle downward adjustments in spending or the amount of money available to the economy. Moreover, inflation only occurred once full employment had been achieved and resulted from demand that was in excess of what the economy could produce at full employment. So any reduction in demand would decrease prices but not employment, at least in theory. In terms of policy, the liberal growth agenda rested upon three main pillars.

Current production had to be balanced with existing productive capacity. This was accomplished by expanding demand via the Kennedy-Johnson tax cuts. Costs were kept in line with wage-price guideposts and the use of presidential authority to convince union leaders to mediate their wage demands. This was known as "jawboning." Finally, growth was stimulated by encouraging investment. Policy instruments included accelerated depreciation and Kennedy's famous investment tax credit. In the 1960s their policies resulted in impressive outcomes. Unemployment rates were less than 4% by 1966, real (or inflation-adjusted) gross national product grew at 5% per year, and the inflation rate remained low. The number of Americans in poverty fell from 22.4% of the population in 1960 to 14.7% in 1966. The boom was driven by an increase in investment, with inflation-adjusted gross private fixed investment rising from $270 billion in 1959 to $391 in 1966. The only stubborn inconsistency was the degree of inequality, with the US distribution of income being more than four times as unequal as that of Sweden and twice as unequal as that of the Soviet Union. But policies of growth were to take precedence over those of distribution. President Lyndon Johnson believed that redistribution policies were doomed to failure because they were counter to the Puritan work ethic, they would be a political disaster, and they were counter to the growth agenda. Consequently the direction of the war on poverty was toward productivity enhancement of the poor rather than toward income maintenance programs. Despite the lingering inequality, conditions for many of those traditionally left out of prosperity did begin to improve with growth. Before the war Black men earned only 41% of the incomes of White men, and Black women earned only 36 percent of White women. By 1960 the figures had risen consistently with Black men now earning 67% of White men and Black women earning 70% of the wages of White women. The postwar prosperity was built on a series of growth coalitions with organized labor, the civil rights movement, and the women's movement basing their strategies of reaching the top on economic expansion sufficient to include them. It was a time when a far greater proportion of the population believed that the wise actions of the government could benefit them than is commonplace in the early twenty-first century [27].

As long as the material conditions of prosperity, international hegemony labor peace and rising productivity, cheap oil, and the domestic limitation of cutthroat competition, remained in place, expansionary monetary and fiscal policy could produce growth with stable prices. However, by the 1970s the very success of earlier action led to the demise of the postwar social structure of accumulation. By the 1970s domestic oil production peaked, Europe and Japan caught up in terms of productivity, inflation gripped the nation in conjunction with rising unemployment, wages began to fall, and jobs began to leave as the economy became both globalized and more competitive.

10.3.5 Peak Oil and Stagflation

A great deal has already been written about the era of stagflation, some of which will be reviewed in this chapter. However, what tends to be missing in an economics literature that concentrates primarily on social forces and the internal limits to accumulation and growth is the advent of external biophysical limits. It was in the 1970s that the biophysical limits, in the form of peak oil, began to affect world economics and politics. As per M. King Hubbert's prediction, domestic oil production peaked in 1970. Yet demand for oil to fuel transportation, heating and continued to grow at about 3% per year. The era of rapid and sustained economic growth based upon cheap oil came to a temporary end, giving rise to a decade of malaise in the United States and elsewhere, characterized by not only economic stagnation and high unemployment but rising prices as well. The oil shock did not come all at once, but in 1973 a series of events that had been building throughout the postwar period came to crescendo in the first energy crisis that affected the United States seriously.

In the 1950s the world oil industry was destabilized by the same force that destabilized historically the oil industry in the United States: large new discoveries, glutted markets, and falling prices. Crude oil production in the non-socialist world rose from 8.7 million barrels per day in 1948 to 42 million barrels per day in 1972, mostly as a result of discoveries in the Persian Gulf area. Consequently, although US production increased, the share of US production fell from 64% to 22% in the same time period. World proven reserves increased from 62 billion barrels to 534 billion barrels, excluding the socialist nations. In addition, greater quantities of Soviet oil entered the world market. By 1960 Soviet production was nearly 60% of the Middle East. This exceeded domestic demand, and the oil entered the world market, putting additional downward pressure on market prices. In April 1959 huge new discoveries of high-quality, low-sulfur oil (light sweet crude) were made in Libya, and by 1965 Libya was the world's sixth largest oil producer. The result was more cutthroat competition and falling prices. But oil companies had to pay royalties to producing nations on the basis of the official posted price, which was not falling with the increased supply. Consequently, their profit margins fell. In August of 1960 Standard of New Jersey unilaterally cut the posted price by 7%, enraging the oil-producing nations. Spurred on by the oil ministers of Saudi Arabia and Venezuela, the producers met with the intention of forming a body similar to the Texas Railroad Commission which would prorate shipments and allow them to control the decrease in prices. In September, the Organization of the Petroleum Exporting Countries (OPEC) was born [28].

Political turmoil hit the Mideast in the late 1960s. In 1967 Israel invaded Egypt. Saudi Arabia withdrew their oil from the world market in an attempt to create shortage and economic discord among Israel's supporters in Europe and the United States. However, the strategy was ineffective and led primarily to declining revenues for the Saudis. There was still sufficient spare capacity in world oil production and in the United States to make up for the difference. That was soon to change. Moreover, in 1969 a coup led by Colonel Muammar al-Quaddafi overthrew King Idris. The new government demanded a large increase in the posted price and ordered oil companies to cut production. With the Suez Canal still out of service, the quick trip across the Mediterranean enhanced the power of Libya. Furthermore the Trans-Arabian Pipleline (or Tapline) was ruptured by a bulldozer making oil transportation even more difficult. This set the stage for competitive price increases among producing nations. Iran increased its price in 1970 followed by Venezuela and Libya again. By the time negotiations came to an end, the posted price had increased by 90 cents per barrel. By 1970 the United States was essentially powerless to control the situation, as it no longer possessed the spare capacity to overcome events in the Middle East. US oil production peaked in 1970 at slightly more than 11 million barrels per day, never again to increase, despite increased drilling effort, new discoveries, and tremendous political pressure.

10.3.6 The Fateful Year of 1973

In September, Colonel Quaddafi nationalized 51% of the remaining oil companies not expropriated in the original coup. He worried little about retaliation as the spare capacity to overcome his moves no longer existed. Europe was simply too thirsty for Libya's light sweet crude. But this effort was dwarfed by the events of the following month. Still hurting from the humiliation of the 1967 defeat, new Egyptian President Anwar Sadat, in conjunction with Syria, launched a surprise attack on Israel during the holy month of Ramadan in the Islamic World but also on the highest religious holiday of Yom Kippur in Israel. Sadat's forces were on the verge of defeating the Israelis, who were running short of munitions and materiel. If they were not resupplied, they would lose the war militarily. The United States attempted to keep the resupply effort low key, but cold-war logic called for resupply, seeing as the Soviet Union had armed, and was resupplying, the Egyptian and Syrian forces. The plan was to land their huge transports under the cover of darkness. However, adverse weather in the refueling station in the Azores delayed the operation, and the American planes landed in broad daylight. Israel regrouped and staved off defeat. But events were soon to grow in scale. Outraged by the American resupply efforts, Saudi Arabia called for a boycott of oil to the supporters of Israel, particularly the United States and the Netherlands. The Saudis called for production cuts of 5% per month for the entire world and a complete cutoff to the Americans and the Dutch. They threatened the partners in Aramco with loss of the concession if they sent as much as one drop of oil home. So interestingly enough, it was the US oil companies themselves who carried out the mechanism of the boycott, not the Saudi State. As recently as 1967 the removal of oil from the world market had not worked as a political weapon for the Arab States, as sufficient spare capacity existed in the world market to overcome their efforts. This was no longer the case once the production of the world's major swing producer, the United States, had peaked. The Saudis withdrew about 16 million barrels per day from the world oil supply, and other producers had insufficient spare capacity to make up the difference. Iran increased oil exports by some 600,000 barrels, but they, and some others, could not compensate for the Saudi withdrawal. All in all, the world's oil supply fell by about 14%.

In the United States, gasoline prices quadrupled as the world price of oil increased with success of the Saudi boycott. To begin with, the nation's oil imports nearly doubled from 3.2 million barrels per day when domestic production peaked in 1970 to 6.2 million barrels per day in 1973. Before the October war, the posted price was $5.40 per barrel. By December oil was selling for as much as $22 per barrel. Gas lines became a feature of American life, as motorists would wait for hours to buy gasoline often to find the station had run out by the time they reached the pump. Calls for action to increase the supply abounded from all corners of the nation. However, the oil companies were no longer just American enterprisers but multinational corporations who tried to apportion the hardship equally among their various markets. There would be no special treatment for any particular nation, especially the United States. Patriotism did not include the potential loss of the Saudi concession for the American partners in Aramco. The United States president, Richard Nixon, was essentially powerless to do much of anything, embroiled as he was in the loss of his own job owing to the revelations of the Watergate Scandal. However, the effects of the oil price run-up wreaked havoc with his New Economic Policy, intended to break the specter of stagflation that had been emerging for years and abetted seriously by the increase in energy prices [29].

10.3.7 The End of the Liberal Growth Agenda

The liberal growth agenda was based on the idea that the government should stimulate economic growth but also had the power to "fine-tune" the economy to manage unemployment and inflation. This ideology and set of policies fell apart in the 1970s. The 1973 energy crisis was not the only force that crippled the US economy. In fact the pillars of postwar prosperity were all crumbling. The rising power of the oil-producing nations was only one sign of the end of *Pax Americana*. There were many others. Europe and Japan, once war-torn nations in a state of shock, caught up to, and even surpassed, the United States in terms of industrial output. The terms of trade, or the ratio of export prices to import prices, rose from near parity to 1.3 to 1 in 1972. They plummeted to less than 1.1 to one by 1979. Despite the rising cost, imports increased from 4% of GNP in 1948 to 10% in 1972. The US share of total world exports in 1955 was 32%. They stood at only 18% in 1972. The postwar monetary system was based on fiat money, where the value of a nation's currency depends upon productive power and political stability. American productivity growth, which averaged 2.7% per annum in the 1950s, fell to 0.3% per year in the 1970s. As the rise in oil prices attests to, the United States no longer bought in a buyer's market and sold in a seller's market. Moreover, the expansion of cold-war military spending plus the outflow of funds directed toward direct foreign investment worsened the US balance of payment situation. The Bretton Woods Accords mandated the United States to convert holdings of foreign currencies to gold at the price of $35 per ounce. By 1973 outstanding claims exceeded the American gold stocks. Richard Nixon "closed the gold window," and the Bretton Woods Accords collapsed, thereby ending the dominant position of the dollar and all its benefits. Soon, the world was to open up to an unprecedented increase in global oligopolistic rivalry. The days of the insulated oligopoly position of US business were nearing their end, and the demise was reflected in the decline of corporate profits. After-tax corporate profits for the nonfinancial sector, which averaged 10% in 1965, dropped to less than 3% in 1973.

With productivity growth on the decline and international dominance eroding, American corporations could no longer "afford" the expensive mechanisms of labor peace erected in support of the capital-labor accord. Less access to energy was a primary factor in the decline of productivity growth, and the capital-labor accord depended upon growing productivity. An "open shop movement" began in housing construction by the 1980s, and myriad consulting firms specializing in "managing without unions" also emerged. As a result, wages began to fall. Hourly income, which grew at 2.2% per year in the long expansion of 1948–1966, grew only at 1.5% year in the time period between 1966 and the 1973 oil boycott. Unemployment rates, which had been as low as 3.6% in 1968, began to rise as well, reaching 5.6% by 1972. Pressures on the economy had been building since the long boom and took the form of classic Keynesian "demand-pull inflation." With the economy at nearly full employment, rising military expenditures, coupled with increased consumption and investment, began to increase the claims on national output beyond the capacity to produce it. Federal budget deficits increased from $2.8 billion in 1970 to $23.4 billion in 1973. The Federal Reserve System accommodated the booming economy by keeping interest rates low and credit readily available. The government also reduced business taxes to keep the economy expanding and spur further investment. There was simply "too much money chasing too few goods," and inflation began to rise from 1.3% per year in 1964 to 3% in 1966. By 1965 President Lyndon Johnson's advisors were recommending either a tax increase or a decrease in spending. Neither strategy fit with Johnson's political or economic objectives. The Federal Reserve did briefly tighten credit, but the strategy was quickly abandoned after the "credit crunch" devastated industries that were dependent upon credit such as automobiles, the construction trades in general and housebuilding in particular.

Upon his election Richard Nixon began to engineer a mild recession in order to decrease inflation, according to the Phillips curve, and unemployment began to rise. However, the recession was short-lived. Having other problems to deal with (e.g., the troubles in international finance and an impending oil crisis), Nixon once again pursued an expansionary fiscal policy. Government deficits rose from $11.3 billion in 1971 to $23.6 billion in the quarter preceding the 1972 election. Unemployment declined and Nixon was reelected, proclaiming, to the chagrin of his conservative supporters, that he was now a Keynesian. However, the brief and mild recession did not wring the inflationary pressures from the economy. Prices continued to rise, but a new phenomenon was about to occur: rising prices in the context of high levels of unemployment. Upon succeeding Richard Nixon as president in 1974, Gerald Ford and his advisors pursued a contractionary policy under the guise of "Whip Inflation Now." Spending was reduced and taxes were increased to produce budget surplus which exerted a downward force on aggregate demand. In addition, the oil price increases (commonly referred to as the OPEC tax) removed another $2.6 billion of purchasing power from the economy. Despite the reduction in spending, prices continued to rise, with inflation averaging 11% by 1974. The Federal Reserve tightened credit as well. The inflation rate abated slightly, to 9.2% in 1975, and then further to 7.8% by 1978. But unemployment increased to 7.7% in 1976 in response to the contractionary policies [30].

Traditional demand management practices were no longer working. If the government expanded the economy inflation worsened without achieving full employment. If the government conducted contractionary policies, unemployment soared without eliminating inflation. Political economists concluded that the economy was suffering from an entirely different form of inflation known as cost-push, where rising prices led to rising business costs, which were passed on to consumers in the form of higher prices. Oligopoly power remained strong and business was able to pass on rising energy costs as higher prices. The last vestiges of the capital-labor accord took the form of cost-of-living adjustments (COLA) provisions in union contracts. When business passed on costs as higher prices, workers received an automatic increase in wages. In addition, oligopolies had long ago stopped relying upon the market to determine prices. Rather, they set a target profit rate and marked up costs in an attempt to achieve their targets. When the nation's monetary authorities raised interest rates, the businesses were able to simply raise prices. Consequently, restrictive monetary policy and high interest rates exacerbated the inflationary spiral rather than reducing it [31]. The decade of the 1970s remained a stagnant one. The ineffectiveness of the policies of the new economists did not change with the election of a Democratic president, Jimmy Carter, in 1976. Unemployment fell to 6.1% by 1978, but this level at a cyclical peak was higher than the rates found in the troughs of recessions in the 1950s and 1960s. Carter attempted to deal with the problem of structurally embedded inflation by deregulating the airline industry, hoping to unleash the forces of competition. Yet inflation varied between 5.75 and 7.6 percent until 1978, which were, themselves, historically high levels in the postwar era. But things were to change rapidly, once again driven by oil prices, in 1979.

10.3.8 The Fateful Year of 1979

Since 1953, when the Central Intelligence Agency helped engineer the overthrow of Prime Minister Mohammed Mossadeq, the Shah of Iran engaged in a rapid modernization program. This modernization led to many of the economic problems associated with rapid growth: traffic-clogged streets, rising prices, urban pollution, and income inequality. By 1979 the Shah's empire crumbled. Initially a moderate social democratic form of government emerged, but it was quickly replaced by the charismatic cleric (or Ayatollah) Ruhollah Khomeini who subsequently proclaimed the Islamic Republic of Iran. In the waning days of the Shah's regime, Iranian oil workers struck, disabling production. Exports fell from 4.5 million barrels per day to less than 1 million. By Christmas 1978 oil exports stopped entirely. Oil prices increased by 150%, stimulating a panic which led to further speculative increases. Saudi Arabia and other OPEC nations increased their own production, but the shortage was real [32]. When the Saudis worried that the increased production would damage their wells, and reduced production, prices spiked again. Iranian students seized the American Embassy. The responsibility for a failed rescue attempt fell upon President Carter, who had tried to govern in the center while imposing an austerity plan. He spoke to the American people that "life was not fair," placed solar panels on the White House roof, turned down the thermostat, and urged his fellow citizens to do the same. Many were in no mood to listen. Earlier in the year, Carter had to deal with the partial core meltdown of a nuclear power plant in the Susquehanna River on the outskirts of Harrisburg, Pennsylvania, at Three Mile Island. America's energy future was highly uncertain, and the economy was on the verge of plunging into another oil-price-driven recession. Carter was to be a one-term president, learning the hard lesson that, in times of austerity, the center moves to the right. The 1980 election pitted the incumbent Carter against former actor and governor of California, Ronald Reagan. Reagan won in a landslide, promising the return of "Morning in America." His economic plan was one designed to restore the lost American hegemony, control labor and energy costs, and boost corporate profits.

10.3.9 The Emergence of Supply-Side Economics

The Reagan-era economic program was designed to raise corporate profits, reduce inflation, and restore American power in the world. Their record was mixed. Profits never increased and the price was an explosion of public and private debt and an increase in inequality. The focus was to be on the supply side of the economic balance. Stimulating aggregate demand alone had led to inflation without reducing unemployment. The idea was that if business costs were reduced and access to capital increased, the increase in aggregate supply would expand output while reducing prices. To accomplish this goal the Reagan Administration launched an interrelated five-point program. The conservative social program, in conjunction with the latest oil price run-up, touched off the worst recession since the Great Depression in 1981–1982. The elements of the supply-side program included:

- The use of a restrictive monetary policy to generate high interest rates and engineer another recession, largely in order to raise unemployment to discipline labor.
- Further intimidate or eliminate labor unions in order to reduce wage-based inflation and enhance the ability of business to appropriate the gains of productivity.
- Deregulate business, especially finance, in order to restore competition. This also entailed the elimination of environmental laws and worker safety laws to further reduce costs to business.
- Increasing the degree of inequality in order to redistribute income and wealth toward the wealthy and corporations. This was accomplished by means of changing the tax code.
- Remilitarization and the return to an aggressive, unilateral, anti-communist military policy.

Jimmy Carter had appointed a conservative central banker, Paul Volker, to the Federal Reserve Chair in an attempt to restrain inflation and prop up the value of the dollar he moved to increase interest rates. In 1978 the rate that banks charge one another for overnight loans (called the Federal Funds Rate) stood at 7.9%. Other presidents, for example, had toyed with contractionary monetary policy (also known as tight money) but had abandoned the experiment when unemployment rates increased. But during the Reagan era, tight money was not abandoned. By 1981 the Federal Funds Rate rose to 16.4%, and rates for home mortgages rose to nearly 20%. Unemployment increased from 5.8% in 1979 to 9.5% in 1982. Failures per 10,000 businesses rose from 27.8 to 89.0 in the same time period. Economists Samuel Bowles, David Gordon, and Thomas Weisskopf termed this policy "the Monetarist Cold Bath."

The Reagan Administration also continued the Carter era experiments with deregulation, launching a public campaign to convince the nation's citizens that regulations were outmoded and cumbersome. As we saw in the previous chapter, the older regulatory agencies, such as the Interstate Commerce Commission, were created at the behest of business to control cutthroat competition. During the Great Depression, the nation's banks were regulated in an attempt to stem the financial crisis. The Reagan Administration turned to the dismantling of the newer regulatory agencies such as the Environmental Protection Agency (EPA) and the Occupational Safety and Health Administration (OSHA) which they believed to be a primary cause of the increases in business costs. Their staffs were cut and Reagan appointed James Watt, who believed that environmentalism was "dangerous radicalism," to head the Department of Interior. Spending on regulation declined by 7% from 1981 to 1983, and staffing was reduced by 14%. The high interest rates that resulted from the monetarist cold bath led to the problem of "financial disintermediation". During the Great Depression, thrift institutions such as savings and loans were allowed to pay higher interest rates on deposits than were commercial banks (Regulation Q). In return they were to loan money only for purchases of homes and apartment buildings. But the increase in interest rates made Regulation Q irrelevant, as deposits left the savings banks to find more lucrative returns in other financial markets. The Garn-St. Germain Depository Institutions Act of 1982 allowed savings banks to pay market interest rates and to invest their funds in more speculative housing projects. The system came to a crashing halt during the presidency of Reagan's successor, George H.W. Bush, necessitating the need for a multi-billion dollar bailout. In addition, the banking sector accounted for more mergers than any other industry by 1986.

As a candidate Ronald Reagan stood on the steps of the State Capitol in Concord, New Hampshire, and proclaimed that for America to get richer, the rich need to get richer. This was to be accomplished by reducing the progressivity of the tax codes, whereby the wealthy pay a proportionately larger share of their income in taxes. The effective corporate tax rate dropped from 54% in 1980 to 33% in 1986. The Economic Recovery Act of 1981, better known as the Kemp-Roth tax cut, reduced the top marginal tax rates of top income earners from 70% to 50% and cut overall taxes by 23% over the course of 3 years. It also reduced estate taxes, allowed for accelerated depreciation, and reduced corporate taxes by some $150 billion. Government revenues fell by $200 billion. As a result the income distribution of the United States changed, becoming more skewed toward the top. The Gini coefficient, which measures overall income inequality, rose from 0.406 to 0.426 over the course of the Reagan Administration. The higher the coefficient, the greater is the degree of inequality. The share of income accruing to the top 1 percent of the population rose from 8.03 in 1981 to 13.17 in 1988. The share that went to the top one hundredth percent rose from 0.65% to 1.99%. This was supposed to free up funds for investment in the newly deregulated economy. Unfortunately, the surge in investment was not forthcoming.

Finally, the last component of the supply-side agenda was an increase in military spending. This was hardly supply-side economics but rather old-fashioned demand expansion by means of increased government spending. Massachusetts Institute of Technology economist and *Newsweek* columnist Lester Thurow went as far as to call Reagan "the ultimate Keynesian." Military spending as a percent of gross national product peaked at 9.2 percent during the height of the Vietnam War but had declined since then. But between 1979 and 1987, inflation-adjusted military spending increased by 57%. The Reagan Administration had clear cold-war objectives. They believed that the Soviet Union would bankrupt itself trying to keep up with American spending. They were correct. Increased military spending plus declining oil revenues were the primary economic cause of the collapse of the Soviet Union at the end of Reagan's presidency. But the increases in military spending were also designed to increase US power in an increasingly militant world. The United States conducted military operations, for instance, in Grenada which had elected a mildly socialist president (Maurice Bishop). It was hoped that the increased military power would restore the days of *Pax Americana* and bring the benefits of a strong dollar and low raw materials prices back to the country [33]. Thus while the economy expanded during the 1980s, it was not possible to attribute it to either reducing the tax burden on the rich or on Keynesianism. The road to prosperity was instead paved with low oil prices.

Given these objectives, the macroeconomic performance of the Reagan years produced mixed results. Inflation rates fell, dropping into the 3–4% per year range by the mid-1980s from a high of 13.6% in 1980. Much is made of the effectiveness of the assault on labor unions in lowering the rate of wage growth and the decline in interest rates since the zenith of the cold bath policy. What is rarely mentioned, but rather important, is the role falling oil prices played in both controlling cost-push inflation and bringing about the demise of the Soviet Union. In the mid-1970s, additional sources of oil were discovered in Mexico and in the North Sea between the United Kingdom and Norway, all beyond the control of OPEC. The first oil from the North Sea flowed into England in 1975. In the period from 1972 to 1974, oil was discovered in the Bay of Campeche in Eastern Coastal Mexico. The wells were prolific enough such that Mexico met its own needs and began to export to the world market. The Trans-Alaskan pipeline, on hold since the late 1960s, was completed in 1977. With the completion of the pipeline Alaskan oil, production soared from a mere 200,000 barrels per day in 1976 to slightly more than 2 million barrels per day in 1988. Since the 1988 peak Alaskan oil production has subsequently fallen to only 700,000 barrels per day as of 2008. Further downward pressures on price came from the development of alternative energy sources: nuclear power in Europe, natural gas and coal, and the conservation that resulted from increased energy prices. By the mid-1980s, a spare capacity of 10 million barrels per day emerged. These forces caused OPEC to reduce its prices. By 1985 the price of oil had fallen to $10 per barrel, reducing the pressure of cost-push inflation [34]. The Soviet Union, deprived of oil revenue, which accounted for a third of its income, could no longer maintain its military spending, especially after its defeat in Afghanistan. The end of the Soviet system was soon to follow.

As a result of decreased income support and an anti-union climate, the growth rate of worker compensation did fall, averaging on 0.6% per year from 1979 to 1990. Unfortunately, productivity growth (or growth of output per worker hour) also grew nearly as slowly, achieving annual growth levels of only 1% during the same time period. So while corporate profits rebounded from their 1981 trough, they were essentially no higher at the end of the Reagan Administration than they were in the beginning of the stagnant 1970s. The real growth in profits would have to wait until the era of Bill Clinton [35]. Perhaps the most negative consequence of Reagan-era economic policy was the explosion of debt.

The Federal budget deficit increased dramatically over the course of the 1980s, driven by the reduction in tax revenues, the high interest rates associated with the monetarist cold bath, and the expansion of Federal spending. Between 1981, the Kemp-Roth tax cut became a law, and 1988 (the last year of the Reagan Administration) when tax revenues as a percentage of gross national product fell from 15.7% to 14%, in the same time period, military spending increased from 5.3 to 6.1 percent of GNP, while interest obligations rose from 2.3 to 3.2 percent. Federal spending on education and infrastructure declined. Given the increase in military and interest spending, the size of the Federal government did not decline, as per the neoliberal goal. Rather, it increased from 20% of GNP in 1979 to 22% in 1981, where it stayed until 1987. The deficit itself, which had ballooned to $221 billion in 1990 from a base of $79 billion in 1981, now represented 2.5% of gross national product by itself [36]. Furthermore the push toward financial deregulation allowed banks and other financial institutions to increase their own indebtedness, although the structural changes of the Reagan Administration would give way to a much greater financial explosion by the early twenty-first century. The relaxation of the antitrust laws, falling inflation, and declining interest rates, once the cold bath shock treatment was completed, provided the incentives for another merger movement in the 1980s. From 1970 to 1977, merger activity averaged $16 billion per year. The value of merger activity increased to $70 billion per year in 1981–1983 and $177 billion from 1985 to 1987. Eleven of the top twenty-five mergers involved oil companies as either buyer or seller. In fact, the top five mergers of the decade were oil company mergers, the largest being the 1984 acquisitions of Gulf Oil by Standard of California for $13.4 billion, and Texaco's purchase of Getty Oil for $10.1 billion in the same year. Other mergers were concentrated in the food products industry, retail trade, and insurance. Cross-border mergers increased in volume and size, as exemplified by the acquisition of Texasgulf, Inc. by the French oil giant Elf [37]. The economy in general, and the oil industry in particular, emerged from the 1980s as a more concentrated economy, better able to withstand the competitive pressures of falling prices without sacrificing unduly their current and future profitability.

Reagan's successor, George H.W. Bush, attempted to carry on the same policies, especially in the area of keeping taxes low. However, deficits kept mounting, and the new president was constrained further by the passage of the Gramm-Rudman-Hollings Balanced Budget and Emergency Deficit Control Act of 1985. The Act imposed binding constraints upon Federal spending and limited the creation of further deficits. Bush campaigned on the promise of no new taxes, but the military spending needed to pursue a war in oil-rich Iraq threatened to expand the deficit beyond the Gramm-Rudman-Hollings limits. Reluctantly Bush agreed to raise taxes, and consequently the conservative wing of the Republican Party abandoned him. This set the stage not only for the election of Democrat Bill Clinton but also for the resurgence of the conservative influence upon the Republican Party. Clinton was destined to carry out the legacy of the Reagan Revolution. Running as a liberal, Clinton campaigned on the basis of renewing economic growth by means of supply-side measures to increase labor productivity. Primary among them were public investments in education and infrastructure. However, there was a competing agenda among the Clinton advisors to reduce the size of the budget deficit in order to protect the integrity of the nation's financial markets, increasingly susceptible to international demands and pressures. The deficit hawks argued that large deficits limit long-term growth, appropriate scarce international capital, and result in rising interest rates and a greater portion of the Federal budget being devoted to interest payments. The deficit hawks won the day. No large-scale fiscal stimulus by means of public investment would be forthcoming. Although the title of Clinton's campaign pamphlet was entitled *Putting People First*, his policies put the needs of the bond markets first. The growth path was to be fine-tuned by monetary policy alone, and the Federal Reserve pursued an essentially "accommodative" expansionary "easy money" policy.

In Clinton's second term, the deficits turned to budget surpluses, rising from $69.3 billion in 1998 to $236.2 billion in 2000. In 1999 Clinton also signed the Financial Services Modernization Act, which repealed the Glass-Steagall Act of 1933. Commercial banking was no longer separated from investment banking. The act provided the impetus for yet another merger movement, this time involving the consolidation of financial services. Citibank merged with Travelers Insurance to form Citigroup. Wells Fargo merged with Norwest to provide myriad financial services, and American Express expanded their product line into nearly every aspect of money management. The bill also insured that hedge funds would remain unregulated forever! As a result of the deregulation of banks and financial services, debt began to expand. Wage growth remained low, averaging only 0.5% per year throughout the 1990s. Moreover, the economy was expanding on the technological changes brought by computerization and the early days of the internet, commonly referred to as the ► dot.com bubble. Most technology stocks were traded on the National Association of Securities Dealers Automated Quotation Index (or NASDAQ). In 1994 the NASDAQ index stood below 1000. By 2000 it had climbed to over 5000.

However, the expansion of debt begun in the Reagan years continued to climb. When wages and incomes of the vast majority of the population are growing slowly, the only way to increase spending is to increase access to credit. From 1990 to 2000, gross domestic product increased from $5.8 trillion to $9.8 trillion. However, outstanding debt increased from 13.5 trillion to $26.3 trillion. Household debt nearly doubled during the period, from $3.6 trillion to $7 trillion, but financial firm debt more than tripled from $2.6 trillion to 8.1 trillion. The economy seemed to be running on financial speculation fueled by easy access to credit, as well as by relatively cheap oil. Oil prices generally remained stable throughout Clinton's years as well as relatively cheap allowing for revenues to be directed toward deficit reduction rather than increasing oil costs. Oil prices were less than $20 per barrel when Clinton took office and remained at the $30 per barrel level when he left. Oil production also remained high, ranging between 25 and 30 million barrels per day. Clinton's years saw neither spikes in gasoline prices nor energy crises.

Clinton also pledged to end "welfare as we know it" and did so by signing the Personal Responsibility and Work Opportunity Reconciliation Act of 1996. The act essentially ended welfare (or Aid to Families with Dependent Children) as an entitlement program. AFDC was replaced by Temporary Assistance for Needy Families (TANF), and recipients needed to work in order to earn their checks. The new law was supposed to restore America's work ethic and was also helpful in deficit reduction. Average monthly welfare payments (AFDC or TANF) adjusted for inflation in 2006 dollars fell from $238 per month in 1977 to $154 in 2000. Not surprisingly with increased financial mergers, a technology bubble in the stock market, rising access to debt, reduced welfare benefits and slowly growing wages the degree of inequality increased as well. The Gini index increased from .454 to .466 to .479 in 2015. This meant that *every* year of the Clinton Administration exhibited greater income inequality than *any* year of the Reagan Administration. The share of aggregate income accruing to the top 5% increased from 21% to 22.4% over the same time period, while the share going to the top 1/100 of a percent rose as well, from 1.74% when Clinton began his term to 2.4% when he left office. Increasing income inequality has become a trend. Inequality was higher in every year of the George W. Bush Administration than in any year of the Clinton Administration. In addition, there was more inequality in most years of the Obama Administration than in the Bush years.

A recession began shortly after Clinton left office, in 2001, driven by the buildup of excess capacity in the computer industry and the subsequent fall in NASDAQ values known as the ► dot.com bust. During the first term of George W. Bush, who narrowly won a contested election, the unemployment rate increased from 4% in 2000 to 6% in 2003. Following attacks on the World Trade Centers and the Pentagon in September of 2001, the Bush Administration pursued wars in Afghanistan and Iraq. Oil prices rose from approximately $30 per barrel in 2003 to nearly $150 per barrel in 2008, driven largely by the dislocation of war on oil-producing countries.

10.3.10 Warning Signs in the Early Twenty-First Century

The economy of the twenty-first century grew, albeit at a slower rate than in the non-depression years of the twentieth century. Without cheap oil as a basis of economic growth, other factors must be called upon to explain economic performance. We believe the primary drivers of economic growth were the creation of demand by means of advertising and ever increasing levels of debt, enabled by central bank policy of low interest rates and financial deregulation, cheap natural gas and coal, and high levels of military spending. The limits of this strategy became clear in 2008 when the financial system virtually collapsed, to be saved only by a trillion-dollar infusion into the reeling financial sector, known as the Troubled Assets Relief Program (TARP), based on the Reconstruction Finance Corporation of the depression era. The financial panic translated into the real economy and unemployment rates rose into the 10% range, while

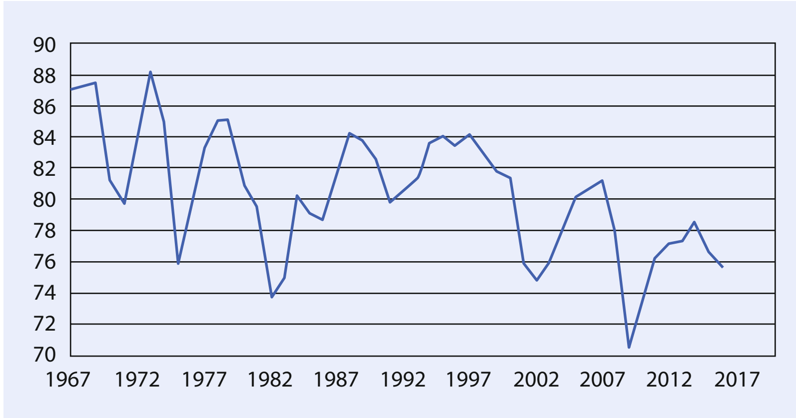


Fig. 10.1 Percent utilization of industrial capacity

capacity utilization fell from 81.3% in 2007 to 70.0% in 2009, before "recovering" to 74.2% in 2010, and rising to 76% by the end of 2016 [39] (Fig. 10.1).

The housing sector was particularly hard-hit—just as it was in the Great Depression. Housing values collapsed by as much as 40% in particularly speculative markets such as Las Vegas, Miami, and the major cities of Southern California. Unemployment in the building trades rose to 20%. The third phase of the crisis began in 2010 and can be found in a fiscal crisis among states. Most states have a balanced budget provision in their constitutions, and the fall in revenue from lost housing values and taxes upon financial assets has created the need to cut costs. Layoffs of public employees are soaring, and some states in the Great Lakes region such as Wisconsin, Ohio, and Michigan have pursued policies of removing the rights to collective bargaining for public employees. More such attempts to reduce cost to state and local governments and instill the "flexibility" of having employees pay for the effects of the economic downturn are likely to occur in the future.

10

10.3.11 The Housing Bubble, Speculative Finance, and the Explosion of Debt

The economic downturn of 2008, the most severe economic recession since the Great Depression, began much as did the Great Depression of the 1930s: with a major hurricane and a collapse of speculative housing. While the events of the late 1920s were centered in Florida, the antecedents of the 2008 crisis were truly global. Throughout the latter years of the twentieth century, a global pool of money, or a glut of savings, were building from sources as diverse as sovereign wealth funds based on petroleum profits, to Chinese trade surpluses, to individual accumulations in high-saving nations. By the middle of the first decade of the twentieth century, this fund had grown to the order of $70 trillion. Traditionally these funds had been invested in safe assets such as US Treasury securities. However, by 2004 the Federal Reserve Board of the United States had driven interest rates down to the 1% range by purchasing Treasury securities from banks, thereby releasing more money into the system following the collapse of the high-tech bubble. Investors were forced to look elsewhere for better rates of return. One location they found was the housing market in the United States, as well as other housing markets. Prices were rising and the structures created in the Great Depression such as insured long-term, amortized mortgages and the creation of a secondary market where mortgages could be bundled and sold as short-term securities made the market appear safe from risk. Rates on mortgages of 5–7% were far more appealing than were 1% returns on Treasury bonds. The demand from global investors was sufficient that standards for qualification based on income, assets, and employment stability were systematically lowered. Yet when all the qualified potential buyers who met the rigorous traditional standards were exhausted, standards were simply lowered to find more customers to meet the rising demand of the global pool of surplus savings. By 2008 mortgage brokers were no longer asking for documentation of income, employment, or other assets. The famous NINJA loan, or liar's loan, was born: no income, no job or assets [38]. By 2006 fully 44% of mortgage loans required no documentation. In addition, the average loan-to-value ratio increased to 89% by 2006, as the number of no-down-payment (100% financing) mortgages climbed from 2% in 2001 to 32% in 2006 [40].

The process was abetted by the general climate of financial deregulation that had characterized the US economy since the 1980s. The secondary market, created during the depression at the insistence of banks, allowed for the pooling of mortgages into mortgage-backed securities. As long as the potential for default was low, because the standards for qualification were high, these securities were fairly risk free, as they had been historically. However, the emerging, and unregulated, sectors of the financial security industry created even more exotic instruments by which to finance housing. Groups of mortgage-backed securities were themselves bundled into collateralized debt obligations (CDOs), and they were further divided into slices (or, to use the French word, *tranches*). Rating agencies, acting on historical data, declared these CDOs to be investment grade (AAA). On the basis of investment grade rating, mortgage security investors were able to purchase insurance policies against possible default known as credit default swaps. In the deregulated climate of 2007–2008, one did not even need to own an asset in order to purchase an insurance policy. The existence of global surplus savings and a lightly regulated climate served as an incentive for mortgage brokers, who would sell the loan immediately, to offer more mortgages to more people who simply did not have the income to pay the loans. But the risk would be managed further up the chain, by regional banks and in the money center banks in the world's financial districts. The nation's central bankers (e.g., Alan Greenspan, Ben Bernanke, Timothy Geithner) assured the public that the new financial innovations would reduce systemic risk. However, a problem was brewing beneath the surface, the problem of unsustainable levels of debt.

The purchase of everything from innovative financial instruments to bundles of loans was highly leveraged, that is, purchased with borrowed money, often at a ratio of 20:1. The system remained solvent as long as housing prices kept rising. Consumers could treat their houses as automatic teller machines. From 2004 to 2005, Americans withdrew $800 billion in equity each year. This allowed for the purchases of more home improvement products, automobiles, and exotic vacations, as well as mundane purchases of daily life. More than 7000 Walmarts and 30,000 McDonalds were constructed to meet the growing demand. New television shows such as "Flip This House" advised potential real estate speculators as to which improvements would result in easy financial profit. Wharton School senior strategic planner James Quinn estimates that without these withdrawals, economic growth would have been no more than 1% annually between 2001 and 2007. Homebuilders followed suit, constructing 8.5 million homes in 2005, about 3.5 million more than could be justified by historical trends [41]. However, by 2006 home prices began to fall. This touched off the downward cascade typical of a positive feedback loop. As homeowners found themselves "underwater," or owing more on their mortgage than the house was worth, mortgage defaults began to increase. Cable News Network estimated that by the last quarter of 2010, 27% of all homeowners were in this situation. As defaults escalated, increasing 23% from 2008 to 2009, the bundled securities that were constructed from these pools of seemingly safe, investment grade, securities began to lose value. Since so many of them were highly leveraged, the falling prices of homes and bundles of mortgages created a panic. Since the financial instruments were so complex, even banks could not figure out what their portfolios were worth. Consequently, the mortgage crisis could not be isolated in the riskier "subprime" market but spread to the entire economy. Major investment banks were crippled as well. Two Bear Stearns hedge funds collapsed, precipitating the general financial panic, Lehman Brothers went bankrupt, and Merrill-Lynch was absorbed by Bank of America, under considerable pressure from the Treasury.

The debt problem, however large, was not limited to housing. As the data in Table 10.1 indicate, debt was expanding in all sectors of the economy. By 2008 household debt, including mortgage debt and consumer credit, amounted to $13.8 trillion, equivalent to the nation's Gross Domestic Product, and far greater than the energy backing it. By 2005 consumer debt exceeded income after taxes, standing at 127% of disposable income. Debt service ratios, or the percentage of disposable income used to pay principal and interest on contracted loans, rose from about 11% in 1980 to nearly 14% in 2005 as the crisis loomed. The burden was felt highly

Table 10.1 Domestic debt and GDP (trillions of dollars)

		Debt by sector				
Year	Gross domestic product	Total debt	Household	Financial firm	Non-fin'l business	Gov't (local, state, and federal)
1976	1.8	2.5	0.8	0.3	0.9	0.3
1980	2.8	3.5	1.4	0.6	1.5	0.4
1985	4.2	7.3	2.3	1.3	2.6	0.7
1990	5.8	11.2	3.6	2.7	3.8	1.0
1995	7.4	14.3	4.9	4.4	4.3	1.0
2000	9.95	19.1	7.2	8.7	6.6	1.2
2005	12.6	28.2	11.9	13.7	8.2	2.6
2010	14.7	37.1	13.5	15.3	10	3.0
2015	17.9	45.2	14.2	15.2	12.8	3.0

► https://www.federalreserve.gov/releases/z1/current/coded/coded-2.pdf

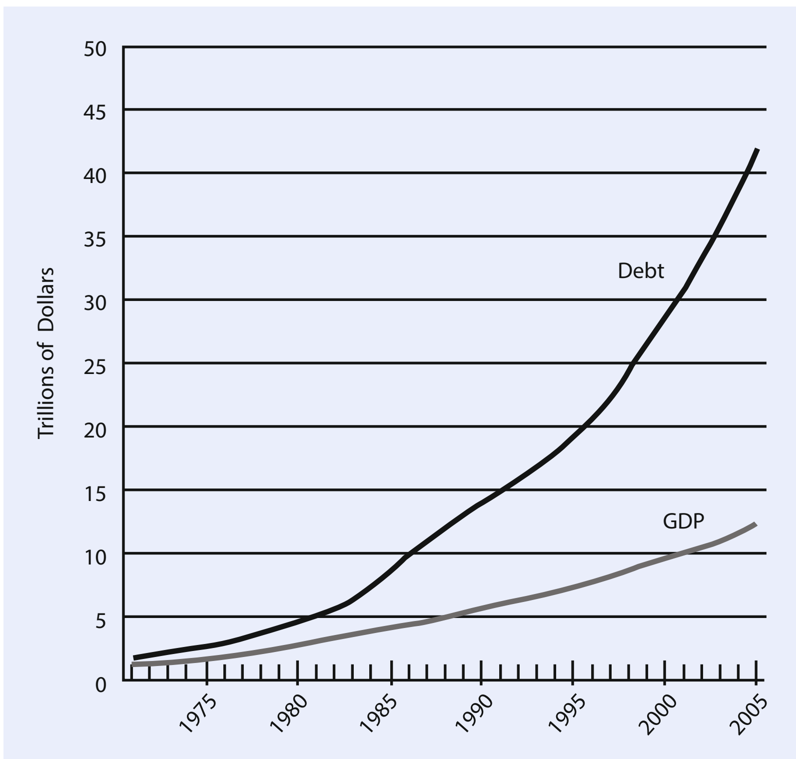


Fig. 10.2 GDP and total debt

unevenly. Those in the top fifth of the income distribution paid only 9.3% of their incomes in debt service by 2004, while those in the middle two fifths paid between 18.5% and 19.4% [42]. The debt of nonfinancial corporations increased 20-fold between 1970 and 2007, while the debt of financial firms (banks, insurance companies, mortgage brokers, etc.) expanded by a factor of 160! (Fig. 10.2).

Banks had long been seen as recipients of deposits and lenders of money, as safe and conservative in their outlook. But in the new world of deregulated finance, the financial service industry became the largest borrower in the economy. It was this leverage that transformed the financial structure and made it vulnerable to disruptions. John Maynard Keynes made the assertion that "speculation does no harm as a small portion of enterprise. However, once the amount of speculation overtakes that of enterprise, the danger of this position becomes serious" [43].

10.3.12 The Deficit and the National Debt

The Federal government was not immune from the increase in debt. Budget deficits climbed from $3 billion annually in 1970 to $1.414 trillion in 2009. The primary drivers of these increased deficits were a reduction in taxes, especially at the top of the income distribution, and the expansion of government spending, primarily for the military and for entitlement programs such as Social Security, Medicare, and Medicaid. Military spending in 1970, the year of peak domestic oil production and the beginning of the era of stagflation, the government brought in $192.8 billion in receipts and spent $195.6 billion, for a deficit of $2.8 billion. In the last year of the Reagan Administration, whose economic policy was built upon increased military spending and tax cuts, the annual deficit soared to a historically unprecedented $155.2 billion. The last 2 years of the Clinton Administration actually saw modest budget surpluses, as the growth rate of military spending declined and tax receipts increased with the high-tech boom. Deficits began to climb again with the second Bush Administration, rising to $458.6 billion in 2008. By 2009 the annual difference between receipts and outlays was $1.4 trillion. Income tax revenue dropped from $1.635 trillion in 2007 to $898 billion in 2010 as the bush administration reduced taxes, in an unsuccessful attempt to stimulate the economy. Military spending, which stood at $294 billion per year when the Bush Administration took office, rose to $616.8 billion in 2008. It continued to climb during the Obama years, reaching the level of $693.6 billion in 2010. As of 2017, Congress is prepared to fund military spending to a greater degree than even the Pentagon has asked for. The Office of Management and Budget estimates that 2011 military spending will exceed $768 billion. In 1970, at the height of the Vietnam War, military expenditures were 8.1% of gross domestic product, while total government spending was 19.3%. By the end of the Clinton Administration, military expenditures had fallen to 3%, while total spending remained about the same, at 18.5%. By 2010 military spending stood at 4.8% of GDP, and total spending rose to nearly 24%. Mandatory expenditures, such as those on health care (Medicare for the aged and Medicaid for the poor), along with Social Security and other income support programs (unemployment insurance, supplemental security income for the disabled, Food Stamps, etc.) increased from $60.9 billion, or 6% of GDP, in 1970 to more than $2 trillion, or 14.7%, in 2009 [44]. Despite the increase in mandatory expenditures for entitlement programs, the income distribution grew more skewed, largely as a result of a stock market boom and subsequent bailouts, along with tax cuts at the top of the income distribution, along with stagnant wages at the bottom. In 1980, at the beginning of the neoliberal economic strategy, the Gini coefficient was 0.403 and the top 20% of the income distribution claimed 16.5% of aggregate income. The top 1% received 8% of income and the top 0.01% 0.065%. By the end of the second Bush Administration, the Gini coefficient increased to 0.466, indicating a greater degree of overall inequality, the top 20% claimed 21.7% of aggregate income, while the share of the top 0.01%, which amounts to about 14,000 families out of a population of 300 million, rose to 3.34% [45].

The 2010 congressional elections saw a large enough segment of the population expressing concern that the Democratic majority was unseated by conservative activists who see as their top agenda item the reduction of the budget deficits and the return of the glory days of the neoliberal agenda in the 1980s. In the first edition of this book, we asked whether we are reaching peak debt as well as peak oil? At the time, the credit system had largely frozen and few loans were granted. Not surprisingly, the total debt outstanding fell. We wondered whether this condition would be permanent. A quick glance at Fig. 10.2 shows that it was not. Debt began to grow again after 2010 and reached historic highs by September 2017, once again showing the role of debt as a driver of economic growth. The political will to expand more debt within the United States is clearly shrinking, and the willingness of other economies and investors to purchase Treasury securities is also in decline. But what are the potential effects of declining government participation in the economy? If one believes that the market economy is resilient and self-regulating, then a decrease in government spending will simply free money for spending in the private sector, and the economy will prosper. If, on the other hand, one believes the explosion of financial speculation and debt was due to investors seeking financial profits in an otherwise stagnant real economy, as indicated by declining rates of industrial capacity utilization, then the reduction of government spending, coupled with the rise of inequality, might cripple the economy by reducing its overall level of demand. The second scenario is far more likely. The growth of inequality may well be an important factor in the slowing of growth over the past few decades. Theoretically, if the rate of return on capital exceeds the rate of economic growth income concentrates at higher income levels [46]. The age of peak oil may well be the age of degrowth as well. It is likely to turn into an age of austerity.

10.4 Conclusion

The world economy collapsed into depression in the 1930s. Governments faced few ways out: fascism, communism, or social democracy. John Maynard Keynes wrote his classic text, *The General Theory of Employment, Interest, and Money*, as a guidebook for "saving capitalism from itself" in order to avoid the other outcomes which he detested. In the United States, the program took the form of "the New Deal." While Franklin Roosevelt was able to restore confidence among a shattered population and put millions back to work, the New Deal did not engineer an economic recovery. It took the Second World War to do that! The United States exited the war in a clear position of economic and military power. Pent-up consumer demand, low gasoline prices, and a very productive factory system insured the economic surplus could be absorbed. Its corporations expanded into former colonies, and the terms of trade were positive and the prospects bright enough that corporations could share the gains from rising productivity with workers, insuring adequate income to buy their products while increasing profits at the same time. Oil was cheap and plentiful, and the American consumer could utilize the rising quantities of cheap and available oil to live the American Dream of a house in the suburbs, good schools, a steady job, and a cornucopia of consumer goods.

All that was to change in the 1970s. Domestic oil production in the United States peaked, and it no longer possessed the spare capacity to cushion events in the world oil market. The 1970s saw two disruptions, in 1973 and 1979, and citizens of the developed world saw rising prices and constricted supplies. At the same time the era of stagflation commenced, and mainstream Keynesian policies no longer worked. Efforts to expand the economy resulted in rising inflation, while efforts to control inflation made unemployment rise to politically unacceptable levels. After a period of impasse, a "neoliberal" agenda was consolidated during the Reagan years, consisting of a belief in small government, deregulation, low taxes, and a strong military, although Reagan, despite his rhetoric, generated far more deficit spending than did Franklin Roosevelt. The economy did in fact recover by the end of the 1980s, but the price was ever increasing levels of inequality and a rising debt burden. The neoliberal approach, hiding a lot of Keynesian government spending, continued through the Democratic administration of Bill Clinton, where it was consolidated further. The bill came due at the end of the second Bush Administration as the soaring debt burden and lax regulatory climate led to a near collapse of the world financial structure. It was during this period that the "undulating plateau" began to assert itself. The recession reduced overall demand and brought down oil prices. The recovery pulled oil prices up, and the monopolized structure of the economy allowed business to pass the higher cost onto customers. This helped reduce the overall level of economic activity and was a primary cause of the both slow growth and the next recession (Fig. 10.3).

Most explanations of the postwar social order focus on the internal dynamics of the world economic system: its overall demand, technology, and the distribution of income. How do these factors affect the aspirations of the world's population for a decent income and a meaningful life? But we contend that the world economic system is limited not only by its internal dynamics but also by the external biophysical conditions posed by the availability of energy and the consequences of using it.

When the first edition of this book went to press, the Middle East was afire with democracy

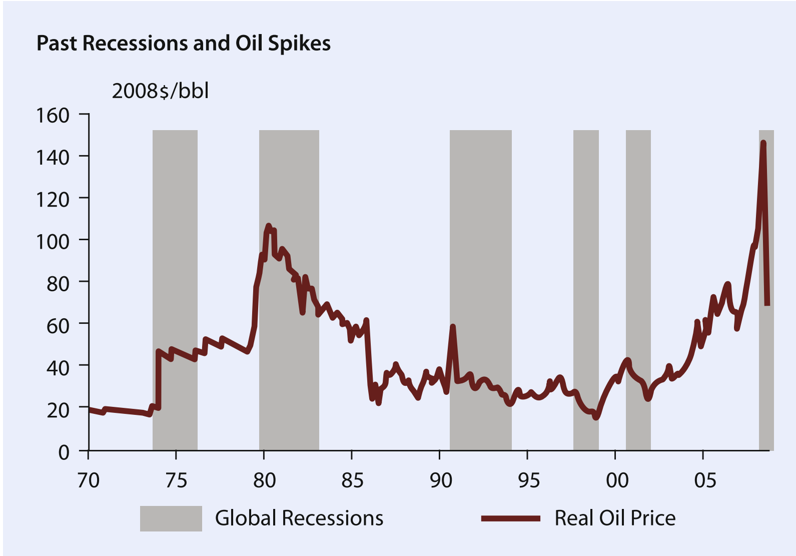


Fig. 10.3 Past recessions and oil price spikes (From Hamilton 2009)

movements. Unfortunately, the hope of the Arab Spring turned into a nightmare of dictatorship and perpetual war. Oil prices, which had climbed to nearly 150 per barrel in 2009, fell to around $50 today. We contend that the fall in oil prices was a primary reason for the economic recovery in the United States, and the impoverishment of the oil-producing regions in the Middle East and in North Africa. On March 11, 2011, an earthquake of magnitude 8.9 on the Richter Scale, the most powerful one in recorded history, struck the northern coast of Japan. The nation was devastated by the quake and subsequent Tsunami. The lack of electricity shut down the cooling systems of the Fukushima nuclear reactor complex. The latent heat from the fuel rods boiled away the water, resulting in a partial core meltdown. The heat also liberated the hydrogen from the oxygen leading to the buildup of flammable hydrogen gas. On March 14, 2011, the second of the reactors exploded. The viability of the third is in question. If the Japanese abandon their commitment to nuclear power and switch to oil or natural gas, what will be the effects on the world markets? Can an economy that has been stagnant for two decades recover? If the Japanese heed the advice of their American advisors and increase consumption in order to grow their way out of economic disaster, what will happen to the world's fossil fuel resources and the quality of its atmosphere?

The Obama years were blessed with low oil prices and a commitment to stimulative policy. The Federal Reserve Bank kept short-term interest rate close to zero for the entire period, and government spending remained high. According to the Bureau of Economic Analysis, government spending stood at nearly $5.6 trillion at the beginning of 2012, increasing to $6.256 trillion by the end of 2016. Given the extension of the Bush-era tax cuts, receipts fell short of expenditures, thereby increasing the Federal budget deficit. Contrary to popular opinion, the deficit did not increase consistently throughout the Obama years. It was lower in 2016 (−$873 billion) than it was in 2012 (−$1.3 trillion). The commitment to a neoliberal policy of war and free trade did not end with the inauguration of a Democratic administration. The US continued to have a military presence in the oil-producing countries of the Middle East and Central Asia. A health-care reform bill (the Affordable Care Act) was passed without a single Republican vote in the first year of the administration, based on the plan that Obama's rival, Mitt, implemented while serving as the Republican governor of Massachusetts. Unemployment fell from nearly 10% of the labor force at the beginning of President Obama's term to less than five at the end, while inflation remained negligible despite the monetary and fiscal stimulation. Unfortunately, the good news was not spread evenly across the population. The job loss across the nation's heartland remained above average, as the jobs that were lost in the 1970s never returned.

As it turns out, the resentment was long-lived and multi-generational. In October of 2016, Republican candidate Donald Trump won a surprise victory on a platform of "Making America

Great Again," largely by restricting immigration; subsidizing the energy industry; rescinding regulations, especially environmental regulations; and encouraging the expansion of fossil fuel use. The early indications are that the integrity of the environment will be a very low priority for the Trump administration. The president has appointed Scott Pruitt to head the Environmental Protection Agency who made his reputation as Attorney General of Oklahoma by suing the EPA for "overreach" as regards regulating greenhouse gases. The former head of Exxon-Mobil, Rex Tillerson, is now the Secretary of State. Furthermore, the US delegation to the most recent Conference of Parties, designed to implement the Paris Climate Accords, is headed by coal company executives. Whether resistance to this program leads to mobilization and a greater attention to Earth's biophysical systems remains to be seen.

At some point the production of oil on a world basis will peak, and will begin to decline. Problems of instability and rising prices will cease to be just cyclical and political but will become secular and geological. What does that portend for the economic system? Will peak oil exacerbate the inherently stagnationist tendencies of the monopolized economy as Baran and Sweezy argue? How can we generate employment and reduce poverty, advocate democracy, and rebuild after natural disasters when the energy base to do so is in decline? If every scientific measurement, from ecological footprinting to biodiversity loss, to peak oil, and to carbon dioxide concentrations in the atmosphere, shows that humans have overshot the planet's carrying capacity, then how can we grow our way into sustainability? We can't, but how do we deal with the consequences of a nongrowing economy which have historically manifest themselves as periodic depressions? We will return to these questions in the final section of our book.

Questions

1. What was the "Treaty of Detroit?" How did it impact postwar labor relations in the United States?
2. What were the four "pillars of postwar prosperity?" Explain how each helped set the stage for the long economic expansion of the 1950s and 1960s.
3. What was the New Deal? What problems did it try to address, and what was its major legislative accomplishments? How successful was the New Deal in restoring American prosperity?
4. What was the role of the Second World War in transforming the US economy?
5. How did the world oil industry, and the US role in it, change in the years after the Second World War?
6. Why could the period from the end of the Second World War be characterized as the era of economic growth?
7. Why was the "New Economics" of the 1960s successful in stimulating economic growth?
8. What is stagflation? What was the role of peak oil in bringing about stagflation in the United States?
9. What other factors led to the erosion of the pillars of postwar prosperity?
10. Why was the "New Economics" unsuccessful in eliminating stagflation?
11. What are the major tenants of the conservative growth agenda, also known as neoliberalism?
12. To what degree was the neoliberal program of the Reagan era successful? What were the economics and social costs of this success?
13. How did the Clinton Administration carry on the neoliberal agenda? How did low oil prices during the 1990s affect US economic performance?
14. How much did debt expand in the first decade of the twenty-first century? What were the economic outcomes?
15. How might biophysical limits affect economic performance as we enter the second half of the age of oil?

In the 1950s and 1960s economic growth was driven by cheap oil, as the oil ceased to become less cheap something else had to drive economic growth—cheap money and the expansion of debt.

References

1. Chandler, Michael, Judith Ehrlich, Rick Goldsmith, and Lawrence Lerew. 2009. *The most dangerous man in America: Daniel Ellsberg and the Pentagon Papers*. San Francisco: Kovno Communications.
2. Editors of Business Week. 1979. The decline of U.S. Power. *Business Week*. March 19. pp. 37–96.
3. Bowles, Samuel, David Gordon, and Thomas Weisskopf. 1990. *After the Wasteland*. Armonk: M.E. Sharpe.

4. Kennedy, David. 1999. *Freedom from fear*, 77. London: Oxford University Press.
5. *Business Week* quoted in Kennedy 1999. P. 84.
6. Franklin D. Roosevelt quoted in Kennedy 1999, p. 123Collins, Robert M. 2000. *More: The politics of economic growth in postwar America*, 5. New York: Oxford University Press.
7. Dighe, Ranjit. 2011. Saving private capitalism: The U.S. bank holiday of 1933. *Essays in Economic and Business History* 29: 41. Kennedy. 1999. p. 136.
8. Klitgaard, Kent A. 1987. "The Organization of Work in Residential Construction. (Unpublished Ph.D. Dissertation) University of New Hampshire.
9. Kennedy, David. 1999. *Freedom from fear*, 131–159. London: Oxford University Press. Ch. 5.
10. Yergin, Daniel. 1991. *The prize: The epic quest for oil, money, and power*, 244–259. New York: Simon and Schuster. Ch. 13.
11. Deffeyes, Kenneth S. 2001. *Hubbert's peak: The impending world oil shortage*, 4–5. Princeton: Princeton University Press.
12. Kennedy, David. 1999. *Freedom from fear*, 252–253. London: Oxford University Press.
13. ———. 1999. *Freedom from fear*, 249–287. London: Oxford University Press. Ch. 9.
14. ———. 1999. *Freedom from fear*, 641. London: Oxford University Press.
15. Yergin, Daniel. 1991. *The prize: The epic quest for oil, money, and power*, 368–388. New York: Simon and Schuster. Ch. 19.
16. ———. 1991. *The prize: The epic quest for oil, money, and power*, 381. New York: Simon and Schuster.
17. Kennedy, David. 1999. *Freedom from Fear*, 624–626. London: Oxford University Press.
18. DuBoff, Richard. 1989. *Accumulation and power*, 150–155. Armonk: M.E. Sharpe.
19. Bowles, Samuel, David Gordon, and Thomas Weisskopf. 1990. *After the Wasteland*, 109. Armonk: M.E. Sharpe.
20. US Bureau of the Census. 1975. *Historical statistics of the United States, colonial times to the present. Bicentennial Edition*. Series G 416–469, N 156–169, P 68–73, Q 148–162.
21. *InflationData.com* July 21, 2010.
22. Yergin, Daniel. 1991. *The prize: The epic quest for oil, money, and power*, 409–430. New York: Simon and Schuster. Ch. 21.
23. ———. 1991. *The prize: The epic quest for oil, money, and power*, 431–449. New York: Simon and Schuster. Ch. 22.
24. Cleveland, C., R. Costanza, C.A.S. Hall, and R. Kaufmann. 1984. Energy and the U.S. economy: A biophysical approach. *Science* 211: 576–579.
25. Bowles, Samuel, David Gordon, and Thomas Weisskopf. 1990. *After the Wasteland*, 27. Armonk: M.E. Sharpe.
26. Collins, Robert M. 2000. *More: The politics of economic growth in postwar America*, 21. New York: Oxford University Press.
27. ———. 2000. *More: The politics of economic growth in postwar America*, 40–67. New York: Oxford University Press. Ch. 2.
28. Yergin, Daniel. 1991. *The prize: The epic quest for oil, money, and power*, 499–540. New York: Simon and Schuster. Ch. 25–26.
29. ———. 1991. *The prize: The epic quest for oil, money, and power*, 588–652. New York: Simon and Schuster. Ch. 29–31.
30. Bowles. 1991. Bureau of labor statistics, inflationdata.com
31. Wachtel, H., and P. Adelsheim. 1977. How recession feeds inflation: Price mark-ups in a concentrated economy. *Challenge* 20 (4): 6–13.
32. Yergin, Daniel. 1991. *The prize: The epic quest for oil, money, and power*, 674–698. New York: Simon and Schuster. Ch. 33.
33. Bowles, Samuel, David Gordon, and Thomas Weisskopf. 1990. *After the Wasteland*, 121–135. Armonk: M.E. Sharpe. Ch. 8.
34. Yergin, Daniel. 1991. *The prize: The epic quest for oil, money, and power*, 745–768. New York: Simon and Schuster. Ch. 36.
35. Kotz, David, and Terry McDonough. 2010. Global neoliberalism and the contemporary social structure of accumulation. In *Contemporary capitalism and its crises*, ed. Terry McDonough, Michael Reich, and David Kotz, 93–120. Mishel, Lawrence, Bernstein, Jared, and Schmitt, John. 1999. *The state of working America 1998-1999*.
36. Bowles, Samuel, David Gordon, and Thomas Weisskopf. 1990. *After the Wasteland*. Armonk: M.E. Sharpe. Ch 10: 146–169, Ch 13: 199–216.
37. DuBoff, Richard. 1989. *Accumulation and power*, 134–235. Armonk: M.E. Sharpe.
38. Davidson, Adam and Blumberg, Alex. 2009. "The giant pool of money." *This American life*, Program # 355. National Public Radio.
39. *Economic Report of the President*. 2011. "Capacity utilization rates, 1962–2010." Table B-54.
40. Quinn, James. 2008. *The great consumer crash of 2009*. New York: Seeking Alpha.
41. ———. 2008. *The great consumer crash of 2009*, 6. New York: Seeking Alpha.
42. Foster, John Bellamy, and Fred Magdoff. 2010. *The great financial crisis*, 30–31. New York: Monthly Review Press.
43. Keynes, John Maynard. 1964. *The general theory of employment, interest, and money*, 159. New York: Harcourt Brace.
44. *Economic report of the president*. Table B-79. Congressional budget office. "Historical budget data. Tables F 10 and F 11.
45. U.S. Bureau of the Census. 2014. *Historical income data, and top incomes database*. Paris: Paris School of Economics.
46. Piketty, T. 2014. *Capital in the twenty-first century*. Cambridge, MA: The Belknap Press of Harvard University.

Globalization, Development, and Energy

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_11

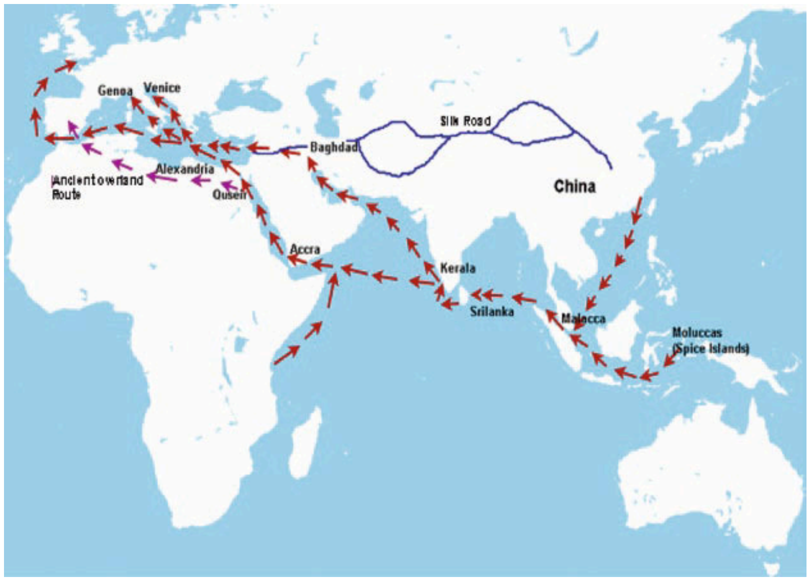


Fig. 11.1 The spice route (Source: Dyfed Lloyd Evans)

Young adults today have grown up in a world where globalization is a ubiquitous reality and where most of our politicians accept its supposed virtues or did so until recently. There are fierce discussions, or more often positions, about whether or how globalization is losing (or gaining) us jobs, or whether we have globalized too much or not enough, but for most people it is just a fact represented by the labels on their clothes or electronic devices indicating production or manufacture from all around the world. This was not the case when the authors of this book were young—at that time nearly everything we ate, wore, or drove was "made in America." Anything from overseas—except specialized luxury goods—was normally viewed with great suspicion. Thus globalization, at least at the scale we see it today, is a relatively recent phenomenon. That makes it useful and important to understand why globalization has become so important, what are the perceived and actual gains and costs and how these are related to energy use [1].

Before we consider this from a modern perspective, however, we think it's important to emphasize that trade has been important for at least as long as written human history and much earlier as indicated by many foreign artifacts found in archeological digs going back tens of thousands of years. People have long wanted luxury goods from abroad and have always sought interesting and different tools, amusements, foods, and experiences not found locally. One of the clearest examples of long-range trade is the "spice route" connecting Europe and the Middle East to all parts of Asia (Fig. 11.1). Spices were very important in ancient times for their own sake and also to hide the sometimes tainted smell and taste of rotting food in the days before refrigeration. Spices were good items of trade because they were exotic, relatively light, and non-bulky and could be carried for thousands of miles by camel and donkey and still make a profit. We were amazed when, at an archeological dig near Stockholm, Sweden, we watched the excitement of the excavators of an ancient Viking site when they found a coin that was from Constantinople, a very long way away. Obviously, the Vikings, often more traders than plunderers, had traveled thousands of miles on European Rivers. Many archeological digs of Native Americans find, for example, arrowheads made from stone quarried hundreds or thousands of miles away. With the advent of European colonization and imperialism in Africa, Asia, and the Americas, trade took on a whole new dimension. As we saw in chapter 2 the mercantilists (fifteenth through eighteenth century), believed wealth was measured in gold or silver and promoted trade and imperialism to obtain these metals. Nevertheless, the day-to-day lives of most people, including Europeans, remained based on materials that rarely traveled more than a few tens or rarely hundreds of kilometers from their growth or extraction.

While Adam Smith highlighted the benefits of free trade and "the system of perfect liberty," his successor, David Ricardo, developed the first formally enunciated theory of trade. This theory, known to the world as comparative advantage, argued that everyone benefited from internationalization and trade. A careful reading of history reveals that the term *comparative advantage* did not originate with Ricardo. He talks of comparative costs and the comparative value of money, as well as speaking of more advantageous employments in his famous chapter "On Foreign Trade" in his 1817 *Principles of Political Economy and Taxation*, but never pens the phrase, comparative advantage [2]. Ricardo's argument was forged in his debate with Thomas Malthus over the repeal of the Corn Laws (see pages 36–37). The Corn Laws prohibited the importation of cheaper grains from Continental Europe. As England's population increased, additional, and lower quality, land had to be put into production in order to meet subsistence needs. Landlords' benefited from this policy, as they were able to charge additional rents when poorer lands went into production. Moreover, food became more expensive as more labor was required to grow food on poorer quality, and nutrient poor, land. According to Ricardo, capitalists were doubly squeezed as rising rents and rising wages both diminished profits.

Ricardo was a shrewd politician, and a Member of Parliament, as well as a prominent political economist. He argued that everyone would be better off if international trade were freed from restrictions such as the Corn Laws. He created a highly abstract, and historically unrealistic, example of the production and trade between England and Portugal for wine and cloth. In his example, Portugal possessed an absolute cost advantage. They could produce both wine and cloth with fewer labor hours. England, however, had relatively cheaper costs of producing cloth, or a lower ratio of labor hours embodied in the production of cloth to wine. Ricardo argued that international specialization in production would result in more commodities being produced for fewer labor hours. Everyone would be better off by trade liberalization. Ricardo also insisted that only finished commodities would be traded internationally. Capital and labor were immobile. If they were not, then capital would flow to where labor was cheaper. Trade between England and Portugal would be no different than trade between London and Yorkshire.

The example paid little attention to history. Portugal had enlisted England's help in a war with Spain. The price of the aid was to open the economy to English cloth imports. Since the application of waterpower to large-scale textile production made English cloth much cheaper, the nascent Portuguese textile industry withered and Portuguese capital flowed towards the vineyards. English imports of cloth far outweighed Portuguese exports of wine, and the trade imbalance was paid for by the gold produced by means of slave labor in Brazil.

The term "comparative advantage" comes from the sanitization of Ricardo's doctrine in the 1930s by Eli Heckscher and Bertil Ohlin. Working from a framework of general equilibrium theory (or neo-Walrasian economics), Hecksher and Ohlin replaced ratios of labor hours with ratios of opportunity costs, which are subjective valuations of the cost of the best-foregone alternative. Normally, opportunity cost increases, and increasing opportunity cost is synonymous with diminishing marginal returns. In Heckscher and Ohlin's model, opportunity cost remains constant. So Ricardo's greatest theoretical contributions, the labor theory of value and diminishing marginal returns, are missing from the modern theory of comparative advantage. Now, comparative advantage depends upon "resource endowments." Rich countries should continue to specialize in finance and research, while poor countries should specialize in mineral extraction, labor-intensive agriculture, and the manufacture of mass production goods such as clothing and electronics. Moreover, in the model, all industries are perfectly competitive, and no nation has any technological advantage. From this set of assumptions, it is an easy mathematical exercise to derive mutual gains from trade, despite an empirical record that the poor parts of the world are becoming far poorer as trade relations become less restricted, and that the terms of trade favor the already rich nations, who capture the highest amount of value-added through the supply chain [3].

11.1 Trade and Imperialism

The advantages of trade were often conflated with those of raw exploitation of others and with imperialism. During the sixteenth, seventeenth, eighteenth, and nineteenth century, most European powers laid claim to territory in Africa and the Americas. We have already discussed the raw exploitation of natives in these areas by the Spanish as they sought gold and silver, the English for tea from India and Ceylon, and especially sugar from Barbados and so on. Much of the labor energy for the production of these products came from actual or virtual slaves. Few consumers of cotton clothes in 1860 or of rubber tires in 1900 (or even of some clothes, diamonds, or cell phone materials today) understood the human slavery that produced the products they purchased, or where the raw materials came from, or the human cost of those entities. We found a particularly chilling account in Hochschild's book *King Leopold's Ghost* about how some estimated ten million Africans were savagely used and killed as Belgians and other Europeans "developed" the interior of Africa for ivory (much used before plastics were available for everything from false teeth to piano keys) and rubber (for tires and many other things). The lies used by Leopold to justify his horrendous abuse of the people living in the Congo basin are a reminder of how so many economic practices are sugarcoated by governments and in the press.

Whatever the virtues or not of globalization, it is clear that it is a fact and that the world has become enormously internationalized in the last decades (Fig. 11.2). Essentially all recent American presidents until Trump have called for either continued or more "free trade," implying a continuation of internationalization. Where there are arguments against free trade, they tend to be that many US factory jobs are moved overseas, resulting in economic hardship in the United States. An obvious example is automobiles, as the United States in 1950 produced some 99% of the automobiles it used but now imports about half. As a consequence the city of Detroit and the State of Michigan, which once had the comparative advantage of relatively easy access (through Great Lakes shipping) to Minnesota iron ore and Pennsylvania coal, plus the early development of mass production of automobiles by Henry Ford, have suffered enormous economic impact. What is less obvious today, at least to the comfort of the developed world, is that increased internationalization of trade means that the processes of exploitation of nature and of manufacturing all require an enormous amount of labor, and the working conditions elsewhere often have fewer safeguards than does labor in the United States. This misuse of others in an attempt to get low production prices extends even to the existence of virtual slavery, as we saw recently in the use of "made in the United States" sweatshop labor in Guam during the trial and conviction of Jack Abramoff and as continues today as chronicled by groups such as Amnesty International.

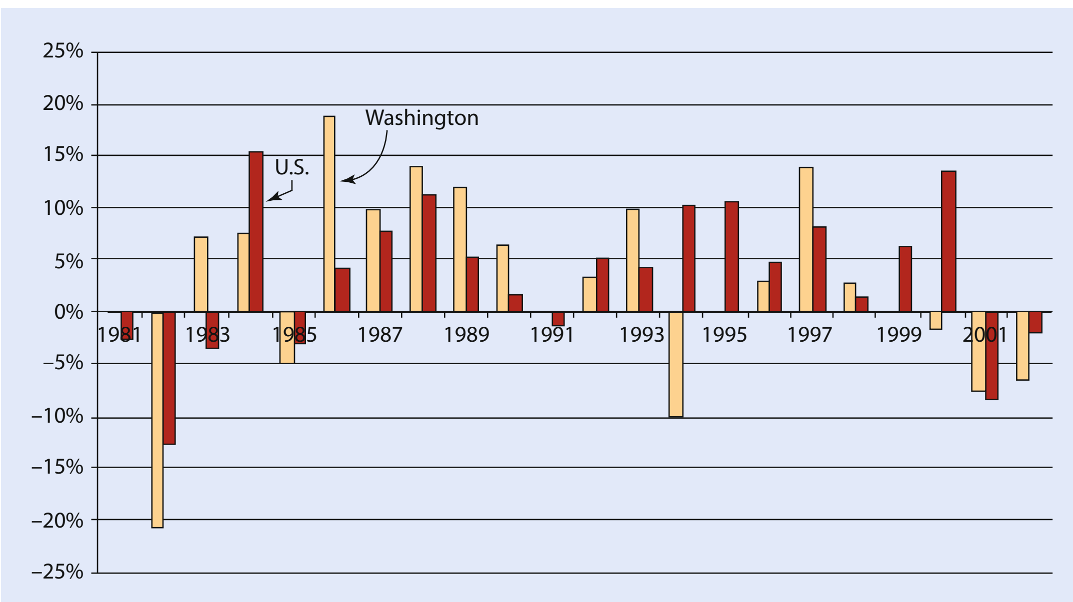


Fig. 11.2 Whatever the virtues or not of the internationalization, it is clear that it is a fact and that the world has become enormously internationalized in the last decades

11.2 The Concept of Development and Its Relation to Trade

Most of the world today is quite poor, at least relative to the affluent nations, such that one billion of the Earth's more than seven billion people live on only one dollar a day, and some three billion people live on less than $2.50 a day [4]. A very general concept is that there are very large pressures for the poorer countries to develop in order to become less poor and that this development is often done in accordance within the concept of comparative advantage, that is, a search for some kind of product that might be produced well in that location. While Ricardo originally devised his concept of comparative advantage around the idea of various growing conditions, one comparative advantage that poor developing nations almost always have is cheap labor. As the process of industrialization continues to require unskilled rather than skilled labor, there seems to be always new areas where wages are so low that people are willing to work hard in terrible conditions for very little to get some small piece of the global economic pie. Of course each area would like to get a bit larger piece of that pie. Thus around the world people do not wish to stay poor, and consequently there are extreme pressures for nations to "develop," which normally means to increase economic activity. The way that is usually used to do this is in accordance with the neoclassical model, although of course development in fact requires the land, capital, other biophysical resources, transformations, and processes to occur if it is going to work. The pressure to develop comes from many sources including governments attempting to help or placate their constituents, idealistic foreign aid or NGOs from the developed world, various business and economic interests who are interested in a cut of the hopefully increased action, and, of course, the people themselves who may be quite tired of an economically restricted life. What is rarely mentioned is that the real force behind much or perhaps most development is simply an increase in the number of people over time (a biophysical aspect), so that if some kind of development does not keep pace, people will get poorer, which nobody wants. This economic activity and its changes normally are measured as GNP or GDP or sometimes per capita GNP or GDP.

The principal tool, or more accurately suite of tools, used to guide development is, as in most things economic these days (or at least was until quite recently), neoclassical (or free market or neoliberal or "University of Chicago") economics. The ascendancy of the neoliberal model occurred over the first half of the twentieth century as economists sought to generate a "scientific," "neutral" model that would focus on improving welfare of the economy in general and leave the issue of the distribution of that wealth (properly in their view) to governments, hence absolving economists from any responsibility pertaining to that issue. The logic, summarized nicely in Palley [5] and Gowdy and Erickson [6], is that free markets will lead to "Pareto optimization" where, due to market pressures for lower prices from suppliers, the various factors of production (i.e., land, labor, capital, and so on) are being used so "efficiently" that they cannot be combined in any other way that would generate greater human satisfaction. The logic continues that if markets are completely "free" (e.g., from governmental interference) at each step of the production chain, each producer will be seeking the lowest possible prices, and each potential supplier will be seeking to cut his or her costs (ideally through "efficient" use of resources) so that the total net effect is that the final demand product will be generated as cheaply in that economy (which means increasingly the global economy) as possible. This should lead to lowest possible prices, which is the objective of many economists. It should also lead to low prices for people in poorer countries. Most economists argue that this process works very well and generates substantial net benefits (e.g., Bhagwati [7]). Likewise most economists are enthusiastic about the free market system because, at least in theory, it is *efficien*t, that is, economic resources are generating as much personal well-being as possible from their limited resources.

An important part of this is that there should be trade, and an important component of trade is that there should be more trading partners, including less developed regions where there are resources that the developed world increasingly needs and where there is "unmet demand" for the products of the industrial countries [8]. This is a, or the, mantra of most neoclassical economists and has guided how we undertake trade and our relations with the less developed world and, increasingly, government itself over the last half of the twentieth century. Thus development should lead to more wealth for both the nation becoming developed and for the developed country increasingly trading with it. In theory this should lead to efficiency, that is, that all parts of the economy are generating what consumers desire at a maximum rate given the resources at their disposal. Thus, at least in theory, development should lead to both improved conditions in the nation being developed and in the developed nation supplying the funds for that development through foreign aid. Yet the degree to which this does occur is not at all clear from objective analyses of the behavior of real economies, and the converse is often true (See [4, 9]).

11.2.1 The Leverage of Debt

In Latin America and Africa, especially, there have been pressures for development promoted by development agencies of the developed nations, internal elites, foreign NGOs, and the World Bank for many decades. These efforts have been motivated by genuine humanitarian concerns as well as (often) by the self-serving desires of the development agencies themselves. More recently there have been enormous pressures to repay debts associated with development (and other reasons) and for revisions in how economics are undertaken, according to the neoliberal model from outside entities including especially the World Bank and the International Monetary Fund (IMF). The pressures have come from the leverage these institutions have because of outstanding international debt from many countries in Latin America and elsewhere. Given the nearly impossible demands on governments due to poor and growing populations, and the difficulty in extracting taxes from rich elites who are often the same as those running governments, the easy solution has been and continues to be debt, which is a tax on future citizens. When governments can no longer afford to pay their debt service, which often exceeds 10–25% of total GNP and perhaps all tax incomes, a not surprising result is that from time to time governments have defaulted. Default has generally meant that the banks and their agents are able to impose their sometimes draconian "structural adjustment" programs which has meant, basically, reducing government expenditures, eliminating tariffs that have protected home industries (such as agriculture), and basically opening countries to globalization. The basis for this is usually neoclassical economics as codified in the "Washington Consensus." The results are mixed at best but often horrific, and are insightfully reviewed in Kroeger and Montanye [10].

Most structural adjustment programs also include policies and incentives for development, normally of industries that will generate foreign exchange (after all the bank's objective in structural adjustment is to get dollars or euros to repay the debt owed to them). For example, as part of the structural adjustment program implemented in Costa Rica for the mid-1990s, there were large incentives to encourage the development of "non-traditional" agricultural crops for everything from Macadamia nuts to cut flowers. Since these crops tend to be as dependent upon expensive imported agrochemicals as are bananas, it is not surprising that they did not have any significant effect on resolving debt. Meanwhile rising oil costs add greater balance-of-payment strains on most economies. In Costa Rica population growth has meant more food imports and the need for more agrochemicals for domestic crops, also making the resolution of debts more difficult [11]. The failure of many past development concerns, generally fueled by neoclassical economic concepts of growth, to deal with the issue of population growth binds developing countries into pursuing economic growth, whether real growth is possible or not.

11.2.2 The Logic for Liberalizing Economies

In the United States, especially, during the Reagan and Bush years, conservative leaders were extremely successful in convincing many formerly apolitical or even labor union people that their own personal conservatism in issues such as family, society, religion, gun ownership, and so on could be best met through making an alliance with economic and political groups whose agendas were quite different. These groups and their representatives in government were very much opposed to government in general and any interference with individual "freedom," especially intervention in the market. Thus they opposed, for example, government programs to generate energy alternatives (such as solar power or synthetic substitutes for oil), believing that market forces were superior for guiding investments into energy and everything else. They also tended to be opposed to restrictions on economic activity based on environmental considerations and even mounted campaigns to discredit scientific investigation into environmental issues such as global warming.

The authors wish to point out that they use the term "liberal" and "conservative," as they tend to be used regularly and loosely in the United States, to refer to the role for government—usually larger by the Democratic Party and smaller by the Republican Party (at least in theory—the data are quite a bit more mixed). The terms themselves are often very misleading—for example, many conservative people are extremely interested in conservation of nature and the concept of free trade is advocated by many liberals too—and in fact as we pointed out earlier in many countries such as Argentina, "liberal" means liberal free trade and is often associated with business interests.

These new conservative or neoliberal forces tended to oppose government policies that restricted free trade. This view contributed to the movement of many American companies or their production facilities overseas where labor was cheaper and pollution standards often less strict. As developed in ▶ Chap. 7, by 2000 the United States seemingly had recovered from the stagnant 1970s and the recessions of the early 1980s and early 1990s. Stock values began to increase steadily, and the general economic well-being of many Americans led to a general sense of satisfaction in market mechanisms. The end of communism in Eastern Europe and Russia effectively ended the cold war, and the free market approach to economics came to be the only game in town with respect to economics. The presidential administrations of Republican George H. W. Bush and Democrat Bill Clinton alike pressed a free trade agenda. These programs included for many foreign lands reduced spending on social programs and the reduction of government ownership and enhanced international trade. The terms of trade greatly improved for the United States as markets became "liberalized," and prices of basic commodities from coffee to cotton to oil declined by more than 100%. Unfortunately poverty rates often soared in Africa and Central America as a consequence. For example, the price paid to a farmer for a pound of coffee in Costa Rica (about a dollar per pound) was essentially barely changed from 1980 to 2005. These issues are discussed in depth by, e.g., Annis [12] and Bello [13], and reviewed in Hall [11]. Fundamentally the arguments go back to the "Ricardian" concept of comparative advantage, as previously discussed, and to the concept that free trade will lead to efficiency. An implicit assumption of those who promote international trade and the advantages to all that are supposed to flow from it is that the players have equal power in the face of the supposedly neutral market. Of course, this is patently absurd—a small coffee grower in Costa Rica does not have equal power in the face of some coffee buyer for a large U.S. supermarket chain.

11.2.3 We Need to Test Our Economic Theories About Globalization, Development, and Efficiency

A recurrent theme of this book is that if economics is to be accepted as a real science, we must expose the main ideas to empirical testing. For example, Gowdy has undertaken this by reviewing the work of those who have subjected the basic tenets of our dominant economic paradigms using the scientific method "one cannot help but be impressed with the rigor of modern social scientists" [14]. There is a crying need to subject more of our economic theories to broad, unbiased, and thorough assessment of whether they deliver on what they promise (Bromley [15], Gintis [16], Hall et al. [17] Sekera [18]). There may be no trusted, or at least broadly accepted, concept within economics with a greater need of such testing than that of "efficiency" because efficiency is the principal argument used to promote the neoliberal model and of its application to international development and unrestricted international trade. Economists themselves have increasingly questioned the effectiveness of their development models. A particularly fine example of this is William Easterly's book, *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics*. Easterly reviews the use of economic theory (basically neoclassical) as applied to development, especially development in the tropics. Easterly did what few economists do: he actually tested whether the models of economists that had been the backbone of billions of dollars of aid had accomplished what they were supposed to do. In particular Easterly asked whether the main development model, the Harrod-Domar model, as used by contemporary neoliberal development economists, is as sanitized as is the model of "Ricardian" comparative advantage. The model abstracts Harrod's equations for savings and the capital labor ratio from the psychological propensities that produce instability. Domar's concept of the "dual nature of investment" is ignored completely. As a result, the rate of income growth is a function of the national rates of savings and investment. Governments of starving countries should thereby increase forced savings, and further impoverish their citizens, for the overriding goal of economic growth. The Harrod-Domar investment model, had, when applied, resulted in a perceptible increase in GNP as it was supposed to. His answer was that there was a perceptible increase in GDP for only for 4 of 88 cases where it had been tried. In other words, when tested, these models were a disaster with respect to achieving their goals. LeClerc [19] arrived at a similar conclusion while testing a broader array of economic models as applied to development. Likewise Sekera [18] found for many examples in the United States that private entities did not deliver services more efficiently than the governmental institutions that they replaced in the name of improved efficiency. Anyone involved in the broad world of investment economics should read these three studies.

Sometimes it is not terribly difficult to test certain economic models yourself even though it is often said that real economies are too complex, and the difficulty of undertaking proper tests and controls is so daunting that you should not expect economic concepts to be explicitly testable. As an example Hall's former student Dawn Montanye asked whether the (neoliberal) structural adjustment model imposed upon Costa Rica by USAID (Agency for International Development) in the early 1990s had achieved its own clearly stated objectives when the subsequent behavior of the economy was examined [20]. This was a seemingly straightforward and reasonable thing to do that, although, curiously, seems not to have been undertaken by USAID. Her results were yes for two and no for four out of their six principal objectives. In addition, there were a number of quite important but unanticipated "bads" that occurred even for the cases where the objectives were met. If in fact there is such a large void between theory and application then one wonders whether or not there should be so many routine pronouncements on how to run real national economies based on conventional theory and models [19].

If efficiency is the main reason that neoclassical economics is promoted, and if, to our knowledge, this efficiency has been tested barely or not at all, how then might we go about testing efficiency? One can argue that since many Latin American countries have been under tremendous pressure from roughly 1990 to 2005 to "liberalize" their economies according to the neoliberal "Washington consensus" models, especially those countries such as Costa Rica that have been subject to structural adjustment, a program often imposed upon debt-laden countries that are desperate for loans and who must turn to "the lender of last resort" (the World Bank and especially the International Monetary Fund), then if indeed structural adjustment does lead to efficiency, this should be obvious from the data comparing pre- and post-structural adjustment. That this is not observed (except arguably in Chile), it seems to us hard to argue that structural adjustment and neoclassical economics do in fact lead to economic efficiency.

11.2.4 Definitions of Efficiency

The first thing to consider about the word "efficiency" is that it is often confused with "*efficacy*," which means "getting the job done," without regard to efficiency. The engineer's definition of efficiency measures output over input. But a second difficulty with the meaning of efficiency is that it is hard to find a consistent definition of output of what? And input of what? Economists usually think of efficiency (of, e.g., an economy) as the output of all desirable goods and services over the input of all resources available for production, usually referring to money or capital or labor. Perhaps the best way to explain efficiency, as economists use the word, is by giving the counterexample of economics that is, supposedly, *not* efficient. This is because the economist's definitions of Pareto efficiency and allocative efficiency are, in essence, immeasurable. Pareto efficacy means trading to the point where no one individual can be made better off except at the expense of another. But well-off is entirely subjective. Allocative efficiency can occur only at the output level where price = marginal cost. This can happen only in the market structure of perfect competition, which does not exist in the real world. In the socialist states of Eastern Europe and the Soviet Union from roughly 1920 to 1990, the determination of how much of a good and service was produced (i.e., the allocation of productive resources) was decided in large part by *central planning*, that is, by government economists whose jobs were to decide how many tractors, carrots, chickens, or other commodities were needed. There were some famous fiascos resulting from this (or at least good stories), so that, for example, in the 1950s in Russia and Poland, too many tractors and far too few refrigerators were ordered by the central planning committee, so that there were mountains of unused tractors, while people were very unhappy because they needed refrigerators. To most Western economists, this was a tragic example of how it was far better to leave the decisions of what to make up to markets, i.e., Adam Smith's invisible hand of supply and demand. In other words, in the centrally planned economy, the productive resources of the nation, steel mills, labor, and factories themselves, had been used *in*efficiently, that is, they had produced too much of one thing that was not needed or wanted and not enough of another that was. In addition it had required a large, perhaps expensive government bureaucracy to do the allocation decisions. It is this argument about efficiency that is used most commonly by neoclassical economists to argue for free markets and free trade. Centrally planning a large industrial economy is a daunting task, and it was even more difficult before the age of large-scale electronic computing. This is why most centrally-planned economies made no attempt to plan all aspects of an economy. Rather they focused on the most important sectors, known as "the commanding heights."

A problem with estimating efficiency by this method is that it is very difficult to decide just what the inputs are that should be considered as the inputs to a particular economic activity since the economy is very complex. Despite the constant use of the word efficiency by economists, you would be hard pressed to find where that has been measured or tested explicitly (except for some very general international comparisons using often rather arbitrarily defined quantifications of such terms as "level of financial development" and "improvements in efficiency," e.g., King and Levine [21]).

Engineers often use a very explicit measure of efficiency: simply the ratio of energy out of a process to the energy in. For example, coal is converted to electricity at about 40% efficiency in a modern power plant and gasoline to road transport at about 20% efficiency. Humans, too, generate work at roughly 20% efficiency. Some of the energy loss is inevitable such as losses to the second law of thermodynamics, some is related to needing to run the process at a more rapid rate than would generate maximum efficiency, and some is caused by poor design or poor housekeeping (i.e. not keeping the tires properly inflated).

A kind of combined ratio is often used to measure efficiency of economies within biophysical economics: the GDP output over the energy input, usually for a country. We call this the *biophysical economic efficiency*. The economic output must be corrected for inflation to compare different years. The ratio does not mean anything explicitly (as the engineering one does) but rather is one relatively unambiguous way that we can measure the efficiency of an economy—e.g., test explicitly the hypothesis that more free trade leads to greater efficiency, something that, as we said above, is not possible to do with the more nebulous "productive resources" perspective usually given by economists. It is useful mostly for comparative purposes—either for different countries or for one country over time—which we do here. The idea is that since the economies of many countries were explicitly or implicitly (via the general spread of neoclassical economic concepts) "converted" at least partially to less government restrictions and more market freedom in the 1990s and early 2000s, then our hypothesis is simply to test whether national economies in general (especially in those countries like Costa Rica and Chile which were subject to explicit structural adjustment consistent with neoclassical economics) became more efficient during the 1990s. If efficiencies are increasing then this would tend to support the hypothesis and the contrary.

11.2.5 Testing the Hypothesis that Freer Trade Leads to Economic Efficiency

The actualization of neoclassical economics in Latin America and elsewhere was carried out with great enthusiasm, some would say relentlessly, through a program called "the Washington Consensus" that was administered to, especially, countries that could not pay interest on their debts to the World Bank or to the International Monetary Fund (IMF) [22]. These programs of increasing free trade and reducing governmental spending ("stabilize, liberalize, and privatize") were thought to be good and tough medicine for the debtor nations. And they were supposed to lead to economic efficiency. Since we could not find any data by economists about whether economies had in fact become more efficient after liberalization, we undertook this ourselves by examining simple time trends in biophysical economic efficiency.

Our methods were very simple: for developing countries plot the biophysical efficiency (i.e., real GDP/energy used, agricultural output per unit of fertilizer, and so on for various countries) and see if there is any trend toward increasing efficiency. Explicitly we test the hypothesis that following the implementation of neoliberal policies (either in the country or more generally worldwide after 1990), there will be subsequently an increase in the biophysical efficiency of nations. We undertook this explicitly for 4 countries in each "developing" continent and for 133 countries more recently [23].

11.2.6 Results of Testing for Biophysical Efficiency Following Liberalization

We found in both studies that when the energy use and the GDP for all countries in the world are plotted on the same graph, the results are basically linear, indicating that energy is required, or at least associated with, increases in the production of GDP for essentially all nations (Fig. 11.3). We also found that for those countries that were increasing in per capita wealth that energy use increased at approximately the same rate as the GDP (Fig. 11.4).

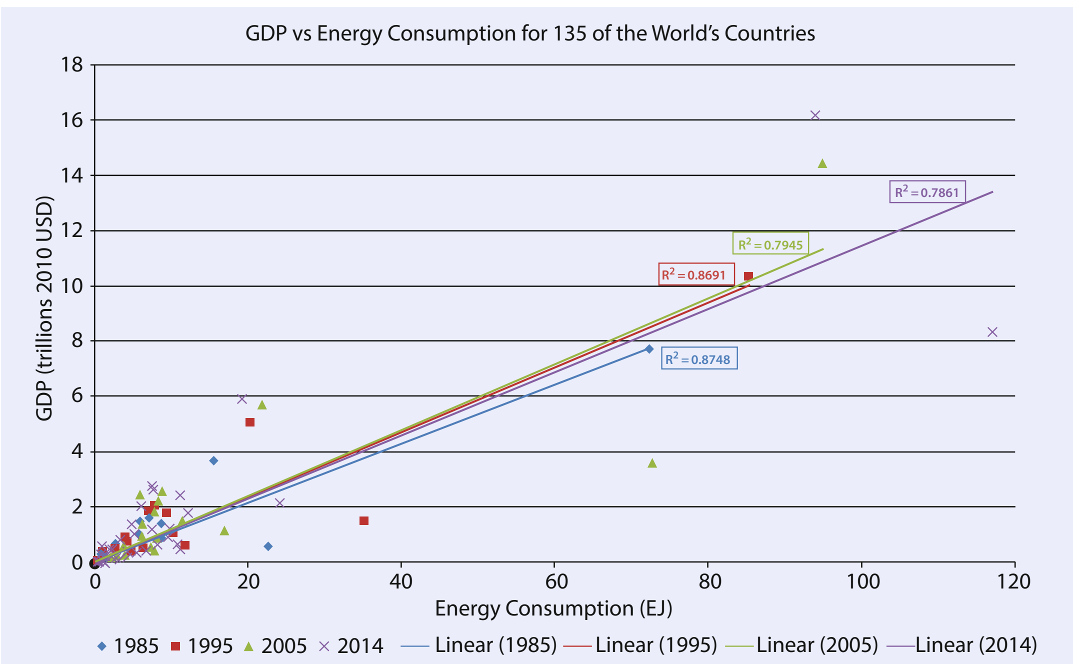


Fig. 11.3 The relation of energy use and GDP for 127 countries in 1980, 1995, 2005, and 2014. The basically linear results indicate that energy is required for, or at least associated with, increases in the production of GDP for essentially all nations, and that whatever (small) increase in efficiency that may have occurred (i.e., an increase in the slope of the line) tended to occur before the increased global liberalization of markets that began usually in the 1990s. Primary electricity is multiplied by 2.6 relative to fossil fuels to reflect their quality (Source: Ajay Gupta)

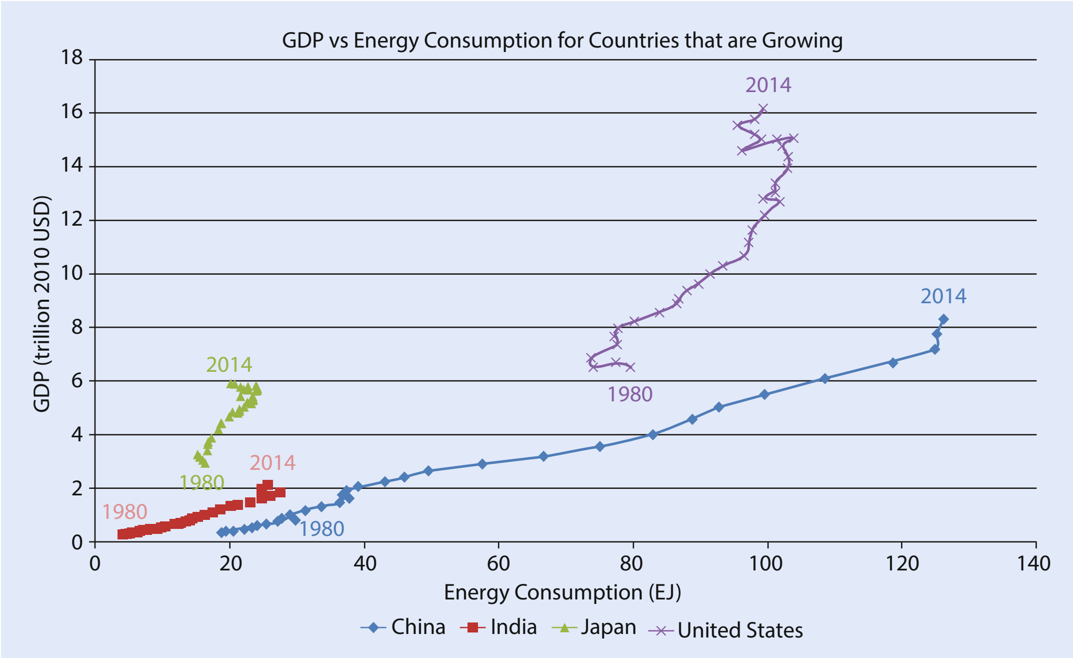


Fig. 11.4 The relation of GDP and energy use for four economically growing nations of the world. Rapidly developing nations including China and India tend to have a strong correlation between their increase in energy use and their GDP, implying that energy is needed for economic development. This appears less so recently for the developed nations such as Japan and the United States which have increased their official GDP with little or no increase in energy use (Source: Ajay Gupta)

Today there is an enormous variation in the wealth of different national economies, from the poorest, where people tend to live on but 38 cents a day (or 140 dollars a year) to the wealthiest in the developed world where annual mean incomes varied from 50,000 to 87,070 dollars annually in 2008 (World Bank 2009). Not surprisingly, from our perspective, the energy use by these different countries varies similarly from about 0.32 GJ per capita to nearly 800 GJ per capita for 2005 (Fig. 11.5). Additionally as countries have developed, they have tended to use more energy over time, generally in rough proportion to their increase in wealth (Fig. 11.4). When we examine the relations of GDP and energy use for developing countries in Africa and Latin America (the region especially impacted by "liberalization of markets"), we find no evidence at all that biophysical efficiencies have increased in the developing countries analyzed in response to the liberalization trends of the 1990s and early 2000s. When all countries are considered, biophysical efficiency has tended to, if anything, remain the same or decrease, both since 1970 and also since 1990 (Fig. 11.6). Colombia, relatively unaffected by neoliberal policies, may be an exception. Similar results were found for many other countries [11, 24, 25] . Hence the hypothesis of this chapter that the increasing use of "liberal," "free market," "neoclassical," or "Washington Consensus" approaches to economics in the last decade of the twentieth century in the developing countries will necessarily bring increased efficiency of economies is not supported, and we must seek some other explanation for economic growth besides increased efficiency (as derived from neoclassical policies or anything else). These results are consistent with the increasing view of many development economists themselves [25].

Our results do show, however, that efficiencies have increased in many developed nations (Figs. 11.4 and 11.7). Whether this is because highly developed countries are capable of becoming more efficient through pure technology, or rather have basically exported their frequently polluting and energy-intensive heavy industries to the rest of the world is another question. For some countries, many energy exporting countries, efficiency is lower (Fig. 11.8). Further analysis in which the embodied energy associated with imports and exports is added or subtracted to the energy use (denominator) of the equation suggests that when this is done, the val-

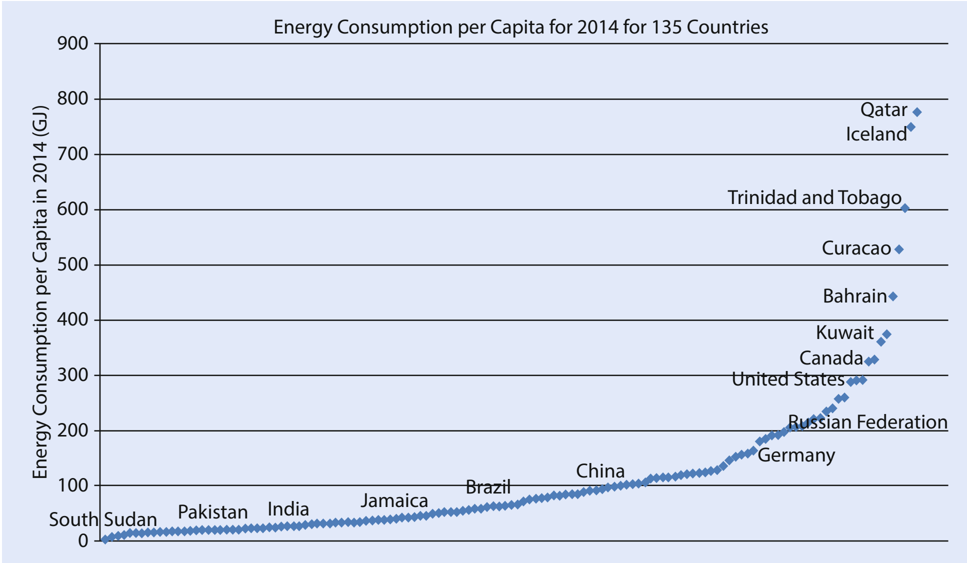


Fig. 11.5 One hundred and thirty-five nations of the world ranked in order of increasing energy use per capita (Source: Ajay Gupta)

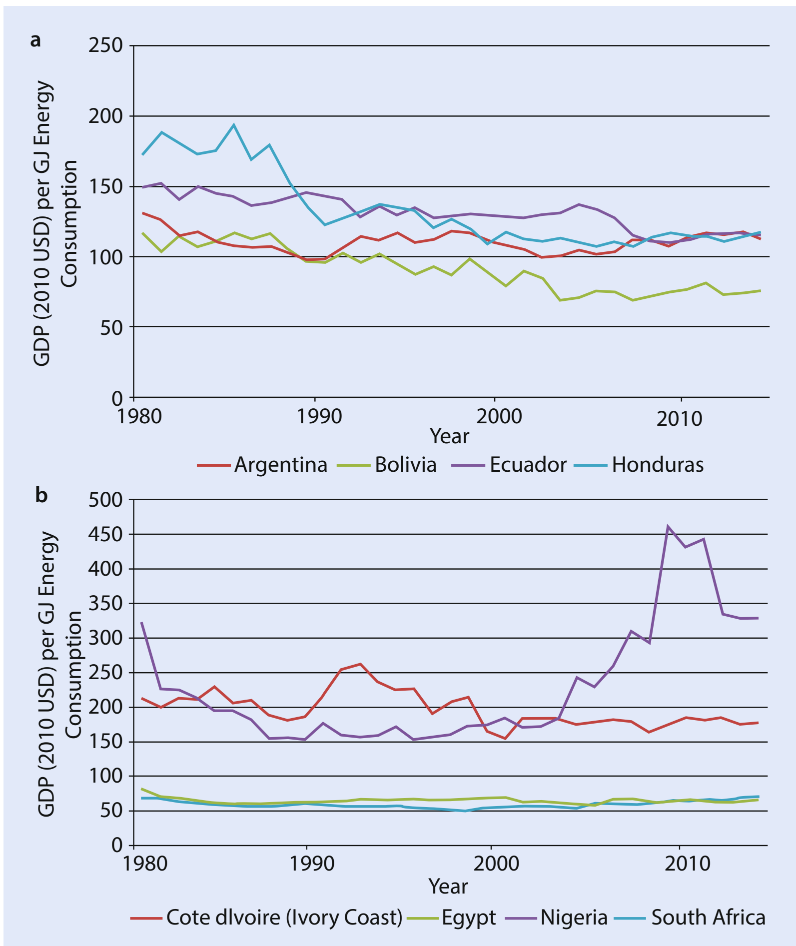


Fig. 11.6 **a** The ratio of GDP to energy use and for four countries in Latin America from 1971 through 2001. The flat or decreasing lines for all countries after 1980 are not consistent with the hypothesis that liberalizing markets increase efficiency (Source: Ajay Gupta). **b** The ratio of GDP to energy use and for four countries in Africa from 1971 through 2001. The flat or decreasing lines for all countries after 1980 are not consistent with the hypothesis that liberalizing markets increase efficiency (Source: Ajay Gupta)

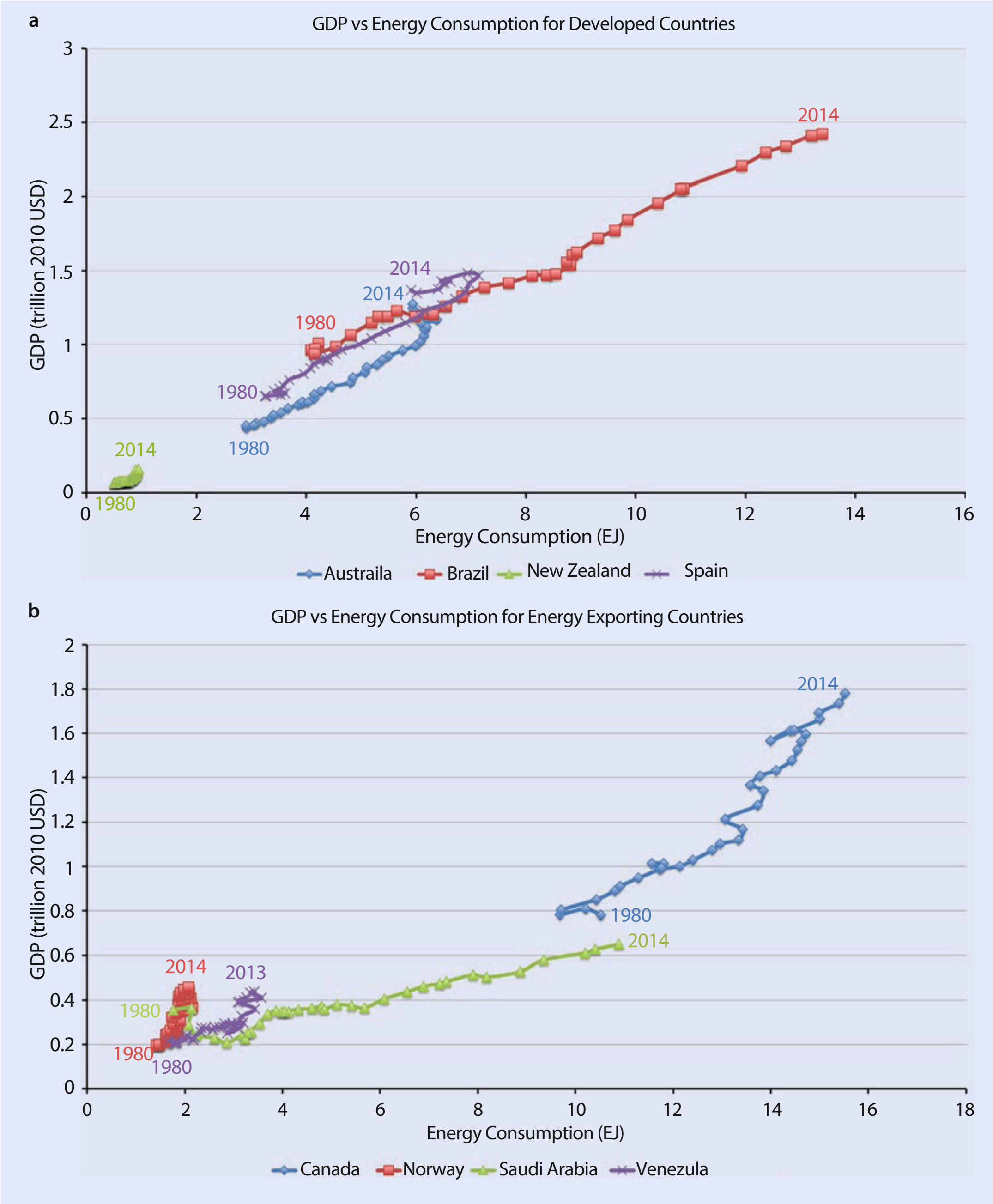


Fig. 11.7 **a** Energy use and GDP for several well-developed (the United States and Japan) and rapidly developing (China and India) nations. Note the relatively high apparent efficiency of the United States and Japan. Most relatively developed but still growing nations increase energy use in proportion to economic growth (Source: Ajay Gupta). **b** The same relation as Fig. 11.6a for large energy-producing nations, which use much more energy per unit GDP produced. This figure suggests that much of the reason for the apparent increase in efficiencies in, e.g., the United States and Japan (Fig.11.4) is due to the import of energy-intensive components (such as energy itself) to the economy

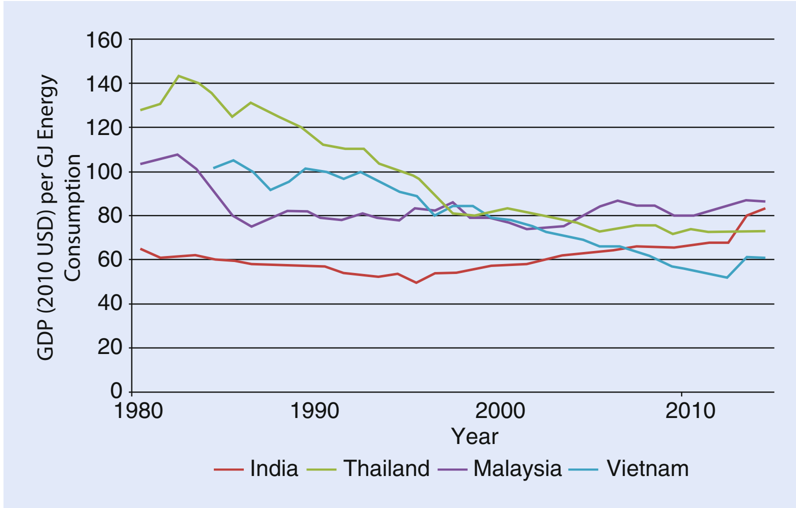


Fig. 11.8 The efficiency increases for some Asian countries, India, and decreases in others, Malaysia and Vietnam (Source: Ajay Gupta)

ues tend to become much more similar, i.e., that the main result for the disparity of the different countries is the degree to which each country is associated with undertaking the "heavy lifting" for others [26].

In our original paper, we also examined the efficiency of GDP vs. water and forest products used as well as agricultural output vs. inputs of fertilizer and found that there was always a strong relation between economic development, as indicated by GDP, and the use of resources with no indication of an increase in efficiency over time [24].

11

11.2.7 Development as an Increase in Energy Use

Our main conclusion from these and many other results is that for the vast majority of countries, there has been no increase in efficiency as measured by GDP produced per unit energy or other resources since "liberalization" of the economy. Instead it is clear that whatever economic growth has occurred as a consequence of (or at least highly correlated with) increasing the rate of the exploitation of energy and other resources. We conclude that neoclassical economics does not increase wealth by increasing efficiency in any sense that a scientist or engineer would recognize, at least by the relatively crude assessments we use, but only by increasing the rate of resource exploitation. These resources can be domestic or imported if there are other resources, including specialized human resources, that allow one to pay for them. If wealth comes from resource exploitation and not from efficiency, then the concept of development must be very tightly tied to the soils, climate, agricultural potential, mineral resources, and other biophysical resources that tend to be given short shrift in conventional economic analysis. If the human condition is to be made better, as has been the case in Costa Rica, it apparently requires as much attention to the biophysical as well as the political and monetary environment, and it must be within biophysical possibilities.

If neoliberal economics does not seem to be in agreement with empirical tests and in addition violates the basic laws of physics and is not consistent with its own assumptions (e.g., ▶ Chap. 3), then what alternative do we have to guide development or to attempt to operate our economies by, at least in the macro sense?

Our partial answer is *biophysical economics*, a rather imperfect but growing approach to economics that is based upon the recognition that wealth is fundamentally generated through exploitation of natural resources and that recognizes that economic policies are mostly about directing how energy is invested in that exploitation. The fundamental approach to biophysical economics can be found in our books (Hall [11], LeClerc and Hall [11]) and, of course, this one. This approach leads to some rather different views as to how we can improve the average economic plight of the poor of the world. In particular the biophysical model of development puts a real onus on the availability of affordable energy for successful development to occur.

There are many models of development including the Harrod-Domar (focusing on the importance of savings), the Rostow (focusing on

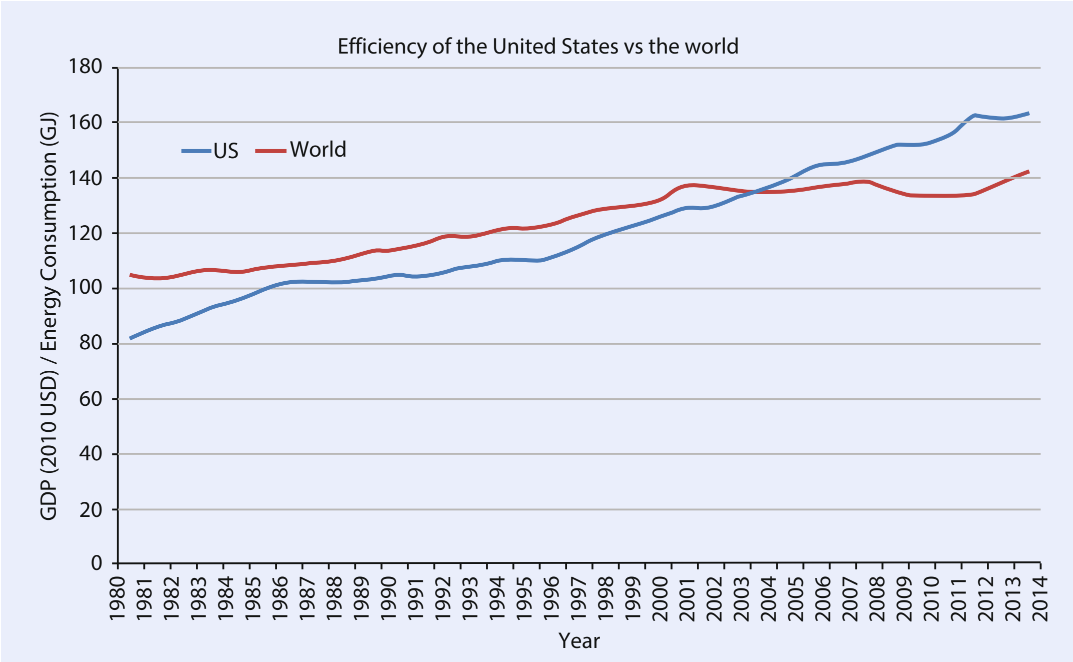


Fig. 11.9 While there has been an increase in apparent efficiency (GDP/unit energy) in a few highly developed countries such as the United States there has been little or no increase for the world as a whole, suggesting that much of the energy used to generate wealth by the highly developed countries has been outsourced

"stages of development"), and others. These are reviewed in LeClerc [19], and it is no secret that data that supports their importance in generating or even explaining development is pretty thin. We suggest another model, specifically the *biophysical model of development* which says that real material development (i.e., an increase in wealth) occurs only when the ratio: energy resources/number of people increases. This can be seen in Fig. 11.4, where per capita wealth increased only in those countries where per capita energy use increases (e.g. China and India).

While there are many theories of development (see review in LeClerc [11]), few of them are very powerful in predicting success or failure of development, and not surprisingly, few of them connect development directly to energy use. Hall [11] found many examples where development, at least as expressed as real GDP per person, is correlated very closely to the energy used per person. Where energy use has increased more rapidly than populations, people have become wealthier; where energy use has increased less rapidly than population, people get poorer.

While it is true that the United States and some other developed nations (Fig. 11.9) [11] have become more efficient in turning energy into GDP, according to Robert Kaufmann (personal communication), about half of that is due to the increased use of higher-quality inputs, and much of the rest is due to the change in the economy from industrial production (much of which has been exported) to services (or even, strangely, consumption!). The degree to which this can occur for other countries is not clear. The GDP produced per unit energy for the world as a whole has remained nearly constant or increased only slightly, suggesting that gains in the developed countries are matched by decreases in the less developed countries that often are undertaking more of the heavy industrial work for the developed nations [23]. Thus our explanation for increases in economic activity is that quite simply if more resources, and explicitly more energy, can be developed economic activity can occur. This energy is used to fuel the productive process which in the contemporary world is more dependent upon energy than either capital or labor [11]. While this is hardly news to most energy scientists, it is quite remarkably the degree to which it is a concept foreign to economists. Wealth comes from nature and the exploitation of nature and much less so from markets or their manipulation.

In all fairness it should be mentioned that it is not just the neoclassical model that seems to be having trouble generating economic growth. According to a web-based review (cepa.newschool 2004), there have been various models for encouraging development over time, and each has basically been abandoned when it had failed to generate much in the way of the desired development. This is in agreement with LeClerc's perspective. The CEPA review, and our own, concludes that any rationale for the dominance of the neoliberal model today and the evidence for its effectiveness is "ambivalent," at best, yet this perspective continues to be shoved down the throats of many developing nations.

If it is not clear that neoliberal policies have resolved the persistent economic problems of the developing world, why then are they pursued so continuously? The cynical view is that they serve rather nicely the interests of those who impose them by maintaining cash flow to the banks of the developed countries and their shareholders. Whether in fact net benefits always, or even generally, occur, or occur in a way that leads to net human welfare of all affected, is a much more contentious issue within the broader world of those who think about these issues than most economists are likely to agree to. If, for example, at the most trivial level, one can generate low prices by paying laborers as little as possible or by paying as little as possible for environmental cleanup then there will be strong pressures for this to happen. Such pressures are behind much of the move for globalization, although there are those that argue that by bringing developed world standards of labor treatment and pollution control to the developing world that too generates net benefits. We are unaware of much in the way of thorough systems-oriented case history research that examines whether or not this is true except for Brown et al. and Kapilinsky [27, 28].

11.2.8 Development in More Detail: Assessment of Sustainability in Costa Rica

We certainly recognize that the assessment above can be criticized for superficiality, although we believe the results are nevertheless basic and important. But we have, with our colleagues, undertaken such analyses in much more detail in the past. The most important of these studies has been for the economy of Costa Rica, which we have examined in great detail (e.g., Hall [11] a 761-page book published in 2000 with explicit, data-intensive chapters on each of the major segments of the economy) from a biophysical and conventional economics perspective. Our original purpose for undertaking this analysis was to determine how a sustainable society and economy might be developed. Subsequently we view the book as a model for undertaking biophysical economics. But to our surprise, our study (also given in LeClerc and Hall) found at least 19 reasons that Costa Rica (often the poster child of sustainability) could not possibly be considered sustainable. Many of these reasons were based upon the interaction of energy and resource use to create a situation of decreasing efficiency (as defined in this paper). These 19 reasons include:

1. Impossible debt loads which have been approximately constant since the 1970s and which drain the government of substantial precious revenue each year.
2. There are too many people to feed, especially without fertilizers and other industrial inputs to agriculture, which can hardly be made in Costa Rica. Even with these Costa Rica now imports about half its food, requiring even more foreign exchange.
3. This results in a need to generate foreign exchange for the necessary agricultural and food inputs.
4. Even with increasing inputs, the yield per hectare for most crops has not increased since about 1985 due to erosion, depletion of nutrients, and a saturation of response to fertilizers.
5. Costa Rica, as a nation with no fossil fuels, has been, continues to be, and almost certainly will become even more dependent upon imported fossil fuels. This is true despite the very great efforts that Costa Rica has undertaken to exploit its natural advantage it has with many renewable energies: hydropower, wind, and geothermal, all a consequence of its extensive and high mountains.
6. Therefore Costa Rica is extremely vulnerable to an increase in oil prices and eventual oil depletion. All oil-importing countries are very susceptible to decreasing future availability of oil. The continued population growth makes this problem more severe year after year. Attempts at a growth economy mostly have been negated by population growth, much of it from immigrants.

7. Despite enormous efforts there have been no "silver bullets" (i.e., magic solutions to problems), and probably, the concept of sustainable development has no utility, at least so far, except, perhaps, to make the user feel good and to attract tourists.
8. Nevertheless, Costa Rica has generated an extremely good society on a relatively small resource base. There is a great deal that the rest of the world can learn about the efficiency by which Costa Rica generates good government services on a relatively small monetary and resource base.

11.2.9 Discussion

Our main conclusion from these and many other results is that there has been no increase in efficiency as measured by these criteria since "liberalization" of economies. Such economic growth as has occurred is usually by continuing the general increase in the rate of resource exploitation and use, especially energy. Neoclassical economics does not increase wealth by increasing efficiency in any sense of a scientist or engineer, at least such as we can see, but only by increasing the real work done in economies including the rate of fuel use and resource exploitation, although of course these resources can be imported (if there are other resources, including specialized human resources, that allow one to pay for them). If wealth comes from resource exploitation and not from the economists' view of free market efficiency, then the concept of development must be very tightly tied to the soils, climate, agricultural potential, mineral resources, and other biophysical resources that tend to be given short shrift in conventional economic analysis. If the human condition is to be made good, as has been the case in Costa Rica, it requires apparently as much attention to the political as economic environment, but it all must be within biophysical possibilities.

So, in summary, what must we do if we seek economic development that works?

1. Examine neoclassical economics with suspicion.
2. Use the scientific method!
3. Build a real biophysical model of the actual economic possibilities based on the real resources of a nation and its population level.
4. Consider reducing demand through, e.g., population control as an at least equally viable development strategy to increasing economic activity and, hence, the need for fossil fuels.

Questions

1. Why do you think the world economy has been so globalized?
2. What early economist might be especially interested in seeing the degree of globalization that has taken place?
3. What was the "spice route"? What replaced, in part, its function? Can you give an energy argument for that?
4. What has been the relation between imperialism and foreign trade?
5. What does "development mean"? What are some of the groups that encourage development today? (I am meaning government foreign aid, NGOs, and local investors.)
6. Many say that economic globalization is a two-edged sword with positive and negative aspects. What are some of the positive aspects? Negative?
7. Have most development models been tested? Why or why not? If so what results were found?
8. Do you think it is always difficult to test whether economic models work? Why or why not?
9. How is efficiency different from efficacy?
10. Define several uses of the word efficiency related to global issues.

References

1. Hall, C.A.S., and J.Y. Ko. 2004. The myth of efficiency through market economics: A biophysical analysis of tropical economies, especially with respect to energy, forests and water. In *Forests, water and people in the humid : Past, present and future hydrological research for integrated land and water management*, ed. M. Bonnell and L.A. Bruijnzeel, 40–58. UNESCO: Cambridge University Press.
2. Ricardo, D. 1970. *Principles of political economy and taxation*. Cambridge: Cambridge University Press.
3. Klitgaard, K. 2006. Comparative advantage in the age of globalization. *The International Journal of Environmental, Cultural, Economic and Social Sustainability* 1(3):123–129.

4. United Nations. 2010. http://www.un.org/esa/socdev/rwss/docs/2010/chapter2.pdf.
5. Palley, T.I. 2004. From Keynesianism to neo-liberalism: Shifting paradigms in economics. In *Neoliberalism – A critical reader*, ed. D. Johnston and Shad Filho. London/Ann Arbor: Pluto Press. tpalley@osi-dc.org.
6. Gowdy, J., and J. Erickson. 2005. The approach of ecological economics. *Cambridge Journal of economics* 29 (2): 207–222.
7. Bhagwati, J. 2004. *In defense of globalization*. New York: Oxford University Press.
8. Stiglitz, J. 2002. This perspective is a near mantra for many in both political parties in the US. In *Globalization and its discontents*. London: W.W. Norton, Ltd.
9. Bromely, D. 1990. The ideology of efficiency. *Journal of Environmental Economics and Management* 19: 86–107.
10. Kroeger, T., and D. Montanye. 2000. Effectiveness of structural development policies. In *Quantifying sustainable development: The future of tropical economies*, ed. C.A.S. Hall, 665–694. San Diego: Academic Press.
11. Hall, C.A.S. 2000. *Quantifying sustainable development: The future of tropical economies*. San Diego: Academic Press ; LeClerc, G., and C.A.S. Hall, eds. 2008. Making development work: A new Role for science. Albuquerque: University of New Mexico Press.
12. Annis, S. 1990. Debt and wrong way resource flow in Costa Rica. *Ethics and International Affairs* 4: 105–121.
13. Bello, W. 1994. *Dark victory: The U.S., structural adjustment and global poverty*. London: Pluto Press.
14. Gowdy, J. 2005. Toward a new welfare foundation for sustainability. *Ecological Economics* 53: 211–222; Gowdy, J.M. 2004. The revolution in welfare economics and its implications for environmental valuation and policy. *Land Economics* 80: 239–257.
15. Easterly, W. 2001. In *The elusive quest for growth: Economists' adventures and misadventures in the tropics*. Cambridge: MIT Press. We cannot emphasize enough that anyone who wants to understand economic efficiency should read the paper by Bromley, 9 above.
16. Gintis, H. 2000. Beyond Homo economicus: Evidence from experimental economics. *Ecological Economics* 35: 311–322.
17. Hall, C.A.S., P.D. Matossian, C. Ghersa, J. Calvo, and C. Olmeda. 2001b. Is the Argentine National Economy being destroyed by the department of economics of the University of Chicago? In *Advances in energy studies*, ed. S. Ulgaldi, M. Giampietro, R.A. Herendeen, and K. Mayumi, 483–498. Padua, Italy: Servizi Grafici Editoriali.
18. Sekera, J. 2016. *The public economy in crisis: A call for a new public economics*. New York: Springer.
19. LeClerc, G. 2008. Chapter 2. In *Making development work: A new role for science*, ed. G. LeClerc and Charles Hall. Albuquerque: University of New Mexico Press.
20. Montanye, D. 1994. Examining sustainability: An evaluation USAID's agricultural export-led growth in Costa Rica. Master's Thesis. State University of New York, College of Environmental Science and Forestry.
21. King, R.G., and R. Levine. 1993. Finance and growth: Schumpeter might be right. *The Quarterly Journal of Economics* 108 (3): 717–737.
22. Williamson, J. 1989. What Washington means by policy reform. In *Latin American readjustment: How much has happened*, ed. John Williamson. Institute for International Economics: Washington.
23. Gupta, A.J. et al. Estimating biophysical economic efficiency for 134 countries (in preparation).
24. Ko, J.Y., C.A.S. Hall, and L.L. Lemus. 1998. Resource use rates and efficiency as indicators of regional sustainability: An examination of five countries. *Environmental Monitoring and Assessment* 51: 571–593.
25. Tharakan, P., T. Kroeger, and C.A.S. Hall. 2001. 25 years of industrial development: A study of resource use rates and macro-efficiency indicators for five Asian countries. *Environmental Science and Policy* 4: 319–332.
26. Wiedman, T.O., H. Schandl, M. Lenzen, S. Suh, J. West, and K. Kanemotoc. 2012. The material footprint of nations. *Proceedings of the national academy of sciences of the United States of America* 112(10):6271–6276.
27. Brown, M.T., H.T. Odum, R.C. Murphy, R.A. Christianson, S.J. Doherty, T.R. McClanahan, and S.E. Tennenbaum. 1995. Rediscovery of the world: Developing an interface of ecology and economics. In *Maximum power*, ed. C.A.S. Hall, 216–250. P.O. Box 849, Niwot, CO 80544: University Press of Colorado Press.
28. Kaplinsky, R. 2005. *Globalization, poverty, and inequality: Between a rock and a hard place*. Maldin, AAA: Polity Press.

Are There Limits to Growth? Examining the Evidence

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_12

In recent decades there has been considerable discussion in academia and the media about the environmental impacts of human activity, especially those related to climate change and biodiversity. Far less attention has been paid to the diminishing resource base for humans. Despite our attention, resource depletion and population growth have been continuing relentlessly. The most immediate of these issues appears to be a decline in oil production, a phenomenon commonly referred to as "peak oil," because global production of conventional oil appears to have reached a maximum and may now be declining. However, a set of related resource and economic issues are continuing to come home to roost in ever greater numbers and impacts—water, wood, soil, fish, gold, and copper—so much so that author Richard Heinberg [1, 2] speaks of "peak everything." We believe that these issues were set out well and basically accurately by a series of scientists in the middle of the last century and that events are demonstrating that their original ideas were mostly sound. Many of these ideas were spelled out explicitly in a landmark book called *The Limits to Growth*, published in 1972 [3]. In the 1960s and 1970s, during our formative years in college and graduate school, our curricula and our thoughts were strongly influenced by the writings of ecologists and computer scientists who spoke clearly and eloquently about the growing collision between increasing numbers of people—and their enormously increasing material needs—and the finite resources of the planet. The oil-price shocks and long lines at gasoline stations in the 1970s confirmed in the minds of many that the basic arguments of these researchers were correct and that humans were facing some sort of limits to growth. It was extremely clear to us in 1970 that the growth culture of the American economy had limits imposed by nature, such that, for example, the first author made very conservative retirement plans based on his estimate that we would be experiencing the effects of peak oil just about the time of his expected retirement in 2008. In fact it was a wise decision, as many less conservative plans lost one third to one half of their value in the crash of 2008.

These ideas have stayed with us, even though they largely disappeared from most public discussion, newspaper analyses, and college curricula. Our general feeling is that few people think about these issues today, but even most of those who do so believe that technology and market economics have resolved the problems. The warning in *The Limits to Growth*—and even the more general notion of limits to growth—is seen as invalid. Even ecologists have largely shifted their attention away from resources to focus, certainly not inappropriately, on various threats to the biosphere and biodiversity. They rarely mention the basic resource/human number equation that was the focal point for earlier ecologists. For example, the February 2005 issue of the journal *Frontiers in Ecology and the Environment* was dedicated to "visions for an ecologically sustainable future," but the word "energy" appeared only for personal "creative energy"—and "resources" and "human population" were barely mentioned. But has the limits-to-growth theory failed? Even before the financial collapse in 2008, newspapers were brimming with stories about energy- and food-price increases, widespread hunger and associated riots in many cities, and various material shortages. Subsequently, the headlines have shifted to the collapse of banking systems, increasing unemployment and inflation, and general economic shrinkage. A number of people blamed at least a substantial part of the current economic chaos on oil price increases earlier in 2008. Although many continue to dismiss what those researchers in the 1970s wrote, there is growing evidence that the original "Cassandras" were right on the mark in their general assessments, if not always in the details or exact timing, about the dangers of the continued growth of human population and their increasing levels of consumption in a world approaching very real material constraints. It is time to reconsider those arguments in light of new information, especially about peak oil. ▪ Figures 12.1, 12.2, 12.3, 12.4 and 12.5 give a vivid perspective on how some of these issues play out at the local level.

12.1 Early Warning Shots

A discussion of the resource/population issue always starts with Thomas Malthus and his 1798 publication *First Essay on Population*:

> I think I may fairly make two postulata. First, that food is necessary to the existence of man. Secondly, that the passion between the sexes is necessary, and will remain nearly in

Fig. 12.1 The global population has doubled in the last four decades, as exemplified in this crowded market in India. Although some regions suffer from poverty, the world has avoided widespread famine mostly through the increased use of fossil fuels, which allows for greater food production. But what happens when we run out of cheap oil? Predictions made in the 1970s have been largely ignored because there have not been any serious fuel shortages up to this point. However, a reexamination of the models from 35 years ago finds that they are largely on track in their projections (Source: American Scientist)

Fig. 12.2 A village on one of Bangladesh's coastal islands was devastated by a cyclone in 1991, in which a total of more than 125,000 people were killed. Large storms had caused destruction in 1970 and would again in 2006. Although people in areas such as these are aware of the risk, overcrowding often prevents them from moving to safer regions (Source: American Scientist)

its present state…. Assuming then, my postulata as granted, I say, that the power of population is indefinitely greater than the power in the earth to produce subsistence for man. Population, when unchecked, increases in a geometrical ratio. Subsistence increases only in an arithmetical ratio. Slight acquaintances with numbers will show the immensity of the first power in comparison of the second.

Fig. 12.3 In 1979 motorists were forced to line up for rationed gasoline during a period of oil price shocks and reduced production. Such events were compelling support for the argument that the world's population could be limited by a finite amount of natural resources (Source: American Scientist)

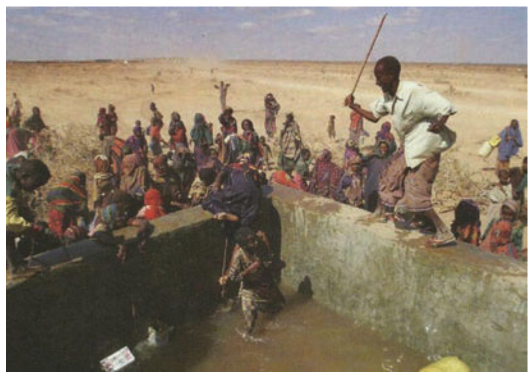


Fig. 12.4 In drought-stricken Southeast Ethiopia, displaced people wait for official distribution of donated water. Children who try to make off with the resource hours ahead of the appointed time are chased off with a cane. Such incidents demonstrate that water is another resource often available only in limited quantities (Source: American Scientist)

Fig. 12.5 Oil is not the only resource that may have peaked, with use outstripping the Earth's ability to support the level of consumption. In Sardinia, off the coast of Italy, commercial fishermen's catches are down by 80% compared to what their fathers used to haul in (Source: American Scientist)

Malthus continued with a very dismal assessment of the consequences of this situation and even more dismal and inhumane recommendations as to what should be done about it—basically to let the poor starve. Most people, including ourselves, agree that Malthus' premise has not held between 1800 and the present, as the human population has expanded by about seven times, with concomitant surges in nutrition and general affluence—albeit only recently. Paul Roberts, in *The End of Food* [4], reports that malnutrition was common throughout the nineteenth century. It was only in the twentieth century that cheap fossil energy allowed agricultural productivity sufficient to avert famine. This argument has been made many times before—that our exponential escalation in energy use, including that used in agriculture, is the principal reason that we have generated a food supply that grows geometrically as the human population has continued to do likewise. Thus since Malthus' time, we have avoided wholesale famine for most of the Earth's people because fossil fuel use also expanded geometrically.

The first twentieth-century scientists to raise again Malthus' concern about population and resources were ecologists Garrett Hardin and Paul Ehrlich. Hardin's essays in the 1960s on the impacts of overpopulation included the famous *Tragedy of the Commons* [5], in which he discusses how individuals tend to overuse common property to their own benefit even while it is disadvantageous to all involved. Hardin wrote other essays on population, coining such phrases as "freedom to breed brings ruin to all" and "nobody ever dies of overpopulation," the latter meaning that crowding is rarely a direct source of death but rather results in disease or starvation, which then kill people. This phrase came up in an essay reflecting on the thousands of people in coastal Bangladesh who were drowned in a typhoon. Hardin argued that these people knew full well that this region would be inundated every few decades but stayed there anyway because they had no other place to live in that very crowded country. This pattern recurred in 1991 and 2006.

Ecologist Paul Ehrlich [6] argued in *The Population Bomb* that continued population growth would wreak havoc on food supplies, human health, and nature and that Malthusian processes (war, famine, pestilence, and death) would sooner rather than later bring human populations "under control" down to the carrying capacity of the world. Meanwhile agronomist David Pimentel [7], ecologist Howard Odum [8], and environmental scientist John Steinhart [9] quantified the energy dependence of modern agriculture and showed that technological development is almost always associated with increased use of fossil fuels. Other ecologists, including George Woodwell and Kenneth Watt, discussed people's negative impact on ecosystems. Kenneth Boulding, Herman Daly, and a few other economists began to question the very foundations of economics, including its dissociation from the biosphere necessary to support it and, especially, its focus on both growth and infinite substitutability—the idea that something will always come along to replace a scarcer source. These writers were part and parcel of our graduate education in ecology in the late 1960s. More recently Lester Brown [10] and others provide convincing evidence that food security is declining, partly because of distributional issues and partly because of declining soil fertility, desertification, and a decrease in the availability of fossil fuel-derived fertilizer.

Meanwhile Jay Forrester, the inventor of a successful type of computer random-access memory (RAM), began to develop a series of interdisciplinary analyses and thought processes, which he called system dynamics. In the books and papers he wrote about these models, he put forth the idea of the coming difficulties posed by continuing human population growth in a finite world [11]. The latter soon became known as the limits-to-growth model (or the "Club of Rome" model, after the organization that commissioned the publication). These computer models were refined and presented to the world by Forrester's students Donella Meadows and Dennis Meadows and their colleagues [3]. They showed that exponential population growth and resource use, combined with the finite nature of resources and pollution assimilation, would lead to serious instabilities in basic global economic conditions and eventually a large decline in the material quality of life and even in the numbers of human beings. At the same time, geologist M. King Hubbert predicted in 1956 and again in 1968 that oil production from the United States would peak in 1970. Although his predictions were dismissed at the time, US oil production in fact peaked in 1970 and natural gas in 1973.

These various perspectives on the limits to growth seemed to be fulfilled in 1973 when, during the first energy crisis, the price of oil increased from $3.50 to more than $12 a barrel. Gasoline increased from less than $0.30 to $0.65 per gallon in a few weeks, while available supplies declined, because of a temporary gap of only about 5% between supply and projected demand. Americans became subject for the first time to gasoline lines, large increases in the prices of other energy sources, and double-digit inflation with a simultaneous contraction in total economic activity. Such simultaneous inflation and economic stagnation were something that economists had thought impossible, as the two were supposed to be inversely related according to the Phillips curve. Home heating oil, electricity, food, and coal also became much more expensive. Then it happened again: oil increased to $35 a barrel and gasoline to $1.60 per gallon in 1979.

Some of the economic ills of 1974, such as the highest rates of unemployment since the Great Depression, high interest rates, and rising prices, returned in the early 1980s. Meanwhile, new scientific reports came out about all sorts of environmental problems: acid rain, global warming, pollution, loss of biodiversity, and the depletion of the Earth's protective ozone layer. The oil shortages, the gasoline lines, and even some electricity shortages in the 1970s and early 1980s all seemed to give credibility to the point of view that our population and our economy had in many ways exceeded the ability of the Earth to support them. For many, it seemed like the world was falling apart, and for those familiar with the limits to growth, it seemed as if the model's predictions were beginning to come true and that it was valid. Academia and the world at large were abuzz with discussions of energy and human population issues.

Our own contributions to this work centered on assessing the energy costs of many aspects of resource and environmental management, including food supply, river management, and, especially, obtaining energy itself [12]. A main focus of our papers was *energy return on investment* (EROI) for obtaining oil and gas within the United States, which declined substantially from the1930s to the 1970s. It soon became obvious that the EROI for most of the possible alternatives was even lower. Declining EROI meant that more and more energy output would have to be devoted simply to getting the energy needed to run an economy.

12.2 The Reversal

All of this interest began to fade, however, as enormous quantities of previously discovered but unused oil and gas from outside the United States were developed in response to the higher prices and then flooded into the country. Most mainstream economists, and a lot of other people too, did not like the concept that there might be limits to economic growth, or indeed human activity more generally, arising from nature's constraints. They felt that their view was validated by this turn of events and new gasoline resources. Mainstream (or neoclassical) economists presented, mostly from the perspective of "efficiency," the concept that unrestricted market forces, aided by technological innovation, seek the "most efficient" (generally meaning the lowest prices) at each juncture. The net effect would be a continued satisfaction of consumer demand at the lowest possible prices. This would also cause all productive forces, including technology, to be optimally deployed, at least in theory.

Economists particularly disliked the perspective of the absolute scarcity of resources, and they wrote a series of scathing reports directed at the scientists mentioned above, especially those most closely associated with the limits to growth. Nuclear fusion was cited as a contender for the next source of abundant, cheap energy. They also found no evidence for scarcity, saying that output had been rising between 1.5% and 3% per year. Most importantly, they said that economies had built-in, market-related mechanisms (the invisible hand of Adam Smith) to deal with scarcities. An important empirical study by economists Harold J. Barnett and Chandler Morse in 1963 [13] seemed to show that, when corrected for inflation, the prices of all basic resources (except for forest products) had not increased over nine decades. Thus, although there was little argument that the higher-quality resources were being depleted, it seemed that technical innovations and resource substitutions, driven by market incentives, had and would continue indefinitely to solve the longer-term issues. It was as if the market could increase the quantity of physical resources in the Earth.

The new behavior of the general economy seemed to support their view. By the mid-1980s the price of gasoline had dropped substantially. The enormous new Prudhoe Bay field in Alaska came online and helped mitigate to some degree the decrease in production of oil elsewhere in the United States, even as an increasing proportion of the oil used in America was imported. Energy as a topic faded from the media and from the conversations of most people. Unregulated markets were supposed to lead to efficiency, and a decline in energy used per unit of economic output in Japan and the United States seemed to provide evidence for that theory. We also shifted the production of electricity away from oil to coal, natural gas, and uranium.

In 1980 one of biology's most persistent and eloquent spokesmen for resource issues, Paul Ehrlich, was "trapped," in his words, into making a bet about the future price of five minerals by actuary Julian Simon, a strong advocate of the power of human ingenuity and the market and a disbeliever in any limits to growth. The price of all five went down over the next 10 years, so Ehrlich (and two colleagues) lost the bet and had to pay Simon $576. The incident was widely reported through important media outlets, including a disparaging article in the *New York Times Sunday Magazine* [14]. Those who advocated for resource constraints were essentially discredited and even humiliated.

So indeed it looked to many as though the economy had responded with the invisible hand of market forces through price signals and substitutions. The economists felt vindicated, and the resource pessimists beat a retreat, although some effects of the economic stagnation of the 1970s lasted in most of the world until about 1990. (They live on still in places such as Costa Rica as unpaid debt from that period [15].) By the early 1990s, the world and US economies basically had gone back to the pre-1973 model of growing by at least 2 or 3% a year with relatively low rates of inflation. Inflation-corrected gasoline prices, the most important barometer of energy scarcity for most people, stabilized and even decreased substantially in response to an influx of foreign oil. Discussions of scarcity simply disappeared.

The concept of the market as the ultimate objective decider of value and the optimal means of generating virtually all decisions gained more and more credibility, partly in response to arguments about the subjectivity of decisions made by experts or legislative bodies. Decisions were increasingly turned over to economic cost-benefit analysis where supposedly the democratic collective tastes of all people were reflected in their economic choices. For those few scientists who still cared about resource-scarcity issues, there was not any specific place to apply for grants at the National Science Foundation or even the Department of Energy (except for studies to improve energy efficiency), so most of our best energy analysts worked on these issues on the weekend, after retirement or *pro bono*. With very few exceptions, graduate training in energy analysis or limits to growth withered. The concept of limits did live on in various environmental issues such as disappearing rain forests and coral reefs and global climate change. But these were normally treated as their own specific problems, rather than as a more general issue about the relationship between population and resources.

12

12.3 A Closer Look at the Arguments

For a distinct minority of scientists, including ourselves, there was never any doubt that the economists' debate victory was illusory at best and generally based on incomplete information. For example, Cutler J. Cleveland, an environmental scientist at Boston University, reanalyzed the Barnett and Morse study in 1991 and found that the only reason that the prices of commodities had not been increasing—even while their highest quality stocks were being depleted—was that for the time period analyzed in the original study, the real price of energy had been declining because of the exponentially increasing use of oil, gas, and coal, whose real prices were simultaneously declining [16]. Hence, even as more and more energy was needed to win each unit of resources, the price of the resources did not increase because the price of energy was declining.

Likewise, when the oil shock induced a recession in the early 1980s, and Ehrlich and Simon made their bet, the relaxed demand for all resources led to lower prices and even some increase in the quality of the resources mined, as only the highest-grade mines were kept open. But in recent years energy prices increased again, demand for materials in Asia soared and the prices of most minerals increased dramatically. Had Ehrlich made his bet with Simon over the past decade, he would have made a small fortune, as the price of most raw materials, including the ones they bet on, had increased by 2–10 times in response to huge demand from China and declining resource grades.

Another problem is that the economic definition of efficiency has not been consistent. Several researchers, including the authors, have found that energy use—a factor that had not been used in economists' production equations—is far more important than capital, labor, or technology in explaining the increase in industrial production of the United States, Japan, and Germany. Recent analysis by Vaclav Smil found that over the past decade the energy efficiency of the Japanese economy had actually decreased by 10%. A number of analyses have shown that most agricultural technology is extremely energy intensive [17]. In other words, when more detailed and system-oriented analyses are undertaken, the arguments become much more complex and ambiguous and show that technology rarely works by itself but instead tends to demand high resource use.

Likewise oil production in the United States has declined by 50%, as predicted by Hubbert. The market did not solve this issue for US oil because, despite the huge price increases and drilling in the late 1970s and 1980s, there was less oil and gas production then, and there has been essentially no relation between drilling intensity and production rates for US oil and gas since. (See ◘ Fig. 12.6 for an update.)

There is a common perception, even among knowledgeable environmental scientists, that the limits-to-growth model was a colossal failure, since obviously its predictions of extreme pollution and population decline have not come true, at least for the world as a whole. But what is not well known is that the original output, based on the computer technology of the time, had a very misleading feature: there were no dates on the graph between the years 1900 and 2100 (◘ Fig. 12.7). If one draws a timeline along the bottom of the graph for the half waypoint of 2000, then the model results are almost exactly on course some 35 years later in 2008 (with a few appropriate assumptions about what is meant by, e.g., "resources") (◘ Fig. 12.7). Of course, how well it will perform in the future when the model behavior gets more dynamic is not yet known. Although we do not necessarily advocate that the existing structure of the limits-to-growth model is adequate for the task to which it is put, it is important to recognize that its predictions have not been invalidated and in fact seem on target. We are not aware of any model made by economists that is as accurate over such a long time span (◘ Fig. 12.8).

12.4 Avoiding Malthus

Clearly, even the most rabid supporter of resource constraints has to accept that the Malthusian prediction has not come true for the Earth as a whole. Human population has increased some seven times since Malthus wrote his article, and in many parts of the world, it continues to grow with only sporadic and widely dispersed starvation (although often with considerable malnutrition and poverty). How has this been possible? The most general answer is that technology, combined with market economics or other social-incentive systems, has enormously increased the carrying capacity of the Earth for humans. Technology, however, is a double-edged sword, whose benefits can be substantially blunted by *Jevons' Paradox*, the concept that increases in efficiency often lead to lower prices [Jevons found in the middle of the nineteenth century that more efficient steam engines were cheaper to run so that people used them more—as today more fuel-efficient automobiles tend to be driven more miles in a year] and hence to greater consumption of resources [18].

And technology does not work for free. As originally pointed out in the early 1970s by Odum and Pimentel [7, 8, 19], increased agricultural yield is achieved principally through the greater use of fossil fuel for cultivation, fertilizers, pesticides, drying, and so on, so that it takes some 10

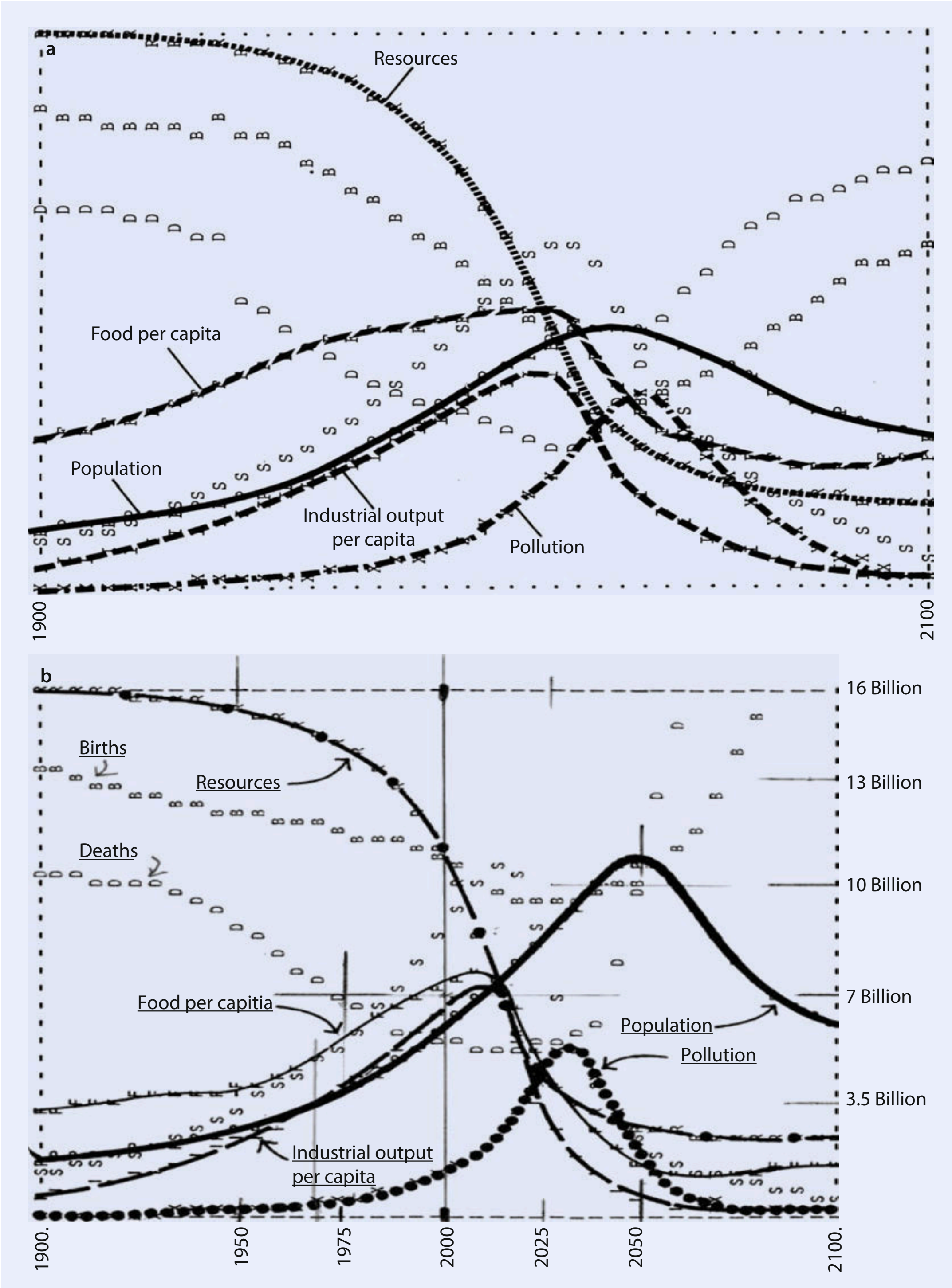


Fig. 12.6 **a** The original projections of the limits-to-growth model examined the relation of a growing population to resources and pollution, but did not include a time scale between 1900 and 2100 (Source: Hubbert 1968) **b** If a halfway mark of 2000 is added, the projections up to the current time are largely accurate, although the future will tell about the wild oscillations predicted for upcoming years (Source: Meadows et al. [3])

Parameter	Predicted	Actual
Population	6.9 bllion	6.7 bllion
Birth rate per 1000 people	29	20
Death rate per 1000 people	11	8.3
Values vs. 1970 levels (set at 1.0)		
	Predicted	**Actual**
Resources	0.53	
Copper		0.5
Oil		0.5
Soil		0.7
Fish		0.3
Pollution	3.0	
CO_2		2.1
Nitrogen		2
Per capita industrial output	1.8	1.9

Fig. 12.7 The values predicted by the limits-to-growth model and actual data for 2008 are very close. The model used general terms for resources and pollution, but current, approximate values for several specific examples are given for comparison. Data for this long time period are difficult to obtain; many pollutants such as sewage probably have increased more than the numbers suggest. On the other hand, pollutants such as sulfur have largely been controlled in many countries (Source: American Scientist)

calories of petroleum to generate each calorie of food that we eat. The fuel used is divided nearly equally between the farm, transport and processing, and preparation. The net effect is that roughly 19% of all the energy used in the United States goes to our food system.

Malthus could not have foreseen this enormous increase in food production through petroleum. In fact Malthus, associated with and supported by the landed gentry, tended to view machines in general not as pushing back the collision but rather as threatening the position of the landed class.

Similarly, fossil fuels were crucial to the growth of many national economies, as happened in the United States and Europe over the past two centuries and as is happening in China and India today. The expansion of the economies of most developing countries is nearly linearly related to energy use, and when that energy is withdrawn, economies shrink accordingly, as happened with Cuba in 1988. (There has been, however, some serious expansion of the US economy since 1980 without a concomitant expansion of energy use). This is the exception, possibly due to the US outsourcing of much of its heavy industry, compared to most of the rest of the world. Thus, most wealth is generated through the use of increasing quantities of oil and other fuels. Effectively each person in the United States and Europe has on average

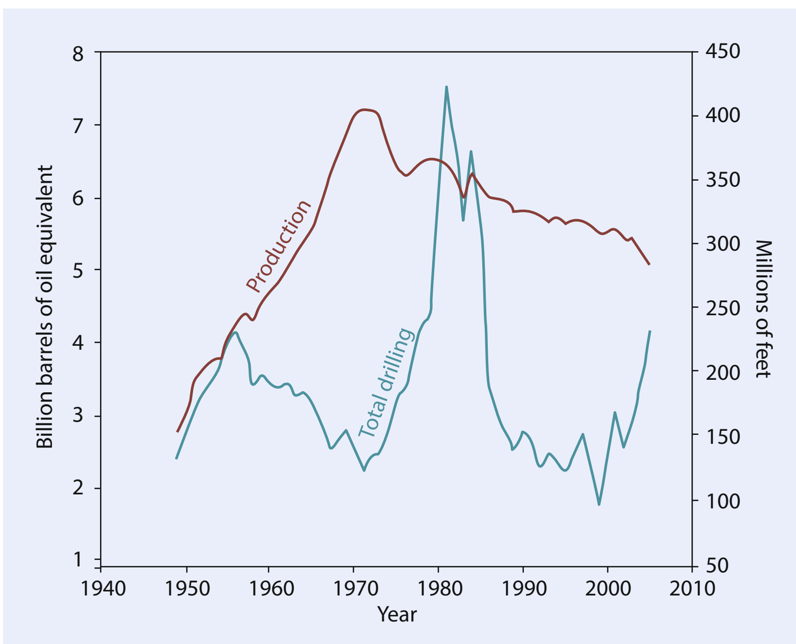


Fig. 12.8 The annual rates of total drilling for oil and gas in the United States from 1949 to 2005 are shown with the rates of production for the same period. If all other factors are kept equal, EROI is lower when drilling rates are high, because oil exploration and drilling are energy-intensive activities. The EROI may now be approaching 1:1 for finding new oil fields (Source: American Scientist)

some 30–60 or more "energy slaves," machines to "hew their wood and haul their water," whose power output is equal to that of many strong people.

Thus a key issue for the future is the degree to which fossil and other fuels will continue to be abundant and cheap. Together oil and natural gas supply nearly two-thirds of the energy used in the world, and coal another 20%. We do not live in an information age, a post-industrial age, or (yet) a solar age, but a petroleum age. Unfortunately, that will soon end: it appears that oil and gas production has reached, or soon will reach, a maximum. We reached that point for oil in the United States in 1970 and have also now reached it in at least 18, and probably the majority, of the 50 most significant oil-producing nations. The important remaining questions about peak oil are not about its existence, but, rather, when it occurred for the world as a whole, what the shape of the peak will be, and how steep the slope of the curve will be as we go down the other side.

The other big question about oil is not how much is left in the ground (the answer is a lot) but how much can be extracted at a significant energy profit. The EROI of US petroleum declined from roughly 100:1 for discoveries in 1930 to 30:1 in 1970, to about 10:1 in 2006. Even these figures are relatively positive compared to EROI for finding brand new oil in the United States, which, based on the limited information available, appears likely to approach 1:1 within a few decades.

Historically most of the oil supplies in the world were found by exploring new regions for oil. Very large reservoirs were found rather quickly, and most of the world's oil was found by about 1980. According to geologist and peak-oil advocate Colin Campbell [20], "The whole world has now been seismically searched and picked over. Geological knowledge has improved enormously in the past 30 years and it is almost inconceivable now that major fields remain to be found." Thus increased drilling appears to not be a viable approach to getting more petroleum as the finding and production of oil in the United States at least is not influenced by the amount of drilling above some very low rate. Meanwhile the world uses two to four times more oil than it finds (Fig. 8.3), and the EROI for most alternatives is much less than what we have been used to with fossil fuels (Fig. 12.9).

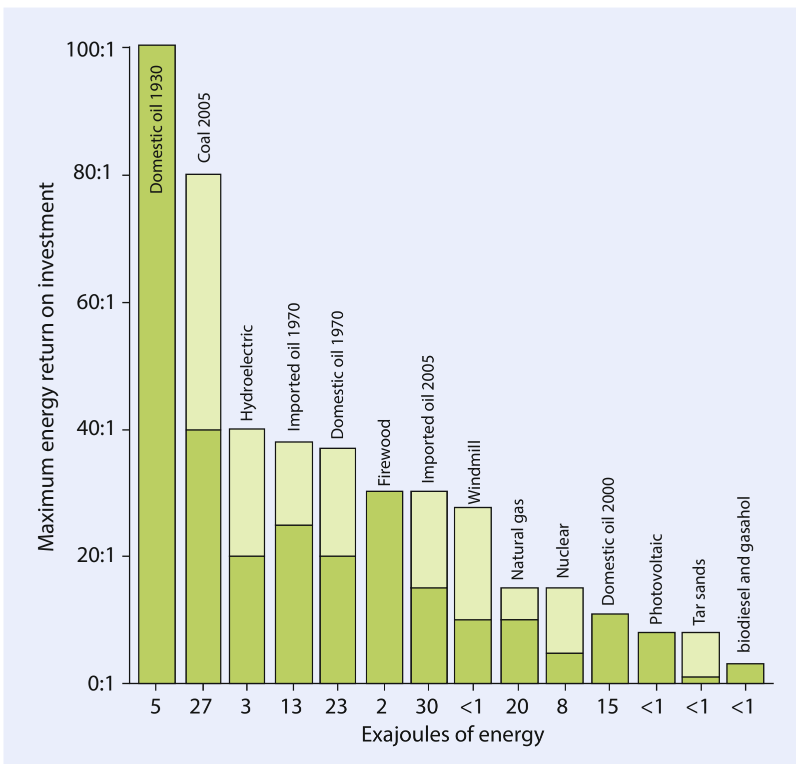


Fig. 12.9 Approximate values of energy return on investment (EROI) are the energy gained from a given energy investment; one of the objectives is to get out far more that you put in. The size of the resource per year for the United States is given below the bar. Domestic oil production's EROI has decreased from about 100:1 in 1930 for discoveries to 30:1 in 1970, to about 10:1 today, the latter two for production. The EROI of most "green" energy sources, such as photovoltaic, is presently low. (Lighter colors indicate a range of possible EROI due to varying conditions and uncertain data.) EROI does not necessarily correspond to the total amount of energy in exajoules produced by each resource. For updated values, including a discussion of electricity vs. thermal output [21] (Source: American Scientist)

12.5 Energy Scarcity

The world today faces enormous problems related to population and resources. The idea that there might be "limits to growth" from the physical availability of resources as well as the adverse effects of using them was discussed intelligently and for the most part accurately in many papers from the middle of the last century. Eventually they largely disappeared from scientific and public discussion, in part because of an inaccurate understanding of both what those earlier papers said and the validity of many of their predictions. An important issue is timing: what if these concepts are correct but merely delayed, for example, by the development of ways to exploit lower-grade resources or by economic slowdown? Most environmental science textbooks focus far more on the adverse impacts of fossil fuels than on the implications of our overwhelming economic and even nutritional dependence on them. The failure today to bring the potential reality and implications of peak oil, indeed of peak everything, into scientific discourse and teaching is in our view a grave threat to industrial society. The concept of the possibility of a huge, multifaceted failure of some substantial part of industrial civilization is so completely outside the understanding of our leaders that we are almost totally unprepared for it. For large environmental and health issues, from smoking to flooding in New Orleans, scientific evidence of negative impacts has historically preceded general public acceptance and policy actions by several decades.

There are virtually no extant forms of transportation, beyond shoe leather and bicycles, that are not based on oil, and even our shoes and bicycle tires are now often made from oil. Food production is very energy intensive, clothes and furniture and most pharmaceuticals are made from and with petroleum, and most jobs would cease to exist without petroleum. On our university campuses, one would be hard pressed to have had any sense of that beyond complaints about the increasing price of gasoline. Most Americans believe that new technologies have generated a large surplus of oil, when in fact nearly half of our oil used is still imported, and most of our fracked oil provinces are already past their peak in production.

No substitutes for oil have been developed on anything like the scale required, and most are very poor net energy performers. Despite considerable potential, renewable sources (other than hydropower or traditional wood) currently provide less than 2% of the energy used in both the United States and the world, and the annual increase in the use of most fossil fuels is generally greater than the increase in electricity from wind turbines and photovoltaics. Our new sources of "green" energy are simply increasing along with (rather than displacing) all of the traditional ones.

If we are to resolve these issues, including the important one of climate change, in any meaningful way, we need to make them again central to education at all levels of our universities and to debate and even stand up to those who negate their importance, for we have few great intellectual leaders on these issues today. We must teach economics from a biophysical as well as a social perspective. Only then do we have any chance of understanding or solving these problems.

Questions

1. What is "the limits to growth"? Give two interpretations.
2. Who was Thomas Malthus? What, basically, did he say?
3. Has Malthus predictions held? For what time period? Why or why not?
4. There was relatively little scholarly writing about Malthus' ideas until the 1960s, when some people again paid a lot of attention to them. Can you name any of the people who rediscovered Malthus' ideas? What fields did they come from? Why do you think this was the case?
5. What was the Club of Rome?
6. What happened in the 1970s that seemed to support the concept that there might be "limits to growth"?
7. How was this related to economic issues? Why did this have such a widespread impact?
8. What happened in the 1980s that completely changed the perspective of many on limits to growth?
9. Discuss the concept of markets in relation to this changing perceptions.
10. What additional insight to the influential work of Barnett and Morse was undertaken by Cutler Cleveland?

11. Why did many people think the limits-to-growth model failed? What do you think?
12. Why do many people think that technology can generate solutions to resource problems?
13. What is Jevons' paradox?
14. What is the general relation between energy availability and limits to growth? Do you think that technological advances change that relation?

References

1. Hall, C.A.S., and J.W. Day Jr. 2009. Revisiting the limits to growth after peak oil. *American Scientist* 97: 230–237. See also Turner, Graham M. 2008. A comparison of the limits to growth with 30 years of reality. *Global Environmental Change* 18: 397–411, which arrives at much the same conclusions.
2. Heinberg, R. 2007. *Peak everything: Waking up to the century of declines*. Gabriola Island B.C: New Society Publishers.
3. Meadows, D.H., D.L. Meadows, J. Rander, and W.W. Behrens III. 1972. *The limits to growth*. Washington, DC: Potomac Associates.
4. Roberts, P. 2008. *The end of food*. New York: Houghton Mifflin.
5. Hardin, G. 1968. The tragedy of the commons. *Science* 162: 1243–1248.
6. Ehrlich, P. 1968. *The population bomb*. New York: Ballantine Books.
7. Pimentel, D., L.E. Hurd, A.C. Bellotti, M.J. Forster, I.N. Oka, O.D. Sholes, and R.J. Whitman. 1973. Food production and the energy crisis. *Science* 182: 443–449.
8. Odum, H.T. 1973. *Environment, power and society*. New York: Wiley Inter science.
9. Steinhart, J.S., and C. Steinhart. 1974. Energy use in the US food system. *Science* 184: 307–316.
10. Brown, Lester R. 2009. Could food shortages bring down civilization? *Scientific American* 300 (5): 50–57.
11. Forrester, J.W. 1971. *World dynamics*. Cambridge: Wright-Allen Press.
12. Charles Hall. Webpage: http://www.esf.edu/efb/hall/
13. Barnett, H., and C. Morse. 1963. *Scarcity and growth: The economics of natural resource availability*. Baltimore: Johns Hopkins University Press.
14. Tierney, J. 1990. Betting the planet. *New York Times Magazine*. December 2: 79–81. See also Passell, P., M. Roberts, and L. Ross. 2 April 1972 "Review of limits to growth". New York Times Book Review.
15. Hall, C.A.S. 2000. *Quantifying sustainable development: The future of tropical Economies*. San Diego: Academic.
16. Cleveland, C.J. 1991. Natural resource scarcity and economic growth revisited: Economic and biophysical perspectives. In *Ecological economics: The science and management of sustainability*, ed. R. Costanza. New York: Columbia University Press.
17. Smil, V. 2007. Light behind the fall: Japan's electricity consumption, the environment, and economic growth. *Japan Focus,* April 2.
18. Hall, C. 2004. The myth of sustainable development: Personal reflections on energy, its relation to neoclassical economics, and Stanley Jevons. *Journal of Energy Resources Technology* 126: 86–89.
19. Hamilton, A., S.B. Balogh, A. Maxwell, and C.A.S. Hall. 2013. Efficiency of edible agriculture in Canada and the U.S. over the past 3 and 4 decades. *Energies* 6: 1764–1793.
20. Campbell, C., and J. Laherrere. 1998. The end of cheap oil. *Scientific American*. 78–83.
21. Hall, C.A.S., J.G. Lambert, S. B. Balogh. 2014. EROI of different fuels and their implications for society. *Energy Policy* 64: 141–152.

12

The Petroleum Revolution III: What About Technology?

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_13

A main concern of this book has been that—given the extreme dependence of most contemporary societies and economies on energy and the finite nature of fossil fuels—what kind of a future can the young readers of this book expect if our economic needs and expectations face severe constraints in the future availability of fossil fuels? As we have developed previously in this book (▶ Chap. 8) and frequently elsewhere, the two principle concerns we have about future availability and affordability of fossil fuels have been absolute supplies (e.g., "peak oil," the idea that oil will reach a peak in production and then inevitably decline) and declining EROI. But what if these issues were not to occur or to do so only so far in the future that they would have no meaning to anyone alive today? Certainly there have been economists who have argued that technology and substitutions will indefinitely hold off the effects of depletion [e.g., 1]. Could they be right?

At the time of this writing (mid-2017), it is ambiguous whether resource limitations in the petroleum industry, and for energy more generally, are causing any severe constraints in economic activity, at least for the relatively wealthy parts of the world. Gasoline is not cheap, but it is not as expensive as it had once been either, as had been predicted by many. In inflation-corrected terms, it is more expensive than in the past, which may be related to economic impact. There is no shortage; if you have the money, you can buy as much as you want. There are two principle reasons for this: the production of oil in the United States, which had been declining (as predicted by M. King Hubbert) since 1970, began an upward surge again in 2008, causing an increase in global oil production which was otherwise flat (◘ Fig. 13.1). And the increase in consumption, which had been increasing in both the United States and especially the world for 150 years, began to level off. We examine the first reason in some detail here.

Starting in about 2008, there has been a great deal of excitement and debate about whether "unconventional" oil from, for example, the Bakken Formation in North Dakota and the Eagle Ford field in Texas and natural gas from shales, such as the Marcellus shale, can provide or were providing an energy renaissance for the United States. While the amount of oil in these formations is enormous, the rocks (as likely to be sandstone or limestone as shale) have low porosity (pore space) and permeability (ability of oil to flow through the formation) so that only 5% of the oil in place can be extracted even by new, energy-intensive "heroic" efforts. This compares with an average of 38% from conventional fields. New technologies were required, including horizontal drilling, shooting a series of holes into the horizontal pipe, and the shattering or "fracking" of the rocks with very high-pressure water. Then special sand is added to keep the strata apart, and then the water is withdrawn, allowing the oil to drip from the rock and go back through the pipe to a collection location at the initial vertical hole (◘ Fig. 13.2). This is obviously a very sophisticated procedure, and it requires a great deal of very sophisticated and expensive equipment. Lateral extensions up to 2 miles long are now routinely used.

The concept of fracking is actually an old one, initially done, with sometimes fatal results, with nitroglycerin, which had been occasionally used to enhance oil production as long ago as about 1930. Likewise, horizontal drilling has been around for a very long time. What is new is their combination along with the use of special sand to prop open the cracks made by high-pressure water sent downhole. This new technological package has been used to exploit extensive horizontal source beds of oil in Texas, North Dakota, and elsewhere (◘ Figs. 13.2 and 13.3). While there is a lot of oil in these formations, and it is often of good quality, the low porosity and permeability are problems. As a result, production in these wells tends to decline very rapidly compared to conventional oil wells. On the plus side, there are few dry holes because the oil is in massive, relatively uniform "source rock" formations rather than more concentrated "trap rock" reservoirs.

The result has been a dramatic reversal in the long-term decline of oil production in the United States and the world (◘ Figs. 8.5 and 13.1). A peak nearly at the level of the 1970 peak was achieved in 2015, although production since then has declined but then trended upward again. Whether the decline was due to the depletion of the "sweet spots" or the decline in effort due to the low price of oil is not quite certain. In 2016, the United States again began importing about half the oil it consumes. An analysis of top counties in fracked plays such as the Bakken and Eagle Ford shows that average well productivity has begun to

13

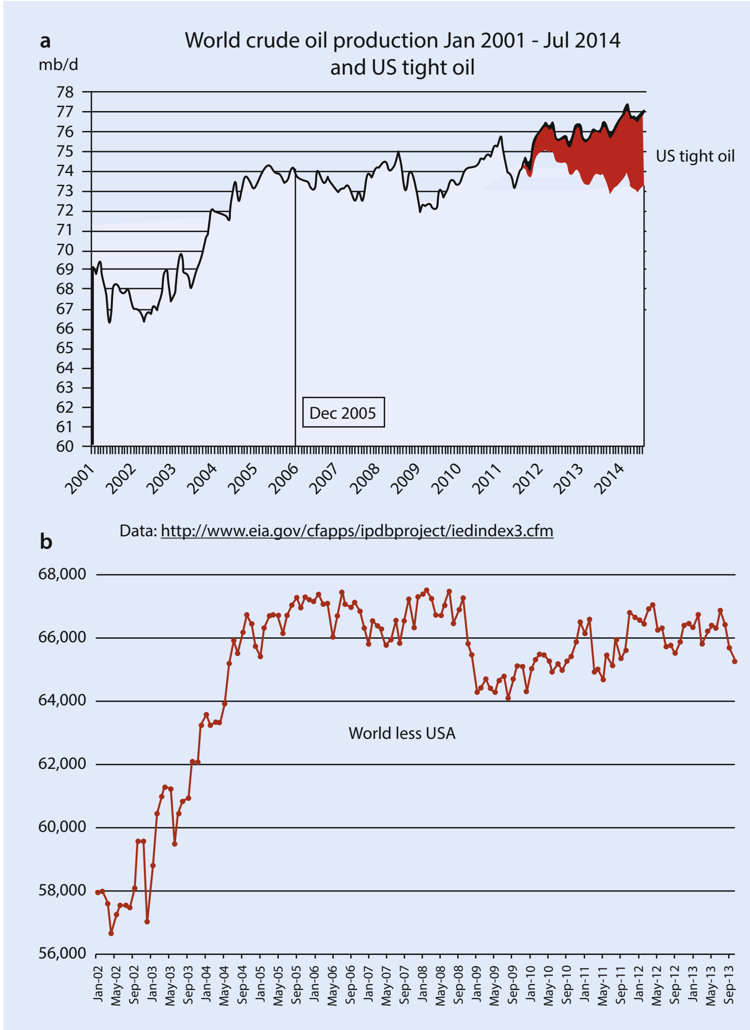


Fig. 13.1 **a** Time series of oil production in the United States 1983–2017. The peak in US oil production of 10,500 thousand barrels per day occurred in 1970 (Source U.S. EIA; ► http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=WCRFPUS2&f=W). **b** Production of world conventional crude oil without including the oil from the United States

decline (Fig. 13.4), meaning that the best locations have been exhausted along with possible well interference (from "infill drilling"—wells being drilled too close together—robbing Peter to pay Paul). David Hughes, a geologist with a lot of experience and credibility analyzing fracked oil and gas resources, estimates that the total quantity of oil that will be recovered from these "fracked" fields will be in total no more than about 25% of remaining conventional oil—significant but not really a game changer [2].

Curiously, according to Art Berman [3], almost no companies that have invested in fracking have turned a profit, and at least at about $50 per barrel, the main reason, that most companies continue to frack is to remain in business, and to generate enough revenue to repay the debts incurred to commence drilling, not make a profit (as we saw with oil in its early history in the United States in ► Chap. 10). There is much being made about "increasing efficiency" but that turns out to be mostly due to decreasing the payments to the oilfield servicing companies. It is quite curious as to why investors (and the service companies) keep investing in companies that are not making money.

The horizontal drilling-fracking technique has been applied in China and England with uncertain results and in Poland where it has been abandoned as too low yielding and expensive. It is unlikely that fracking will do other than delay the inevitable US peak and decline by more than about a decade. On the other hand, it seems that so far every time the United States is about to have a catastrophic decline in oil production, something

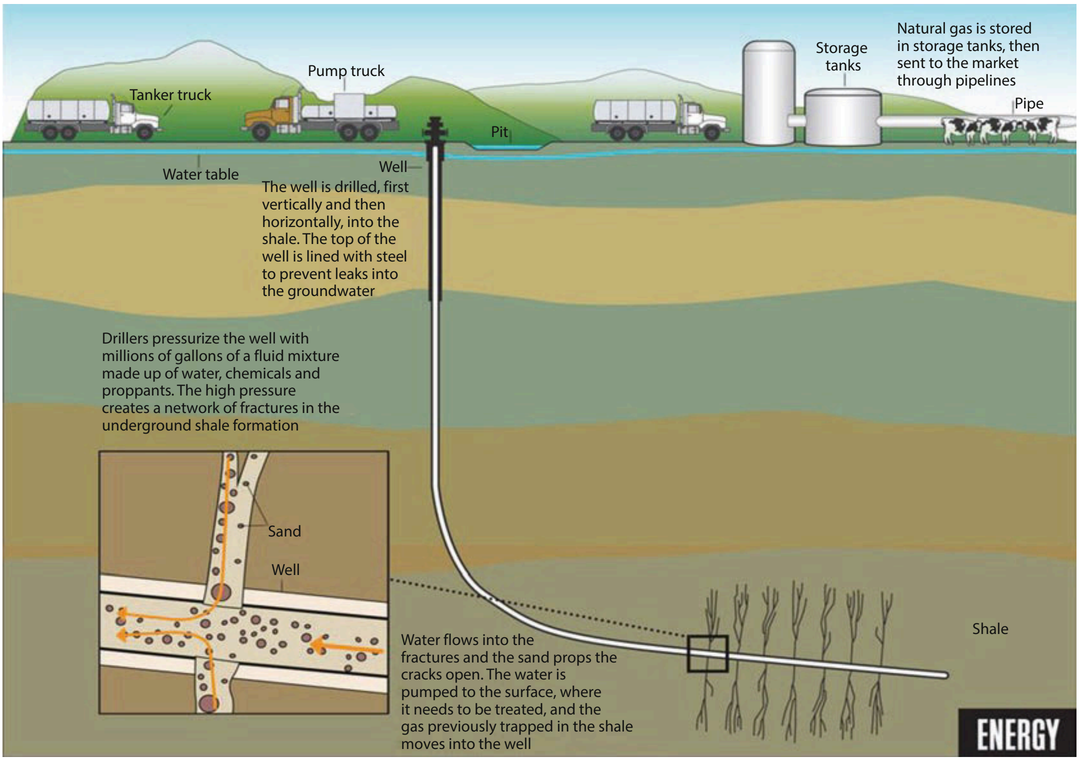


Fig. 13.2 Fracking procedure (From Kaufmann and Cleveland 2016)

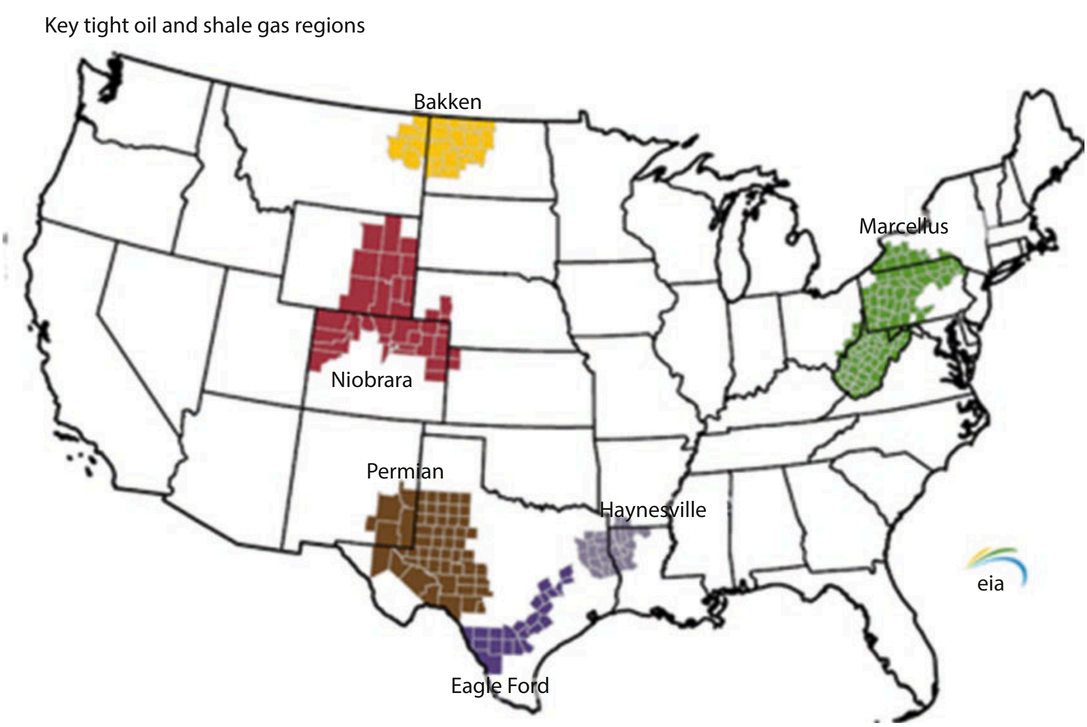


Fig. 13.3 Most important locations for fracking oil and gas (By county)

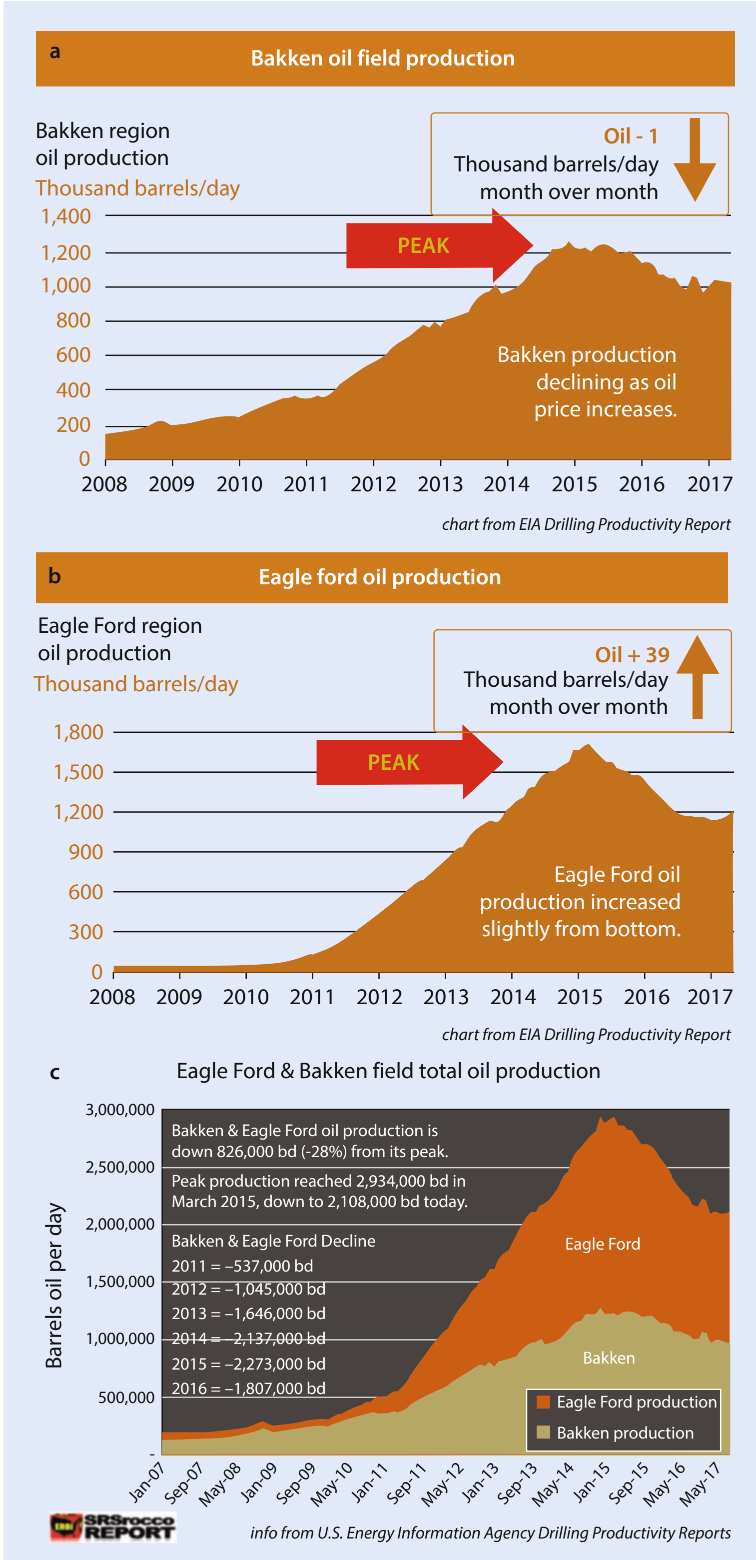


Fig. 13.4 **a–c** Increase and then decline in the major US shale oil plays (From ► *SRSrocco Report*. April 20, 2017)

comes on line to give us at least temporary relief: Alaska in 1975 and fracking in 2008. For how long can we count on such miracles? Will the world shift to non-carbon sources of energy if relatively cheap oil is still available? The next quarter to half century is likely to be extremely interesting with respect to energy, although we have hardly a clue to predict the unknown technologies.

Meanwhile conventional gas production in the United States has peaked and dropped off to less than half the peak, while unconventional gas of all kinds is mostly compensating for the decline in conventional gas while increasing production slightly [4]. The United States continues to import large quantities of oil and some gas from abroad. Thus, while fracked oil and natural gas are likely to be very important as conventional oil production declines, the decline in conventional resources is at least as important a story.

13.1 New Hubbert Peaks?

There is another huge change occurring with respect to the production of oil and other fossil fuels. Over the past five decades, those of us who have been thinking about "peak oil" would perceive of oil peaking first, then natural gas a decade or two later, and then coal much later [e.g., 5]. In our perception the amount of coal was very large and would more than compensate for the decline in oil and gas that would occur by, say, 2025. But starting in about 2012, new assessments suggested that while the peak in oil might take a little longer to unfold, there might be a peak much sooner in coal! Three different and independent assessments have concluded that a peak in all fossil fuels might occur as soon as 2025 [6–8].

13

13.2 New Technologies to the Rescue?

It is possible that a production peak could come even earlier, based on human efforts to reduce CO_2 emissions [9]. Several authors are very enthusiastic about the possibility of a relatively "carbon-free" future [10, 11]. But that would not be easy.

Energy authority Vaclav Smil [12] has summarized the challenges of moving off mostly fossil fuels:

> There are five major reasons that the transition from fossil to non-fossil supply will be much more difficult than is commonly realized: scale of the shift; lower energy density of replacement fuels; substantially lower power density of renewable energy extraction; intermittence of renewable flows; and uneven distribution of renewable energy resources.

Trainer [13] likewise found that costs of renewables (including wood and hydropower) in the United States were extremely large.

Whether renewable energies, such as wind, biomass, and solar PV, could replace some large part of the fossil fuels anytime soon seems to be highly unlikely, although advocates suggest that it is possible (In *Our Renewable Future*, Heinberg and Fridley believe the need to avoid climate catastrophe will make the investment worth it even at the enormous cost it would entail (estimated on page 123 as 20 times the present rate of all investments in renewables for many decades).

David MacKay [14] concluded: "we must have no delusions about the area required for large-scale solar power; about the challenge of transmitting energy over large distances; about the additional costs of handling intermittency; and about the need for breakthroughs not only in the whole-system costs of photovoltaics but also in the cost of systems for storing energy. CSP (concentrating solar) plants need to be in safe locations, and the ultra-high voltage direct current transmission (UHVDC) system required in order to transport the electricity to points of final use must be built. This is not currently feasible in North Africa, for example."

To give an example of the difficulties, today most renewable energy comes from hydropower and biomass, and the contribution of the latter is declining, so that the total contribution from all renewable in the United States has barely increased from 11% in 2010 to 12.6% in 2016 to a projected 16.1% by 2040 (U.S. EIA). Meanwhile, the EIA projects that all fossil fuels will continue to increase in absolute terms. So much for reducing CO_2 emissions! Some solar advocates project a much higher transformation rate, in line with what these authors see as necessary, as prices for, e.g., PV-generated electricity decline. We shall see. Whatever happens the fossil energy cost of

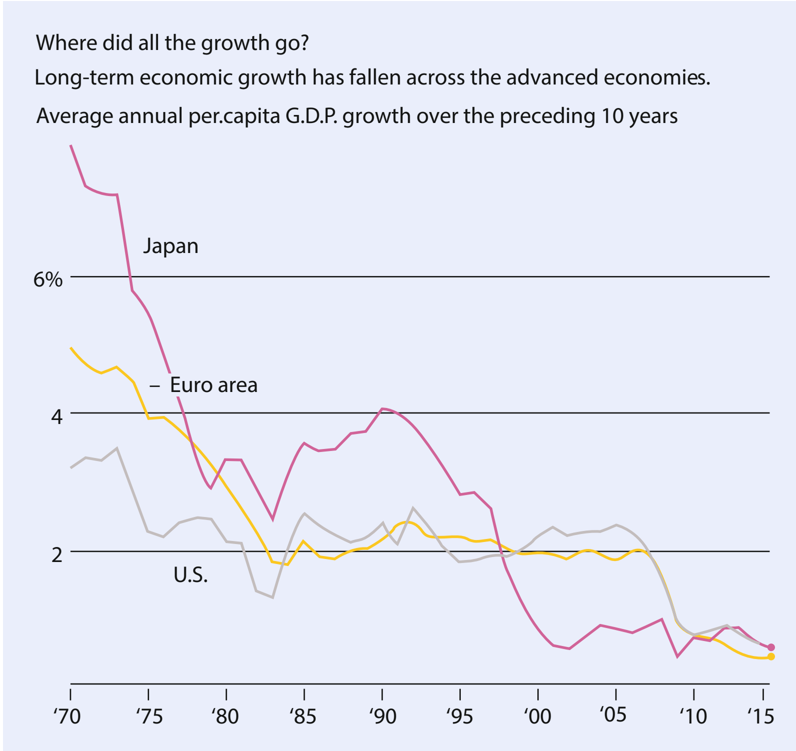


Fig. 13.5 Decline in economic growth rate for the United States, Europe, and Japan

the transition to solar will be enormous. Simply contemplating transportation without oil is almost impossible to imagine [15].

13.3 EROI

Of perhaps greater concern than the quantity of oil and other energy sources is their declining EROI. The world will not run out of hydrocarbons. Instead it has, and will increasingly, become difficult to obtain cheap petroleum, because what is left is an enormous amount of low-grade hydrocarbons which are likely to be much more expensive financially, energetically, politically, and environmentally. As conventional oil becomes less available, society probably will make investments in different sources of energy and improvements in energy use efficiency, in theory reducing our dependence on hydrocarbons but possibly decreasing the EROI of our overall fuel mix. Meanwhile much of the world's economy has essentially stopped growing, perhaps in response to increasing resource limitations, implying a very different future that might greatly change our projections and options (Fig. 13.5).

13.4 Conclusion

The number of possible scenarios for future energy production is very large. Some indicate a major decline in energy availability; others suggest that renewables can take up the slack. Our assessment is that it is unlikely that we can build alternatives rapidly enough to fill in for declining oil and possibly other fossil fuels once serious declines begin, which seems inevitable [16].

References

1. Adelman, M.A., and M.C. Lynch. 1997. Fixed view of resource limits creates undue pessimism. *Oil & Gas Journal* 95: 56–60.
2. Hughes, David. December 12, 2016 Revisiting the U.S. Department of Energy Play-by-Play Forecasts through 2040 from Annual Energy Outlook 2016.
3. Berman, Art. Personal communication and website.
4. US EIA weekly reports.
5. Hall, C.A.S. 1975. The Biosphere, the industriosphere and their interactions. *The Bulletin of the Atomic Scientists* 31: 11–21.
6. Maggio, G., and G. Cacciola. 2012. When will oil, natural gas, and coal peak? *Fuel* 98: 111–123.

7. Mohr, S.H., J. Wang, G. Ellem, J. Ward, and D. Giurco. 2015. Projection of world fossil fuels by country. *Fuel* 141: 120–135.
8. Laherrere, Jean. ASPPO France website.
9. McGlade, C., and P. Ekins. 2015. The geographical distribution of fossil fuels unused when limiting global warming to 2°C. *Nat* 517 (7533): 187–190.
10. Jacobson, M.Z., M.A. Delucchi, G. Bazouin, Z.A. Bauer, C.C. Heavey, E. Fisher, S.B. Morris, D.J. Piekutowski, T.A. Vencill, and T.W. Yeskoo. 2015. 100% clean and renewable wind, water, and sunlight (WWS) all-sector energy roadmaps for the 50 United States. *Energy and Environmental Science* 8 (7): 2093–2117.
11. Heinberg, R., and D. Fridley. 2016. *Our renewable future: Laying the path for one hundred percent clean energy*. Richard: Island Press. see Hall 2016. Review in *BioScience* 66: 1080–1081.
12. Smil, V. 2011. Global energy: The latest infatuations. *American Scientist* 99: 212–219.
13. Trainer, T. 2013. Can Europe run on renewable energy? A negative case. *Energy Policy* 63: 845–850. Trainer, T. 2012a. A critique of Jacobson and Delucchi's proposals for a world renewable energy supply. *Energy Policy* 44: 476–481.
14. MacKay, D. 2010. *Sustainable energy without the hot air*. Cambridge: UIT Cambridge.
15. Friedemann, A.J. 2016. *When trucks stop running: Energy and the future of transportation*. New York: Springer.
16. See web site of Matt Mushalik in Australia for intelligent update of fossil fuel and other data.

Energy and Economics: The Scientific Basics

We have made the argument in Part I of this book that it is important to bring more science, including the natural sciences but also the social sciences and all under the wing of the scientific method to economics. This section is meant as a review of the scientific method and as the basic information (mostly from the natural science) necessary to properly understand economics from the biophysical perspective and as represented in ◘ Fig. 3.3. We start with the necessary history and facts about energy, then relate that to economics, then introduce basic mathematical skills, and finally ask whether economics today can properly be considered a science.

Contents

What Is Energy and How Is It Related to Wealth Production?

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_14

14.1 Energy: The Unseen Facilitator

Energy is, at best, an abstract entity for most contemporary people. Only rarely does it enter our collective consciousness, generally in those relatively rare times when there are particular shortages or sharp price increases in electricity or gasoline. In fact, as this book will demonstrate, energy and its effects are pervasive, relentless, all-encompassing, and responsible for not only each process and entity in nature and our own economic life but also for many aspects of the basic nature of our psyches and many of the ways that world history has unfolded. Few understand or acknowledge this because the pervasive impact in energy shown in this book does not usually enter into our collective training and education, and it does not enter into our educational curricula. Why is this so? If energy is as important as we believe then why is that not more generally known and appreciated? The answers are complex. One important reason is that the energy that is used to support ourselves, our families, or our economic activity generally is used at some other location and by other people, often in order to reduce environmental impacts on people, or by quiet, automatic machines whose fuel tends to be relatively cheap. After all, coal, oil and gas, our principal sources of energy, are basically messy, smelly, dangerous, and unpleasant materials. The energy from food that we need to fuel ourselves surrounds most of us abundantly and is available readily and relatively cheaply. Society has gone to great lengths to isolate most of us physically and intellectually from the energy sources upon which our food, our comfort, our transportation, and our economy depend. It is convenient to ignore energy because many facts about it are uncomfortable to know.

Perhaps a more important reason for our failure to understand the pervasive role of energy is that most uses of energy are indirect. Humans are conditioned, both evolutionarily and in their social education, to want and need the goods or services rendered by energy, but not energy itself. In fact energy per se, with the exception of food energy and warmth in winter, is hardly ever desirable or useful directly. This conditioning, however, does not diminish the requirement for energy for virtually everything that we do, nor compensates for the fact that our use of energy has become enormous in contemporary life. Today each average American has some 60 to 80 slaves toiling tirelessly to keep us at about 70 degrees, well fed, mobile, entertained, and so on. Where are these slaves? We can see the car engine, the furnace, or the air conditioner, but who is aware of the electric pump supplying water or running the refrigerator or the massive electrical and fossil-fueled devices digging up the Earth to bring us the energy to run these devices. Who thinks about the energy required to make the metals and plastics in our car; the timbers and concrete in our homes, offices, and schools; or the paper in this book? But they all require it, and a lot of it.

Another reason that we do not think much about energy is that energy today remains enormously cheap relative to its value. If we want water delivered to our house, we might hire a person to do the job. A very strong person can work at a rate of about 100 watts so in a 10 hour day could do 1000 watt-hours (1 kilowatt-hour) of work, say hauling water from a well to our sink or shower. If we paid that strong person at minimum wage, he or she would charge about 80 dollars for the 10 hours' work. But if we installed an electric pump, we can get the same work done for about ten cents per kilowatt-hour. Since humans work at about 20% efficiency but electric pumps perhaps 60%, the relation is tipped even more in favor of the electric pump. So, to do the same physical work with a person that a pump could do would cost about 800 times 3 or 2400 times more with the worker compared to the electric motor! So this is the main reason that an average American or European today is far richer than the richest king of old—we have cheap energy to supply us with the necessities and luxuries in life. A problem is that we have become dependent in many ways upon this cheap energy and the goods and services it supplies. The value of this energy is far more than what we pay for it, because the services energy provides is far more valuable than its monetary cost. Additionally its potential abundance is much more limited than our dependence would imply.

14.2 A History of Our Understanding of Energy

Two hundred years ago, no one understood energy as a concept, although they certainly understood many practical consequences such as plants needing sunlight for growth and the need for wood to do many economic things such as cooking and making metals or cement. Any concept of energy was tied up in confusion, often mystical, about the actual results of energy use, because energy cannot be seen or felt and only its effects can be observed. Fire was thought of as a basic substance (as in earth, air, fire, and water) rather than as the energy released from the destruction of chemical bonds generated earlier by photosynthesis and the formation of new bonds with oxygen. How could people then possibly understand energy if they did not have any concept of chemical bonds, oxygen, or chemical transformations? How could they possibly understand that the growth of plants, the work of a horse, the erosion of water, the heat generated by fire, and their own exertions had some common something that tied them all together? To them they were independent entities. The failure of educated people to understand energy comprehensively as the principles of economics were being developed was probably the principle reason that economics developed as a social, rather than more powerfully as a biophysical, science.

As in most other things in their life that they did not understand, ancient people attributed energy, or at least some aspects of it, to a god or gods: the sun, of course, was worshiped by many cultures who understood clearly its importance for their food and warmth, but there were many other energy gods or special energy-related entities: Prometheus, Hephaestus, Pele, Vesta, Hestia, Brigid, Agni, and Vulcan to name a few. These people had no possible way to see that there were common concepts linking the sun and the fire resulting from burning wood, nor could they understand that so many other processes that they also attributed to different gods (wind, rain, agriculture, the existence of wild creatures, and so on) were also connected to the sun. The knowledge that a sharp sixth grader today has about energy and science in general would be far beyond what the most learned person would understand 400 or even 200 years ago. We have learned an astonishing amount about how the world really works through science. Even today, however, we cannot measure energy directly but only its effects! But we have become much better at that and in understanding how all of this is related.

During 200 years, from roughly 1650 to 1850, a series of remarkable discoveries and experiments, mostly from French, Scottish and English scientists, allowed us to understand in a comprehensive way the essentials of energy. First and foremost among these were the remarkable discoveries of Isaac Newton. Newton discovered the three laws of motion, and in the more than 350 years since then no fourth law has been discovered! He also derived the law of universal gravitation and wrote critically important books on optics. Nevertheless by his own admission, he did not understand economics, and he lost most of his money on an ill-advised investment scheme. He said "I can calculate the motion of the heavenly bodies, but not the madness of the crowd."

14.3 Newton's Laws of Motion

The first law says that a body in motion (and this includes no motion, i.e., rest) tends to stay in that state unless acted upon by an outside force. This is completely counterintuitive, as most moving things come to a stop! But Newton realized that it was an outside force, friction, that caused them to stop, and if there were no friction they would continue in their path indefinitely. The first law explains many things we experience—the momentum of an automobile when we put in the clutch, the path of a baseball (although we need to include gravity), and even centrifugal force.

The second law says that the acceleration of an object, say a baseball being hit, equals the force applied to the object divided by the mass of the object. It is familiarly written as:

$$F = MA$$

which can be rewritten as:

$$A = F/M$$

Thus a powerful baseball hitter, such as the legendary Babe Ruth, was capable of applying great force (F) to a baseball with his bat, accelerating it greatly (A), and giving it enough velocity (sometimes) to travel out of the ball park. The force that he applied could be measured by measuring the mass (M) of the baseball and the amount that the ball was accelerated. If one could make a baseball twice as heavy with a lead core, it would, other things being equal, be accelerated only half as much.

Newton's third law of motion says that for every force, there is an equal and opposite force. This is evident when you are in a small boat and move your body one way and the boat moves in the opposite direction. It is obvious to anyone who has fired a rifle that the gun moves back against your shoulder when the bullet is accelerated forward. It was also obvious to early designers of ship-borne cannon that if proper arrangements were not made the recoil of the cannon could do more damage to the ship shooting it than to the target!

Newton also determined the "universal law of gravitation" that two bodies would attract each other as the product of their masses and the inverse of their distance squared. This law is so powerful that it can be used to explain essentially perfectly the orbits of planets around the sun, the movement of a hit baseball, or the relation of electrons to the nucleus of molecules.

Probably the most important results of Newton's work are that it showed that the physical world followed definite laws that appeared (and still appear) to never be broken no matter where or when applied and that many of these laws could be expressed by simple mathematical equations. Although the concept of energy was not yet known to Newton, we now understand that energy was related to matter by the relation of force to mass. In the hundreds of years of science that has followed, many have tried to find simple, elegant mathematical laws that were as powerful as Newton's, but with the exception of Albert Einstein and James Clerk Maxwell, few succeeded.

14

The essence of what "force" (as in Newton's second law) was, where it came from, and how it changed over time remained elusive. The next important step in our understanding was the understanding of the relation of physical energy to heat. It was obvious that over time a lot of fuel wood was needed to run any major production process. Also, it was certainly apparent to observers that many physical actions were associated with heat, as was obvious by the heating of turning wagon wheels or a hard working horse or person, or the drilling of the hole in a cannon. But why this should be or what it meant remained elusive.

14.4 The Mechanical Equivalent of Heat

Many early scientists and engineers, seeing and understanding the tremendous force made possible when water was heated to form steam, were interested in building engines to do mechanical work. Thomas Savery built the first heat engine as early as 1697. Although his and other early engines were crude and inefficient, they could do a lot of work, and they attracted the attention of the leading scientists of the time. Classical thermodynamics as we know it now evolved in the early 1800s with concerns about the states and properties of everyday matter including energy, work, and heat. Sadi Carnot, the "father of thermodynamics," published in 1824 the paper that marked the start of thermodynamics as a modern science. Its title was "Reflections on the Motive Power of Fire, a discourse on heat, power, and engine efficiency" which outlined the basic energetic relations among the Carnot engine, the Carnot cycle, and motive power. This small volume gave for the first time the basic relations between input energy and the necessary transformation of a part of it to heat as work was done. It also derived a means of calculating the maximum efficiency that a machine could obtain as a relation between the temperature of the input energy and the temperature of the environmental sink to which the final heat was exhausted:

$$W = (Ts - Te) / Ts$$

The Carnot efficiency of heat-to-work conversion of an ideal heat engine that receives heat of high absolute temperature, Ts, from a source (e.g., a furnace) and rejects heat of lower temperature $Te < Ts$ to a sink (e.g., a river). By definition, it cannot exceed 1. This equation explains why despite the vast amount of heat stored in, for example, the surface of the North Sea in summer so little work can be done from it: the difference between the surface temperature (30 degrees) and the deepest water (2 degrees) is too small compared to, say, the temperature difference in an oil fired power plant, where temperatures at the turbine entrance may reach 817 degrees C and the cooling water might be from 6 degrees (winter) to 17 degrees (summer). It also explains why power plants are slightly more efficient in winter.

Only a few hundred copies were published, and Carnot died thinking his work had had no impact. But a few copies were found by others who developed the concepts further. The term *thermodynamics* was coined by James Joule in 1849 to designate the science of relations between heat and power. By 1858, "thermodynamics," as a functional term, was used in William Thomson's paper "An

Account of Carnot's Theory of the Motive Power of Heat" [1]. The first thermodynamic textbook was written in 1859 by William Rankine, originally trained as a physicist and a civil and mechanical engineering professor at the University of Glasgow.

The quantitative study of the relation between heat and mechanical work was undertaken further by Joule and Benjamin Thompson (also known as Count Rumford) who was astonished when he found that by immersing newly cast brass cannon into water while boring the hole in them using horse drawn power, he could actually make the water boil. He and other onlookers were astonished that they could generate heat without fire. The fact that the water would boil for as long as the horse kept turning the drill killed the earlier dominant "phlogiston" theory that heat was a substance that flowed from one object to another—because it never ran out! Great progress was made in understanding energy relations by Julius von Mayer and James Joule who measured "the mechanical equivalent of heat" by taking a pulley and rope, attaching a weight to one end and wrapping the other end of the rope around a shaft that went into an insulated water chamber where it operated a paddle wheel (Fig. 14.1). As the weight dropped (doing so many kilogram meters of mechanical work), the increase of temperature inside the chamber could be measured. By doing so, Joule found that one newton-meter of work (or 7.2 foot pounds) was equivalent to 1 joule of heat energy. We now use the joule as the preferred unit of energy. It is equal to about the energy of picking up a newspaper.

More commonly we use larger units. The kilojoule (kJ) is equal to one thousand joules. The average amount of solar energy received per second by one square meter of the Earth's surface is 239 joules (i.e., the solar constant ± the albedo (reflectance) divided by 4, the ratio of Earth's surface to Earth's cross section). Thus, one kilojoule is about the amount of solar radiation received by one square meter of the Earth in about 4 s. The megajoule (MJ) is equal to one million joules or approximately the kinetic energy of a one-ton vehicle moving at 160 km/h (100 mph). The gigajoule (GJ) is equal to one billion joules. A gigajoule is about the amount of chemical energy in 7 gallons of oil. A barrel of oil has about 6.1 GJ.

Also in England and France, another very important discovery was made in the 1770s, that of oxygen. Probably Joseph Priestly in England discovered oxygen a little earlier than did Antoine Lavoisier in France, although the latter probably understood its significance better while quantifying its abundance and reactions (Fig. 14.2). Both derived oxygen by heating oxides of mercury. Lavoisier discovered that the atmosphere contained oxygen or "eminently breathable air" by showing that an animal lived longer in a container of pure oxygen than in a container of air. He also clarified the role of oxygen in combustion and the rusting of metal and its role in animal respiration, recognizing that respiration was "slow burning." He also came up with the basis for the law of conservation of matter by showing that after a chemical reaction the elements always weighed the same as they did before the reaction.

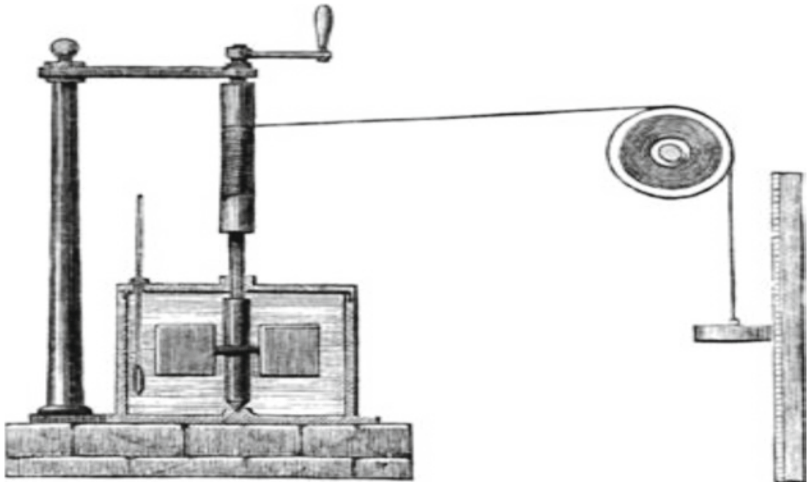


Fig. 14.1 Joule's machine for measuring the mechanical equivalent of heat or perhaps better said as the quantity of heat released per unit of mechanical work done (Source: 2009 ► citizendia.org)

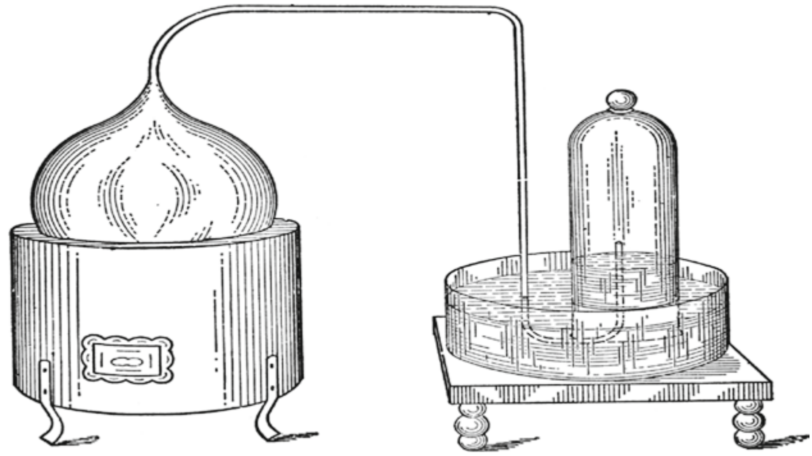


Fig. 14.2 Lavoisier's experimental approach to measuring the oxygen content of the atmosphere (Source: Florida Center for Instructional Technology)

These earlier investigators of energy turned what had been a complete mystery into a well understood and quantifiable science, and we owe a great deal to their work. Except for Albert Einstein's discovery of the equation for turning mass into energy (and vice versa, as in the Big Bang) and the development of quantum physics, there has, arguably perhaps, not been any comparable discoveries of the basic physics of energy, especially that can be represented readily by simple equations. However, as we shall see, perhaps the most important discoveries came with applying basic energy laws and ideas to more complex systems, including ecology and economics.

14

14.5 What Is Energy?

A definition of energy turns out to be more difficult than what one might think. The high school physics definition "the ability to do work" does not take us very far. Robert Romer wrote a good physics textbook which was about using energy concepts to understand all the conventional material of physics because "all physics is about energy." Yet even he admitted that he was unable to give a satisfactory definition of energy. He said we can see its effects, we can measure it, but we don't really know what it is. Usually we detect energy being used because something is moved, a car, a basketball player, chemicals against a gradient, and so on. For our day-to-day purposes, energy is mostly either photons coming from the sun or chemically reduced (i.e., normally, hydrogen-rich) materials such as wood or oil that can be oxidized to generate work (i.e., move something) at some point in space and time, i.e., the process of combustion. In general, *reduced* means hydrogen rich and oxygen poor, so that a fuel is generally a hydrocarbon like oil or occasionally a carbohydrate such as alcohol (the "ate" on the end refers to the presence of oxygen, so that a carbohydrate will have somewhat less energy than a hydrocarbon per gram but still enough to be used as a fuel). When a reduced fuel is oxidized, energy is released, and the hydrogen is released as water (H_2O) and the carbon as carbon dioxide (CO_2). The general equation for combustion of a hydrocarbon is:

$$C_nH_{2n} + O_2 \rightarrow H_2O + CO_2$$

The exact numbers required to balance the equation depend upon the exact form of the hydrocarbon burned but are for oxidation of common biological foods about:

$$C_6H_{12}O_6 + 6\,O_2 \rightarrow 6\,CO_2 + 6\,H_2O + \text{energy}$$

The equation for photosynthesis is the same but runs from right to left.

Most of our energy comes into the Earth originally in the form of photon flux from the sun. Some small part of this energy is captured by plants in chemical bonds and then passed through food chains. Thus, we are able to use the energy in a hamburger by oxidizing the reduced matter in the animal tissue, through digestion within the body of the consumer. This energy initially was obtained by the cow when it ate grass that had in turn captured that energy from the photons, and then passed it as chemical bonds to the cow and then to us. Even when we are driving a car, we are oxidizing formerly reduced plant material (oil) that is constructed of high-energy chemical bonds originally made with energy captured from the sun by algae. All life is powered by organisms capturing energy by photosynthesis (where electrons are activated) or eating energy that originally came from the sun and passing those activated electrons along trophic (food) pathways, using some of that energy to run life processes, to a terminal electron acceptor, usually oxygen. It is analogous to electrons activated by a generator or a battery running along wires to a terminal electron acceptor (the ground or the pole of the battery).

Power refers to the rate at which energy is used. For example, a light bulb is rated in kilowatts, a unit of power, so that a 100 watt light bulb uses 360 kilojoules in an hour, equivalent to the energy in about 10 milliliters of oil. An automobile engine is rated in horsepower, roughly the rate at which a horse can do work, which was used to estimate the power of early steam engines. Since automobiles today typically have 100–200 horsepower engines, one can see how much fossil fueled engines have increased the ability of humans to do work (Table 6.2). If we want to know the total energy used, we multiply a measure of power (e.g., 100 watts) times the time of use (say 10 hours) to get the total energy use, in this case 1000 watt-hours or 1 kilowatt-hour).

The use of different terms to describe energy (e.g., calories, kcal, BTU, watt, joule, therm) may seem very confusing, but they all measure one thing: the quantity or rate of heat produced when all of the energy has been converted to heat. Table 14.1 gives many energy conversions as well as the metric prefixes that establish magnitude.

Table 14.1 Energy conversions as well as the metric prefixes that establish magnitude.

One calorie	4.1868 joule
One BTU	1.055 KJ
One KWh	3.6 MJ
One therm	105.5 MJ
One liter of oil	37.8 MJ
One gallon of oil	145.66 MJ
One barrel of oil	6.118 GJ
One ton of oil	41.9 GJ = 6.84 barrels

14.6 Quality of Energy

When considering energy as a resource in a general way, there are several critical things to think about. First of all, there is the *quantity* of it, how much there is at the disposal of the species or human society using it. For example, there is several times more coal in the world compared to oil. Second is the *quality* of that energy: that is, the form that it is in, which has a great deal to say about that energy's utility. The most obvious example is food. The energy in corn has obvious utility to us as food where the energy in wood or coal does not. There are many other aspects of quality. Corn, a grass, is a very productive crop so where the land is crowded people often eat nothing but corn (or other grasses such as wheat or rice) because it gives the most food production per hectare. But corn lacks a critical factor absolutely required for humans: the amino acid lysine. If the corn is fed to a cow, then the energy bonds in the corn will be transferred to energy bonds in the flesh of the cow. This animal protein has a full complement of amino acids and hence is a higher-quality food, at least from that perspective. Many relatively poor humans in Latin America (and elsewhere) eat mostly rice and beans. This is actually a very good diet because the rice and beans are cheap and they complement each other: the amino acid lysine is missing in rice but found abundantly in beans, while rice is basically carbohydrates, a good energy source, and beans are protein rich. Thus rice and beans provides an excellent diet for humans, although it is still missing one critical ingredient: vitamin C. Fortunately vitamin C is abundant in chile peppers, which is often used as a condiment by people who have a rice and bean diet. So cultural selection appears to be often associated with real dietary needs, all of which insures that the energy that fuels humans has the required quality.

We often say that the energy in the protein-rich beans, or a chicken that is fed rice, is of a higher *quality* than the rice because the animal food contains more protein, a food type absolutely necessary for humans and most animals that is in insufficient supply in many plant foods. Many would say it tastes better too. Thus people may feed rice or other grain to an animal to get a smaller quantity of higher-quality chicken. Likewise coal or oil can be burned to generate a smaller quantity (as measured by heating ability) of electricity. But this electricity has a higher quality in that it can be used to do things such as light a light bulb or run a computer that one cannot do with the oil or coal. We are willing to take roughly three heat units of coal or oil and turn it into one heat unit of electricity because it is more useful to us, that is, it can do more work and hence is more economic, in that form. We say the quality of the electricity is higher, and a special term, called *emergy*, has been derived to represent quality of energy in a comprehensive fashion [2].

A related aspect of energy is its ability to do work defined by physicists in a very careful, specific way. The term used here is exergy, which is that component of energy that can actually do the work, as opposed to being transferred into heat due to the minimal second law requirements for some to be turned into heat. In formal second law analysis and technical thermodynamics in physics and certain engineering, the terms exergy and enthalpy are used to measure quality [3].

There is a third component of energy, also related to its quality, which relates to the energy required to get that fuel. We normally measure this property as EROI, or energy return on investment, and this issue will be explored in much greater detail in ▶ Chap. 18. We often hear very bullish statements about the tremendous amounts of energy that are all around us just waiting for us to exploit. But there is a catch. The energy has to be of a high enough quality to make it worthwhile to exploit, and real fuels must have a very high EROI. For example, we normally can get only about a third of the energy out of an oil field simply because the remaining oil sticks tightly to the substrate. If we really wanted that oil, we could get it—we could dig a 2 mile deep hole and shovel it out of the ground and heat it in a giant pot. But obviously that would require far more energy than one would get from the oil. In fact we use steam, pressure chemicals and pumping, and to some degree it works. But at some point, getting more of the remaining oil out simply costs too much money for the energy to do it, and the well is closed off. Reduced carbon, a potential fuel, is extremely abundant in shale rocks throughout the world, and as such it represents, some say, a tremendous energy source. In certain very carbon-rich rocks, it is possible to get oil or gas out with a substantial energy profit. But for the majority of these rocks, more energy would be required to get this dilute carbon out of the rocks than the energy contained within it, so that rock cannot be considered a fuel. Similarly, the oceans contain a tremendous energy potential in the hydrogen found in the water molecule. But that hydrogen is not a fuel, for it takes more energy to separate it from the oxygen it is combined with than can be recovered by later burning it. EROI comes into play more generally when we examine our commonly used fuels. For example, the petroleum that flowed out of Spindletop probably had an EROI of far greater than 100:1. The EROI of all oil and gas production was initially low, about 20:1, then increased to about 30:1 in the 1950s, and then has declined to about 10:1 today. Finding and developing a brand new barrel of petroleum today (vs pumping out an existing stock) require perhaps one barrel for each three to five barrels found. Similar patterns have held for other fuels, such as for coal, over time, although for coal the numbers, although decreasing, are much higher. Thus in general as time goes by, the highest-quality fuels are used first, and the EROI declines. While it is true that occasionally brand new very high-grade petroleum resources are found, the probability for most of our main resources is vanishingly small because, according to Colin Campbell, the whole world has been seismically and otherwise explored and picked over for many decades.

Similarly we have used up our highest-grade copper ores, so that the average grade mined fell from about 4% in 1900 to 0.4% in 2000. This lower-grade copper requires more energy to get a kg of pure copper out, and we can say that its RoE ([material] return on energy investment) is declining. Humans, usually being no economic fools, tend to use high-grade resources first, high-grade meaning more concentrated or easier to access. This important concept is called the *best first principle*, and it is very important as we consider the possibilities before us. The principle was also derived very clearly in economics by David Ricardo two centuries ago.

14.7 What Are Fuels?

Fuels are normally energy-rich, reduced compounds of hydrogen and carbon which we call carbohydrates if they also contain some oxygen or hydrocarbons if they do not. We often think of fuels as energy *carriers*, for they store and allow energy to be moved from its source to where we wish to use it. Oil in the ground or even in the gas tank is not useful. Rather it becomes useful when it releases energy in the process of transfer from the reduced state to the oxidized state. Thus the utility of a fuel depends upon having a redox (reducing-oxidizing) gradient between the fuel and the final electron acceptor. A key to the way that organisms have evolved is that life has tended to break this process down into a series of tiny steps that captures or releases some of this energy step by step. Thus electrons are passed through energy capture devices, such as a membrane or a whole plethora of oxidized-reduced chemical compounds, which cycle from energy-rich reduced forms to energy-poor oxidized states and, as appropriate, the converse. In a way this flow of energy through biological food chains is not unlike the flow of electrons in a wire that we call electricity. Some energy sources, the sun for biology or a generator fueled by falling water or the combustion of fossil fuel, gives the electrons a boost, a kick in the pants as it were. In electricity the wire provides a circuit for the electrons to travel along, and the energy represented by their excited state can be used by a device such as a motor or light bulb put in the path way of the electrons flowing from the source to what we call a sink, which represents a place that the low-energy electrons can return, generally to be kicked into a high-energy state again. The energy provided by the kick is simply moved to the place where it is utilized in a light, motor or whatever. Similarly electrons that have received a "kick" from the sun in photosynthesis pass through the complex "wires" of biological circuits carrying the energy derived from the photon to reduced carbon compounds in a plant and then through food chains to various animals and decomposers. So when you eat your corn flakes or a hamburger, remember that the energy that allows you to run, jump, or just exist came from the sun through the magic of photosynthesis.

14.8 Why Energy Is So Important: Fighting Entropy

When we think about energy, it is normally from the perspective of going somewhere, or keeping warm in the winter, or some friend's high-energy level. But the reach of energy is far more pervasive. The principal reason is due to what is normally called *entropy*. Entropy is often used inaccurately or vaguely. Physically it describes, essentially, the tendency of the components of a physical system to spread as evenly as possible in space and over all states of motion. Entropy is the physical measure of disorder, that is, randomness. The concept of molecules arranged in a definite pattern (such as in a building or an animal) is the opposite of those molecules being spread out randomly. The natural tendency is for molecules to be arranged randomly, that is, to have high entropy. Some have called this property "the entropy law." While this concept may seem far removed from our day-to-day existence where we live surrounded by ordered structures (such as in the computer, I am using as I write this), it is in fact critical, for everything with which we deal is impacted by entropy, and everything that we own tends to degrade (i.e., become more random) over time: our cars (that's why we need to take it to the shop), our homes (that is why we need fire insurance, repainting, plumbers, termite controllers, and so on), our food (that is why we need refrigerators), our closets, and even ourselves (that is why we need to eat and why most of us require medical intervention at various times in our life). What all these things are—cars, houses, computers, and ourselves—is bits of negentropy, or negative entropy, that is an ordered structure of molecules, something that is by itself extremely unlikely. Life must be nonrandom to exist, that is, life consists of very specific aggregates of molecules that are completely different from the general environment within which it resides. But although by chance alone negentropy is extremely unlikely, in fact it is common around us, and this is due principally to natural selection that has generated life plans that extract energy from the environment and invests that energy into creating, maintaining, and reproducing life forms.

Thus the creation of negentropy requires energy to concentrate and organize molecules as well as a plan as to what reorganization will work. For life, the plan is a species's DNA, and analogously for a mechanic or plumber, it is the wiring or piping diagrams, shop manuals and so on, or his or her training and experience that allows the car or house to continue to exist. But without energy the plan is useless for it requires energy to take metals or other materials out of the ground and air to make new biomass or new brake drums or pads, cylinder blocks, pipes, faucets, and so on.

It even takes our personal energy expenditure to reduce the daily entropy of our closets. More generally life, including civilization, is about very specific structures, or construction according to a plan, and then the maintenance of that structure. Both of these things are energy demanding, the degree of which is a function of the complexity, size, and makeup of the plan. That is why we eat, why plants photosynthesize, and why modern civilization requires coal, gas, and oil: to get the energy necessary to maintain and in some cases build the very specific structures that we are and that characterizes all life and also our economies. An organism's DNA gives it the pattern or plan for the very specific structures, physiologies, and behaviors that have worked well in the environment—in which it is found—or at least have worked well up to the present. Those patterns that did work in the past may or may not work in the future, depending upon whether there are environmental changes or whether some other species has figured out a new way to exploit that environment. But all organisms are, in a sense, betting that what they have will work well enough for what life is all about—propelling genes into the future. It is a wonderful process, and the results are magnificent!

A simple example will help to think about this. Both a ham sandwich and your own self are extremely unlikely, nonrandom structures of molecules of carbon, nitrogen, phosphorus, and so on that have been developed by taking the elements and materials of nature, initially scattered more or less at random over the surface of the Earth, and concentrating these elements and their compounds into structures that would be extremely unlikely—except for the investment of energy into a plan—a wheat plant, a pig, and ourselves for starters and then additionally all that goes into a ham sandwich. Once the structure is made, energy must be continuously invested, or the materials of which it is composed will go back on their own toward entropy—i.e., a more random assemblage—and the structure will fall apart. A simple example is your ham sandwich. If that sandwich is put into a refrigerator, a device that uses energy to maintain the structures of its contents, the integrity of the sandwich will be maintained for some time. Pull the plug on the refrigerator (i.e., cut off the energy), and the sandwich begins to go into a more random assemblage, first smelly organic residues and then eventually carbon dioxide and simple nitrogen compounds such as ammonia. Pull the energy plug on yourself by not eating, and the same will happen to you, eventually. Likewise a car will not run without both fuel and the energy required for its repair; a city cannot run without its fuel supplies, power plants and many kinds of repair personnel, or an entire civilization without all of these things, which must be supplied essentially daily. Most past civilizations that have lost their main energy supplies became extinct, as we will develop later.

The practical meaning of this is that it is always necessary to find new energy resources to construct and maintain whatever structures we have, including houses, cars, civilizations, and ourselves. This is familiar to us in the shop costs, medical bills, and taxes we must pay to maintain our cars, ourselves, and our roads and bridges against the entropic forces of nature that would otherwise result in time in cracked and broken roads and bridges rusted to pieces. Curiously it is necessary to generate additional entropy to maintain areas of negentropy. The refrigerator must take high-grade electricity and turn it onto lower-grade (more entropic) heat in order to maintain the ham sandwich in its desired configuration, and each of us must take low-entropy food and turn it into high-entropy heat and waste products in order to maintain ourselves. Even the creation of this book, which hopefully represents highly ordered information, requires the generation of excess entropy around us, as a look at either of our offices will confirm.

14

14.9 Laws of Thermodynamics

Thermo means heat (or energy), and *dynamics* means changes. *Thermodynamics* is the study of the transformations that takes place as energy or fuels are used to do work. *Work* means that something is moved, including a rock or your leg lifted, a car driven, water evaporated or lifted up in the atmosphere, chemicals concentrated, or carbon dioxide transformed from the atmosphere into a green plant. There are two principle laws of thermodynamics, called the *first law of thermodynamics* and the *second law of thermodynamics*. Quite simply the first law says that energy (or for some particular considerations energy matter) can never be created nor destroyed but only changed in form. Thus the potential energy once found in a gallon of gasoline but then used to drive a car, say, 20 miles up a hill, is still found somewhere, as the momentum of the car, as heat dissipated by the radiator or where the tires met the road, or in the increased potential energy of the car at the top of the hill. Most of the original energy will be found as heat dissipated into the environment, where it is essentially impossible to get any additional work out of it. (Technically you could capture that waste heat and use some of it, but it would require the use of even more energy to do so.) But some fraction of the work done can be used again, for example, the automobile could be rolled back to its original downhill position using the force of gravity. The second law says that all real-life processes produce entropy. At every energy transformation, some of the initial high-grade energy (i.e., energy that has potential to do work) will be changed into low-grade heat barely above the temperature of the surrounding environment. In other words, the first law says that the *quantity* of energy always remains constant, but the second law says that the *quality* is degraded over time. The practical meaning of this is that with the exception of the reliable energy input from the sun, it is always necessary to find new energy resources to construct and maintain whatever structures we have, including houses, cars, civilizations, and ourselves. The implications of this have had overwhelming impacts upon all human enterprises and histories and constitute the remainder of this book.

To our knowledge, there are no examples of any action occurring on earth, or anywhere else for that matter, that is not subject to the laws of thermodynamics. The only possible exception, given in the first part of this chapter, is that that the law of conservation of energy needs to be expanded to the law of conservation of mass energy when nuclear reactions (in a star, nuclear bomb, or nuclear power plant) are considered. This is because mass can be converted to energy (and the converse) according to Einstein's famous equation: $E = MC^2$, which says that under special circumstances energy created equals mass times the speed of light squared. In other words in a nuclear conversion, a small amount of mass can be converted to a huge amount of energy, although this can take place only under very special conditions. This is an example of how science often moves forward. The first law of thermodynamics might seem to have been violated when we learned about nuclear reactions, but with Einstein's help, we only had to understand that while the first law works very well for everyday conditions, we have to expand it to include mass for the very special circumstances of a star, that is, we need to learn how to expand our law.

14.10 Types of Energy

Although energy is critical to all of our daily activities an actual definition, as we have said, is hard to come by. Energy is usually defined as the capacity to do work, where work implies something is moved (a rock or animal from here to there, chemicals concentrated, and so on. The most important routine work activities that take place within the human realm are driven by the sun (solar energies). These activities include the evaporation and lifting of water from the sea to provide us with rains and rivers that flow from mountains, the concentration of low-energy carbon from the atmosphere into higher-energy tissues of a plant through photosynthesis, the passing of this energy through food chains (e.g., with a deer or a cow eating plants), the generation of winds that moves atmospheric water from the ocean to the land and cleanse the local skies of pollutants, the generation of soils through complex processes of forests and grasslands, the running of many complex processes in natural ecosystems, and so on (Fig. 14.3). Increasingly we also use fossil (meaning old) fuels such as coal, oil, and natural gas. Energy that is being used at the time in question to undertake work is called *kinetic* energy, and energy that has the possibility to do work but that is not doing it now is called *potential* energy. Examples include a rock at the top of a hill, the energy in a pile of firewood, the concentrated energy within a flashlight battery not being used, and the chemical energy in a gallon of gasoline sitting in a gas tank. When the gasoline is used to move a car, the potential energy of the gasoline is changed into the kinetic energy of the automobile in motion and into heat. Most energy that we use is derived directly from the sun either at present (i.e., wind, the power of dry air to evaporate, and so on) or in the past (i.e., the gasoline came from petroleum that was once solar energy captured by small plants (phytoplankton) drifting in the sea. Other sources of energy besides the sun include the energy of planetary motions

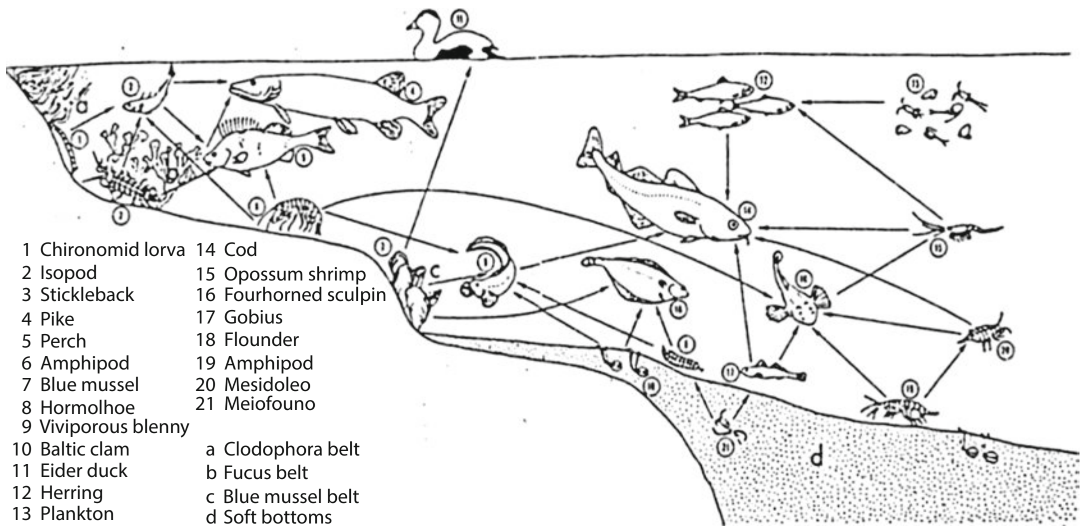


Fig. 14.3 Energy flow through a Baltic ecosystem. The energy that enters from the sun is captured by green plants and then is passed to herbivores and then carnivores through food chains (sometimes called food webs) (Source: Bengt-Owe Jansson)

(which causes tides), geological processes such as volcanoes and crustal movements and that of nuclear decay (which causes the interior of the earth to be hot).

Solar energy is especially important as it runs the whole "heat engine" of the earth (see next chapter). It also runs local weather. For example, when steady winds are forced upward by a mountain in their path, the air masses cool, generating a rainy region on the windward side (think Seattle, Washington) and a dry or even desert area on the leeward side (think Yakima, Washington). Thus the unequal interception of solar energy on different parts of the Earth's surface generates the world's winds, its wet and dry areas and, more generally its climatic zones. Solar energy also evaporates water from the surface of the ocean, lifting and purifying it in the process, moves it onto land masses while causing it to rain as solar-powered winds push the air masses up mountains, and in so doing generating the world's rains and rivers. While we may not appreciate a particular rainy day, the rains are essential to our purified water supplies and the growth of plants upon which all animal life, including our own, depends. An understanding and appreciation of the world's hydrological cycle and the critical role of energy in it is perhaps one of the most fundamental things we can learn about how the Earth, and hence our economy, operates. Curiously this process is not considered a part of most economics even though it is probably the most important step in the world economy, that is, the purifying of water and the lifting of it to the land and to the mountains that supply most of the world with its water for agriculture, for all economic activity, and for life itself. It is not considered by conventional economics because it is free, i.e., it does not enter into markets. But being free and indispensable makes it more, not less, valuable to our economy, and we need to think of it that way especially as we must pay more to compensate for the pollution and other abuse of water that is increasingly part of the hydrological cycle.

14.11 Energy and Life in More Detail…

Life, in all of its manifestations, runs principally on contemporary sunlight that enters the top of our atmosphere at approximately 1400 watts (1.4 kilowatts or 5.04 MJ per hour) per square meter for a point perpendicular to the sun's rays. Roughly one quarter of that amount reaches the Earth's surface. This sunlight does the enormous amount of work that is the thermodynamic consequence of this energy input and that is necessary for all life, including human life even when isolated from nature in cities and buildings. The principal work that this sunlight does on the Earth's surface is to evaporate water from that surface (evaporation) or from plant tissues (transpiration) which in turn generates elevated water that falls eventually back on the Earth's surface as rain, especially at higher elevations. The rain in turn generates rivers, lakes, and estuaries and provides water that nurtures plants and animals. Differential heating of the Earth's surface generates winds that cycle the evaporated water around the world, and sunlight of course maintains habitable temperatures and is the basis for photosynthesis in both natural and human-dominated ecosystems. These basic resources have barely changed since the evolution of humans (except for the impacts of the ice ages) so that preindustrial humans were essentially dependent upon this limited, or perhaps more accurately diffuse, although predictable energy base.

In *photosynthesis* energy from the sun is captured by green plants using chlorophyll, a very special compound similar in structure to the hemoglobin in our blood. Chlorophyll appears green to our eyes because it uses (i.e., absorbs) the shorter red and longer blue wavelengths from the sun and reflects back the green wavelengths that it does not use. A thick layer of green plants cover the earth wherever temperatures are moderate and water is abundant The amount of energy trapped by photosynthesis is immense, roughly 3000 exajoules per year, which is about six times larger than the energy use of all human activities (488 exajoules per year). The first step occurs in the center of the chlorophyll molecule where electrons circling the magnesium-nitrogen compound in the center of the molecule are "hit" by a photon from the sun and "pushed" into a larger orbit, which allows them to store more energy and then pass it to special chemical compounds. This is similar to how a professional skater stores energy in her outstretched arms when her partner gives her a well-aimed push, and then uses that stored energy to speed up her spin by pulling her arms back to her sides. Free electrons are normally made available from reduced compounds and move through biological circuits to fuel biotic processes. That energy is first stored temporarily in reduced compounds in plants such as NADP, which are then used to split water to get hydrogen and an excited electron, and CO_2 to get carbon. Plants then combine the carbon and hydrogen to make reduced, energy-rich compounds such as sugar. Eventually, the electron is passed to an electron acceptor, normally oxygen, but occasionally sulfur or some other element. These electrons are reenergized when green plants give a new kick to the electrons when a photon from the sun again drives photosynthesis. And hence the process continues, with the energy from solar-derived photons driving every biological activity including the movement of my fingertips on this keyboard. It is incredible!

The chemistry of photosynthesis is based on the energy from photons being used to split carbon dioxide and also water to get or *fix* reduced carbon and hydrogen which is then used to generate sugars with oxygen as a waste product:

$$6\,CO_2 + 6\,H_2O \rightarrow C_6H_{12}O_6 + 6\,O_2$$

The sugars are then synthesized into the more complex compounds of life. These include cellulose (the basic structural material of wood, which is just a lot of sugars attached one to another into a network of the same materials called a *polymer*) and, with the addition of nitrogen, the proteins of animal and many plant tissues. This same equation is "run backward" by animals and decomposers that use the chemical compounds. When green plants first evolved, some three billion years ago, and especially one billion years ago when plants colonized the land, they changed the atmosphere from an anaerobic one to an aerobic one. This can be seen in, for example, the rocks of Glacier National Park where there are green layers of iron containing rock that were laid down before the oxygenation of the atmosphere and similar "rusted" red rocks that were laid down later after the evolution of an oxygenated atmosphere.

What about animals? Take a look at most wild or domestic animals. Usually they are eating, i.e., getting energy, or trying to position themselves to do so (Fig. 14.4). If they are not eating, they tend to be resting, conserving energy. In the breeding season, obviously things get a bit more complicated. Plants too are spending most of their time dealing with energy: for example, they are photosynthesizing any time the sun is shining and in various ways attempting to protect themselves from energy losses by, e.g., making natural pesticides. Humans are a bit different because food energy is (at this time in our history) so abundant and cheap, at least for the richer half of humanity, that we have to invest relatively little time or

Fig. 14.4 Herbivores grazing in Kenya (Source: Kathy Wooster)

personal energy to feed ourselves. We are also different now because our energy requirements are only about half of what they were when we were more active. For example, early New England farmers had to eat (and drink, especially ale!) about 7000 kilocalories (30 MJ) each day to fuel their hard agricultural work, although many hard workers in poorer countries get by on less than half that. Any of us today who ate that many calories would become huge!

The study of biology, from biochemistry to ecosystem biology, is very much about the study of how energy is passed from one chemical entity to another. Biochemists often focus on the importance of the energy storage materials NADPH and ATP, scientists who study at the level of one organism often consider feeding behaviors, and the physiology of energy transfer within and across the gut wall, whereas ecosystem biologists talk about the transfer of energy from plants to herbivores to predators. The importance of energy in biotic function has captured the attention of many of our great biological thinkers, for example, Alfred Lotka, Harold Morowitz, Max Kleiber, Howard Odum, and others. And what do they conclude? Basically, that life, or more specifically the individual organisms and species that constitute the packages of life, is about capturing as much energy as possible per unit time with as little expenditure or investment as possible per unit gained, using that net energy gained to sequester more energy and other resources, to use to create structures and fuel behaviors to propel their genes into the future.

14

As far as we know, this is entirely the result of the uncaring processes of natural selection, those organisms and, ultimately, genes that were successful at this pattern were those that tended to survive, prosper, and eventually be relatively dominant on the Earth's surface. Some people prefer to use the more general term "resources" rather than just "energy resources" when discussing these issues, and there is occasionally a good case to be made for that. Obviously water is a critical resource for plant growth, and all the solar energy a plant could ask for might be available in an Arizona desert although water is very much limiting. In other situations some specific nutrient, such as phosphorus or nitrogen, may be limiting, but even these limitations can be mitigated by the plant investing more of its energy into growing longer roots to exploit more soil or transferring molecules across fine roots. Thus for most of the earth, the critical issue is energy, and life seems to be very good at expanding to capture as much of the available energy around it as possible.

Two important concepts here are *energy investments* and *energy opportunity costs*. The former means that life must always invest energy into fighting entropy, getting other resources, reproducing, and so on. The second means that since every organism has only a limited energy supply at any one time, and any particular investment into one process means that there is that much less energy to invest elsewhere. If a tree invests more energy in growing long roots to get more water or nutrients, then less will be available for growing tall, and it might be shaded by a competitor or eaten by an insect. If more energy is diverted into making natural pesticides (such as caffeine, mustard oils, or various alkaloids) then less is available for growing roots, and so on. Likewise if a civilization invests more energy into military activities, or expanding office space, or building fancy homes, or looking for oil, then less energy is left for repairing bridges or education. Politics is all about how to make energy investment decisions, although it is done through deciding where to spend money. In both trees and politics, there is a tendency to invest in a way that can capture more energy (through plant or economic growth), but that can work only when there are additional energy resources that can be exploited by using less energy than that gained. If the energy resources become restricted, then investing in growth can be self-defeating, a situation that many of the world's economies now face.

14.12 Energy Storage

Life is of course about much more than simply gaining and using energy, for life must use energy when and where it needs it in order to help the organism adjust to a continuously changing environment. Just as a motor or light needs a switch to turn it off and on as needed, life too must have switches. As a simple example, if our muscles were firing all the time, they would be useless, and in fact such a condition is a pathology called tetanus. Thus life has evolved a whole series of complex controls and switches that use available energy from the sun or food a little at a time, storing it and releasing it as needed and as controlled by hormones and the nervous system operating through very complex biochemistry. The general solution that has evolved for the storage and on/off problem has been through the use of various storage reservoirs. This allows for the capture, storage, transport, and release of the energy made available to the organism by photosynthesis or by ingesting food. The most common such compound for *short-term storage* is adenosine triphosphate (ATP) and its less energized form ADP. Whenever the body needs energy, quickly it calls on ATP to deliver that. These compounds are ubiquitous to life and are critical to all activities that an organism does. *Medium-term storage* is the glycogen in our liver, and *longer-term storage* is all too familiar to us as body fat.

14.13 A Big Jump in the Earth's Energy Supplies for Life

Thus free oxygen increased with the first massive increase in land plants as a waste product of their photosynthesis that split water and carbon dioxide to generate the carbon and hydrogen needed to produce reduced carbohydrates such as sugars. For all the existing plants and animals and microbes on the Earth this free oxygen, itself extremely reactive, was initially a severe toxic threat, a widespread and dangerous pollutant. Some say that the evolution of oxygen-releasing green plants was the greatest environmental impact the earth has ever faced! Some have argued that the mitochondria were initially evolved (or as we said above "captured") to sequester the dangerous oxygen before it destroyed other parts of the organism, and only later developed the capacity to enhance the metabolic activity of the host cell. Over time natural selection created organisms (including humans) with protective skins that require oxygen to live and to use completely their food. But even today there are many environments where oxygen is not present. They are normally obvious to us from their smell of hydrogen sulfide, characteristic of, for example, the mud of a marsh or the inside of our intestines, which would not be a good place to have oxygen for then the energy in our food would be used up before it got to do us any good! In these environments oxygen remains a poison for many of the organisms.

Thus, it appears that evolution has operated in many complex ways, such as incorporating oxygen-using organelles (i.e., mitochondria) in all animals that live in an oxygen-rich environment, to derive means of using energy more powerfully. Apparently the main ways to do this were worked out very long ago in the evolution of life since nearly all life has the same internal energy structures and uses the same basic phosphorus-based chemistry for storage and quick release. Biochemist Paul Falkowski makes an elegant argument that in many ways the biochemistry that life depends upon now is inappropriate for our existing oxidized environment. It can be understood only as a "holdover" from life's anaerobic past—that is, the anaerobic mechanisms that worked in the past were too deeply engrained into the processes of life for life to abandon, and so were retained and modified even if not perfectly suited for the new aerobic environment. Although complete oxidation of food using mitochondria allows the most complete use of food, many different approaches to utilizing food energy have evolved, and these different pathways are still used variously by different species and in response to different environmental conditions. If oxygen is not present, the less thorough but quite adequate energy release process called fermentation still can be used, and this process generates energy-intermediate alcoholic residues which we have exploited to generate beer and wine. The partial transformation of the grain or fruit into usable energy leaves as residues alcohol and CO_2, which generates the bubbles in beer.

A more general perspective is that energy is passed through and among organisms in a series of complex *redox* (reducing-oxidizing) reactions, until the full food value is extracted and some or all of the carbon base is turned into CO_2. Energy is passed from one organism to another through an ecosystem along food chains and food webs. Plants capture the energy from the sun and turn a portion of it into their own tissues, leaves, stems, roots, and so on. Then some of that energy is passed to herbivores (plant eaters) and then carnivores (meat eaters) and decomposers. The word trophic means food, and trophic dynamics is the study within ecology of how energy is passed along food chains within an ecosystem and what happens to that energy. An important thing that happens is that energy is lost (actually turned into heat) at every step as necessitated by the second law of thermodynamics. Most of the energy that is lost was actually used by the organism itself for its own maintenance metabolism. This is due to the necessity for each organism to "fight entropy" through energy investments, and the necessary losses to heat arising from the second law of thermodynamics. Usually only a small proportion, very roughly 10%, is passed from one trophic level (such as plants) to the next (herbivores). This is one reason there are few top carnivores—if there are four or more trophic levels each passing on only 10% of the energy, then only a very small amount of the original energy captured by photosynthesis makes it to the top carnivore.

While it is obvious that an organism must get enough *quantity* of energy to maintain itself, it is also necessary for it to get sufficient *quality* of energy. Most generally the missing ingredient in vegetative material for humans or for other animals is sufficient protein. Kwashiorkor is a common disease of people on an insufficient protein diet, characterized by cinnamon-colored hair and a protruding belly, as well as many personal metabolic problems restricting the ability of people to work. Once in the 1950s well-meaning nutritionists made a large effort to increase protein production of certain groups of people who had this disease, for example, feeding existing grains to chickens and fish, to try to increase the protein available to these people. But the program backfired because the people were actually energy-starved, and their bodies were burning the proteins for fuel, not using them for structural development. In other words our bodies have an even greater need for energy than for structural building and repair. So, feeding energy-rich grains to animals to produce a smaller quantity of protein actually exacerbated the problem by reducing the energy available to the people, even though they got more protein, their desperate bodies had to use it for fuel, not maintenance or growth of new tissues! But where calories are sufficient protein is critical for normal healthy development, thus the quality of energy is often as important as the quantity. Obviously in our food, quality is a much more complex issue than simply protein or not.

Proteins are foods made of amino acids that are based upon nitrogen as well as carbon. One can think of a hamburger: the bun is a carbohydrate made of carbon, hydrogen, and oxygen, and the beef is protein, which has those elements and much nitrogen as well. While we normally think of protein as meat, there are many other sources. For example, ecologists have found that many of the animals of an estuary or a forest are dependent upon detrital food chains, that is, food that has been dead a relatively longtime before being consumed (as opposed to grazing or browsing living materials). Dead plant material is mostly carbon and as such contains little nitrogen, which is critical for the protein needs of the animals that feed upon it. But in estuaries and forest floors much of the decomposition of this material occurs by bacteria, and certain bacteria can do something that most other organisms cannot: they can fix nitrogen from the air and turn it into protein. Thus, the animals that eat microbially mediated food get much better nutrition because there tends to be more protein. This may sound repulsive to humans, but maybe it is much less so if you think about the microbially mediated foods we eat: bread, cheese, beer, wine, salami, sour cream, and so on. In fact most of our party foods are microbially mediated!

14.14 More on Energy and Evolution

Plants and animals in nature have been subjected to fierce selective pressure to do the "right thing" energetically, that is, to insure that whatever major activity that they did, and do, gained more energy than it cost and generally got a larger energy net return than alternative activities. Biology in the last century had, appropriately enough, focused mostly on fitness, that is, on the ability of organisms to survive and reproduce, in other words to propel their genes into the future. While it is a no brainer that a cheetah, for example, has to catch more energy in its prey than it takes to run it down, and considerably more to make it through lean times and also to reproduce, it took the development of double-labeled isotopes and the exquisite experimental procedures by the likes of Donald Thomas and his colleagues [4] to show how powerfully net energy controlled fitness. They studied tits (chickadees) in France and Corsica, and they found that those birds that timed their migrations, nest building, and births of their young to coincide with the seasonal availability of large caterpillars, which in turn were dependent upon the timing of the development of the oak leaves they fed upon, had a much greater surplus energy than their counterparties that missed the caterpillars. They fledged more, larger, and hence more-likely-to-survive young while also greatly increasing their own probability to return the next year to breed again. Those of their offspring that inherited the proper "calendar" for migration and nesting were in turn far more likely to have successful mating and so on. Thomas et al. also showed how the natural evolutionary pattern was being disrupted by climate change, so that the tits tended to get to their nesting sites too late to capitalize upon the caterpillars, who were emerging earlier in response to earlier leaf out. Presumably if and as climate warming continues natural selection will favor those tits which happened to have genes that told them to move north a bit earlier.

14.15 Maximum Power

Howard Odum has taken these concepts one step further by arguing that it is not just the net energy obtained but the *power*, that is, the *useful energy per unit time*, that is critical. Odum argued that there is generally a trade-off between the rate and the efficiency for any given process; that is, the more rapidly a process occurs, the lower its efficiency and vice versa. Under a given set of environmental conditions, it is not advantageous to be extremely efficient at the expense of the rate of exploitation nor to be extremely rapid at the expense of efficiency [5, 6]. For example, in a series of elegant observations and experiments, Smith and Li [5] found that a trout that feeds on drifting food in a rapidly flowing stream will acquire large amounts of food drifting by but at a low net efficiency; i.e., much of the energy surplus created by the consumption of this large amount of food is spent in muscle contraction for the trout so that it can fight the faster current. Likewise a trout in slow water can be very efficient because its swimming costs are lower, but the slower water brings with it less food, and thus the overall energy surplus will be limited by the lower rate at which food is provided. Dominant trout will pick an optimum intermediate current speed, which will result in faster growth and more offspring. Subdominant trout will be found in water moving a little faster or a little slower. In some experiments trout with no competitive power would be found drifting aimlessly in still water slowly starving to death.

Of course life in all of its diversity also has a diversity of energy lifestyles that have been selected for—sloths are just as evolutionarily successful as cheetahs, while warm-blooded animals pay for their superior ability to forage in cold weather with a higher energy cost to maintain an elevated body temperature—the list is endless.

Nevertheless each lifestyle must be able to turn in an energy profit sufficient to survive, reproduce, and make it through tough times. There are few, if any, examples of extant species that barely make an energy profit—for each has to pay for not only their maintenance metabolism but also their "depreciation" and "research and development" (i.e., evolution), just as a business must, out of current income. Thus their energy profit must be sufficient to mate, raise their young, "pay" the predators and the pathogens, and adjust to environmental change through sufficient surplus reproduction to allow evolution. Only those organisms with a sufficient net output and sufficient power (i.e., useful energy gained per time) are able to undertake this process through evolutionary time, and indeed some 99 plus percent of all species that have ever lived on the planet are no longer with us—their "technology" was not adequate, or adequately flexible, to supply sufficient net energy to balance gains against losses as their environment changed. Given losses to predation, nesting failures, and the requirements of energy for many other things, the energy surplus needs to be quite substantial for the species to survive in time.

14.16 Natural Economies

Of course in nature, plants and animals do not exist in isolation but combined in complex arrangements that we call ecosystems, tied together by the movement of energy and materials from one species to another in what is often referred to as food chains or food webs. We call the green plants that capture the solar energy *primary producers*, the animals that eat the grass *herbivores*, the animals that eat other animals *carnivores,* and so on. Eventually all plant and animal material ends up as dead organic material, often called detritus, and this material is then broken down into very simple materials or even elements by bacteria and other decomposers. We call the study of these relations *trophic* (meaning food) analysis and each successive step from the sun trophic levels. It is rather amazing to think that all the energy necessary for all the animals and all the decomposers, and even the plants at night and in the nongrowing season, comes from the photosynthesis during the daylight hours during the growing season.

We can call all of these trophic interactions collectively *natural economies*. In other words, nature too, just like human economic systems, is all about production, exchange within and between species, and eventual degradation. Of course natural ecosystems are different from modern human economies in that there is no money—but the economy exists just fine without the money, as might conceivably ours (i.e., many economies are based on barter alone). This idea that nature too has economies is a very powerful one for it allows us to focus on just what are the essential features of an economy when we strip it of the human additions of money, debt, credit, and so on.

14.17 Summary So Far: Surplus Energy and Biological Evolution

The interplay of biological evolution and surplus energy is far more general, as emphasized a half century ago by Kleiber [7], Morowitz [8], Odum [9], and others. Plants and animals are subjected to fierce selective pressure to do the "right thing" energetically; that is, to insure that whatever major activity that they undertake gains more energy than it costs and beyond that gets a larger energy net return than either alternative activities or their competitors. It is obvious that a cheetah, for example, has to catch more energy in its prey than it takes to stalk it and run it down and considerably more to make it through lean times and also to reproduce. Plants too must produce an energy surplus to supply net resources for growth and reproduction, as can be seen easily in most clearings in evergreen forests where living boughs on a tree that are in the clearing are usually lower down than they are in the more densely forested and hence shaded side of the tree. If the bough does not carry its weight energetically, that is, if its photosynthesis is not greater than the respiratory maintenance metabolism of supporting that bough, the bough will die (or perhaps even be sloughed off by the rest of the tree).

Every plant and every animal must conform to this "iron law of evolutionary energetics": if you are to survive, you must produce or capture more energy than you use to obtain it; if you are to reproduce, you must have a large surplus beyond metabolic needs; and if your species are to prosper over evolutionary time, you must have a very large surplus for the average individual to compensate for the large losses that occur to the majority of the population. In other words, every surviving individual and species needs to do things that gain more energy than they cost, and those species that are successful in an evolutionary sense are those that generate a great deal of surplus energy that allows them to become abundant and to spread.

While probably most biologists tacitly accept this law (if they have thought about the issue), it is not particularly emphasized in biological teaching. Instead biology in the last century focused mostly on fitness, that is, on the ability of organisms to propel their genes into the future through continuation and expansion of populations of species. But in fact energetics is an essential consideration as to what is, and what is not fit, and many believe that the total energy balance of an organism is the key to understanding fitness.

14.18 Energy and Economics in Early and Contemporary Human Economies

Humans are no different from the rest of nature in being completely dependent upon sunlight and food chains for our own energy requirements and nutrition, and on being part of complex interactions among very complex food chains leading to ourselves. Human populations, like those of any other species, must capture sufficient net energy to survive, reproduce, and adapt to changing conditions in the area in which they live. Humans must first feed themselves before attending to other issues. For at least 98% of the million years that we have been recognizably human, the principal technology by which we as humans have fed ourselves, that is, obtained the energy we need for life, has been that of hunting and gathering. Contemporary hunter gathers –such as the !Kung of Kalahari desert in Southern Africa that we introduced in ► Chap. 7—are probably as close to our long-term ancestors as we will be able to understand. Most hunter-gatherer humans were probably similar to other species in that their principal economic focus is on obtaining sufficient surplus energy as food gained directly from their environment. Studies by anthropologists such as Lee [10] and Rapaport [11] confirmed that indeed present-day (or at least recent) hunter-gatherers and shifting cultivators acted in ways that appeared to maximize their own energy return on investment, perhaps 10 joules returned for each one invested. Angel found that agriculture actually decreased the average physical fitness of humans [12].

Human evolution, broadening the definition to include social evolution, is different from other species, for the human brain, language and the written word have allowed for much more rapid cultural evolution. The most important of these changes, as developed in ► Chap. 6, were energy related: the development of energy-concentrating spear points and knife blades, agriculture as a means to concentrate solar energy for human use, and more recently the exploitation of wind and water power and, of course, fossil fuels. What is important from our perspective is that each of these cultural adaptations is part of a continuum in which humans invest some of their energy to increase the rate at which they exploit additional resources from nature, including both energy and nonenergy resources.

For a particularly good example, the development of agriculture allowed the redirection of the photosynthetic energy captured on the land from the many diverse species in a natural ecosystem to the few species of plants (called cultivars) that humans can and wish to eat, or to the grazing animals that humans controlled. It also allowed the development of cities, bureaucracies, hierarchies, the arts, more potent warfare, and so on—that is, all that we call civilization, as nicely developed by Jared Diamond in his book *Guns, Germs, and Steel*.

A human as a machine works at about 20% efficiency, that is, the power output of a human (i.e., his/her muscular work) is about 20% of the food energy input to that machine. Thus, over a 10-hour day, a human can deliver about one half to one horsepower hours or about 5–10% of what a horse can do (and on about 5–10% of the food) [13]. Put another way, the power output of a human at rest is about 60 watts, and at peak performance, a strong worker might generate about 300 usable watts, although that rate cannot be sustained for very long. A very strong person might be able to deliver 100 watts or 1 kilowatt-hour (3.6 MJ) over a 10 hour day. The human machine cannot deliver this power if the temperature is above 20–25 °C, so that other things being equal it is more difficult to generate surplus wealth in the hot tropics [14]. A horse can generate about 3 kilowatts. By comparison a four-cylinder standard automobile engine generates about 1000 kilowatts and a jet turbine engine about one million kilowatts. Clearly the world now has at its disposal a tremendous amount of power compared to the past (Tables 14.2, 14.3, 14.4).

Anthropologist Leslie White once noted that a bomber flying over Europe during the Second World War consumed more energy in a single flight than had been consumed by all the people

Table 14.2 The energy cost of various things. The ratio of energy to GDP changes year to year mostly as a function of inflation but also as the economy appears to becoming more efficient

U. S. approximate energy use per unit of economic activity (in 2016) when GDP was 18.569 trillion dollars:

$$\text{Energy per \$GDP} = -\frac{90.6\ \text{Exajoules}(e18)}{18.569\ \text{trillion}(e12)\text{dollars}} = \frac{4.88\ e6\ \text{joules}}{\text{dollar}} = 4.88\text{Mega joules per dollar}$$

One dollar of economic activity requires	4.88 MJ = 0.03 gallon of oil
One thousand dollars requires	4.88 GJ (0.79 BOE)
One million dollars requires	4.88 terajoules (790 BOE)
One billion dollars requires	4.88 petajoules (790 thousand BOE)
One trillion dollars requires	4.88 exajoules (790 million BOE)
18.57 trillion dollars requires	18.57 e12 * 4.88 MJ/$ = 90.6 EJ (14.8 e9 BOE)

Source: US Dept. Commerce; US EIA BOE = barrels of oil energy equivalent

Table 14.3 Selected fuels and their heat equivalents

Fuel	Heat equivalent (MJ)
Residual oil (1 barrel)	6626.5
Crude oil (1 barrel)	6163.8
Distillate oil (1 barrel)	6139.6
Gasoline (1 gallon)	131.8
Electricity (1 kilowatt-hour)	3.6
Natural gas (1 cubic foot)	1.1

Source: State of Oregon DOE

of Europe during the Paleolithic, or Old Stone Age, who existed when people lived entirely by hunting and gathering wild foods [15]. White estimated that such societies could produce only about 1/20 horsepower per person—an amount that today would not suffice for even a fleeting moment of industrial life. Over time, humans increased their control of energy through technology, although for thousands of years most of the energy used was animate—people or draft animals—and derived from recent solar energy. A second very important source of energy was from wood, which has been recounted in fascinating detail in Ponting [16], Smil [17], and especially Perlin [18]. Perlin estimates that by 1880, about 140 million cords of wood were being used in the

Table 14.4 MJ used per 2005 dollar spent in select sectors of the economy

Sector	MJ
Oil and gas field machinery and equipment	7.36
Petroleum lubricating oil and grease manufacturing	61.30
Cement manufacturing	68.4
Rolled steel shape manufacturing	15.60
Fabricated pipe and pipe fitting manufacturing	9.84
Water transportation	48.80
Other miscellaneous chemical product manufacturing	16.30
Other basic organic chemical manufacturing	21.70
Explosives manufacturing	22.70
Watch, clock, and other measuring device manufacturing	5.65
Oil and gas extraction	9.26
Drilling oil and gas wells	9.87
Support activities for oil and gas operations	6.98

Source: Economic Input-Output Life Cycle Assessment Model developed by the Green Design Institute at Carnegie Mellon University (We do not know exactly how these were calculated so are simply passing them on. We also suspect that the nominal precision used does not reflect reality) world per year. Massive areas of the Earth's surface—Peloponnesia, India, China, parts of England, and many others—have been deforested three or more times as civilizations have cut down the trees for fuel or materials, prospered from the newly cleared agricultural land and then collapsed as fuel and soil become depleted. Archeologist Joseph Tainter [19] recounts the general tendency of humans to build up civilizations of increasing reach and infrastructure and complexity that, again and again, have eventually exceeded the energy available to that society.

People have understood how to get energy from winds or from a stream for millennia to, for example, grind grain, but the technology and incentives to do so increased rapidly from about 1750 onward. Fred Cottrell [13] gives a thorough review of the importance of the increased use of energy by civilization, to which he, rightly in our opinion, attributes most other advances in civilization. Water power was especially important in, for example, New England in the early years of this country. But the real push in the development of "modern" civilization came with learning how to burn coal to do many things, but especially to make iron and to run steam locomotives. With these inventions, mostly in England in the 1800s, industrial development really took off, and this led to what most people call "the industrial revolution." It was not simply the development of the use of coal but a whole suite of financial, chemical, metallurgical, and other developments that accelerated each other and led to the enormous production of wealth that took place in England, Scotland, and Germany during the 1800s. For example, James Watt could not develop his famous steam engine until his friend William Wilkinson had perfected the iron refining and drilling technologies that allowed for the construction of the perfectly round cylinder needed for Watt's steam engine. Even their interactions required the social environment of the Scottish *enlightenment* for their ideas to evolve and to come to fruition as actual components of society. Most thinking people at that time believed that these were wonderful inventions that would finally free people from the drudgery of every day existence and allow them to build a better society through rational thinking.

At the same time, many of the English Romantic poets, notably William Wordsworth, were horrified by the smoke and grime and repetitive jobs of the industrial revolution and pined for the bucolic preindustrial England. Our societies today need such vast amounts of energy that we provide it by mining stocks of solar energy accumulated eons ago, and converted into coal, natural gas, and petroleum. Without these stocks, our populations would be much less, and we could not live as we do. Clearly the world now has at its disposal a tremendous amount of power compared to the past.

In summary, it seems obvious that both natural biological systems subject to natural selection and the cultures and civilizations that preceded our own were highly dependent upon maintaining not just a bare energy surplus from organic sources but rather a substantial energy surplus that allowed for the support of the entire system in question—whether of an evolving natural population or a civilization. Most of the earlier civilizations that left artifacts that we now visit and marvel at—pyramids, ancient cities, beautiful buildings and rooms, monuments, and so on—had to have had a huge energy surplus for this to happen, although we can hardly calculate what that was. Certainly massive works from the past represented small net surpluses from thousands or millions of people carefully organized or brutally forced to do this work. Archeologist and historian Joseph Tainter has written elegantly about the role of surplus energy in constructing and maintaining ancient empires—Mayan, Roman, and so on [19]. Tainter argues that as empires get larger, they can spend more and more energy impressing potential adversaries and that the construction of impressive capital cities in itself shows potential competitors that the empire has so much surplus wealth that it makes much more sense for them to knuckle under, become part, and pay tribute than to fight the empire. The ever-expanding frontiers, however, and the need for ever more surplus energy as the distance needed to bring in food and other resources from increasingly distant provinces, increasingly decrease the net energy delivered to the center. Eventually the empire falls in on itself and collapses from its very need for the complexity, and its energy costs, required to generate the necessary surplus energy. This has happened again and again in antiquity and more recently with the collapse of the German Third Reich, the British Empire, and the Soviet Union. An important question for today is to what degree does the critical importance of surplus energy apply to contemporary civilization with its massive, although possibly threatened, energy surpluses? At what point have we developed so much infrastructure that it requires all the surplus energy we can get just for maintenance metabolism, so that growth is impossible?

Contemporary industrial civilizations are dependent upon the sun and in addition on fossil fuels. Today fossil fuels are mined around the world, refined and sent to centers of consumption. For many industrial countries, the original sources of fossil fuels were from their own domestic resources. The United States, Mexico, and Canada are good examples. However, since many of these industrial nations have been in the energy extraction business for a long time, they tend to have both the most sophisticated technology and the most depleted fuel resources, at least relative to many countries with more recently developed fuel resources. For example, in 2010 the United States, originally endowed with some of the world's largest oil provinces, was producing only about 40% of the oil that it was in the peak year of 1970, Canada had begun a serious decline in the production of conventional oil, and Mexico in 2006 was startled to find that its giant Cantarell Field, once the world's second largest, had begun a steep decline in production at least a decade ahead of schedule. (See previous chapter for an update.) Howard Odum's "maximum power" hypothesis is a very powerful and insightful way to think about the evolution of nature and of human society. Odum explains, for example, how oil-rich nations gained ascendancy over solar-based societies—at least for as long as their oil lasted. But it also suggests that countries that waste their energy or are unable to come to grips with the finite nature of premium energy will not be selected for. A scary thought is that it does not take an enormous amount of energy to generate horrific war – all of World War II, in which more than 50 million people lost their lives and a billion more were seriously compromised, was fought on 7 billion barrels of oil, about the quantity that the United States uses in 1 year at relative peace.

Thus, as we face the inevitable contraction in the availability of our most important fuels and as the difficulties of generating alternatives at the scale required seem to mount day by day, we must face the possibility that our own economy and civilization, which is almost universally based on the concept of continual growth of just about everything, may need a massive rethinking for planning for the future—in other words a new economics. This book is meant to give you the conceptual tools to begin that process [20].

Questions

1. If energy is so important, why are most people unaware of most of the energy that they use?
2. What is meant by "the mechanical equivalent of heat"? How was this demonstrated?
3. Can you explain Carnot's equation: $W = (Ts{-}Te)/Ts$? What implications does this have for the limits with which we can turn fuel into work?
4. Why, if the amount of energy stored in the surface of the North Sea is so great, is it not possible to extract this energy for use by society?
5. What is oxygen? If oxygen is so reactive, why do we have oxygen in the atmosphere?
6. What is the law of the conservation of matter?
7. What is energy? Do you think it has been defined adequately?
8. What is combustion? Can you give an example of an equation representing combustion?
9. What is the relation between energy and power?
10. Energy is often given in different units such as therms, kilowatt-hours, joules, calories, and so on. How are these units different? How are they the same? Which unit should you use? Why?
11. Define the relation between energy quantity and energy quality. Can you give an example where it is important?
12. Explain the differences among energy, exergy, and emergy.
13. What are fuels?
14. What is entropy? What is negentropy? Can you give an example from everyday life? What is the relation between energy and entropy?

15. What is the relation between negentropy and a plan? Can you give several examples?
16. How does negentropy relate to biotic evolution?
17. Why does the maintenance of negentropy generate entropy?
18. What is the first law of thermodynamics?
19. What is the second law of thermodynamics?
20. Can you define the first and second laws of thermodynamics using the words quantity and quality?
21. What might be considered an exception to the laws of thermodynamics? In your opinion, is this really an exception?
22. What is the difference between kinetic and potential energy? How are they related?
23. How is the surface of the Earth a heat engine?
24. Give the basic equation for photosynthesis.
25. What is the relation between energy investments and energy opportunity costs?
26. What are some of the biotic chemical compounds in which energy is stored?
27. Discuss the terms aerobic and anaerobic in relation to the Earth's evolutionary history.
28. According to Paul Falkowski, why do organisms carry within them inappropriate chemistry for today's environment?
29. If a metabolic process produced alcohol or vinegar, what does this tell you about the efficiency of the use of the original plant material?
30. What does redox mean?
31. Define trophic dynamics and give an example. What does this tell us about the efficiency of ecosystem processes?
32. What element characterizes proteins and makes them different from carbohydrates?
33. Relate energy to evolution.
34. What does the maximum power principle tell us about the efficiency of a biological or physical process?
35. Does nature have economies? How so? Do you think it is accurate to describe nature in that way?
36. What is the "iron law of evolutionary energetics"?
37. Relate the principles learned in the earlier part of this chapter to human societies.
38. How did wood use precede the industrial revolution?
39. Summarize your views on how natural and human societies use energy to survive and prosper.
40. Do you think that technology will make the end of the oil era of little concern? Why or why not?

References

1. Kelvin, W.T. 1849. "An Account of Carnot's Theory of the Motive Power of Heat – With Numerical Results Deduced from Renault's Experiments on Steam". Transactions of the Edinburg Royal Society, XVI. January 2.
2. Odum, H.T. 1996. *Environmental accounting: Emergy and environmental decision making*. New York: Wiley.
3. Gaudreault, K., R.A. Fraser, and S. Murphy. 2009. The tenuous use of exergy as a measure of resource value or waste impact. *Sustainability* 1: 1444–1463.
4. Thomas, D.W., J. Blondell, P. Perret, M.M. Lambrechts, and J.R. Speakman. 2001. Energetic and fitness costs of mismatching resource supply and demand in seasonally breeding birds. *Science* 291: 2598–2600.
5. Smith, J.J., and H.W. Li. 1983. Energetic factors influencing foraging tactics in juvenile steelhead trout, *Salmo gairdneri*. In *Predators and prey in fishes*, ed. D.G. Lindquist, G.S. Helaman, and J.A. Ward, 173–180. The Hague: Dr. W. Junk Publishers.
6. Curzon, F.L., and B. Ahlborn. 1975. Efficiency of a Carnot Engine at maximum power output. *M.Phils.* 43: 22–24.
7. Kleber, M. 1962. *The fire of life: An introduction to animal energetics*. New York: Wiley.
8. Morowitz, H. 1961. *Energy flow in biology*. New York: Wiley.
9. Odum, H.T. 1972. *Environment, power and society*. New York: Wiley-Interscience.
10. Lee, R. 1969. ! Kung bushmen subsistence: An input-output analysis. In *Environment and cultural behavior; ecological studies in cultural anthropology*, ed. A.P. Vada, 47–79. Garden City, NY: Published for American Museum of Natural History by Natural History Press.
11. Rappaport, Roy A. 1967. *Pigs for the Ancestors*. New Haven: Yale University Press.
12. Angel, J.L. 1975. Paleoecology, paleo demography and health. In *Population ecology and social evolution*, ed. S. Polar, 667–679. The Hague, Mouton.
13. Cottrell, F. 1955. *Energy and Society*. New York, NY: McGraw Hill.
14. Sundberg, U., and C.R. Silversides. 1988. *Operational efficiency in forestry*. Dordrecht: Kluwer.
15. I thank Joe Tainter for bringing this quote to my attention. Leslie White was a great believer in the importance of energy for human affairs and is well worth reading today.
16. Ponting, C. 1991. *A green history of the world*. London: The Environment and the Collapse of Great Civilizations. Sinclair Stevenson.
17. Smil, V. 1994. *Energy in world history*. Boulder: Westview Press.

18. Perlin, J. 1989. *A forest journey: The role of wood in the development of civilization*, 1989. New York: W.W. Norton.
19. Tainter, J.A. 1988. *The collapse of complex societies*. Cambridge, Cambridge shire, New York: Cambridge University Press.
20. For more comprehensive treatments of energy itself and its relation to economics written by physicists see: Kummel, R. 2011. The second law of economics: Energy, entropy and the origin of wealth. Springer, New York and Ayers, R. 2017. Energy, Entropy and wealth maximization. Springer N.Y.; Hall, C.A.S. 2017. Energy Return on Investment: A unifying principle for Biology, Economics and sustainability.

The Basic Science Needed to Understand the Relation of Energy to Economics

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_15

The fundamental cause of the trouble is that in the modern world the stupid are cocksure while the intelligent are full of doubt. (Bertrand Russell).

This chapter is designed to provide, in a very basic way, enough science so that it is possible for the reader who has not had an extensive background in science, or who simply wants a review focused on the science associated with understanding real economic systems, to do so relatively easily. The contents of this chapter are divided into five main sections: understanding nature, the scientific method, the physical world, the biological world, and is economics science?

15.1 Understanding Nature

15.1.1 What Is Nature?

We start by considering what is nature and the natural world. In the most common view, nature is all of the world that is not explicitly human or human-dominated. In the lovely Rocky Mountain rural environment where this is being written, it seems obvious what nature is: go to one of the National Parks, get on a hiking trail, and hike until there seems to be little human influence. Nature clearly is the rocks, streams, clouds, and animals. But here too, it may be difficult to find pure nature, as there is usually a trail under your feet maintained by other human hikers and the Park Service; many of the plants, including lovely flowers, are introduced pest species; all of the plants are growing in an environment influenced by carbon dioxide increased by human activity; the glaciers we may look at are shrinking; and the rainbow and brown or brook trout you may see or catch were stocked from original populations in British Columbia, Europe, or the Eastern United States. On the other hand, humans are products of natural selection in natural environments, are animals just as much as deer or trout, and are limited in as many ways by their own genetic and physiological capacities as are the wild plants and animals. Humans, like other animals, can die from too much heat or cold, and they need water nearly daily and food regularly—or they die. But humans are different from most other animals in that they can modify their environment significantly. In addition, humans can adapt rapidly through cultural evolution. For the purposes of this book, we do not get very concerned about the nuances and usually say that while humans are derived from, and still part of, nature, *culture* is that which is human-dominated and *nature* is that which is not, including land, oceans, rivers and lakes, soils, rocks and mineral deposits, the natural plants and animals, and microbes of all of these places at scales from the subatomic to the universe and possibly beyond. Nature is also the natural forces that constrain all these things and allow them to operate. Humans of course have always sought to, indeed needed to, exploit nature for their own survival and, often, for the production of wealth. In order to do this, it was necessary to understand nature to some degree. So how have humans gone about understanding nature?

15.1.2 Human Explanation of Nature

Human existence has always been fraught with uncertainty and with great difficulty in being able to understand and predict events. This has been especially true with respect to our economic lives. Early humans understood nature well enough to gather the plants and hunt the animals that were necessary for them to eat and to predict usual seasonal patterns of plant growth and animal migrations, and early farmers certainly understood a lot about plants, soils, water, manure, and so on. But humans have always sought more cosmic or at least comprehensive explanations for the natural events around them and for more power in predicting or influencing whether a particular venture would be successful or not. Early Greeks and Romans, and indeed most prescientific peoples, believed that a god or whole series of gods controlled the day-to-day events in their lives, including the weather, how well their crops grew, and so on. Very often, the ancients would make some sort of a sacrifice—frequently human—as an investment to please the gods and to help insure the success of a planting, a military campaign, or whatever. Similar practices seemed to be characteristic of many other cultures around the world. These practices give humans a sense that there is something they can do to influence important events in their life. But how do we know whether these various approaches, or any others, work at a rate any better than random? In other words, nearly any human endeavor will always have some chance of succeeding and some chance of failing, independent of any divine, governmental or policy intervention, or even, perhaps, whether the endeavor itself is a particularly good idea. How can we increase our odds of getting something right? The answer is to use the *scientific method.* But first, we need to think a little more about why prediction can be so difficult, even with the scientific method.

Most of us have had both good and bad things happen to us, and frequently these have been beyond our control. Why should the events of life be such a mixed bag of successes and failures? Is it just the random or at least unpredictable nature of the universe? Perhaps it is because natural selection itself *must* be based on both failures and successes occurring. In other words, evolution must have both successes to move genes forward in time and failures to help generate the most fit. This was obvious to Charles Darwin [1]. But how can we determine when something good happened that was a result of our good decisions or actions vs. by chance alone? This is where science comes in, for it can help us to determine whether something really works or not, or works just by chance alone. Certainly science cannot resolve all issues, for example, science may have little to say about what values should be pursued by a person or a nation (although it can help in understanding the effects of implementing certain values), but we do believe that the domain of science can and should be expanded, and this includes into economics and indeed the general understanding of our lives.

15

15.1.3 Cause and Effect

Normally in science we seek *cause* and an *effect* and reasons for their linkages. So if we observe an effect, such as an apple falling from a tree, we ask, as did the great early physicist Isaac Newton, "why"? Newton determined that it was the attraction of the Earth to the apple, and the apple to the Earth, that caused this to happen and expressed this idea in beautiful and elegantly simple mathematics: the force between two objects was proportional to the product of their masses divided by the square of the distance between them. This simple law, which works equally for molecules and for the sun and planets in our solar system, has been verified again and again by others. We say that the force is the *independent* variable, that is, it exists whether the apple falls or not, and the falling of the apple is the *dependent* variable, that it occurs when the force is applied in the right direction and at the right distance. Likewise, in economies a dependent event (say the production of some corn) will occur only if the independent variable takes place, that is, the farmer plants the seeds. Of course, the corn production will take place only if other things occur too: the sun must shine to provide energy, rain needs to fall or irrigation water provided, there must be sufficient fertilizing elements in or applied to the soil, and so on. In this case, we would say that the production of the corn is a multiparameter issue, that is, the dependent variable occurs as a result of many independent variables. The various independent variables in turn may be a consequence of other independent variables, such as climate change or a farmer's economic ability to provide fertilizer or willingness to work hard. These factors operating together form a *system*, that is, a series of interconnected causes and effects. Thus, unraveling economic cause and effect is not always easy. This is why we advocate in later chapters a *systems approach* to understanding real energy and economic issues. This may seem impossibly complex to the reader now, but in fact with proper training is quite manageable.

The degree to which energy studies should be based in science has rarely been questioned, as energy analysis in many respects forms the basis of science. In addition, most aspects of energy seem to follow known scientific laws. An important question, however, one to which we have no easy answer, is to what degree *economics* should be a science. While economics is usually identified as a social science, the degree to which its basic assumptions are given using, and subjected to, the scientific method is not quite so clear. Introductory economics books don't put forth their fundamental economic principles as hypotheses to be tested but as truisms to be learned. In addition, there is usually no particular effort to ask, as we do here, whether or to what degree economic principles are consistent with the basic scientific laws. The reason that these issues are important in economics is that real economic systems must operate in the real material world where the laws of science always apply, regardless of whether we or some economist might wish them not to.

15.2 The Scientific Method

15.2.1 Formalizing Our Search for Truth Amidst Uncertainty

How do humans get to know things? How can we know things for sure? The answer is partly that there is no way that we can know anything absolutely for sure, and a common aphorism is that those people who know little often know it with certitude, while those who know a great deal tend to approach that knowledge with great uncertainty and humility. Thus, it is true that we cannot trust finally and forever even those things derived from good science, for there may be special cases or new information that causes us to change our mind or at least to understand how what we thought was true had some limitations. For example, we once thought that matter could not be created or destroyed, although could be changed in form. But the great physicist Albert Einstein found that under special conditions, matter could be transformed to energy according to his famous equation $E = mc^2$. In this case, the advance of science told us that the earlier law of conservation of energy worked under usual conditions but that there are exceptions. This perspective enriches our understanding of the law of conservation of matter, which is now considered the law of the conservation of matter and energy. Angier [2] has written a useful book that summarizes much of what we have learned from the scientific method, and how we have learned it, in a very accessible style.

We believe very strongly that—even if there are many important exceptions to the power of science—if there is any knowledge we can trust, it must be derived from, or at least be consistent with, science and the scientific method. We believe that because the economy must operate in the real world, it cannot operate as if it were a perpetual motion machine, which in fact is the most common way of representing economic systems in introductory economics textbooks. More generally, there is a great deal of information derived in the natural science disciplines that could be of great value for understanding actual economies but that this information is rarely if ever put into economic textbooks. In addition, as we stated in the introduction, we do not believe that the education of our young people should be compartmentalized so that one learns natural sciences only in chemistry, physics, geology, or biology classes or, on the other hand, that you never hear about science in an economics course.

But what is science? How does "scientific truth" agree with or differ from other kinds of truths, including logical truths, economic truths, religious truths, and so on. Before we give more economics, we will focus on more science, going beyond the basic energy needed to understand economics by developing some basic science needed to understand both energy and economics.

15.2.2 The Need for Science to Understand How Economies Work

The more we can increase our scientific understanding of the world, the better we should be able to understand what good economics is and should be. This follows in the same way that our ability to do medicine is improved as we understand better the human body, the environment of humans, the technology of disease prevention and control, and the social interactions between health-care providers and sick people. In other words, we believe in a comprehensive systems approach for all but the simplest problems. Our list of the most important things you need to learn about science includes especially the scientific method and the most basic concepts pertaining to nature including matter, energy, life, and the fundamental interactions of all of these within the biosphere. Economics, if it is to be a real science, *must* be consistent with, and constrained by, these scientific principles, for we know of no exceptions to them. Humans can *want* to do many things, but they are *able* to do only what is possible within the laws of nature and the resources actually available, and if these concepts are not understood, human endeavors are apt to backfire (some might say continue to backfire).

15.2.3 Steps in the Scientific Method

The scientific method is usually taught to undergraduates as a series of experiments, with hypotheses, tests, and controls that a scientist follows in the process of gaining new knowledge. The formalized procedures of the scientific method usually includes observation of phenomena, the formation of hypotheses that are thought to explain those phenomenon, and the rejection of those hypotheses that are not supported by appropriate experimentation. Usually, the procedure requires a "test" and a "control," identical in all respects except for the one factor that is being tested. Thus, to test the hypothesis that phosphorus is needed for plant growth, one might grow two plants in pots with the soil in the pots being identical except that one contains phosphorus and one does not. The use of a control is usually critically important to identify the causative agent.

In fact, the process of science tends to be much more complex and messy, with many different pathways to new scientific understanding. Neither Isaac Newton nor Charles Darwin, probably the two most important and creative scientists that ever lived, particularly followed the scientific method as mentioned above. Rather, they were extremely astute observers and thinkers about what might be behind what they observed. Today, the fundamental criteria by which science is judged are that the mechanisms are consistent with known science and also that the results generated by the science *works*, "works" being defined as generating predictable results that are repeatable by others. For example, when we sent men to the moon, we were able to aim the space capsule based simply on Newton's laws of motion, laws that worked so well that not even small midcourse corrections were necessary—even though we had never tested them previously outside of the Earth's local environment. Surely, we would like to have such predictive power in economics! A problem, however, as any good economist or scientist will tell you, is that it is difficult to make predictions in a multiparameter world, that is, in a situation where many factors in addition to the one you are interested in or have control over might influence the results. Since real economies have many inputs and many outputs, determining which factor or factors may be most important can be quite difficult (but see the next chapter where we show how this can be done well, if not perfectly).

This is a problem faced by much of the rest of science too, and often it has been overcome with the help of statistical analysis designed for that purpose. But first, we need to think more basically about how we seek and sometimes find truth using science. As defined by the scientific methodologist Glymore [3], "science" is that field of intellectual inquiry that is amenable to the scientific method. Exactly what constitutes the scientific method is certainly debatable. Most practicing scientists would agree that most good scientists, natural and social, strive for *rigor*. Generally, rigor means intellectually defensible while using conceptual models that capture both reality itself and the mechanisms that determine the relation between cause and effect. It is often assumed that mathematical rigor means scientific rigor, but as we shall see, this is often not the case.

Within the natural sciences rigor generally means, at a minimum, that the concept, descriptor, or model used (1) is explicitly and unambiguously defined, (2) is consistent with *first principles* (i.e., things we know to always hold—at least so far!), (3) has been tested with adequate controls using some form of the scientific method (where that is possible) and has survived that testing, (4) explains an appropriate and nontrivial set of observed phenomena well, and, perhaps most important, (5) is repeatable by others who also follow the above rules. If *all* of these criteria, and as appropriate others, are *not* met, then we have to consider the theory or approach in question as a theory, or a hypothesis, or a myth, or something else but not yet in any sense a scientific law or even a scientifically supported concept or theory. *Probably the strongest criterion that marks something as science is that the observation and/or experiment is repeatable by others who follow the appropriate directions of the person promoting the hypothesis (and who usually are trying to get it to fail).* While it is very hard to say something is unequivocally correct using the scientific method, and some philosophers of science (most notably Karl Popper [4]) make the point that we can only fail to falsify a hypothesis, the true power of science comes from a theory's ability to withstand very explicit attempts to falsify it and to predict nontrivial outcomes.

Many very exciting new concepts in natural science have fallen when they have failed to satisfy all the points given above. On the other hand, there are some extremely powerful scientific theories, such as plate tectonics and natural selection, that explain a great many observations but that are amenable to experimentation in only a limited way. So it is not always required to meet all of the criteria listed above, but if they do not, then we have some very careful explaining to do. For example, Charles Darwin thought that we would never see natural selection in action, or be able to test it explicitly, because he thought the timescales were far too long and the experimental manipulation extremely difficult, in part due to the complex, multiparameter reality of nature. Nevertheless all other information—such as the fossil record—was so convincing that nearly all biologists came to accept Darwin's theory even without experimental verification. Recently, however, biologists such as Peter and Rosemary and Grant [5] and Dolph Schluter [6] have devised very clever observations and even experiments that have allowed us to observe and even manipulate natural selection, and it works essentially exactly as Darwin had hypothesized. So with very careful attention to scientific methodology and to the system in question, it is possible to undertake experiments to test our hypotheses even when people originally thought it impossible. An amazing thing is that as we learn much more about the mechanisms of how life works with all of our new molecular biology, we confirm that in fact nature behaves very much according to the basic principles that Darwin put down 150 years ago.

We next look at some fundamental physical and biological laws and principles that have been derived by using the scientific method that we believe are most solid and also most important for a good understanding of real economies and of biophysical economics. Most fundamentally, we ask "how does it work?" Whatever our answer, it must be consistent with science and derived by the scientific method; otherwise, we cannot accept its validity.

15.3 The Physical World

The two fundamental divisions of the physical world are energy and materials. Thus, we start our tour of scientific knowledge with a look at energy. Then we will look at materials.

15.3.1 Energy Sources

The principal sources of energy for the Earth are the sun, the movements of the sun and the moon relative to the Earth, and the radioactive decay within the Earth. The movements of the sun and the moon cause tides, and possibly some large-scale movements of portions of the solid Earth. The decay of radioactive elements within the Earth (plus residual heat from early Earth history) causes the interior of the Earth to be warmer than the surface. These factors also cause volcanoes and continental drift. Essentially all other energy, including wind, oil, gas and coal, our food, and that of all nature, comes directly or indirectly from the sun, which is an enormous thermonuclear furnace fed by hydrogen being converted to helium. We do not know exactly what the energy is that comes from the sun, but the effects are obvious. Scientists have more or less settled on calling sunlight "photon flux" and the amount of it "photon flux density." Sunlight tends to be of relatively short wavelength (Fig. 15.1) and because of this has very high energy and can do a great deal of work. The solar

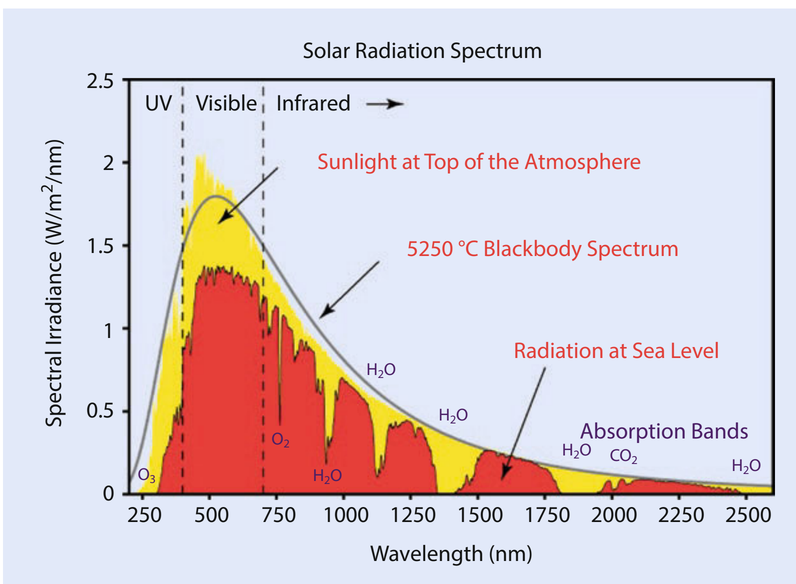


Fig. 15.1 Distribution of sunlight as a function of wavelength. **a** at top of atmosphere and **b** on the Earth's surface. (Source: Wikimedia Commons)

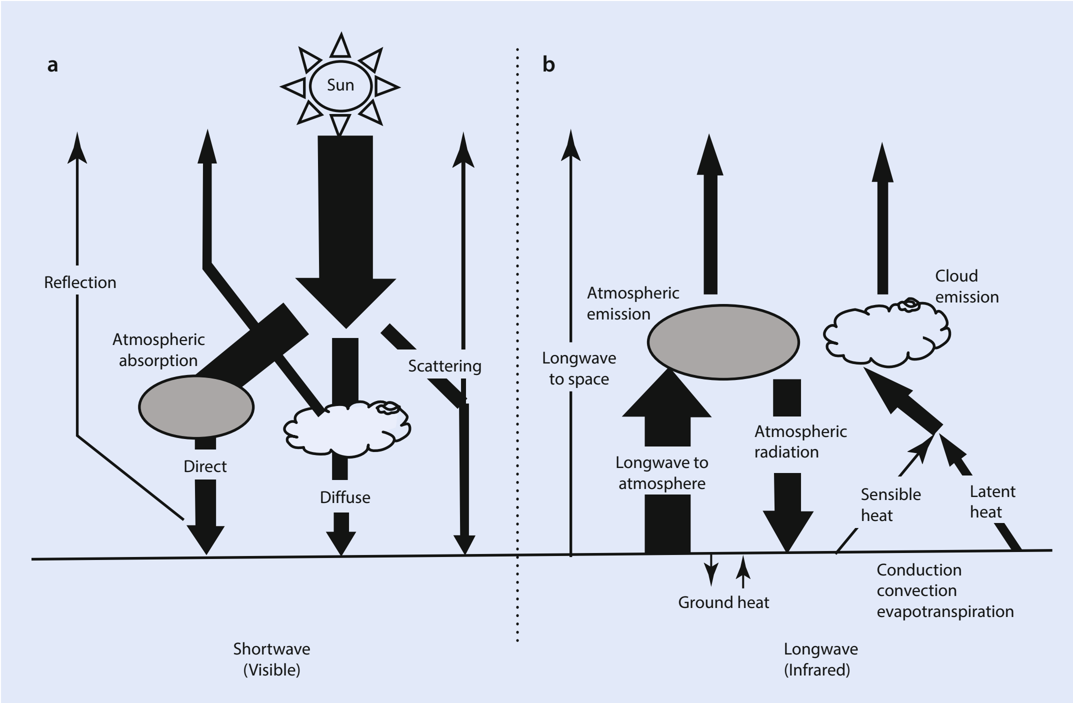


Fig. 15.2 Disposition of incoming solar radiation. Short wave radiation is high energy incoming photons and long wave radiation is outgoing (Source: Amy Chen)

constant, that is, the amount of sunlight received from the sun at the top of the atmosphere, is about 1367 Watts per meter squared perpendicular to the sun, equal to 4.9 KJ per square meter per hour. About one-quarter of this, on average, gets transferred through the atmosphere to the Earth's surface. Figure 15.2 gives the disposition of this incoming solar energy. Some of it is immediately reradiated to space, some evaporates water, and the majority is turned into longer wavelength, less energetic waves that we call sensible (i.e., we can sense it) heat. This transformation is very obvious when you walk barefoot on a black surface when the sun is bright. Sunlight has a broad spectral distribution, meaning that when separated by a prism, it has many different colors. Plants absorb and use for photosynthesis red and blue light but not green and hence reflect green. The sky is blue because the small particles suspended in the atmosphere are at roughly the same size as the blue wavelength, so the other colors go straight through the atmosphere, while some of the blue light is reflected (scattered) from the atmosphere to your eyes.

When the solar energy strikes the Earth's surface, the portion that is not reflected does considerable work. We can feel the effects in the heating of dark surfaces. The largest amount of work that sunlight does on Earth is to evaporate water. Wind and more generally weather is caused by the uneven heating of the Earth's surface by the sun. This operates the great heat engine of atmospheric circulation. Most importantly, the sun heats the Earth more at the equator than toward the poles because the land is perpendicular to the photon flux. This in turn causes the air over the equator to rise. As this air rises, it cools, and the associated loss of energy means that the atmosphere can hold fewer water molecules—which over time fall out as rain. Thus, the equator is a very wet region, and it is here that tropical rain forests are found. The exact place of the greatest rain changes north and south with the seasons, but it is always directly "under" the sun, i.e., at the location where the sun's rays are most nearly perpendicular. The rising air is eventually constrained by gravity and accumulates at about 5–10 miles high over the equator. This causes a

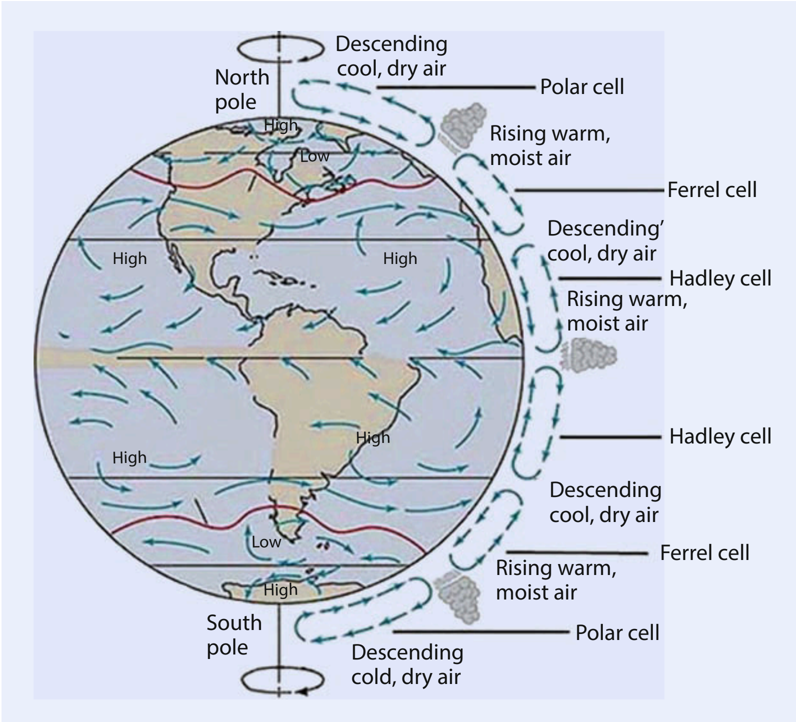


Fig. 15.3 The basic heat engine of the Earth. Electromagnetic radiation, usually considered as traveling in "packets" called photons, enters the Earth's atmosphere after traveling from the sun. Since the Earth's surface is more nearly perpendicular to their entrance path near the equator, they tend to be more concentrated there and subsequently heat the Earth's surface especially well at the equator. This causes warm air masses to rise at the equator and then disperse north and south as described in the text. As the air masses rise they cool, and as cooler air has less energy to keep the water molecules suspended, it rains a lot on the (thermal) equator, which moves north in summer and south in winter. The rising air masses create high pressure above the equator, pushing air masses north and south until they descend at 30° (Source: Kaufmann & Cleveland, 2008)

high-pressure zone there and, because the air masses have been moved upward, a low-pressure zone on the surface at the equator. This high pressure at altitude pushes the air north and south until it cools enough to descend at about 30° north and south. As the air descends, it warms again and hence has the energy to hold more and more water, so that when it comes in contact with the Earth's surface, it literally sucks the moisture out of the soil and vegetation, generating the Earth's great deserts. It also generates another high-pressure area there (at the Earth's surface at 30° north and south) which pushes air back toward the low-pressure air on the equator while being bent to the right in the Northern Hemisphere and left in the Southern Hemisphere by the Earth's rotation (the Coriolis force). This causes the steady *trade winds* characteristic of the tropics which become increasingly moisture-laden as they approach the equator. The high pressures at about 30° also push air masses poleward, and these winds as effected by the Coriolis force cause our familiar westerly winds in the temperate zones of the Northern hemisphere. These are familiar to those living in the temperate regions as they watch storm systems move across the land from west to east (Fig. 15.3). The net result is the very steady trade winds of the tropics.

British meteorologist George Hadley figured out the first (equatorial) cell in 1735 which bears his name. What he did was to explain the wind patterns that savvy ship captains had known since the time of Columbus: use the aptly named trade winds for moving from Europe to the Americas and the westerlies further north to go from the Americas to Europe, while avoiding, where possible, the doldrums on the equator and the horse latitudes at 30° where air masses move vertically rather than horizontally. This is an early example of where scientific knowledge was of great assistance to the economic situation of those who understood it.

15.3.2 Basic Thermodynamics

The basic laws of thermodynamics were given in the previous chapter. Their importance includes the concept that while material can be recycled energy cannot. Once we have used energy, it is, essentially, gone forever as a useful resource. This has enormous implications as civilization plows through its remaining resources of fossil fuels.

15.3.3 Entropy and Its Relation to Human Economies

When we think about energy, it is normally from the perspective of our own personal ability to get something done, go somewhere, or keep warm in winter or cool in summer. But the reach and importance of energy is far, far more pervasive principally because of *entropy*, which we have covered in the previous chapter. The things bought and sold in economies, cars, houses, and food, are bits of *negentropy*, or negative entropy, that is something highly organized or specialized structures, something extremely unlikely by itself. A nation, civilization, or economy must constantly invest money and energy into maintenance; otherwise buildings, bridges, and even entire civilizations will collapse. In addition, if there is to be growth, then that requires additional energy.

A simple example will help to think about entropy. A tuna sandwich, your own self, and an automobile are extremely unlikely, nonrandom structures that have been developed by taking the elements and materials of nature, initially scattered more or less at random over the surface of the Earth, and using energy to concentrate these elements and their compounds into very specific structures. Once the structure is made (i.e., wheat, the tuna fish, yourself, a bridge, a city), energy must be continuously invested or the materials of which it is composed will tend to go back on their own toward entropy—i.e., a more random assemblage, and the structure will fall apart. If that sandwich is put into a refrigerator, a device that uses energy to maintain the structure of its contents, the integrity of the sandwich will be maintained for some time. Pull the plug and the sandwich goes into a more random assemblage, first smelly organic residues and then, eventually, carbon dioxide and simple nitrogen compounds. Likewise, yourself, a car, or a modern city cannot run for long without the energy required for its repair, something sometimes called "fighting entropy." If and as high-quality energy becomes less available, then we may have to choose between maintaining our infrastructure, building more, or other consumption such as driving a car. While this may sound far-fetched as of 2017, the majority of states in the United States are facing severe debt and budget (and hence energy) shortfalls and are having to make painful decisions about which programs and infrastructures to maintain.

Most civilizations that have lost their main energy supplies have collapsed, as Tainter [7] and Diamond [8] have elegantly examined. Mexico is still rich in oil even as its main fields decline and uses much of it to maintain the 20 million people concentrated in Mexico City. The need for a continual input of energy to that city was once made clear to us when we were caught in a 10 mile long traffic jam of bumper to bumper enormous trucks that bring food and fuel into Mexico City every night. Mexico is filled with the ruins of enormous earlier cities and civilizations that, by some accounts, grew beyond their capacity to provide the energy resources that their large populations needed. Will the same fate befall modern Mexico City when oil becomes less abundant, as inevitably it will and which has already begun?

15.3.4 A Little Geology of Importance to Economics

We now shift our focus to materials. Economics is about goods and services. All goods are derived in some way from nature (including minerals, the soil, and the atmosphere), so it is useful to have information about where they came from. Services too are generally derived from nature, for example, the fuel that runs a transportation service or the metals in a bus. Most of the materials that we use in our economic life come from either plants, i.e., agriculture (food or chemical feedstocks), or forests (paper, lumber), or from the ground (rock; sand; cement; minerals such as iron, copper, and aluminum; fossil fuels including coal, natural gas, and oil). Most plastics are derived from fossil fuels, especially natural gas. The conditions under which these materials are found are normally considered the province of geology, agronomy, or forestry.

The first important geological fact about the Earth is that it is very old, roughly 4.5 billion years old. Over this very long time period, mountains were thrown up by volcanic or tectonic activity, continents drifted across the ocean and life evolved, and in the process changed the Earth itself. Some kind of simple life has existed for about half to three-quarters of that time, but fishes, for example, and primitive life on land have existed for only about 500 million years. Land plants evolved and changed the atmosphere from reducing to oxidizing [9]. Humans as a recognizable species have been around for about 1 million years, less than one-thousandth of the time that the Earth has had life. It is thought that very large asteroids from outer space hit the Earth every few hundred million years and change things very much, for example, by eliminating dinosaurs and opening up the environment for the evolution of mammals.

There are three basic types of rocks, igneous (formed by volcanic activity), sedimentary (formed by deposition of sand, silt, or marine skeletons on the bottom of the sea or large lakes), and metamorphic, which are either of the former that have been transformed by crustal movements and pressures. Sedimentary rocks are further divided into sandstones, shales, and limestone, formed specifically from sand, silt, and marine organisms. In areas once covered by the ocean, such as Central New York State, there are often alternating layers of sandstone and shale, representing successive geological eras. Why is there sometimes shale and sometimes sandstone and sometime limestone, sometimes in alternating bands? It is because in the past different types of sediments were found at differing distances from the source materials on the continental shelves. Since sand drops out of moving water relatively rapidly the presence of sandstone implies that the source of the sediments was originally not very far or that ocean currents were strong. The finer silt that constitutes the shales could travel much further from their continental origin before falling out, and limestone represents the remains of active populations of animals that made their shells out of calcium carbonate. Each of these materials can contain a certain amount of organic material (i.e., leftover plant and animal material) that can be the basis of the formation of fossil fuels.

The earth is a very dynamic place if you think in terms of geological time, with large crustal plates moving about its surface. For example, South America is separating increasingly from Africa to which it was once joined. Centers of activity where one plate smashes into another such as along the Andes of South America are characterized by mountain chains, volcanoes, and frequent earthquakes. The continents move about in response to geologic energies (deep "hot spots") that sometimes come up in the middle of the oceans, often causing volcanoes (as in Iceland and Hawaii) and continental drift. These hot spots generate island chains such as Hawaii, where a plate drifting over a single hot spot formed the islands from volcanic activity that is still continuing on the southern tip of the southernmost Island. At other locations, the Earth pulls apart, causing rift valleys. Good examples are found in East Africa, where there are a series of large lakes formed in basins where the land is being pulled apart. Eventually the edges will move far enough apart so that the sea will tumble in and the lakes will become inland seas. This has already happened in the Red Sea where Egypt is separating from the Arabian Peninsula and where Madagascar has separated from Africa. Another example is where Scotland has drifted away from Norway. As we shall see below, these rift areas are very important for the formation of oil.

15.3.5 Concentration, Depletion, and the "Best First" Principle

The most important geological issue relating to economics is that the materials that economies are based on, whether those of antiquity or of today, are not found distributed randomly (as we might expect from our discussion of entropy) about the Earth but rather in various concentrations of widely different purity and quality, the most concentrated are called ores or deposits. This is because past geologic energies, includ-

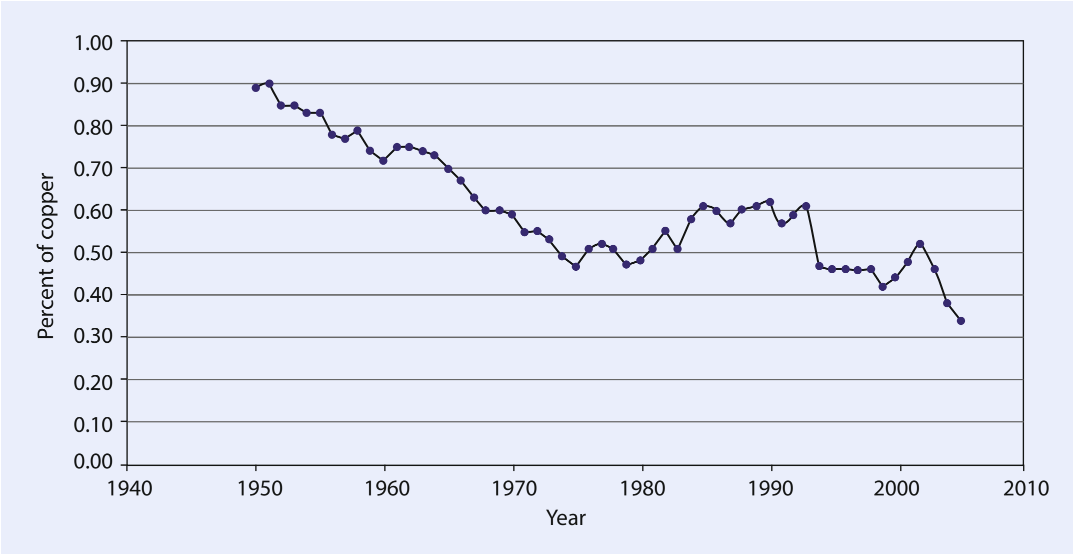


Fig. 15.4 Average grade of copper mined in the United States (Source U.S. Bureau of Mines and the U.S. Geological Survey)

15

ing volcanism, tectonic actions, river transport, microbial actions, or other processes, have tended to concentrate the different elements (and certain compounds) in particular locations where they may be orders of magnitude (i.e., factors of 10) more abundant than the general background "crustal average". Such differences have been obvious for millennia to humans who have tended to exploit, and deplete, the highest grade materials first. The initial copper and tin deposits in Crete, one place humans began the process of mining and smelting metals, were initially at such high concentrations that the metals abundantly flowed in a pure stream out of fireside rocks. When these rocks were all depleted, then humans had to invent mining and much more complex metallurgy to supply the metals. Today, the Earth is a very well-explored place—and with a few exceptions there have been relatively few large discoveries of very important materials for many decades. Now, rich mines are only a memory, and we get most of our metals either from recycling (roughly half) or from huge, relatively low-grade deposits that require enormous machines and very large quantities of energy to extract the metals from the ores.

A good example is for copper in the United States. Over time, the best grades of copper were mined first because it takes less energy (and labor and equipment and hence money) to process these materials into forms that society finds useful. For example, if you go to the end of Main Street in Butte Montana, you will look into a hole several miles across and nearly half a mile deep. This was once a hill, and it had been called the "richest hill on earth." The hill contained copper ore that was up to 50% copper, and once the proper machinery was in place, it was relatively easy and very profitable to mine that hill. Some 20 billion pounds of copper, plus gold, silver, zinc, and other minerals, were taken out of that hill. Some ancient geological processes, we are not sure exactly what but it appears to have involved cooling of mineral rich magma-heated water intrusions, concentrated copper there, where it had lain until the miners dug it up. Now, that rich copper ore is gone, and the huge hole has been slowly filling up with water which has turned to sulfuric acid because of the sulfur deposits that were associated with the copper. It is so acidic that if migrating waterfowl land on the lake in that hole they immediately die.

Today, the average grade of copper ore extracted from US copper mines contains about 0.4% copper [10], in other words, only about 1% of what they were getting out of Butte at the turn of the last century (Fig. 15.4). Consequently, some 100 times more ore has to be dug up, crushed, and processed per kg of copper delivered to society than back then. Thus, an important geological issue that effects economics is that, over time, the best deposits tend to be used first, so that the energy, dollar, and often environmental cost of getting a purified product tends to increase. Of course, technologies tend to improve over time, reducing costs and often energy use. Technology is in a race with depletion, sometimes one "winning," sometimes the other. In the case of copper, it appears that at first, the energy cost of getting a kilogram of pure copper decreased and subsequently it increased [10]. For most materials, it appears that energy costs are increasing, but a much better case-by-case review is needed. It is possible that the depletion of copper could limit the development of solar PV energy [11].

This *"best first" principle* is rarely mentioned in the economic literature (although it seems consistent with the law of diminishing returns). The concept also applies to many other aspects of human and indeed other organismal behavior. This principle has enormous economic implications as we deplete so many resources upon which we depend, especially as we can no longer count on more energy being available to mine ever lower-grade resources.

15.3.6 The Formation of Fossil Fuels

Because oil and gas are so important to our economic life and because there is so much controversy about how much is left to exploit, it is important to consider in some detail the very special circumstances that were required for their formation. Oil and gas are organic materials, that is, they are plant and animal remains composed of mostly carbon (and hydrogen) as is all life. (The word "organic" technically means carbon based; organic chemistry is about the chemistry of carbon and has little to do with the popular use of the term to denote low use of agrochemicals.) As life evolved, a great deal of organic material was formed, most of which was oxidized relatively soon and turned back to carbon dioxide in the atmosphere, becoming available for new plant growth. But some of this organic material found its way to *anaerobic* (meaning without oxygen) basins. For example, coal was formed in great freshwater swamps in what is now Pennsylvania, Ohio, and Wyoming.

Oil was formed in two principle places: rift basins such as once existed between Scotland and Norway or Saudi Arabia, and river deltas such as off the Mississippi or Niger River. Rift basins are formed when the land on one or both sides moves apart (as is the case today with East African lakes), generating deep basins called grabens, often with lakes or invading marine waters within (Fig. 15.5). Phytoplankton, tiny marine or freshwater plants would grow in the water and fall to the bottom of some of the deep rift basins where there was no oxygen and hence little decomposition. This process is greatly assisted when the water cannot mix deeply, i.e., is *stratified* by temperature, a phenomenon familiar to many of you who have dived deeply into a summer lake. Hence, a general requirement is that the oil-forming basins were located in the tropics and/or were active during periods of climate warming. Warm surface water can be mixed with deep water only with great difficulty (such as by a fierce wind). In the tropics, both lakes and the ocean tend to be strongly stratified all year around, so that very often the deeper parts use up all of their oxygen and remain anaerobic.

Under extremely rare circumstances, often related to a warming climate with lots of evaporation, the sinking phytoplankton were protected from oxidation in the deep, non-mixing anaerobic bottom waters for long periods of time, thousands to many millions of years. As time went on and if the climate happened to change from dry to wet, sediments would wash down from the surrounding hills, covering the organic material with layers of sand and silt which, over time, became rock (Fig. 15.5b). If enough sedimentary rock (say 3–5 thousand meters) covered the basin, the pressure would heat up the organic material, and over millions of years, the ancient phytoplankton would be "pressure cooked" at about the temperature of boiling water, breaking the long plant molecules of typically hundreds of carbons tied together into shorter ones, thus forming oil and gas. The familiar word "octane" refers to oil with eight carbon atoms arranged in a ring which is the best formulation for gasoline as it does not combust too easily and hence cause preignition or knocking. Natural gas is what remains when the chains have been broken to lengths of only one or two carbon atoms.

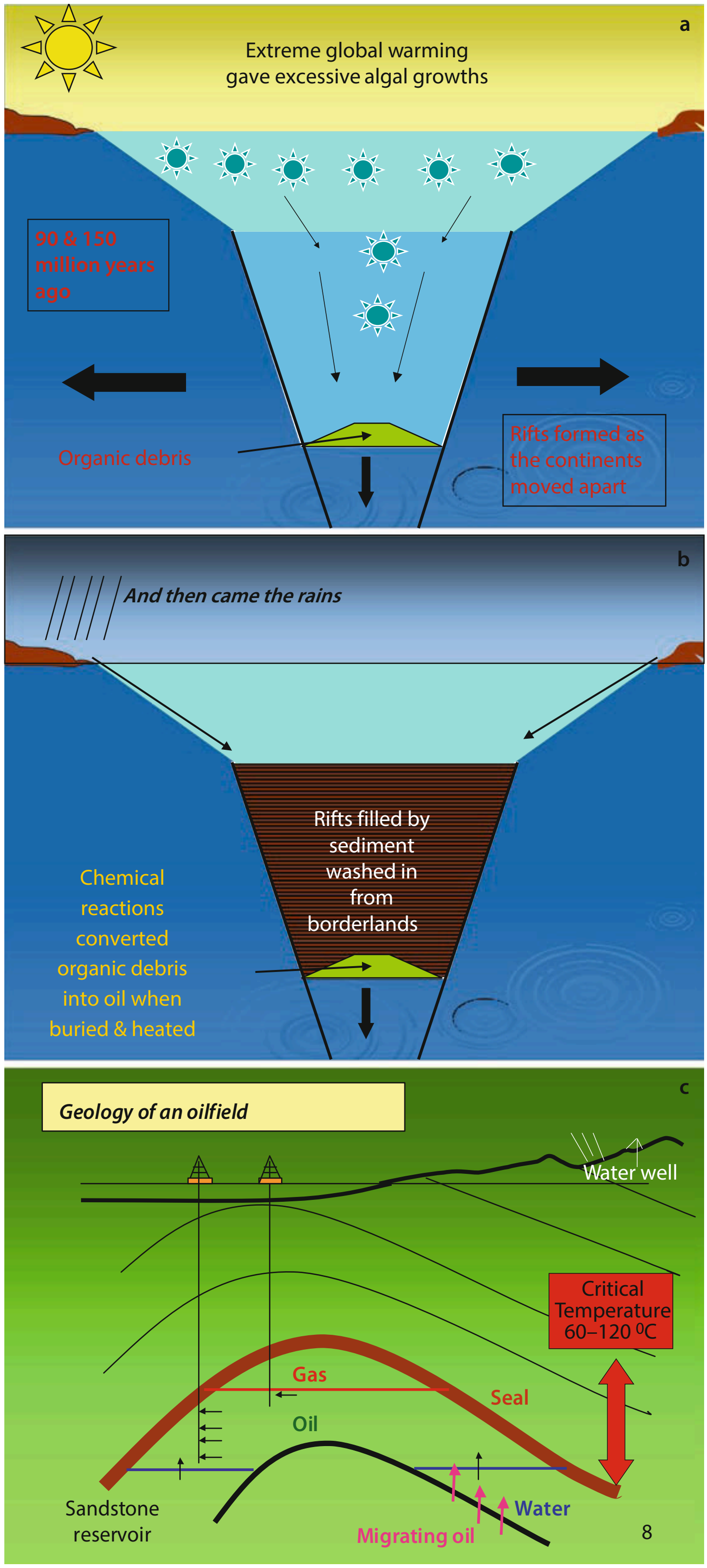


Fig. 15.5 The typical formation of oil. Oil is not formed often or in very many places and requires very special conditions for formation. It was formed on the Earth in only two general geological times, about 90 and about 150 million years ago. In order for oil to form a series of steps must occur in sequence (Source: Colin Campbell): **a** First a very deep lake or marine trench must be formed, such as when the crust moves apart forming a graben, during a period of climate warming. Phytoplankton, whose growth is encouraged by the warm conditions, sinks into the deep anaerobic waters. **b** Then it is necessary to have an extensive period of rains that wash sediments into the basin, covering the organic materials with thousands of meters of sediments. Then the organic material is pressure cooked for many tens of millions of years, breaking down ("cracking") the complex molecules into simpler ones. **c** The relatively light hydrocarbons end up moving upward from the source rocks. Most of it escapes to the atmosphere, but some small part is caught by impervious "trap rocks." This forms the oil and gas deposits we exploit

These very rare and special rocks are known in petroleum geology as *source rocks*. The oil and gas thus formed would then tend to rise upward over geological time as they are less dense than the Earth's sediments within which are found. Some small proportion, perhaps 1%, of the oil and gas migrating upward from the source rocks find their way to particular rock formations impervious to their movement, such as salt domes or sandstone, where they are trapped. These rocks, which may be far above the source rocks, are known as *trap rocks* and are normally the locations that humans exploit (◘ Fig. 15.5c). A good example of where all this took place was the rift valley where Scotland and England left Norway some 100–200 million years ago. The oil that we now exploit from the North Sea was created as a series of grabens were formed and flooded with water. Large phytoplankton growth in the productive water settled into deep basins and eventually were covered by thick layers of sediments. Some of these layers, particularly those made of limestone but sometimes sandstones, formed both reservoirs and traps.

Similar burial of phytoplankton or other organic matter sometimes has taken place within, and off of, river deltas where highly productive estuarine systems such as those associated with the Mississippi, Niger, and Orinoco rivers generate a lot of organic material and where periodic sediment deposits covered over anaerobic basins. The general lesson from these descriptions is that the special conditions required for the creation of exploitable oil and gas fields have been quite rare in the geologic past (occurring mostly some 90 and 150 million years ago in very special and limited environments) and that the time to make oil and gas is extremely long. As a consequence, significant commercially exploitable oil and gas are found in a relatively few regions of the Earth's surface. Coal, requiring similar but far less stringent conditions for its production, is much more common. Gas too is widely dispersed, but the main reservoirs were relatively rare. On the other hand, gas is found widely at low concentrations in "tight" shales and sandstones. Exploitation of these diffuse resources is becoming increasingly important as the large true gas fields found earlier face serious depletion. Whether or not these newer "unconventional" fields can maintain US gas production at the present level for very long is unknown at this time.

As with copper (◘ Fig. 15.4) and another example of the "best first" principle, humans have tended to exploit the large, high-quality and easy oil deposits first. They have exploited deeper, and deeper offshore regions, mainly off the Mississippi River, where there are more than 4000 very expensive offshore platforms that are responsible for much of the United States' remaining oil and gas production. As this was being written, there was considerable excitement about finding the new, possibly large Tiber oil field in the Gulf of Mexico. But the field is 35,000 ft (6 miles) under the Gulf of Mexico and would be extremely, perhaps prohibitively, energy intensive to develop. On the other hand, the high pressures there may force the oil to the surface without expensive pumping or pressurizing. For the United states, we found the most oil in the 1930s and for the world the most in the 1960s (◘ Fig. 8.3). All of these factors have very important implications for EROI (▶ Chap. 18).

15.3.7 A Little Chemistry of Importance to Economics

The world and everything in it, including yourself and your surroundings, is composed of chemicals. Economies generally mine or otherwise obtain source materials for chemicals (called *feedstocks*), refine or transform them, often times combining them with other chemicals, and using or selling the products. The most fundamental chemicals, incapable of being transformed to other chemicals, are called elements; these include such familiar chemicals as hydrogen, oxygen, and carbon. When two or more elements combine, they generate *compounds*, which include most of the common materials of everyday life: hydrogen and oxygen combine to make water, hydrogen and carbon to natural to make gas or oil. The chemistry of the world and of our economy is extremely complicated, but usually it is based mostly on only about 20 or 30 elements and their compounds.

15.3.8 Conservation of Matter: Supplies of Inputs

Perhaps the most important aspect of chemicals, or more generally all materials, for economics is the *law of the conservation of matter*. This law says that while matter (also called mass) can be transformed in many ways, it can be neither created nor destroyed. Again there is the exception that under very special conditions (nuclear reactions) matter can be transformed to energy. There are two reasons that this law is of critical importance to economics. The first has to do with the supply of the materials required by the economy, and the second relates to the disposal of waste materials. In other words, the goods that interest us as consumers or economists are derived from elements "borrowed" from nature. Cars are made from iron, copper, sand (for glass), natural gas (for plastics), and many other things; fish come from the sea; many houses, books, and newspapers come from trees; clothes from plants, animals or, increasingly, petroleum; computers are made from plastics, copper, aluminum, gold, and silicon, and so on. Essentially every good starts as some material extracted from nature somewhere. Energy is required to do step to make the final product.

Take for example plastics, a suite of materials made from hydrocarbons. Plastics are ubiquitous and very useful, they can be formed into many shapes, and they are cheap. A common ketchup bottle may have seven layers of different kinds of plastics to protect the ketchup inside. Chemists have learned to be very clever at manipulating elements and molecules, but the carbon and hydrogen atoms in the plastic still have to start with raw materials from feedstocks, usually natural gas, or sometimes oil or coal. Increasingly plastics are recycled, but they may pollute the environment for 1000 years. As fossil fuels become more expensive, the molecules can come from biomass such as from trees or crop residues.

15.3.9 Carbon Chemistry

Most of our food, fuels, plastics, and many other things are carbon based. Carbon can take many extraordinarily different forms and can be transformed from one form to another relatively easily. Carbon may be found as carbon dioxide gas in the atmosphere; pure carbon in a pencil "lead" or a diamond; combined with hydrogen in *hydrocarbons* such as coal, gas, and oil; with hydrogen and oxygen in *carbohydrates* that includes most of the fuels that we eat; with the element calcium in limestone (from which we make cement), and so on. In general (as stated in ► Chap. 14), compounds with lots of hydrogen and little or no oxygen, such as hydrocarbons, are called *reduced* and serve as excellent fuels, and compounds that have a great deal of oxygen (such as CO_2) are called *oxidized* and are poor fuels. Carbohydrates are mostly reduced but slightly oxidized and hence do not make quite as good a fuel per gram as hydrocarbons. Combining oxygen with a hydrocarbon or carbohydrate releases energy that can be used to propel an athlete, an automobile, a chemical reaction, or a manufacturing operation.

You are made of principally carbon, with quite a bit of hydrogen, some oxygen, and more than a little nitrogen. Natural selection has chosen carbon as the basic skeleton for life because it has the possibility of combining with other atoms in four directions (i.e., it has four electrons in its outer or active ring), allowing the construction of the quite complex compounds that life requires, such as carbohydrates and fats. Because the element carbon is so closely associated with life, the chemistry of carbon, living or not, is called *organic* chemistry. Carbohydrates, fats, and protein are the basic biological compounds and also the basic food groups. Nitrogen too is an element of special importance, with five electrons in its outer shell and room for three bonds, and is also able to make very complex compounds that are often proteins. In its elemental form N_2 nitrogen forms about 78% of the atmosphere. In this state it is very inert, meaning that it does not react with most other elements except under very special conditions. But nitrogen can also be found combined with oxygen and with hydrogen, and in these states (nitrates and ammonia), it is extremely important for life because organisms can take the nitrogen from these compounds and (with carbon) make proteins. Proteins are important because they allow very great specificity, that is, very exact kinds of molecules. Nitrogen is critical for economies because it is the most important fertilizer used in agriculture, because plants need it to make their own proteins, and agriculture is usually one of the, or the, most important sectors in the economies of most nations.

15.3.10 Nitrogen Chemistry and the Haber-Bosch process

Although nitrogen is one of the most abundant elements on the Earth's surface (as N_2 in the air), it is relatively rare in its "fixed" form, that is, combined with hydrogen or oxygen. Fixing is uncommon in nature because it takes a great deal of energy to break the three chemical bonds holding the two nitrogen atoms of N_2 together. This occurs only when great energy is applied to the atmosphere (as in a lightning bolt) or when special organisms (only certain bacteria and blue green algae) invest a lot of their own photosynthetically derived energy into deliberately splitting the two nitrogen atoms apart so that they can get nitrogen for their own purposes, principally to make proteins. Until 1909, the major source of nitrogen for agricultural plants was from manure, and the first author's father remembers spending much of his childhood, as many did in 1920, hauling cow manure from the barn to the fields. Many of the readers of this book would be doing that too except for one great chemical discovery.

Ammonia (NH_3) is an extremely valuable chemical because of its long use in the dye industry and because it was the basis for explosives and fertilizer. Until 1908, however, ammonia was made only by natural process from certain bacteria and blue green algae or from lightning in the atmosphere. As such, its supply was limited. Although the principle by which synthetic ammonia might be made was simple and known for about 100 years, the actual process had eluded many important chemists. The equation is simply:

$$N_2 + 3H_2 \rightarrow 2NH_3$$

The N_2 is readily available as the major component of the atmosphere, although extraordinarily unreactive, and the hydrogen was readily available from coal or natural gas. After failing in several earlier attempts in 1909, the German chemist Fritz Haber discovered how to split the nitrogen molecules of the air industrially by adding a great deal of energy to N_2. He did this by heating a cylinder that was injected with air (the source of nitrogen) and natural gas (the source of hydrogen) while compressing the gases and using a special catalyst (initially osmium) [11]. The result was an output flow of ammonia (NH_3), a chemical very useful to plants and to industrial chemistry. None of Haber's university colleagues understood why he was so excitedly running around the campus shouting that he had done what no other person had done—to create "fixed" nitrogen from atmospheric nitrogen. Nor did they understand, as Haber did, why this was so important.

Haber had been assisted in this by a contract with the German industrial firm BASF, which quickly scaled up Haber's mechanism to commercial scale and under the leadership of Carl Bosch soon built enormous factories. These factories required enormous amounts of energy to run the process. While the early attempts produced some spectacular explosions, it also, once perfected, generated enormous amounts of commercial ammonia which has insured, at least potentially, enough food for all and freed most of us from carrying manure to the fields. It also had some rather different results, as industrially derived ammonium nitrate was and is the basis for gunpowder and other explosives. In 1914, at the start of the First World War, the Germans had only 6 months of gunpowder, derived from Chilean guano (bird dung). Without the industrially produced gunpowder of the Haber-Bosch process, the war would have ended quickly [12]. Thus, the Haber-Bosch industrial fixation of gunpowder is credited with making First World War last for 4 additional miserable years, and, one might add, allowing the second World War to be as devastating as it was. Even terrorists today blow up markets in Baghdad and buildings in Oklahoma City using ammonium nitrate explosives.

15.3.11 Phosphorus

Plants need more than nitrogen fertilizer to survive and grow. Phosphorous and potassium, and in smaller quantities, sulfur, molybdenum, and perhaps a dozen other chemicals are all essential plant nutrients. When the nuclear scientists Goeller and Weinberg [13] examined the entire periodic table, they found that for all elements necessary to civilization at that time, there was a substitute: aluminum wires could substitute for copper, energy could in effect substitute for nitrogen through the Haber process, and so on. But they found one exception: phosphorus. Phosphorus was completely necessary for plant growth and life in general, and there was no substitute. In the approximate words of geochemist Edward Deevey [14] some five decades ago, "there is something peculiar about the geochemistry of the Earth today that life is so dependent upon phosphorus, but it is now in such short supply." In other words, it might seemed that life evolved when phosphorus was more abundant. Today, most phosphorus comes from mines in Florida and Morocco, and much of it goes in a one-way trip from mine to crops to animals to humans to toilets to waterways to the ocean. Thus, the chemistry of phosphorus is of critical concern to modern economies because of its critical importance and non-substitutability for plant growth and because its main sources (in Florida and Morocco) are increasingly depleted. Thus, more energy is required for fertilizer production, and because as a waste product, it causes very undesirable growths of algae in our water bodies.

15.3.12 Conservation of Matter: Wastes

A second implication of the law of conservation of matter beyond the continual need for new supplies is that all of the elements in all of the materials that are ripped out of the Earth and brought into an economy must end up somewhere: as products or by-products, as recycled matter, or as wastes dumped into the environment. So if we manufacture a product, say a cleaning chemical, that material, or at least its elements, will be around indefinitely in some form or another. In the past, and still in many situations, whatever was left over after humans had used something was simply dumped—into the river, into a landfill, into the environment. Additionally, each step in the use of a chemical implies losses at each step, in mining, concentrating, processing, manufacturing, transport, use, and disposal. At each one of these steps, some part, large or small, of the original product is lost to the general environment. When the economies of humans were based mostly on the products of nature directly, their wastes (e.g., food wastes, logging wastes) were normally simply the routine wastes that were part of ecosystems and could be processed like any others—for which billions of years of natural selection had generated the dung beetles, bacteria, and so on to take advantage of these resources (to them) and in so doing keep things "cleaned up." Over the past several hundred years, humans greatly increased the scale of things—of agriculture, of mining, of economies, and of themselves—in cities and, eventually, through industrial and scientific processes. Humans also generated thousands of new chemicals that organisms had no previous experience with and for which there were often few organisms able to process. The net effect was to overload many ecosystems that had previously been able to adapt to humans. For example, the synthetic fertilizers that were generated from the Haber process and from mining phosphorus and potassium tended to be much more abundant in some locations than the natural quantities and hence wash into rivers and lakes, where they often caused serious pollution even though these elements had always been part of nature. While phosphorus and nitrogen are essential requirements for all plant life, an excess amount in waterways caused this once rare element to become abundant in many places, especially in the surface waters of the Earth where it generates undesirable algae growth and low oxygen conditions, a condition known as *eutrophication*.

Over time, nature tends to process human-made chemicals into more innocuous forms, but there is often very serious production of pollution along the way. Humans have become much better at recycling materials in recent years, and this recycling often has reduced greatly the amount of waste materials entering the environment. But recycling does not always reduce environmental impact as much as one might think, and again we need to think using a systems approach. For example, it would seem to be unequivocally good for the environment to recycle newspapers, that is, to make new newspapers from old. But if newspapers are to be recycled first, they need to be deinked, and then the fibers separated from the other materials. When all is said and done, it takes *more* energy to make a ton of newspaper from recycled materials compared to virgin materials, and more wastes are produced, mostly from the old ink. This is a good example where understanding the law of conservation of mass (the materials in the ink) helps us to understand the implications of what might seem initially to be an unequivocally good policy. It may still make sense to recycle newspapers, for example, to save space in landfills, and there are soy inks that are much easier to process, but it is not easy to make that judgment without undertaking a quite complex systems study. Probably the thing that makes the most sense is to reduce our use of paper, for example, a very large component of its use is for advertisements that people do not even look at or that even if they do are for products that may be really quite unnecessary and that also generate pollutants in their manufacture. In fact this is happening now as the internet increasingly does the function once done by newspapers, but one cost of this is that the revenue to newspapers declines so it is harder for them to maintain their staff of investigative reporters, in our view a critical part of society.

Thus, chemistry is extremely important to economies with respect to both our growing dependence upon the use of chemicals and also as pollutants. The natural world is full of complex chemical compounds, most of which are relatively innocuous but some of which are very toxic. In fact many natural compounds are designed to be very toxic. All plant materials represent a potential food resource for a whole plethora of viruses, bacteria, and insects, not to mention grazers such as deer. One response has been for plants to develop over time various chemical defenses to make themselves unpalatable or even to kill their potential consumers. Familiar examples include mustard oils, caffeine, turpentine, and, to some, the alkaloid *Tetrahydrocannabis*. While small amounts of these materials may make interesting dietary supplements, a diet composed of only one or more of them would kill us. That is a problem the insects face when they alight on, say, a mustard plant. Eat up and die! Well, not surprisingly, most insects choose to go somewhere else for lunch and the plant is protected. Animals in turn have developed kidneys and livers to detoxify many of these chemicals, so that they can eat some of the material. Thus, over evolutionary time, there is a sort of cat-and-mouse game of defense and offense, with few clear winners but many clear losers—given that maybe 99.9 or so percent of all former species have gone extinct!

Humans have changed the world around them in many ways through the rapid understanding and application of industrial chemistry. One of the most important examples is DDT, the first synthetic pesticide. DDT, developed in the second world war, was considered a godsend to our soldiers, for it was cheap and nontoxic to humans and eliminated many harmful and irritating pests, such as body lice with a single, simple dusting. Soon it was used on agricultural crops with similar spectacular results in reducing the losses to insects. It seemed to be too good to be true, and it was. Rachel Carson, a marine biologist and gifted writer, wrote one of the most important books of all time *Silent Spring* [15], which documented the very large impact that DDT had on bird reproduction. This book helped launch the environmental movement and for the first time suspicion that not all new inventions, nor progress itself, were necessarily desirable. DDT was especially a problem because it did not break down in nature—put it in the environment and it stays there, cycling through food chains and becoming concentrated as one organism ate another. The case against DDT was made further when it was discovered that the insects had become not only resistant to DDT but that some even required it for their survival! Natural selection can be that powerful and that fast! Chemists have responded to these problems by developing new pesticides that, while often more toxic directly to humans, break down to relatively harmless compounds in a matter of weeks so that it appears that the long-term toxicity problems are solved—as long as the good chemicals are used! But the pests still evolve to them and over time the pesticides lose their effectiveness. Agronomist David Pimentel argues that even as we use far more pesticides that we still loose about the same proportion of our food to pests that we did in the past, before pesticides.

Probably the most important, quantitatively, pollutants worldwide, other than the pathogens that kill humans outright, are the various carbonaceous waste products, including especially the fecal wastes of humans and domestic animals, especially when they are dumped into water bodies. This is a natural process and has happened naturally long before humans. But humans and their cities have completely changed the scale. Most natural bodies of water can handle moderate amounts of carbonaceous wastes through oxidation, changing the carbon materials into relatively harmless compounds such as water and CO_2. The problem occurs when too much polluting material is added. Then, the oxidizing capacity of the water bodies are overwhelmed and all or most of the oxygen is used up, resulting in bad smells, fish deaths, and a generally degraded water body. As developed above, a somewhat similar process happens when too much phosphorus is added to water bodies. Phosphorous compounds are the basis of very effective detergents, but they also encourage the excessive growth of aquatic plants, which then die and use up the oxygen in a process called eutrophication. Fortunately, these problems can be ameliorated or eliminated by relatively modest public expenditures in sewage treatment plants and by using other chemicals in detergents. Such successes in reducing environmental impacts are good examples as to how it is possible to successfully resolve serious externalities through good chemistry, good engineering, good economics, and especially good public policy implementation. But at the same time, treating sewage uses a considerable amount of energy and the total impact of growing human populations, and their growing use of materials leads to increased pollution of the Earth as a whole.

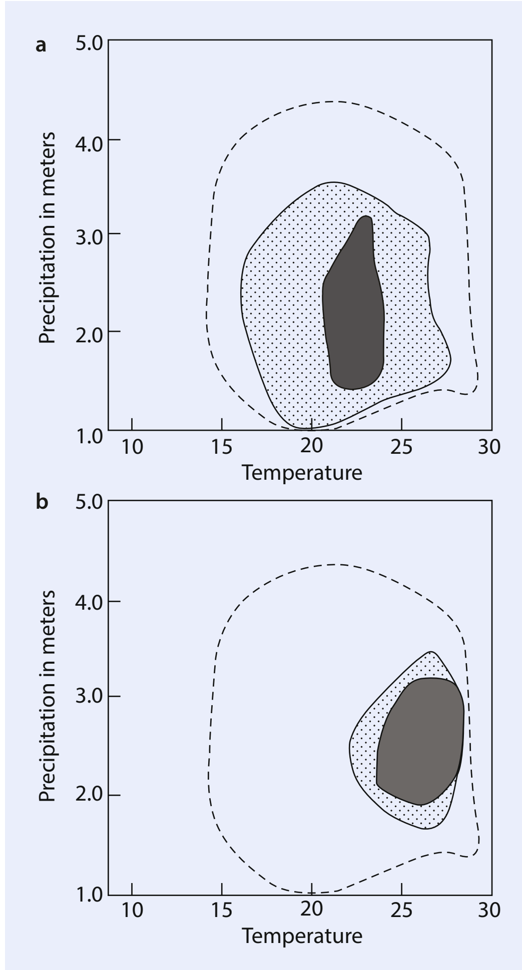
15.3.13 Chemistry and Physics

While chemistry is usually considered independent from physics, in fact, the two interact in many, many ways. Here are a few simple examples:

1. Essentially any chemical reaction is accelerated by *increasing the temperature*. For example, getting food particles to become unstuck from dishes requires work that occurs because you add physical energy by your scrubbing actions and by the chemical reactions in which you emulsify the food particles in a soap or detergent. Using hot water to clean the dishes adds additional energy to the process and accelerates the cleaning. Or leaving the dishes in the sink overnight after first filling them with clean water allows you to use the chemical energy of clean water (the molecules of water have positively and negatively charge ends that attract the materials stuck to your plate) to do work that you would otherwise have to do yourself! Polluted water has less energy to do that as the charged ends are already occupied. Thus, clean water is economically much more valuable than soiled water because it can do more work, such as cleaning work, in industrial or other economic processes.
2. Chemical reactions are usually greatly accelerated by *increasing the surface to volume ratio*, which is usually done by making the reactive particles smaller. This is familiar to most of us when we build a campfire: we must start with small dry twigs or even paper, with a very high surface to volume ratio and hence exposure to oxygen in the air, and then feed in progressively large twigs and the logs once we have a good hot bed of coals. In the process of industrial combustion, some hydrocarbon (such as oil or coal) is combined with oxygen to produce energy that is then used in, say, some economic process. The efficiency with which oxygen combines with the carbon and hydrogen in the fuel depends upon how closely each oxygen molecule comes into contact with each fuel molecule. Humans have been burning coal for a long time, and most of you are familiar with old pictures of factories or locomotives belching black smoke into the air. That black smoke, and the soot and other pollutants that are in it, are a product of incomplete combustion. We have learned to pulverize coal before we burn it, so that the carbon is nearly completely oxidized, providing more energy and less pollution—although even this does not decrease the quantity of CO_2 produced. On a more personal level, it is obvious that if you want to cool your food in a cooler for camping rapidly and completely, then you use crushed ice, but if you want to keep your food reasonably cool for a long time, you want to use a block of ice. All of these examples are related to issues of *surface to volume ratios*, that is, the surface of the material upon which the reaction (cooling or combustion) takes place compared to the total volume of the material. Of course, good engineers learn these and many more basic scientific principles early on, and from them they are able to use resources in a more intelligent and efficient way. What these concepts mean to economics is that there are many ways that simple science can be used to generate more efficient, less polluting economic activities, even when sometimes that costs more.

15.3.14 Climate and the Hydrological Cycle

A basic point of this book is that the basic inputs to the economy are not simply labor and investments but also natural resources, especially energy, and a proper working environment. Of the latter, the most important are soil and water as well as a proper temperature and other attributes of climate. Climate refers to the average temperature, rainfall, humidity, cloudiness, and so on that characterizes a spot or region of the Earth's surface, including the normal variations one might expect. Every species has its own ideal temperatures and humans are no exception. Probably you do not think of it too often but a proper temperature is critical for all of us. Just watch what happens if a classroom deviates very much from the range of about 68–78° Fahrenheit. Students will add or remove clothes, open windows, turn the thermostat or otherwise *thermoregulate.* We don't think about it very much, in part because it seems so natural (it is) and also because we live in a climate-controlled world where often sophisticated clothing and an enormous amount of fossil energy is used routinely to manipulate climates of the spaces that we occupy. If, however, the temperature gets very far from our preferred temperature, people will respond strongly, becoming agitated, work less effectively, become very unhappy, and greatly increase their efforts to try to get comfortable. At the extreme they will die, as happened to many older persons in Chicago and France in the hot summers of the early part of this century. Similarly other plants and animals have a rather restricted range of temperatures and often other environmental conditions they can withstand. There is a rather well-developed science within the discipline of ecology that has undertaken considerable analysis of the response of organisms to *gradients* (i.e., ranges) of temperature and other factors. ◘ Figure 15.6, for example, shows the production of coffee and bananas in Costa Rica as a response to local variations in climate. What this means for economics is that each species of cultivated plant has an optimal place to grow in a country, and once these areas are planted, it is much more difficult to make a good profit on the suboptimal lands. What this means in terms of economics is that the areas of the world in which it is possible to make a good profit from an agricultural crop are far more restricted than all the area that a crop might grow. The take home lesson here is that the physical environment can affect economies in many ways (◘ Fig. 15.6).



◘ **Fig. 15.6** The production of coffee and bananas in Costa Rica as a response to local variations in climate. The vertical axis is precipitation, the horizontal is temperature. While coffee and bananas can grow essentially anywhere in Costa Rica, sufficient yields of coffee **a**, to make it economic are found only in the central circle and likewise for bananas **b** (Source: Hall 2000)

The climate of the Earth is extremely varied, something that might not be quite so apparent to those who have not traveled far or simply hop from one air conditioned airport or resort to another. Most obviously the tropics and subtropics are warmer, or at least they are at low elevations. More importantly, for life, the temperatures there vary less over the year, and most of these areas do not freeze, a critical issue for many plant species. Less obviously there are many areas in the tropics that are very cold. These are at high elevations, and since there are many mountains in the tropics or subtropics (including the Andes, the Himalayas, and many other high mountains), there are large areas of the tropics that are far from warm. For example, Mount Kenya and Kilimanjaro are nearly on the equator, but they have, at least for now, permanent glaciers. Although the tropics at any one location tend to have very little temperature variation over the year, they tend to have much greater rainfall variation, especially in the subtropics (◘ Fig. 15.7). Temperate areas often have much more regular rainfall but greater extremes over the year in temperature. As you go toward the poles, obviously, it gets colder yet and the sea-

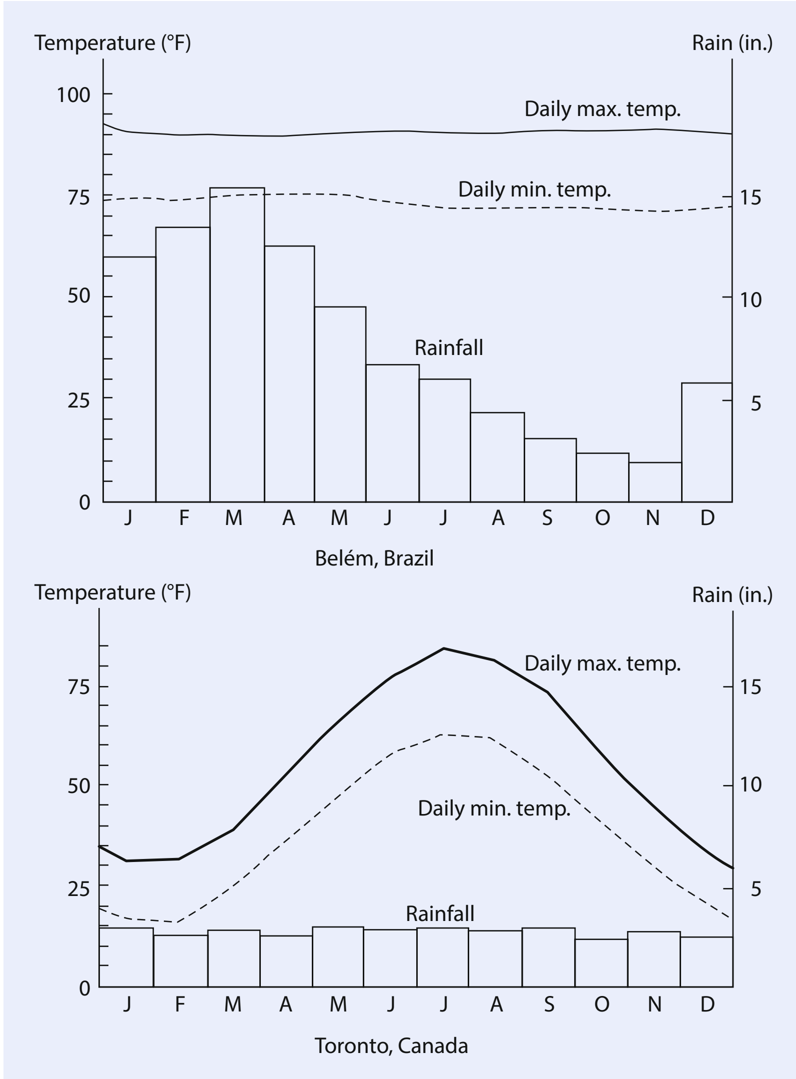


Fig. 15.7 Variability of temperature and rainfall in a typical tropical location (Belem, Brazil) and temperate location (Toronto, Canada) (Source: MacArthur, 1972)

sonal variation in temperature more extreme. Water in general and oceans in particular are much harder to heat than land. We say that water has a very large *thermal mass*. Land areas that are far from oceans or great lakes tend to have *continental* climates, that is, they get warmer in summer and colder in winter than land areas near the water which have what we call *maritime* climates. The West coast of the United States has less temperature variation than the East coast because the winds tend to blow from the west, bringing oceanic influences onto the land. Likewise areas downwind from or close to large lakes have less temperature extremes so that, for example, many wineries are associated with large lakes.

15.3.15 The Hydrologic Cycle

The hydrological or water cycle is closely related to the climate and it, like most things on this planet, depends upon the sun. Solar energy enters our atmosphere and about half of it reaches the Earth's surface where most of it is converted eventually to thermal energy. But first, it does a great deal of work, the most important or at least largest component

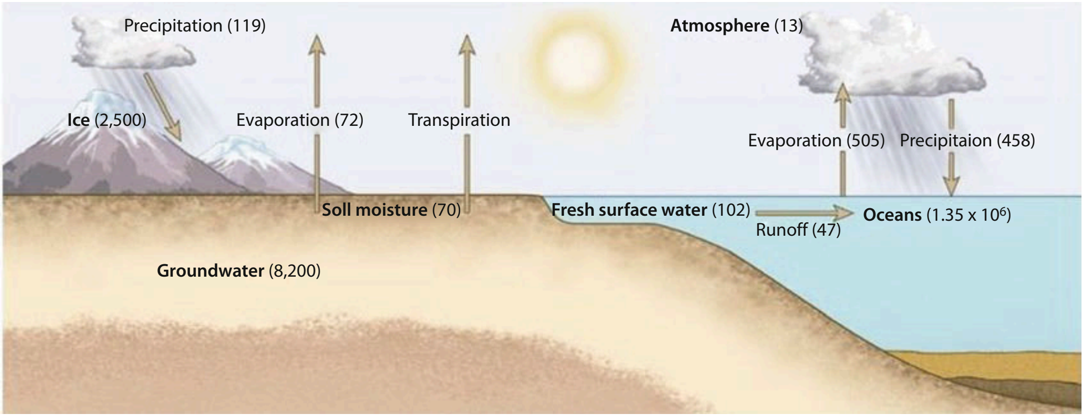


Fig. 15.8 The hydrological cycle. Water is evaporated from the land and (especially) the sea by solar energy, carried to land areas by winds where, if temperatures decline, it is deposited on land. From there, it is evaporated or runs to the sea in river (Source: Kaufmann & Cleveland 2008)

of which is evaporating water. As we will return to from time to time, the proper functioning of the hydrological cycle is probably the most important component of the economy, although it is hardly mentioned in most economic analyses. This issue is increasingly of more than academic interest as water quality and water quantity are increasingly changed and diminished through human activity such as groundwater depletion, deforestation, pollution, and global climate change.

The freshwater cycle starts in the ocean where most of the planet's water resides and most of the planet's evaporation takes place (Fig. 15.8). Evaporation is a truly amazing process that purifies water (since salt and pollutants essentially do not evaporate) and lifts the water into the atmosphere, often higher than the highest mountains. This very pure water then falls onto the Earth's surface, especially over land and most importantly over mountains, providing soils, rivers, lakes, ecosystems, and people with clean water. All of this is very energy-intense work, and it is done completely by energy from the sun. The reader can get some idea as to how much work nature does for us through the hydrological cycle by considering what people living in New York City would have to do to get their clean water were it not done for them by the sun. They would have to go to the Atlantic Ocean, say at Jones Beach, dip out two pails of water and somehow remove the salt. Probably the easiest way would be to build a fire and boil the water, collecting and condensing the steam that would be given off. Then that purified water would have to be put back into the buckets, and you would have to start hiking to, and then up to the top of, the Catskill Mountains, where you would empty your buckets into the streams there. Even if all of the people that lived in New York City did this, it would be but a trickle compared to what comes out of the actual rivers. Nature does a great deal of work for us and for our economy, and we must respect that. This work, however, rarely enters into the economist's calculations because there is no money involved.

The fundamentals of the water cycle can be seen in Fig. 15.8. Water evaporates principally from the sea, travels about the Earth as clouds and also invisibly in the atmosphere, propelled by winds, then falls onto the earth where it is held in the soil and then, if it does not evaporate, travels underground to rivers, and returns to the sea. Rain is more abundant near oceans, especially downwind from them, and at higher elevations due to the *orographic* effect. When air masses are lifted up a mountain, or more usually pushed up by winds, the air cools (Fig. 15.9). Cool air has less energy and so can do less work, including the work of keeping water molecules in suspension. The net result is that more water falls from the atmosphere in mountains, especially tall mountains. As the air masses move over the mountains and descend the other side, they warm and can hold more water, especially as most of the water that once was in the air masses was lost on the

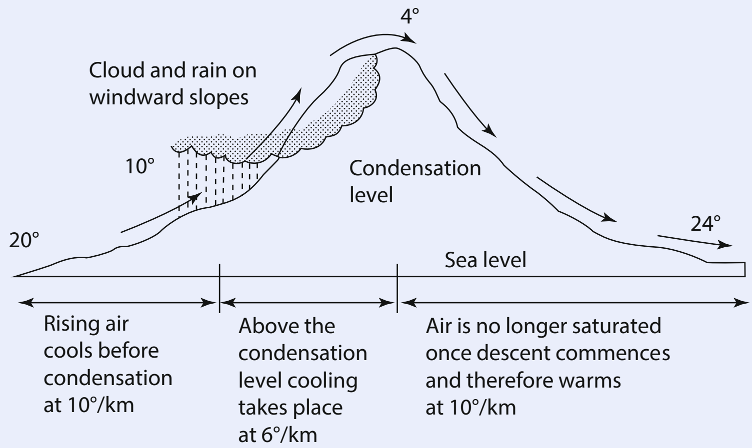


Fig. 15.9 The orographic effect. As air masses are pushed up mountains the air cools and loses energy so that it can no longer keep water molecules in suspensions, and rain occurs. As the air descends on the leeward side of mountains, deserts are created by the dry air (Source: MacArthur, 1972)

windward side. Thus, there tends to be a *rain shadow* downwind from mountains. Most of the rain, both in the mountains and elsewhere, falls onto the ground and filters slowly downhill underground. In general, all of the soil under our feet contains water which is mostly flowing very slowly toward the sea. Some soils can hold considerable water because there are considerable spaces between the soil particles. Gravel and sand hold water much better than silt or clay. Any water-holding belowground substance is called an *aquifer*. The depth below which the soil contains water is called the *water table*, and if you dig a hole to a little below that depth you can have a *well*. Where water flows naturally from where the ground level intercepts an aquifer, we call it a *spring*. More generally where the surface of the ground is below the water table, we find a *river*. Where rivers are dammed by some natural process, such as glacial debris, volcanic flow, or a beaver, we have a *pond* or *lake*. When this blockage is by human activity we call it a *reservoir*.

Rivers do a great deal of work. Since moving water has considerable energy, it can erode and hold in suspension many particles. Very fast water can move boulders, fast water can move gravel, and medium-velocity water can move sand, but slowly moving water moves only silt. Rivers erode landscapes, making valleys and depositing particles alongside the rivers when there are floods and some of the river water slows down outside the river channel. When something is moved by a river, it is called *alluvial*, and the general word for areas next to a river is called *riparian*. Steep upland areas are called *erosional* areas because the action of the river erodes away the rocks there. Where the river slows down in flatter sections, usually downstream, we find *depositional* environments. Hence, riparian or streamside soil tends to be especially fertile for both natural vegetation and for agriculture because new soils are made frequently from floods, and they are also especially important areas for wildlife and even a source of food for many fish. The rain drives the levels of rivers, and the rain varies a great deal. Consequently, the level, width, and volume of flow of rivers vary enormously. Small floods occur yearly, moderate ones at decadal levels, and large ones less frequently. A relatively small flood is a once in a decade flood, a larger one a once in a hundred year flood. A once in a thousand year flood is possible next year, but very unlikely, with a probability of 0.001. A once in 10 million year flood can help to seal off an oil-forming deposit of organic material. Rivers travel through *flood plains*, which are obvious when you look at a topographical map or a river from the right place. Rivers meander back and forth across the flood plain over time, and often shift their position entirely. Much of our societal infrastructure is destroyed each year because people do not respect that eventually rivers will flood floodplains. Misguided Federal flood insurance has encouraged people to live where they should not. An interesting, comprehensive plan for reconsidering how we manage the floodplains of the Mississippi River is given in Mitsch et al. [16].

15

The economic value of the various parts of the hydrologic cycle is immense—even incalculable. Most importantly, it provides rain for our agriculture and rivers to bring water to cities, to industries, and to irrigation. Most of the things that we make require very large amounts of water. Rivers also build soil when they flood in the spring, and the reduced energy of slower flow allows suspended particles to fall out on *floodplains*. It is no accident that human agriculture first started in such fertile riparian regions as the Tigris-Euphrates rivers of present-day Iraq, the Indus river of India, the Yangtze river of China, and the Nile river of Egypt. Another particular advantage of riparian soils, both in the past and also today, is that while most soils over time tend to wear out through erosion and nutrient depletion, the yearly flooding of most natural rivers builds new, fertile soil each year. When rivers are dammed, there are many obvious gains (hydroelectricity, water for irrigation) but also many costs. The costs include the burying of fertile soils under the reservoir and the cessation of the soil-regenerating processes below the dam because the particles tend to sink into the still, low-energy waters of the reservoir and hence are lost from the river.

Natural ecosystems, such as the forests that cover the Catskill Mountain watersheds that are the water supplies for New York City, also do work for the human economy because they clean and purify water as well as regulate stream flow, reducing the flooding potential. Forests and grassland soils and aquifers absorb some of the excess of a heavy rainstorm and then release it slowly over time. Where forests are cut, the water cycles tend to be disrupted and humans must use more energy and more money to correct for these problems, as is also true with river pollution. Rivers will always eventually go where the forces of nature dictate. Humans invest huge amounts of money and energy trying to keep rivers where they want them, but it will always be temporary. Smart economists and smart people more generally will understand what nature will do eventually and will build accordingly. Arrogant ones build many houses where they should not be. If you want to live on the edge of a river or the sea that is your business, but remember there may be a price. The US Federal Government recognizes this and is wisely removing flood insurance from places where human structures do not belong.

Humans have tended to exploit, and often overexploit, whatever water supplies they can find. When there are few people in a region, water is taken from streams or a well, but if over time, more humans move in often the river becomes polluted or the well is pumped dry. Later, people have to go after more expensive water. The city of Los Angeles is a great example. The early explorer John Fremont said of Southern California that it was a lovely spot, but there never would be very many European-Americans living there because it was simply too dry (although many Native Americans were doing just fine there using the relatively small natural rivers). That was changed, however, when the larger rivers in Northern California were diverted through canals all the way to Southern California, allowing the great city of Los Angeles to be developed in a near desert. Water also was diverted from the Owens River far to the East in California and eventually even the Colorado River, several states away. All of this water allowed not only the existence of Los Angeles but also much productive agriculture in Southern California. What is less talked about, however, is the costs of diverting that water, for example, destroying the once very large salmon fisheries of Northern California, causing San Francisco Bay to become much more saline with many adverse effects and completely drying up the Colorado River so it never makes it to the ocean. How do we weigh the costs and the benefits? We will talk about that later. What is clear is that often different people get the benefits and get the costs.

Humans have continued to exploit, develop, manage, pollute, and otherwise influence the natural water supply. Presently, water is an extremely serious issue for much of the world's population. Two especially difficult issues are that human population growth is often greatest where water is least available (such as in the Middle East) and the potentially disastrous effects of climate change. In general, these extremely important issues, or indeed the economic benefits that are a consequence of a well-functioning hydrologic cycle, are not included very well, or often at all, in economic analyses. This is because we do not pay in our markets for the work of nature, but rather just for our cost of exploiting nature. Water is often considered a "free good" and for those who measure the value of things by their price as having little value.

15.3.16 Climate Change

These climatic issues have very large implications for economies, including the crops that can be grown or not grown and the rate of production of those that can be, the amount of energy needed to keep people comfortable, the availability of water over the year, and so on. Of increasing concern is the degree to which climate is changing or may change. So the first question might be: will the climate change? That answer is easy, yes, certainly, for it has always changed! But although climates have always changed, and presumably always will due to natural causes, natural selection has prepared both humans and their important plants and animals for only a relatively small range from within the possible temperatures, soil moistures, water levels, and so on that exist or might exist on the Earth in the future. The reader may have been exposed to various points of view as to whether the climate is changing and what the effects might be. So our second question is: is the Earth warming? Again here there is little disagreement among most environmental scientists: the Earth is indeed getting warmer, glaciers are melting as the first author has seen again and again with his own eyes, the polar ice is probably shrinking, the temperature of the sea and probably the land is warming, and many areas seem to be getting drier.

Our third question is much more difficult: is the present climate change a function of human activities such as putting more and more carbon dioxide into the atmosphere? The answer is probably, and in the minds of very many scientists most certainly. In particular, many of the changes mentioned above are credited to the "greenhouse" effect, the idea that the increase in atmospheric CO_2 caused by the burning of carbonaceous fossil fuels is causing an increase in CO_2 in the atmosphere. What is the greenhouse effect? This is the process where atmospheric gases, principally water vapor and carbon dioxide (CO_2), but also methane, nitrous oxides, and other gases, act as a one way "blanket," allowing high energy, shortwave radiation (i.e., photons from the sun) to penetrate the Earth's atmosphere to a greater degree than lower energy, longwave heat can leave. When the photons strike the Earth's surface, they are transformed to heat (according to the second law of thermodynamics). Since this heat is trapped to some degree by the greenhouse effect, the Earth warms.

The initial lines of argument that said that the Earth was likely to be heating due to human economic activity were theoretical and went back to the great Swedish Chemist Svante Arrhenius who noted the property of CO_2 to absorb thermal energy in the laboratory in the 1880s. He reasoned that since the burning of fossil fuel generated CO_2 that it would inevitably lead to a warming of the Earth's surface. Further logical evidence came from planetary scientists who found that the temperature of the Earth was about 30 degrees centigrade warmer than it "should" be as determined by the position of the Earth relative to Venus and Mars. In other words, the Earth was a little too far away from the sun to be as warm as we are (based on our neighbor planets).

There are at least four main lines of empirical argument that show that the climate is changing: (1) the surface of the Earth is getting warmer, as revealed by thermometers; satellite surveys of, for example, temperatures and polar icecaps; and most critically the temperatures of both deep wells and of the ocean itself (which are very hard to heat!), (2) glaciers and tundra are melting all around the world, (3) many plants and animals are moving poleward and plants and rocks are appearing on the South Pole land mass that have never been previously observed by humans, and (4) the upper atmosphere, robbed of some of its heat from the Earth's surface, is cooling, something that was predicted by climate models before it was observed. Initially, real measurements of temperature change were difficult to interpret, and in the 1960s, temperatures actually seemed to decline! What we understand now is that industrial fuel processes do at least two things to the atmosphere: they increase the CO_2 *and* they release dust, especially sulfate particles, which reflect sunlight and cause a cooling. But the dust settles out in roughly 2 weeks, while the CO_2 is cumulative, that is, once it goes into the air, it stays for a very long time. By the 1980s, the CO_2 effect (in both models and reality) became more powerful than the dust cooling effect, so that the temperatures of the Earth have continued to set new records, more or less year after year and decade after decade. A majority of climate scientists attribute these signs of a warming Earth to the heat-trapping effects of the CO_2 (and water vapor) in our atmosphere. Starting in about the 1970s, computers began to be large and fast enough to run global climate models, and these showed again and again that if we kept increasing CO_2 that temperatures would rise. Many difficulties remain, such as understanding how water vapor and clouds might change, but the trends are clear.

The majority of scientists who work on this problem believe that it is the human-caused release of CO_2 and other "greenhouse gases" that is responsible for the global warming that we have observed. But because the Earth warmed considerably 12,000 years ago as we came out of the last ice age (with no help from human release of CO_2), there are some who say that the warming we are seeing today is just a continuation of that process. Perhaps, the Earth is still responding to whatever caused those changes. Important drivers in this long-term glacial cycling process are thought to be Milankovitch cycles, relating to the distance and tilt of the Earth to the sun, which tend to be repetitive on three very long timescales, changes in solar output (associated with sunspot activity), or something else. The arguments between these two groups are often extremely acrimonious. Thus, we come down on the side that the observed climate change is caused by industrial activity but acknowledge that the case is not quite as air tight as many would like it to be.

15.3.17 How Climate Change Can Affect Human Economies

If in fact global warming continues, the impacts will not all be bad, for example, the movement of many fish species northward in the Northern Hemisphere benefits, for example, Alaskan salmon fishermen at the expense of Oregon salmon fishermen. But overall the effects are expected to be overwhelmingly negative for most economies around the world. For example, Rind [17] predicts that huge areas of the tropics will suffer from serious drying of soils. Considerable information exists that suggests that many tropical and warm-climate diseases and pests are moving Northward in the United States. The Atlantic Ocean is measurably warming, and, because the heat in oceans is the source of energy that fuels hurricanes, the warmer the ocean (probably), the more powerful and possibly frequent the hurricanes and the stronger the hurricanes, the greater the damage to many coastal economies. Bark beetles are moving north in the Rocky Mountains with devastating results on forests because the winters are no longer severe, many birds and ocean fish are moving northward, and Australia and Africa are seeing prolonged and unusual droughts. There are ways that this climate change can enormously impact entire regions and countries: entire cities and island nations, such as the Seychelles, and the Maldives may disappear under the waves as the sea level rises with glacial melt and thermal expansion of oceans. This would displace millions of people inland to regions already stressed by excess populations. Many of the world's great cities in South America and Asia are completely dependent upon the summer melt of glaciers to supply water during that part of the year, and glaciers and sometimes their flows are declining. For example, the glacier that supplied warm-weather water to the city of La Paz, Bolivia, finally disappeared in 2009. These various impacts are clearly occurring now, with some severe economic impacts at this time. The economics of stopping or reversing global warming is overwhelmingly huge, but the consequences of not dealing with it are potentially more serious [18]. If the majority view is correct, then we must make enormous investments into replacing carbonaceous fuels with solar or nuclear power or suffer the consequences. If, on the other hand, the minority opinion is correct or, to further complicate matters, if there were a great increase in the number and severity of volcanoes that throw dust into the stratosphere or a reduction of solar output, the climate could become cooler. The likelihood of this occurring is very small, but the impact is potentially very important. Then it would be a poor use of our resources to change so quickly to expensive and intermittent solar energy sources. What a dilemma! Clearly climate is a very complex and important issue!

Our view is also that making our new energy investments in solar rather than fossil fuel is probably justified for other reasons too, including long-term energy availability, economic and national security issues, making jobs at home rather than abroad, making communities more self-reliant, and protecting the ocean from acidification and the land from the mercury that is released by burning coal. But we also believe that a conversion to mostly renewable resources would be extraordinarily difficult, would require a large proportion of our remaining fossil fuels, and would possibly greatly reduce societies' EROI. The full accounting has yet to be done, and this is a critical area for the application of biophysical economics (see ► Chap. 23).

15.4 The Biological World

15.4.1 Natural Selection and Evolution

We now turn in our quest for the basic science needed to understand economics from the physical world to the biological world. We start with a further consideration of natural selection and evolution. All of life is the product of relentless natural selection operating on our ancestors for millions and even billions of years. This evolution has been a complex process that has resulted in the immense diversity of life as we know it and also our own genetic makeup. It has large elements of chance: will a meteor's path, set by some cosmic forces perhaps a million or billion years ago and light years away, intercept the Earth's orbit or not? Is that meteor large enough to cause a tsunami that wipes out half of Tokyo or an even larger one that might extinguish major components of life? This almost certainly happened some 55 million years ago when, apparently, a huge asteroid struck the Earth, probably near Yucatan, and a large number of species went extinct. These elements of chance have operated in many, many ways and often under what seems to be quite peculiar circumstances. The opposable thumb with which I carried the computer to the table and the stereoscopic vision I am reading the words on the screen are almost certainly an artifact of our ancestor's arboreal existence extending perhaps some 4–20 million years ago.

15

But evolution is nonrandom as well, for we know that a process called *adaptive convergence* generates similar-appearing and similarly adapted plant and animal species in, for example, the different deserts of the world even when starting from completely different raw genetic materials. In each environment of the Earth, there are problems that have to be solved, and only so many ways (thick cuticles, spiny defense of water reserves, and so on for deserts) in which that can be done well. Thus, many different species "converge" in the ways that they solve the problems imposed by a particular environment. For example, the similar-appearing desert plants in Southern Africa and Southern America were derived from very different genetic stocks, Euphorbs and Cactaceae, respectively.

This similarity in life form and function in similar environments is the case in part because the material building blocks available nearly anywhere in nature tend to be the same: the element carbon is especially useful because its valence structure leaves four locations on its outer shell where other atoms can be hooked. Only silica of the abundant elements has this possibility. Carbon has been selected by organisms because it is abundant, lighter, and hence less energy intensive to use as a basic structural material. Nitrogen is also extremely abundant, making up some 70% of the atmosphere and roughly 3–7% of most life excepting water. Its abundance and special properties have been exploited by organisms through evolutionary time for the construction of proteins. Proteins are especially important for life because of their specificity, meaning that the available locations on their outer electron allow for the construction of many complex and very specific compounds. But there is a hitch: atmospheric nitrogen occurs as N_2, and N_2 is characterized by three chemical bonds holding the two atoms together. (Carbon dioxide, e.g., does not have this characteristic.) Thus, there are only a very few groups of organism, essentially the "nitrogen fixers," found only within the "primitive" groups bacteria and blue green algae, that have evolved the energy-intensive means to split the triple bonds and make the nitrogen available for use by these organisms. All other organisms, each of which needs relatively large quantities of nitrogen, get it indirectly from the activities of these two groups. Thus, part of the reason for evolutionary convergence is the relatively limited raw materials from which it makes sense to use for construction and partly because the problems that all life must solve are similar for similar environments.

For example, all around the world, trees must stand up, be anchored, exploit mineral resources form soil, and fix carbon through photosynthesis. This leads to the observation that all around the world trees look basically the same: they have trunks, roots, and leaves to solve the above problems. Where water is rarer, the approach of grass-like organisms works better, and so on. So although evolution is unpredictable, due to the importance of random environmental events and random mutations, to some degree, it is comfort-ing to the experienced biologist that there are many common problems that life in different, or even the same, environment face, and many common ways to solve these problems.

A large component of the randomness that occurs within evolution occurs because of the randomness of environmental change, at least it is random from the perspective of the organisms that are affected. The Earth, which seems more or less stable from the time and space consideration of several human generations, is actually an extremely dynamic and unpredictable place if you wait long enough, with frequent climatic excursions that put very difficult stresses on the organisms that are adapted for the previously normal conditions at each location on the planet. So-called Milankovitch cycles that result from the eccentricities and wobbles of the Earth's orbit relative to the sun are but one of the main forcing functions. The point is that the ecological "theatre," that is, the environmental milieu within which the evolutionary "play" takes place, is a dynamic and changing place, requiring organisms to adapt to those changes, migrate, or die.

15.4.2 How Does Natural Selection Work? The Ecological Theatre and the Evolutionary Play

Charles Darwin made the fundamental observation that populations of reproducing organisms tended to generate many, many more offspring than were necessary to replace the parents. There are three properties of the world that, if true, would necessarily lead to a world in which natural selection must operate. These three properties are first, that there is variation among the genome of a given species (you can see that this is true for our own species simply by riding a bus or teaching a class, especially in any metropolitan area in the United States); second, that these variations are to at least some degree passed from one generation to another (you can see that is true by simply observing that human children tend to look reasonably, but not perfectly, like their genetic parents); and third, that this variation leads to differential survival and reproduction, that is, that from among the variability, some properties of organisms will be more likely, however slightly, to lead to organisms that are more successful at reproducing. The latter is far more difficult to observe today because there is relatively little mortality among children today, especially when compared to, say, one or more hundred years ago when the majority of children would die. Nevertheless it is obvious that a faulty immune system or a less robust physique or physiology can certainly work against the survival and eventual reproduction of people even today, and it would be much more important if medical interventions were not so prevalent and generally successful. It certainly is operational to a great degree for wild plants and animals.

We will examine some additional evidence for this third proposition below. But the logic of this argument is overwhelming: if these three properties of the world are true, then natural selection *must* occur. To our minds and that of most biologists, the evidence is overwhelming and accumulates every year as we find more and more "missing links" in the fossil record, as we watch natural selection work before our eyes as agricultural pests and human pathogens acquire resistance to our once-trusted tools of pesticides and antibiotics in a way that is straightforwardly explained by simple Darwinian selection, and as scientists who study the design and behavior of organisms operating in nature see that those designs and behaviors consistently fit Darwinian predictions. The net result of natural selection has been evolution of life over time and the natural world as we observe it today, including ourselves. If there is a deity that has been in some way responsible for all of this (and science by itself is not equipped to make a judgment on that one way or another), then it is clear that that deity both operates at least most of the time through or in concert with natural selection or that he or she has gone to an enormous amount of trouble to lay down the fossil record, adjust radiocarbon dates, and so on to make it appear that evolution according to Darwinian principles has occurred. If that is the case, one then wonders why. For ourselves and most scientists, it makes far more sense to simply accept the Darwinian explanation. But we recognize that a large portion of our potential readership and many of our friends are strongly religious, and we do not want them to close the book now. Although we personally are not particularly religious, at least in the conventional European-American way, there is nothing inconsistent to us with religious faith and the scientific explanation of the universe. In fact the outline of the creation of the world given in Genesis is pretty close to that as we understand it from science—if we assume that God has long days, and why not? It is time that those arguments are put to rest and we get on with the tremendous problems facing the world.

Natural selection operates on three characteristics of an organism, its morphology (shape), physiology (function, chemical, and otherwise), and behavior. Characteristics of each of these are determined by the genetic plan donated by the organism's parents and by the environmental conditions (e.g., intense exercise will make muscles larger and stronger). But the expression of genes is not perfectly straightforward, for, as Mendel showed, the expression of any characteristic may depend upon how genes from the mother and the father come together, including many issues related to dominance and recessiveness, and because many genes can determine any particular characteristic such as eye or skin color or, at the extreme, personality. We call the genetic makeup of an organism its *genotype* and its actual expression its *phenotype*. Phenotype, that genetic makeup that is outwardly expressed, is what we observe and what natural selection operates on. Thus, an important issue is that natural selection cannot operate simply and directly on genes but only more indirectly through their collective and environmentally contingent phenotypic expression. An important new discovery in biology is that we are finding that traits are not simply determined by genes for that trait but also by other "regulator" genes that turn particular "expression" genes on and off. These genes are also subject to natural selection but the net effect is to make the possibility for more rapid evolution than we had previously thought.

Throughout evolutionary time, evolution has finely tuned organisms to their environment by eliminating those genes that do not contribute to fitness, that is, survival and reproduction. But what is fit is not a constant, for natural selection is chasing a moving target. For example, Jim Brown [19] and his students have unraveled the interaction of climate and the size of pack rats in Colorado and Nevada and found that the size of the rats increased during cooler geological periods and decreased during the warmer periods as the climate cycled over long time periods. While it is clear why it should be advantageous to be large (e.g., in competitive trials for mates), it is not so clear why it should be advantageous to be smaller. These investigators found that during warm periods a large surface to volume ratio, characteristic of smaller organisms, was important for dissipating heat, so large rats would get too warm when the climate was warm. This might not kill the rat directly but would, for example, make it more difficult to forage and hence to get enough food. Without a food energy surplus, females would have a much harder time getting enough energy to reproduce and provide lactation for their young.

15.4.3 Adaptation to Biotic Agents

Probably the biotic components of the environment, including predators, pathogens, and perhaps competitors, are even more important than the biophysical components such as climate in determining the natural selection forces on an organism. These too are related to energy cost. The ultimate example is of course loss to predation, which represents a complete loss of all energy reserves. Other interactions are more subtle, and there is a cat-and-mouse game of energy losses and investments among different species throughout evolutionary time. Trees, for example, are great food for many insects. Since most trees are apparent in the landscape, they can hardly hide from the insects that want to eat them, which of course would rob them of their energy reserves and of their ability to generate an energy profit that would allow for reproduction. The evolutionary response of trees has been to generate what are called secondary compounds, for example, tannins in oak leaves, that defend the trees against most insects. But there is an energetic cost for the tree to make most of these secondary compounds, so through evolutionary time, there has been a trade-off of more vs. less natural pesticides. For oak trees, the "correct" amount of tannins seems to be about 20% of the dry matter of the leaf.

Pathogens too impose an energy loss on organisms even when they do not kill them. A particularly nice study was done by Moret and Schmid-Hempel [20] who trained bumblebees to feed on small glass spheres, which the bees mistook for pollen. When the bees were fed this diet, they would die from lack of energy in about

5 days. When the investigators infected the bees with a bumblebee pathogen, the bumblebees would survive if they had real food but would die in only 3 days when fed the glass spheres. This shows that when challenged with the pathogens, the bumblebees need to use their own energy reserves to fight them.

Finally, competitors decrease the energy flow and sequestration in organisms either by forcing the organism to invest energy or lose exploitable resources in response to a toxic material (butternut trees do this). More commonly, they reduce the light, nutrients, or food available to the competitor or increase the energy cost of sequestering it. Examples are common in any forest. For one example, it is easy to see where evergreen trees grow next to a path or clearing that branches that are shaded die (or are thrown off) sooner than branches that are not shaded. If a branch does not pay the energy cost of its maintenance metabolism through sufficient photosynthesis, it is sloughed off.

15.4.4 Ghosts of Natural Selection Past

Within each species, there is a trade-off between being well adapted to today's particular conditions and maintaining contingencies for more extreme but rarer events. An example is all around those who live in the more Northerly latitudes. The trees that live in these locations obviously must be well adapted to the conditions that exist there today. Each year each adult tree produces on average hundreds to thousands of offspring of which far fewer than one can survive. Those many young will tend to have some genetic variation among themselves, and if the region is a bit drier or wetter, warmer, or colder, subject to more or less impact from a certain herbivore, then some genetic properties are likely to be a little more frequent in future years. In any case, there is genetic selection for the tree to send well-equipped seeds into the world, for a young tree with large food reserves (think of an acorn or a beech nut) would, other things being equal, be more likely to "make it" in the world. But there is a cost too—heavy seeds tend not to travel far.

But at the same time, all of these trees "remember" the ice age, when only those trees with long-range migratory capacity (e.g., smaller seeds that could travel better on the wind, or at least fall further from the parent in a heavy wind) were able to migrate and hence survive better. This ability to migrate is well represented in present day trees in New England, for the region was entirely under ice 12,000 years ago and no trees were found within thousands of miles. And since there were at least five major ice ages, then there was a strong premium against those genetic groups that "forgot" how to migrate. Thus, there may be less selective pressure on organisms to be able to disperse their seeds widely today, but many trees retain that capacity, for once it was extremely valuable. For another example, the common salt marsh grass, *Spartina alterniflora*, is found along most seacoasts in the temperate regions. Each fall, this plant produces millions of seeds at great energy expense. Nevertheless the plant rarely reproduces through these seeds, but rather through the use of underground stems or rhizomes. Why then should the plant produce seeds? The answer is that the seeds are necessary to colonize new areas, and new areas were constantly being formed as the sea rose against the land following the cessation of the past glacial period. Thus, those *Spartina* plants that did not produce seeds were drowned out as the sea level rose, and those that did were able to colonize new areas as they occurred. With climate change again increasing the level of seas, those "migratory" genes are likely to again be advantageous.

15.4.5 The Units of Selection

Natural selection works most obviously on individuals, for individuals survive or not and those that live are obviously the only ones that contribute to future generations. Perhaps it is more accurate to say that organisms that survive and leave the most surviving offspring are the ones that are more likely to be represented in the future. Organisms are selected to do whatever it takes to propel their genes into the future. But the situation is a bit more complicated, for we have found increasingly that evolution works in complex ways. At one extreme, Richard Dawkins [21] talks of *The selfish gene* that what survives or not over a longer period of time is not the species (for after all most species that have been on this Earth are extinct) but rather genes. To Dawkins, the genes are "selfish" in that they "use" organisms and species as their temporary receptacle to carry them forward in evolutionary time. Again it is not that they are deliberately doing this through some kind of cognitive process but that there will be selection for the patterns that cause this to occur. Perhaps it is more accurate to say that from this perspective genes are molecules capable of reproducing, and they exist in populations to the degree that they are successful in doing that.

At the other extreme, there are many who argue that the units of selection are larger than the organism. The simplest and clearest example is that parents will often risk their lives for their offspring: this is obviously a behavior that has been strongly selected for. The late William D. Hamilton argued that there has been selection for organisms to look after relatives not their offspring, cousins, for example, because whereas an offspring has half the genes from a particular parent, a cousin has one quarter and so on. According to Hamilton, an organism should be willing to take on average half the risk to help a nephew or niece that it would for its own offspring other things being equal. The idea is that this is completely consistent with a Darwinian perspective of propelling one's genes into the future. A more complex situation has been argued by Robert Trivers [22]. Reciprocal altruism is the situation where an organism will do something that appears to cost it something (hence reducing its own fitness) in order to assist an unrelated organism—but with the expectation that the one being helped will return the favor at some future time. A clear example of this is a herd ungulate defending the young of another unrelated animal from a predator. Again this seems to have a clear Darwinian genetic basis with direct recompense to the genes of the organism doing the activity, and in fact all may benefit with relatively small costs.

It gets more complex with interspecies interactions, but these are very common and are generally called coevolution. The idea is that a close interspecies interaction often benefits both species. The most common example is honeybees and apple trees: the bee gets its food, and the apple tree gets pollination services. More complex examples exist where the role of a predator in regulating the numbers of a prey can keep the prey from overexploiting its food resources. The more we look, the more of these we find, but an important point is that this does not occur through pure altruism on the part of an organism but apparently only via a tit for tat where the interaction, no matter how complex, is always of direct (or occasionally indirect) benefit to the organism engaging in the activity.

Finally, the most complex issue is to what degree does coevolution occur at the level of an entire ecosystem. Anyone studying ecosystems is impressed with the apparent "harmony" of the system: although there may be important fluctuations in populations or overall structure, one gets the sense that year after year the system continues to "keep itself together," adapt to, and bounce back from incoming stressors such as variable climates or storms while maintaining and even incrementing its basic structure. Herbivores tend to keep plants in check, but not cause their extinction, dead material is degraded into soil increasing its utility for other species, nutrients are maintained within the system, predators and prey increase and decrease but not to the extremes they might be capable of, and so on. To what extent is this "balance of nature" a case of many, complex coevolutions vs. simply "every organism for itself?"

Or, perhaps, are ecosystems regulated by the principle of Le Chatelier:

$$(A + B \longleftrightarrow B + C)$$

This principle, derived in chemistry, says simply that as a chemical (or other) reaction goes forward, it will tend to be limited eventually by the depletion of the source materials that allowed it to occur in the first place or the accumulation of products. For example, plant biomass grows and grows until it has used up the nutrient inventory, and then further growth must await the death, decay, and mineralization of earlier plants. We cannot answer this question of regulation at the level of an ecosystem very well at this time, but one thing is clear: a natural ecosystem is a wonderful and mostly self-regulating thing, whatever the mechanisms that control it might be. They run themselves for free off the energy of the sun. Human-dominated ecosystems, such as agriculture, require our constant intervention and management to be maintained in the form we wish.

15.4.6 Energy and All Biology

Take a look at most wild or domestic animals. What are they doing? Most of the time they are simply eating, if they are able to, or they are trying to position themselves to eat. If they are not eating, they tend to be resting, when they must use energy for their own maintenance metabolism to fight entropy—at least when it is not the breeding season when, obviously, things get more complicated. In other words, animals tend to be either trying to gain energy while using it for necessary maintenance while trying to diminish its loss or to use past energy surpluses to breed. Plants too are spending most of their time dealing with energy: for example, they are photosynthesizing any time the sun is shining and it is warm enough, and at night, they must use some of their energy reserves for maintenance metabolism. Thus, all organisms on this planet are very much about energy every minute of every day. Humans are a bit different because food energy is (at this time in our history) so abundant and also because our energy requirements for the more usual sedentary lifestyle of today are only about half of what they were when we were more active.

15.4.7 Ecology

Both ecology and economics are derived from the Greek word *oikos*, which means pertaining to the household. This is quite appropriate, for conceptually in this book, we are talking about managing both our immediate and also our larger household, and we believe that proper management makes both ecological and economic good sense in the long run. *Ecology* refers most specifically to an academic discipline "the study of interactions among plants, animals and their physical environment within the natural or human-dominated environment" or "the study of environmental systems" [23]. The suffix "-logy" is derived from the Greek *logos* referring to a discourse. This definition is a very different from the popular or newspaper definition of ecology that emphasizes the normative or value-laden "protect the environment" or "concerns for human health" perspective and includes the perspective of values. While most professional ecologists certainly do not mind the word ecology being used to refer to environmentalist issues, and they may in fact be focused professionally on protecting the environment, most would agree that the word "environmentalist" or "environmental" is probably a better word to use than "ecologist" or "ecological" for the activist or protectionist or other values-associated perspective. This retains "ecology" for the more academic or technical one. Finally, the words *environmental scientist* refers to many different people, hydrologists, atmospheric scientists, ecologists, economists, activists, and others, who study the environment from many perspectives using the scientific method. It may refer to a person that is a pure scientist or one oriented towards advocacy or policy. We believe it very important that all people involved in studying the environment and making policy judgments based on such studies use regularly and explicitly, or at least be very aware of, the scientific method for we find that many people have very strong opinions about the environment that are not, in fact, supportable by research to date.

We love the concept of ecology as a basis for thinking about economics because ecology is about interactions among the many physical and biotic components of a section of the Earth's surface, often natural but also including all systems with varying degrees of human influence, up to and including cities. Additionally, real ecosystems are constrained by the laws of nature and the energy inputs and material circumstances of their environment, as are, ultimately, economic systems. We believe that academic ecology has suffered somewhat by being taught too often as principally a biological science with a focus almost entirely on natural plants and animals and with humans too often ignored except as a provider of insults to natural ecosystems. More accurately ecology is about the science of all environmental relations and interactions, both biotic and abiotic, including, when appropriate, humans as part of those systems. It is about *how environmental systems work*, principally natural systems but also cities, counties, and other human-dominated ones. Economic systems are very similar to natural systems in that energy must be used to exploit resources from the Earth and atmosphere and to move and recycle materials through the systems to build structures and to provide energy for maintenance metabolism to fight entropy and to reproduce individuals, cities, and all systems. Humans are dependent upon complicated interactions among many natural and economic energy and material flows.

The important ecological concepts that an economist needs to know to be a good economist are quite extensive and beyond the short coverage that we provide. But there are a few important issues that we can summarize. First, ecologists have tended to study ecology at many levels, at the level of the *individual* organism, or of a *population* of individuals, or of a *community* of different populations (i.e., of all of the species) and finally of *ecosystems*, which includes all of the living and nonliving components of a landscape or a waterscape whether natural or human-influenced. The ecosystem perspective is most useful for understanding economics. Within any of these levels ecologists tend to study the *structure* and the *function* and the *controls* of ecosystems. Structure might include the physical nature of the ecosystems (i.e., size of individual plants), the abundance of different species (or kinds of plants and animals) (collectively known as the *diversity*), the number of individuals of a species (e.g., number of white tail deer per square mile), or biomass, meaning the total living weight of a species, or of all species, again usually expressed per unit area. Function can mean the rate of energy capture from the sun, the use of energy by various components, the transfer of energy from one group to another, the decomposition rate, the way nutrients are recycled, and so on. The controls can include external or climatic controls (temperature, rainfall, catastrophic events, and so on) and internal controls (self-regulating population control, nutrient limitations etc.). Ecologists have tended to focus on these four levels in studying their discipline.

Thus, an ecologist interested in *individual organisms* may look at how individual organisms interact with their local environment, for example, at the effect of temperatures, sunlight, or plant nutrients on the growth of individual plants. Thus, we find that each species tends to do more or less well (i.e., grow, be abundant, or some other factor) along *gradients* of conditions [24] (see ◘ Fig. 15.6). As we have discussed, this climate dependence has very large implications in limiting the types of organisms that can or cannot live in different regions, for example, different agricultural crops can be grown with a good profit only where climatic conditions are rather favorable for them. Another consequence is that each general region of the Earth has only a relatively few species (at least as a proportion of all species) that can live there. One practical consequence is that as various parts of the world are destroyed for economic gain often times many species are lost because they are found nowhere else.

An ecologist using the second approach (called *population dynamics*) might look at how populations change over time and what the controls might be. There have been long and acrimonious arguments about the relative importance of *density dependent* (i.e., influenced by the density of the population being considered, i.e., self-regulation) and *density independent* (i.e., influenced principally by external factors) throughout the history of ecology, a debate that continues today. Ecologists interested in *community* ecology might examine the interactions among all the different species and populations of an ecosystem. The community approach often asks what determines the number of species collectively in a given location, and how these different species control how that ecosystem operates. Finally, ecologists interested in the *ecosystem* approach often focus on energy flow or *trophic* (i.e., food) relations. We can, for example, follow the flow of energy from the sun through the food chain of an ecosystem. Primary producers (mostly green plants) are able to capture solar energy and use that to turn CO_2 and water (with a little help from mineral fertilizing elements) into biomass. *Herbivores* (such as deer or grasshoppers) eat plant material. *Carnivores*, such as wolves or an insect-eating bird, eat other animals, and top carnivores, such as a tiger, eat other animals including carnivores. Detritus is dead plant or animal material, and *detritivores* eat, well, detritus, meaning dead organic material and the microbes within it. That concept may sound disgusting to you, but remember every time you are eating bread, cheese, most crackers, pepperoni, or beer or wine (i.e., every time you have a party), you are, essentially, a detritivore! Because ecosystems science tends to be more focused on energy than other approaches, we will go into it a little deeper here. An energy-based approach is conceptually very useful to think about evolution from a systems perspective [25, 26].

At each transfer of energy from one trophic level to another, about 80 or 90% of the energy is lost as heat, mostly for the energy that is required to support the living organisms and the growth of each trophic level. Tuna fish may require at least seven trophic levels to concentrate the energy of tiny phytoplankton into packages such as sardines or flying fish large enough to be food for a tuna. The low efficiency of transfer from one trophic level to the next (10% or so) is usually considered a manifestation of the second law of thermodynamics, although it also reflects the need for maintenance metabolism at each trophic level. Omnivores are animals such as bears and humans that eat both plant and animal material. The implications of this for economics is principally related to food chain length. Where human population densities are relatively small or agricultural production is high relative to the number or people, then people can afford to eat meat at every meal. Where people are crowded, poor and/or agricultural production is low then people must eat only plants. So, for example, although rich people in India or China may have a considerable amount of meat in their diet, the many poor people there and in many other countries must eat principally rice or other plant materials. There would not be enough plant material to afford the 80 or 90% that would be lost as heat if the food were transformed into another trophic level. Energy is also often the basis for understanding more fully evolutionary issues, as it appears that essentially all aspects of natural selection are at least in part about energy costs and gains [24, 25].

Ecologists are often called upon to help understand and mitigate particular environmental problems by studying important environmental relations among the parts of an ecosystem, including those of one species to others or of the movement of different chemicals, such as nitrogen or phosphorus, through ecosystems. These have become important issues economically in many different ways. For example, as developed above, too much phosphorus (from fertilizers or laundry detergent) tends to make many water bodies *eutrophic*, meaning excessively rich. Probably most readers are familiar with water bodies that should be blue and inviting but are instead green and stinky. Intense algae blooms are often associated with human activities and remain a large and important and often very expensive issue economically. Acid rain is another important issue related to economics. Since most of the energy to run economies comes from fossil fuel, and many fossil fuels are roughly 1% sulfur, the burning of fossil fuels creates sulfuric acid, and this then creates a condition called acid rain that has killed many plants and fish. Acid rain can also be generated from nitrogen from air when air is used to provide oxygen for combustion. Sometimes the issues bring up serious regional issues. For example, acid rain produced in power plants in Ohio has been implicated in fish kills, and economic losses associated with loss of tourism, etc. in the Adirondack mountains of New York State, and the same problem relates cause in England and effect in Sweden, where there has been a huge loss of crayfish, a very popular item in the traditional Swedish diet. In other words, the ecological and economic cost of the activity falls on others who do not take part in the economic gain from burning the fuel. This is called an *externality*, that is, a cost that is not included in the price. Fortunately, it has been shown possible to stabilize and even reduce acid rain, but again it is an expensive process. Because acid rain itself creates many environmental costs, we can say that there are large costs to not mitigating acid rain. Because we have been fairly successful in reducing acid rain, at least in the United States and Europe, we can say that this is a fairly successful example of internalizing an externality.

An important applied area of ecology that we cover here is that of biodiversity losses and more generally what is called *conservation biology*. Almost all human economic activity destroys at least some natural ecosystems, and often the organisms and even species that live therein. In about 1980, a varied group of ecologists, conservationists, and naturalists came together and pooled their different approaches to what they viewed as a global crisis: the global loss of very many species or of what they called biodiversity. There has been a great deal of effort since that has put into attempting to understand and reducing this loss. Since many species are very important for humans (e.g., for food, for pollination of plants, for the many different medicines that come from tropical rain forests, and for regulatory aspects of many ecosystems), there have been many studies of the economic importance of these issues.

15.4.8 Ecological Stability

We end our discussion of ecology with a less precise but extremely important aspect of ecology, that of stability and control. Undisturbed natural ecosystems tend to be broadly the same from year to year. When they are subjected to enormous impacts from changing weather, landslides, invasions, and human impacts, they tend to have within them a tremendous resilience or ability to spring back once the impacts are relaxed. When we study the vegetation on the slopes of Mount St. Helens, Oregon, which was eliminated when the volcano exploded in 1980, we find that the forests are being reconstructed relatively rapidly. Likewise when humans cut tropical rain forests, new forests will form within years or decades if given a chance (if the soil is not destroyed). Again, and again, we find a certain stability to many ecosystems even as they are impacted by natural or human-directed processes. We might think of nature as having a great deal of resilience. This is sometimes called the "balance of nature," although "balance" is not exactly right as there are many fluctuations. But the fluctuations tend to be, within broad limits, within certain ranges, and ecosystems tend to return their base conditions if they are left alone, at least on the scale of human lifetimes. One exception to this can be when new species are introduced that are very different from the original species, such as brown snakes on Guam or starlings on Hawaii. Because the original species have not encountered anything like these species, the ecosystems can be heavily impacted.

In contrast, human societies appear much less resilient, and as Tainter [7] and Diamond [8] point out the historical and prehistorical record is full of the collapse of once proud and dominant cultures and economies. How are these different from the much more stable natural ecological systems? A great deal of this resilience, at least compared to human systems, is that the energy sources (mostly the sun but also inputs from other ecosystems) tend to be constant and predictable in natural systems. So if and as the species change, the amount of primary productivity tends to be limited by the amount of sun and the climate, both of which tend to change little from 1 year or decade or even century to the next. Nutrients are potential limiters to plant growth, but since they are tightly recycled in undisturbed ecosystems, they rarely limit a natural ecosystem. Even floods and droughts tend to come and go within long-term ranges to which the ecosystems are adapted. Humans, through technology and their own too-clever minds, tend to exploit and then overexploit the basic energy and other resources upon which they are dependent. Of course, this leads to the great question facing humanity today: are we exploiting the Earth at a level beyond what the Earth can provide, and if so, do we have the ability to be as resilient as natural systems tend to be?

While this consideration of ecology, like the other sections in this chapter on science, has been very brief, we think it will help the reader understand many contemporary economic issues and, we hope, the need for an ecological basis for understanding economics. There is much more to be learned, and we encourage you to take additional courses in ecology and indeed in science in general, not simply to train yourself for your professional understanding but also, like art or music, to enrich your life by helping you to understand the world around you.

15.5 Is Economics Science?

This chapter has been a review of what we think are the basics of natural science and the scientific principles that we believe are important for understanding real economic systems. A question that must be in the mind of many readers is "to what degree does existing economics follow these rules of science?" We address this issue in the following chapters.

Questions

1. How have humans explained and tried to predict events traditionally?
2. Are humans part of nature?
3. Explain the difference between an independent and a dependent variable.
4. What does multiparametered mean? Can you give an example?
5. Would you, or how would you, reformulate the question: "The scientific method leads to truth?"
6. Give the steps of the scientific method.
7. How do we know when science "works"?
8. What does scientific rigor mean? Can you give five characteristics of scientific rigor?
9. Is it possible to test the theory of natural selection?
10. What are the energy sources for the Earth?

11. What work does solar energy do on Earth?
12. What is a Hadley cell? How does it work?
13. What is continental drift? Where is it occurring?
14. What is the "best first" principle?
15. What, technically, does "organic" mean?
16. Can you give the geological steps usually associated with the formation of oil?
17. What is the difference between source rocks and trap rocks?
18. What are the characteristics of the oil deposits that we have tended to find and exploit first?
19. What is the law of the "conservation of matter"?
20. What does "reduced" mean? How is that different from something that is called "oxidized"?
21. Why is it so difficult for plants to get nitrogen?
22. Who was Fritz Haber and what did he do?
23. Why is phosphorus important?
24. Explain eutrophication.
25. What is pollution?
26. Discuss some characteristics of an environment that increase the reaction rate of a chemical.
27. Why does the West coast of the United States have a more regular temperature than the East coast?
28. Draw the basics of the hydrological cycle
29. Define and explain the reason for the orographic effect and rain shadow.
30. Give an example of how natural ecosystems provide services to cities?
31. Will the climate change? Why or why not?
32. Give four observations consistent with the idea that the world is getting warmer. What are some other processes that might cause the Earth to get cooler?
33. What are three observations that, if true, must lead to organic evolution? Do you think these apply to humans?
34. What are the three general characteristics of organisms that natural selection effects?
35. Discuss the units of selection.
36. What is the principle of Le Chatelier and how might it effect ecosystems?
37. What is the difference between the usual public use and the academic meaning of the word "ecology"?
38. Why does ecology make a good basis for thinking about economics? Why might it not?
39. Discuss structure, function, and control with respect to an ecosystem.
40. What does the word trophic mean?
41. What is an externality? Can you give several examples?

References

1. Darwin, C. 1859. The origin of species. John Murray, London.
2. Angier, N. 2007. *The canon. A whirligig tour of the beautiful basics of science*. New York: Houghton Mifflin.
3. Glymore, C. 1980. *Scientific evidence*. Princeton: Princeton University Press.
4. Popper, Karl. 1934. *Logik der Forschung*. Vienna: Springer. Amplified English edition, Popper (1959).
5. Grant, Peter. 1986. *Ecology and evolution of Darwin's finches*. Princeton: Princeton University Press.
6. Schluter, D. 2000. *The ecology of adaptive radiation*. Oxford: Oxford University Press.
7. Tainter, J.A. 1988. *The collapse of complex societies*. Cambridge: Cambridge University Press.
8. Diamond, J. 2005. *Collapse: How societies choose to fail or succeed*. New York: Viking.
9. Falkowski, P. 2006. Evolution: Tracing oxygen's imprint on earth's metabolic evolution. *Science* 311: 1724–1725.
10. Hall, C.A.S., C.J. Cleveland, and R. Kaufmann. 1986. *Energy and resource quality: The ecology of the economic process*. New York: Wiley Interscience.
11. Smil, V. 2001. *Enriching the earth: Fritz Haber, Carl Bosch and the transformation of world food production*. Cambridge: The MIT Press.
12. Tuchman, B. 1962. *The guns of August*. New York: MacMillen.
13. Goeller, H.E., and A.M. Weinberg. 1976. The age of substitutability. *Science* 191: 683–689.
14. Deevey, E.S., Jr. 1970. Mineral cycles. *Scientific American* September: 148–158.
15. Carson, R. 1962. *Silent spring*. Boston: Houghton Mifflin.
16. Mitsch, W.J., J.W. Day, J.W. Gilliam, P.M. Groffman, D.L. Hey, G.W. Randall, and N. Wang. 2001. Reducing nitrogen loading to the Gulf Mexico from the Mississippi River Basin: Strategies to counter a persistent ecological problem. *Bioscience* 51: 373–388.
17. Rind, D., R. Goldberg, J. Hansen, C. Rosenzweig, and R. Ruedy. 1990. Potential evapotranspiration and the likelihood of future drought. *Journal of Geophysical Research* 95: 9983–10004.

18. Stern, N. 2007. *The economics of climate change. The Stern review*. Cambridge: Cambridge University Press.
19. Smith, J., J.L. Betancourt, and J.H. Brown. 1995. Evolution of Body Size in the Woodrat over the Past 25,000 Years of Climate Change. *Science* 270: 2012–2014.
20. Moret, Y., and P. Schmid-Hempel. 2000. Survival for immunity: The price of immune system activation for bumblebee workers. *Science* 290: 1166–1168.
21. Dawkins, R. 1976. *The selfish gene*. New York: Oxford University Press.
22. Trivers, R.L. 1971. The evolution of reciprocal altruism. *Quarterly Review of Biology* 46: 35–57.
23. Any basic textbook on ecology.
24. Hall, C.A.S., J.A. Stanford, and F.R. Hauer. 1992. The distribution and abundance of organisms as a consequence of energy balances along multiple environmental gradients. *OIKOS* 65: 377–390.
25. Thomas, D.W., J. Blondel, P. Perret, M.M. Lambrechts, and J.R. Speakman. 2001. Energetic and fitness costs of mismatching resource supply and demand in seasonally breeding birds. *Science* 291: 2598–2600.
26. Brown, J.H., C.A.S. Hall, and R.M. Sibley. 2018. Equal fitness paradigm explained by a trade-off between generation time and energy production rate. *Nature* (January 8, 2018).

The Required Quantitative Skills

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_16

16.1 The Basic Mathematics You Need to Know to Understand Economics

If you glance at an advanced economics journal, you will see that many of the pages are filled with dense mathematical equations. Quite often these equations are presented as "proofs" of some economic idea. It is really quite an intimidating experience, even for those of us with a fairly good background in quantitative analysis! How can you, aspiring perhaps to become an economist, or at least to be able to understand others' economic conclusions, do so if your skills in mathematics are limited? We are not sure, but there may be some hope for you. This is because although we believe in the critical importance of a good *quantitative* understanding of economic systems, we are not so sure as to the utility of very strong *mathematica*l analysis and skills except perhaps to understand others' work that uses them. As the wonderful economist Joan Robinson said, we should not be bamboozled by those who are hiding their simple or even ludicrous ideas behind impenetrable mathematics. If you are confused by this sentence, *quantitative* means referring to the use of numbers and basic, although well thought out, arithmetic usually relating to some kind of data and the relations among data. *Mathematical* has many meanings but in this context generally means the ability to use advanced paper and pencil mathematics, often called "mathematical analysis," "analysis," or "deriving closed form solutions," often for theoretical work related only loosely to quantitative analysis. We believe that economics has suffered from the excessive use of complex mathematics, sometimes linked to poorly formulated problems—and sometimes under the misleading assumption that the basic understanding of economics as given in most economics textbooks is always an accurate representation of real economies. We think far too many good minds have spent far too much time undertaking such "mental masturbation" when instead we should be examining much more carefully our basic assumptions of what it is that economists should be doing and how—empirically—actual economies operate. We are not alone in this view. For example, Paul Krugman, a Nobel Prize winner in economics and a very thoughtful and productive economist, said in 2009 while referring to the enormous financial crash of 2008 (and alluding to the famous poem by John Keats "Ode to a Grecian urn"):

> The economics profession went astray because economists, as a group, mistook beauty, clad in impressive-looking mathematics, for truth. …the central cause of the profession's failure was the desire for an all-encompassing, intellectually elegant approach that also gave the economists a chance to show off their mathematical prowess.

There are many reasons for our view that we have used too much mathematics and not enough quantitative empirical analysis in economics, and we give some of them below. Nevertheless we also believe that there is a very legitimate use of many kinds of mathematics in economics, although we have more than a little trouble conceiving of good analysis without also the use of quantitative analysis and data. It may seem like we are talking out of both sides of our mouths but bear with us and we can show you.

Generally all scientists and economists agree that their analyses should be *rigorous*, meaning that it is thoroughly researched and done well according to the standards of those who usually undertake similar analyses. There are at least two very different types of rigor important here, however, *scientific rigor* and *mathematical rigor*. There is often confusion between them. *Scientific* rigor refers to whether or not the formulation of a problem, such as in an equation, is consistent with the known laws and processes of nature; the problem is well understood, including which factors influence which other factors; and the degree to which the actual phenomena are accurately represented by the equations used. *Mathematical* rigor usually means whether or not the equations are solved correctly and less frequently whether they are well formulated or, in the words of engineers, "properly specified" by the use of analytic (pencil and paper) means. While for many problems both scientific and mathematical rigor are required, we find too often that there has been too much attention paid to mathematical rigor and not enough to scientific rigor. Examples of this have been given for Ecology in Hall (1988) [1] and for economics in ▶ Chap. 5.

In the past, before the invention and ready availability of high-speed computers, analytic approaches were generally required because there was no other means (besides impossibly tediously repetitive paper and pencil calculations) to solve even moderately complex mathematical relations. Hence Isaac Newton had to invent calculus, perhaps the foundation of most analytic approaches, to solve the motions of the planets around the sun. For certain mathematical relations where the functions are smooth (i.e., regular) and not too complex, such as the motion of the planets, one has a lot of analytical and predictive power through the use of analytic mathematics. This is because the analytic solutions represent the real phenomenon accurately and their behaviors are often well understood. Unfortunately today much of the reason for this overuse of mathematics is that for most people mathematics is very difficult and intimidating, so that the user of complex mathematics appears very smart (and hence his or her analysis must be right!). Additionally there is often a certain elegance in the use of analytic approaches that many people, including ourselves, admire. Thus tenure is often given to assistant professors who undertake complex mathematical analysis of equations that have little, or even rather absurd, relations to real economic systems, even over other assistant professors who might undertake less elegant quantitative analysis of some real and important issue or data. Or so we have heard.

In fact, most real problems cannot be solved using complex mathematical analysis (i.e., paper and pencil mathematics) (reviewed in [1–4]). The reason is that economics is about, or should be about, many processes that are occurring simultaneously. Analytic mathematics can usually solve more than a few equations simultaneously (think back to your high school algebra when you were taught to solve one, then two, but not three, equations simultaneously). The problem becomes more difficult when the equations are nonlinear (i.e., the basic factors are not represented by a straight line but, rather, a curved line) or when partial differential equations are required. In fact, most real economies are about many nonlinear things occurring and interacting simultaneously. If the price of one major commodity (say oil) changes, it is likely to influence many other aspects of the economy, not just one or two. So a lot of the fancy-looking mathematics has to simplify these complex real problems into simpler "analytically tractable" forms so that fancy solutions can be found through analytic means. The results may look impressive (and indeed often are), but we have to ask very carefully whether the mathematical solution is in fact representative of the real solution or rather some simplified, and hence "analytically tractable," formulation. The answer is sometimes yes, sometimes no. The good news is that there has been developed recently very powerful quantitative tools in computer models and even spreadsheets that allow people of good intuition but relatively modest mathematical skills to undertake extremely quantitative analysis of economies. But there are no spreadsheets that can test whether your concepts are accurately representing the phenomena analyzed.

16.1.1 What Mathematics Then Are most Useful?

The basic mathematics needed to provide a basic understanding of quantitative aspects of economics can be summarized using the following words: functions, linear, nonlinear, exponential growth and decay, limits, age structure, forcing functions, statistics, and calculus, and we are sure some other words we forgot to add. While of course each of these requires one or many full courses in mathematics for their real understanding, we offer a simple although critical overview here to provide a sufficiency of understanding for many, to help beginning students understand why they might wish to take further courses in mathematics if they wish to be an economist or resource analyst, and to provide a review for some. If you are good at mathematics, this chapter will, of course, be far too elementary for you, so skip it.

16.1.2 Mathematical Functions and Forcing Functions

First a few definitions: mathematics is usually about constants and variables. *Constants* never change, such as the value of pi (3.14159, the approximate energy content of an average barrel of oil (6.118 GJ plus or minus some small amount) or something we define as a constant. *Variables* can take on different values over time, for example, the GDP of the United States, the number of people unemployed, the number of salmon caught in a year, and *x* and *y* as conventionally used. Applied mathematics generally deals with numbers measured from nature or a human economy. We call a measurable characteristic of a system a *parameter*, and this includes the constants and

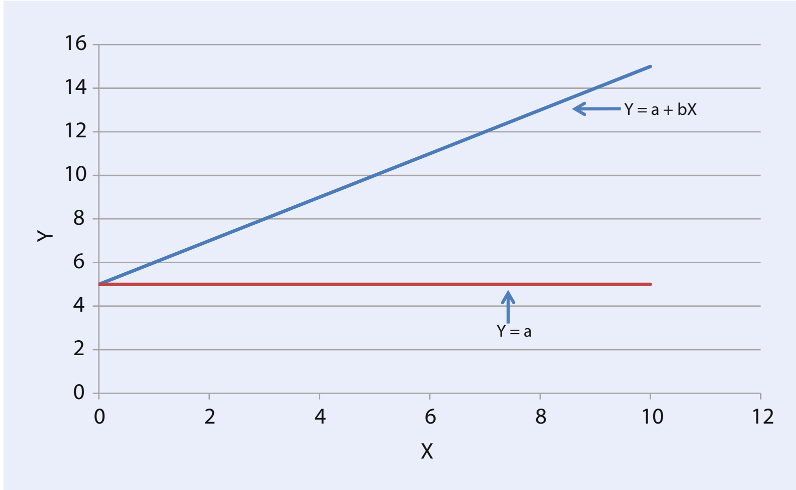


Fig. 16.1 *Red Line*: Linear with $y = a$ constant (in this case 5) and *Blue line*: $y = a$ constant plus a linear term of x

variables mentioned above. This can be confusing because we also call the coefficients in an equation (the a's, b's, and c's in $y = a + bx + cx^2$) parameters! They are just two uses of the same word and have no particular or at least no necessary connection.

Probably the most fundamental way to begin to think about mathematics is to learn about *functions*. When we use the word functions, we say that "y occurs as a function (i.e., in response to) x." Thus plant growth (y) is a function of sunlight (x), profits (y) are a function of investments (x), and so on. When we express the relation mathematically, which we usually do, we speak of *mathematical functions*. We speak of *independent* factors, usually plotted on the X (horizontal) axis, which varies independently of other terms, and the *dependent* variable which varies in response to the independent variable. Probably the basic question for all economic analysis is on a case-by-case basis: is y really a function of x? What kind of function? Is it linear or nonlinear? Will the function I have measured continue to hold into the future? Can I bank on it? What else might y be a function of?

Thus most generally:

$$y = f(x)$$

The simplest use of this equation is for a flat straight line, where a represents a constant:

$$y = a \qquad \text{(Fig. 12.1a)}$$

We can add slope to the line by making the x variable a function of the x variable:

$$y = a + bx \qquad \text{(Fig. 12.1b)}$$

where y is the dependent variable, x the independent variable, and a and b the *parameters* of the equation. In this case, y is the value of the equation as a *linear* (straight line function) of x. The parameters of this equation are a, the y intercept, and b the slope (or vertical rise over the horizontal run—sometimes an m is used) (Fig. 16.1). A more complex relation would be nonlinear (i.e., not a straight line) of which there are many kinds. (Figure 16.2 is one way to generate a nonlinear curve):

$$y = a + bx^2 \qquad \text{(Fig. 16.2)}$$

A more complicated equation is.

$$y = a + bx + cx^2 \qquad \text{(Fig. 16.3)}$$

where y is the *dependent* variable, x is the *independent* variable, and a, b, and c are the parameters. The equation with the squared term is often called the quadratic equation, and the curve it generates is called a parabola (Fig. 16.4). Another common nonlinear curve is a power curve, with a and b again parameters:

$$y = a^x + b \qquad \text{(Fig. 16.5)}$$

The list of possibilities for more complex equations is essentially infinite. These *functional relations* are normally derived by examining statistically how the two variables have been related in the past. So if plants grow more rapidly with more sunlight, we can generally assume that this relation will hold true in the future. Such functional relations work best (often almost perfectly) for physical systems, such as planetary

Fig. 16.2 **a** Growth curve, with $Y = X^2$ and $Y = X^3$. **b** Growth curve, with $Y = X^2$ + plus a cubed function

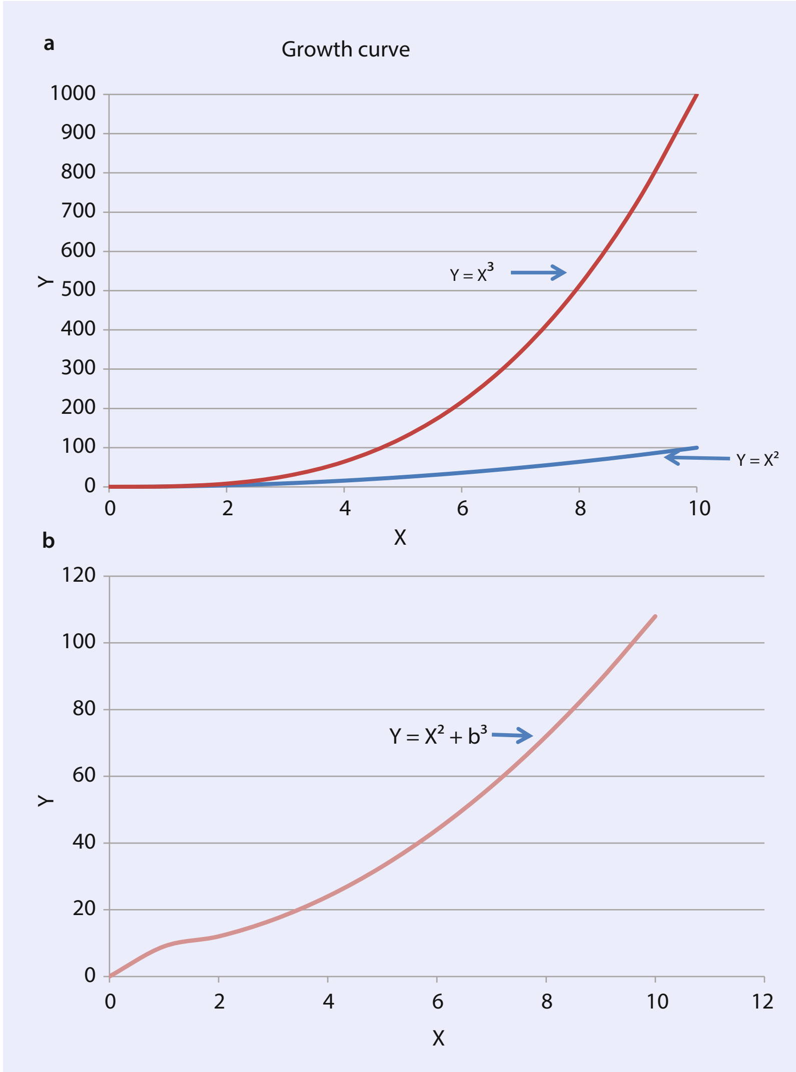
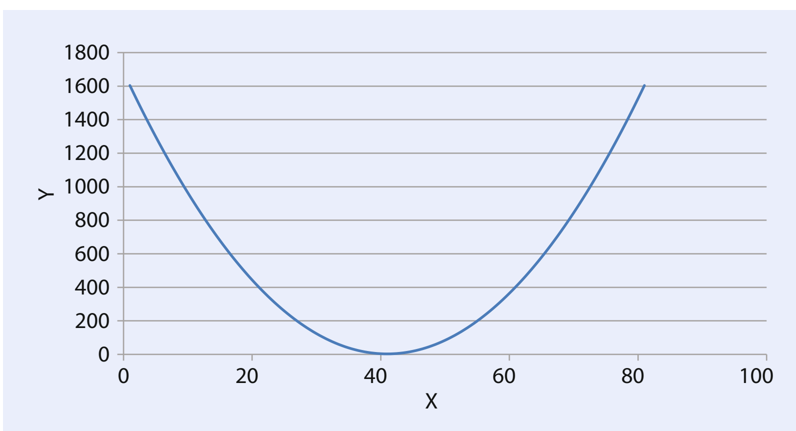


Fig. 16.3 Parabola generated from $y = ax^2 + bx + c$



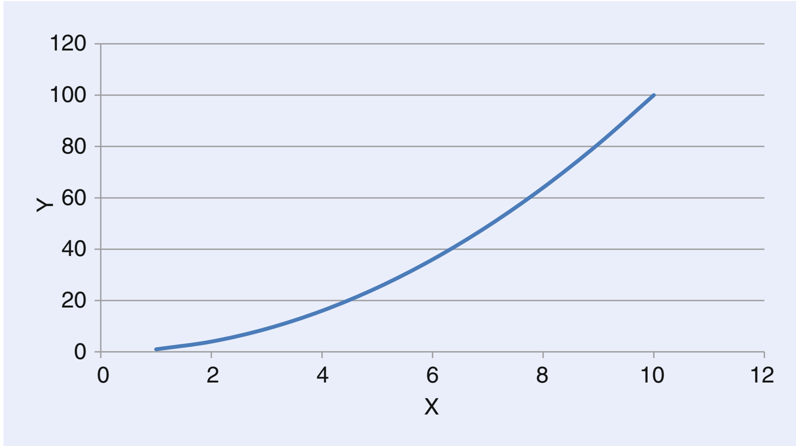


Fig. 16.4 Power curve, where $y = x^a + b$, $a = 2$, and b = zero

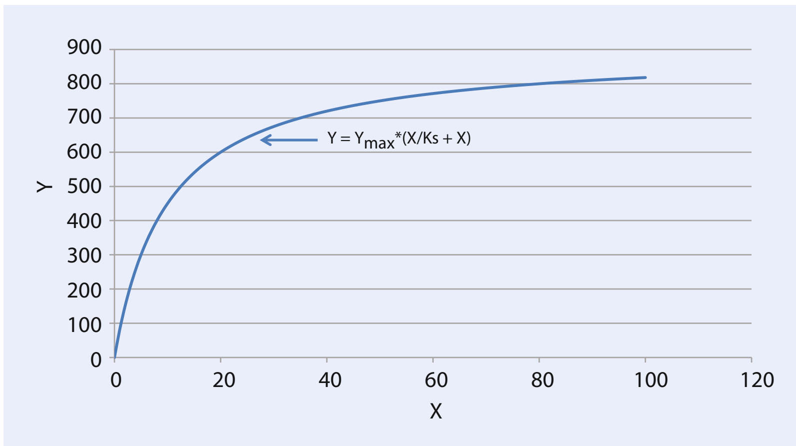


Fig. 16.5 Saturation curve (Michaelis-Menten) where $y = y_{max} * (x/(Ks + x))$, in this case y_{max} equals 900 and Ks equals 5

attractions or the flow of electrons in a wire, pretty well for biological systems and sometimes well but sometimes not well at all for economic systems. Since many, many people are interested in predicting the financial future, various mathematic equations are often used.

There can be many problems with a mathematical approach to economics, for it is not always clear that money or production or whatever the dependent variable is well represented by the mathematics used. But the "mathematical functions" approach does work often enough that many people have become obsessed by it. For example, in economics, an assumption is often made that if investments are made, then a business or output or whatever will grow. No investments, no growth. Now this growth may or may not happen (it could be a bad or even stupid investment), but it is pretty clear that if you do not make investments in, e.g., new productive equipment, that the growth will not take place. If, for example, you find that in the past in your business that the growth of your business is directly proportional to your investments so that twice the annual investments would yield twice the increase in business, then we would say the relation is linear.

But what would happen if the market saturated, that is, if there were not enough potential buyers to buy all that you produced, or if it simply is not possible to find or extract that resource? Then the relation would *not* be linear, and it might *saturate*, meaning that more of the input does not produce *any* more of the output (Fig. 16.5). Another big word for this type of curve is that the response to the input becomes *asymptotic*, or level. A good example of this is that the US economy invested a great deal more dollars into seeking oil in the late 1970s and early 1980s. In earlier times increased investments had in fact generated increased production of oil; however after 1970, the production of oil actually declined each year (and continues to do so for conventional, but not "unconventional", oil). The assumption that investments will generate additional oil simply was not so (or was at least far more complex), for the relation was dictated by geological processes not included in the original economic equations.

16.1.3 Growth: Linear Versus Nonlinear

The first formal real mathematical analysis undertaken in Western economics was based on the concept of linear vs. exponential growth. The English early economist Robert Malthus said that, due to the persistence of "passions between the sexes" the human population will grow exponentially while agricultural production, limited by the availability of good land, could grow only linearly over time. So in our discussion of types of mathematical functions, we too will start with a discussion of these two types of mathematical functions. We will use Q to represent the quantity of land being used for agriculture and the variable N to represent the number of people.

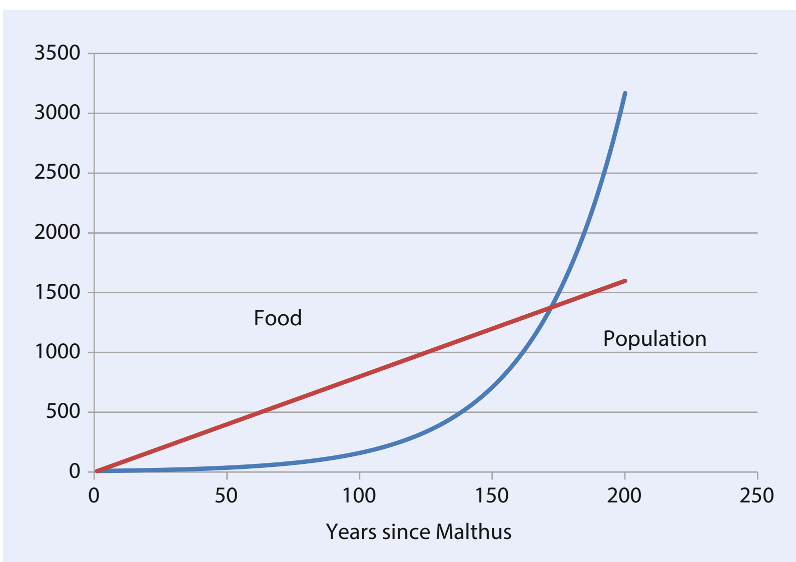
Linear (straight line) growth, or linear most anything, can be represented by a straight line or with the following simple equation:

$$Q_{\text{new}} = Q_0 + t * k$$

In this case, Q_{new} means the new quantity of something, for example, the quantity of farmland in Malthus' argument is simply a function of time and only time. Q_{new} is also known as a *variable*, meaning a number that can change its value while the equation is being solved—in contrast to a *constant*, meaning a number that never changes (such as pi). k is a growth rate, say 8 hectares per year. The variable t refers to time and goes from one to two to three as the equation is solved for one, two, or three (or more) years. Q_0 means the *original* quantity of the farmland under consideration, that is, the value before the analysis is undertaken. This is the key of it being a linear equation—we keep adding the same area each year to the area of farm. So if we had 100 hectares, then after 1 year, we would have 108, after 2 years 116, and so on. When this equation is solved over time, that is, when we solve for many years, the results will look as in ▣ Fig. 16.1, that is, it will be a straight line.

Exponential growth means that each increment (Q_{new}, or N_{new} for population number) is added to the previously determined total quantity, and since the new increment depends on that total quantity, each new dependent value increases over time. This is the common situation of bank deposits growing, in theory, exponentially through compound interest. In that case the solution, Q_{new}, will grow at an increasing rate over time as the amount added in to the quantity becomes more and more. The equation can be solved either recursively (i.e. in a computer):

$$N_{\text{new}} = k * t * N_{t-1}$$ (Blue Line in ▣ Fig. 16.6)



▣ **Fig. 16.6** Malthus: Solutions solving Malthus' linear vs. exponential equations, using approximate values for England in Malthus' time (1800) or it can be solved analytically:

$$N_t = N_0 e^{rt} \qquad \text{(Blue Line in Fig. 16.6)}$$

In this case, N_{new} means the new quantity of something, for example, the number of people is a variable that (usually) increases over time. *e* is the base of the natural logarithm (2.71828), *k* is a growth function or coefficient as before. t refers to time and as before goes from one to two to three as the equation is solved for one, two, or three (or more) years. $N_{t\text{-}1}$ means the population number for the previous time it was solved, which is *not* the same as the original value before the first time the equation is solved. When this particular equation is solved over time, that is, when we solve for many years, the results will look as in Fig. 16.2, that is, it will be a curve increasing at an increasing rate. In both cases, we can solve these equations either analytically or more commonly numerically, that is, in a computer. To do this, we write an *algorithm* (a sequence of mathematical steps) and solve it *numerically.* A simple computer code to solve these equations is given in Table 21.1. Today most complex mathematical equations are usually solved using computer models, which we introduce below.

The difference between the two equations, that is, in the linear vs. exponential, and it is plotted in Fig. 16.6. These differences are essentially what Malthus was talking about: food production would grow linearly, while human populations would grow exponentially for as long as each average family had at least three surviving children. This constant rate of increase would be applied to a larger and larger total number of families over time. In fact since Malthus' time, both the human population and food production have increased exponentially, with (arguably) food production even increasing somewhat more than the human population. The increase in food production is normally attributed to technology, which means plant breeding and better management but especially an increased use of fertilizers, tractors, and so on. Essentially all of these inputs are based on an increasing use of petroleum, of course. Thus what Malthus' equations lacked was a factor for the invention and enormous expansion of petroleum-based agriculture. Of course if petroleum supplies become seriously constrained and good substitutes are not found, then maybe in the long run Malthus' equations were right all along.

16

Exponential growth is very important in economics for at least two reasons. The first is the potential exponential growth of the human population (and hence, in an approximate way, economic activity) increases sharply over time. The second is the exponential growth of money when invested. This concept excites many people who want to make a lot of money, for the potential is huge. A sobering reality check, however, can be found from the Bible. If we were to invest Judas' 30 pieces of silver (worth perhaps \$500 today if they were the size of silver dollars) 2000 years ago at only 2%, then they would be worth

$$X = 500\, e^{0.02*2000}$$

The answer to this simple equation is about 500 quintillion dollars, far more than all the money on Earth now, which the World Bank estimates as 41 trillion dollars. A sobering conclusion from this is that on average investments on this Earth have yielded far less than 2%, which is less than the rate of inflation. That of course does not mean that you cannot do very well in the stock market, as long as the economy grows, anyway! But over the Earth's history, investments have probably failed at least as often as not.

16.2 Statistics

Perhaps the mathematical tool used most commonly in economics is statistics. Statistics are useful in many ways but most importantly:

1. To help understand the degree of uncertainty associated with a number
2. The degree to which different things are, or are not, related, that is, whether indeed y is a function of x and in what way

For examples of two above, we might want to know: is economic growth related to investments? To the number of workers? To the quantity of energy used? To technical innovations? To the exploitation of resources? Which resources? Obviously the answer is not simple. This is very difficult with economic relations. When one is trying to understand a solution of chemicals in the lab, a chemist can usually undertake an experiment with and without a particular material added to the mix to get a pretty strong answer about what does or does not contribute to a particular end product. This is relatively easy because the test and control differ by only one potentially causative variable. With economics, it is normally a lot more difficult to undertake such experiments because you are dealing with a system outside laboratory control and many things may be happening simultaneously. Nevertheless unraveling cause and effect is not impossible, although it is difficult and is increasingly being done for some issues (see ► Chap. 12). So with experiments often difficult or impossible, economists often analyze existing economies over time or compare many different economies, for example, between countries. To do this, the most useful tool generally is considered statistics, although that word covers a great mix of approaches, philosophies, and tools.

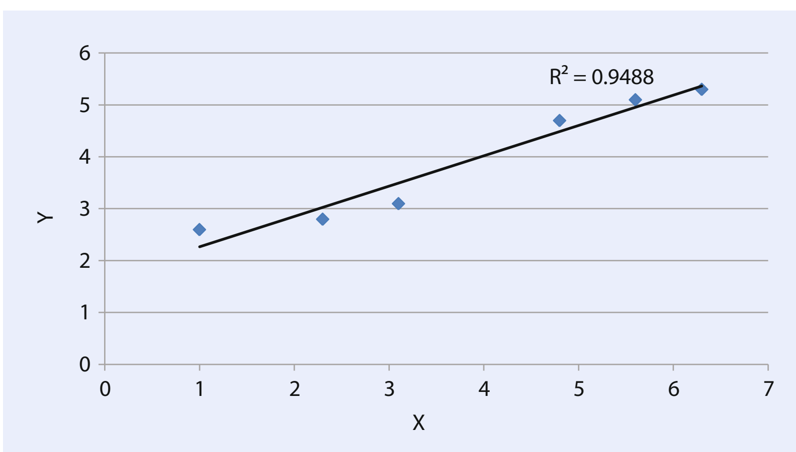
16.2.1 Correlation

Probably the most basic statistical tool is *correlation*. Correlation examines whether when variable a gets larger does variable b? Has economic growth depended upon increased energy use in the Unites States? In this case, we might consider the economic growth the *dependent variable* and the energy use the *independent variable*, independent meaning that it changes without influence of the dependent variable. Plotting the data for 1900–1984 (◘ Figs. 7.3 and 16.7), we would answer "yes it appears that it does." The relatively high r^2 implies that the two are closely related or at least tend strongly to co-occur.

But if we think about it a little bit more, we find at least two problems with what we have done. First of all we cannot say logically whether economic growth depends upon energy use or energy use depends upon economic growth! It is a chicken or egg question with no clear answer. What we can say is that the economic activity and the energy use are *correlated*, or co-related, that is, when one is high, the other tends to be high and the converse. So that is a power (and a weakness) of statistical correlation. It does not tell you something that is not true, but it does not really help you as much as you would like either for determining which is the independent variable and which the dependent or even if that is an appropriate question.

Another problem is that if we look at the relation for 1984–2005, there appears to be considerable economic growth with relatively little increase in energy use (◘ Fig. 11.4). This shows you another characteristic of statistics: what happened in the past may or may not continue into the future. (Or, as we believe, that we have not fully specified the problem, that is, there are some indications that the inflation corrected GDP has been exaggerated relative to the past (see "► shadowstatistics.org") and, of course, the United States has outsourced a lot of its heavy industry since 1984).

A further problem dogging statistical analysis is *covariance*: two parameters may increase or decrease together but in fact have little or no relation to each other. The correlation would suggest that they are responding one to the other, but in fact both may be responding to a third. For example, both temperature and photosynthesis of plants in a field tend to increase during the first half of the day, and one might conclude that one causes the other. But in fact each is responding independently to an increase in sunlight.



◘ **Fig. 16.7** Linear correlation

The issue is further confused by multiparameter issues. Ideally, we would like to be able to study one independent variable and one dependent variable. If we are lucky, we would find a straightforward relation, similar to what we see in Fig. 16.7. But what if some other factor was influencing the dependent variable? For example, we know that plants also need adequate water and nutrients. So if we want to understand, or make a model, of how plants grow, we need to untangle possible effects of each of these. If we are measuring the growth of a natural plant or one in an agricultural field, we would need to collect considerable meteorological and soil data, and probably undertake some careful field experiments, to unravel these effects, and we would then need to use multifactorial statistics to attempt to attempt to understand the influence of each one of them.

16.2.2 Econometrics

Econometrics is defined broadly as statistics on economics but is increasingly associated with analyzing how variables change over time and also testing for causality. Most of these analyses attempt to account for statistical biases that arise when working with time-series variables, and today econometrics is a large academic field with its own textbooks, journals, and so on. These techniques are often very good ways of understanding what is really happening in real economies—as long as the proper factors are entered into the equations. For example, we have been very impressed with Robert Kaufmann's econometrics examining the degree to which the United States is or is not becoming less dependent upon fuels [5] and also where greenhouse gases are going [6].

16.3 Calculus

A great deal of economics is dominated by calculus, a sophisticated approach to quantitative analysis that is concerned with the dynamics, or changes over time, of things: differential calculus with the rate at which things change and integral calculus with the cumulative effect of changes over time. Calculus was invented apparently simultaneously by Isaac Newton in England and Gottfried Leibnitz in Germany. Newton, in particular, needed to understand how to do the mathematics to understand Kepler's laws about planetary motion, and invented calculus to integrate the motion of planets to show how the arcs intersected by planets during equal time intervals but at very different parts of their elliptical orbit intercepted the same area.

While there is a great deal that you can learn about calculus in many mathematics classes, what you need to know about calculus for this book on economics is found in the next two paragraphs. How can that be, you say, when there are semester-long courses in calculus for economics and in college there is calculus I, calculus II, calculus IV, and more. Well that is true, and we do not want to discourage you from taking two or four semesters of calculus if you have not already. But we have found again and again that even if our students have had two semesters of calculus, they do not know or at least remember what calculus means essentially even if they were able to solve many homework questions when they took calculus. We know this by giving our upper division students who have taken calculus a simple calculus test, which is to draw the curve integrating a curve we draw on the blackboard and then the first differential. We also ask them to write down the relation between the speedometer and odometer in a car in terms of calculus. The students get an average of about 25% on the test, the same as at an Ivy League University where one of us previously taught. Most of our science-based college seniors cannot answer these basic questions about calculus, although they have recently passed the course. Some of course can do that and far, far more, but they are not the average. The students have been studying to the test, but in doing so did not learn the most fundamental aspects of calculus. So if you are in that category, here it is, in 3 min, how to think about what is most important in calculus.

Think of the speedometer and the odometer (the little mileage counter usually within the speedometer) in an automobile. In terms of calculus, the odometer *integrates* the speedometer, and the speedometer is the first *derivative* of the odometer (Fig. 16.10). They are inverse functions of each other. So if you drive for 1 hour at 40 miles an hour and 1 hour at 60 miles an hour, after 2 hours the *integral* of your traveling will be 100 miles, that is, you will have traveled 100 miles. Likewise if you work for 1 hour at 10 dollars an hour and 3 hours at 12 dollars an hour at the end of 4 hours, your integrated pay will be 46 dollars. The integral half of the relation is that if you have traveled 100 miles in 2 hours, then by finding the first derivative (and assuming a constant speed), your rate will have been 50 miles an hour. If you vary your speed, then the first derivative of your speed, that is, the rate of change of the speed, you will have a bit harder time deriving the integral, that is, the rate of change. That, in a nutshell, is all that calculus is about, although the essence is that calculus does these calculations for "infinitely small" periods of time. This is not so hard to grasp, for a good odometer is integrating the speedometer at each second (or less) of time, and the speedometer is showing you the instantaneous first derivative rate of integration. Of course the math and the problems can get infinitely more complicated, but this is what is most important that you need to know about calculus for understanding the essence of biophysical economics (Figs. 16.8 and 16.9).

Thus if you integrate your compound interest in the bank, how much will your 100 dollars be worth in 5 years at 10% interest? What will be the integrated cost of global warming-caused sea level rise over 100 years? We encourage you to learn much more about calculus, though, as the concept is really neat and useful. In practice the above examples can be solved easily in a computer using "finite difference" or time step arithmetic. But the answers should still be considered in terms of integrating something over time, and that is what calculus is about. And remember the calculus was invented by Isaac Newton to solve a very practical question: how to understand and predict the motion of the planets around the sun. Calculus is important because it helps us focus not only on the present state of a system but on how it is changing and what the ultimate results of that change will be.

16.3.1 What Is the Proper Use of Mathematics in Economics and Natural Science?

Part of what defines science as science in most peoples' minds (including scientists themselves) is the use of mathematics, and mathematical models, to define and resolve problems. The power of mathematics (in its broad sense) is to make the results of the prediction quantitatively explicit and hence quantitatively predictable. The process of exam-

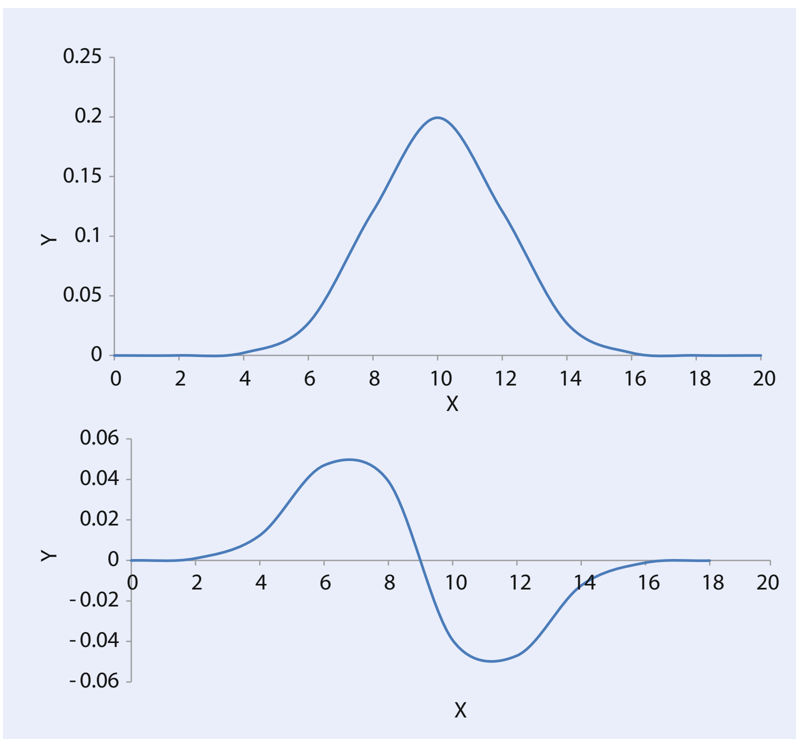


Fig. 16.8 The essence of calculus: The upper line is the first derivative, or instantaneous rate of change, of the lower; conversely the lower line is the integral, or summation of effect, of the upper

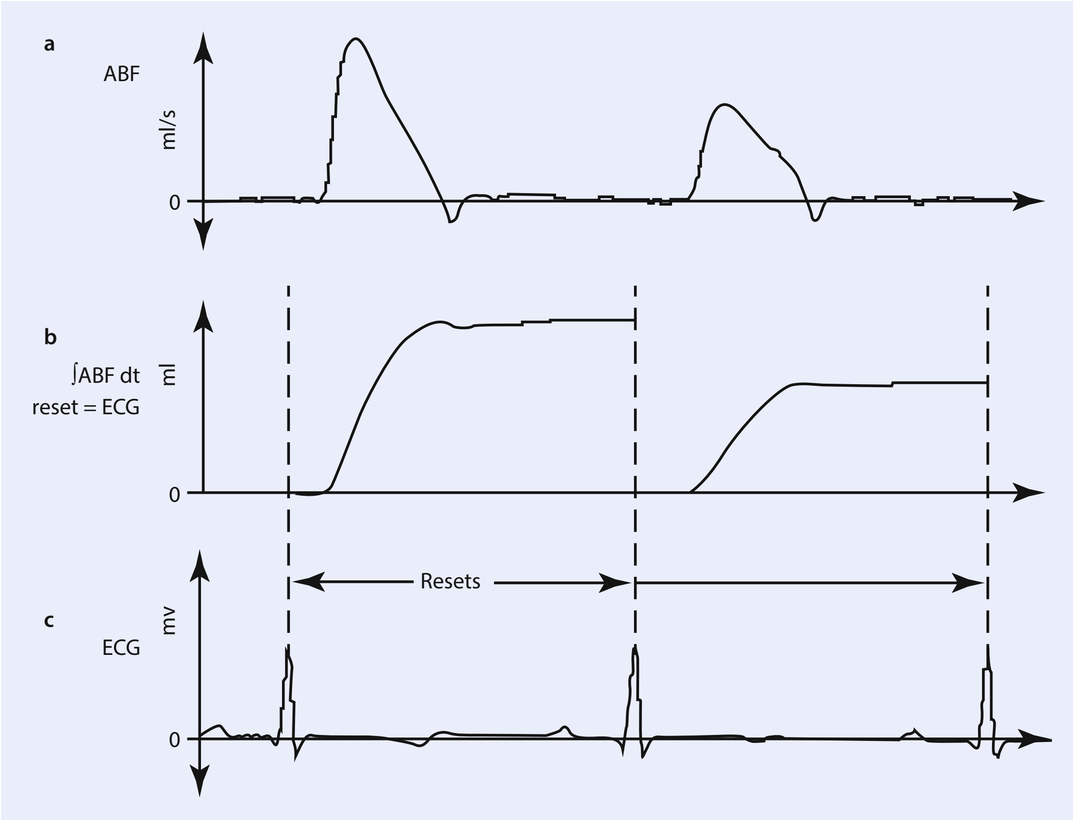


Fig. 16.9 Figure **a** is the integral of figure **b** with periodic resets (figure **c**)

ining whether your model is a correct or at least adequate is called *validation. Sensitivity analysis* is the examination of the degree to which uncertainty in model formulation (how it is structured) or parameterization (what numerical coefficients are assigned) allows one to trust your results or reach certain conclusions. It is through validation and sensitivity analysis that models generate their (sometimes) tremendous power in resolving and even predicting truth, such as that is possible and accessible to the human mind.

But the reader by now has seen our tremendous distrust of many mathematical models. What then is the proper role of mathematics in the scientific process if it is so frequently incorrect? First of all, it is necessary to distinguish *mathematical* from *quantitative*. Quantitative means simply using numbers in an important way in your analysis: three salmon vs seven. This does not require any particular mathematical skill, although getting accurate numbers may require enormous skills of a different kind. *Mathematical* means using the complex tools of quantitative analysis to manipulate those numbers or to study relations among them. It includes algebra, geometry, calculus, and so on. We emphasize here that it is our belief that it is much better to learn good quantitative methods that include understanding well the relation between the real world and the equations you are attempting to use than becoming a mathematical whiz at solving problems poorly connected to reality.

16.3.2 Analytic Versus Numeric

As we have said, there are two principle means of manipulating numbers by any of these means: *analytic* (or closed form) and *numeric* (or simulation). This is basically the difference between paper and pencil and computer mathematics. The analytic approach gives explicit and exact solutions to a rather limited set of equations for a particular point in time or set of conditions using,

Fig. 16.10 Odometer from an automobile. The red arrow indicates speed, the numbers in the middle are the odometer that integrates that speedometer over time. The speedometer is the first derivative of the odometer. This is the essence of integral and differential calculus

generally, rather complex and difficult analyses applied to usually fairly simple (of necessity) equations. Thus what you need to make *analytical* mathematical models work is really very simple systems. This is sometimes described in physics as the *two body* system (easy to solve) vs the *three body system* (very difficult to solve). Real atmospheric or real economic systems are not so simple, and pushing real systems kicking and screaming into a small enough "box" to be analytically tractable is not science. In our opinion there are very few real problems in economics that can be represented adequately by such simple relations, and much of the economics that is done by complex analytic analysis is giving mathematical but not economic results. The use of analytical mathematics, however, does have one major benefit. Through the manipulation of equations, you can transform a cause and effect relation that is stated in a way in which you cannot see the patterns you need to see into an understandable output and derive the patterns you need to understand. In other words, sometimes analytic approaches can help you visualize clearly a concept you are trying to understand.

The second, *numerical*, technique gives approximate answers to an enormously broader set of possible equations using sometimes more complex equations often arranged in complex *algorithms* (or numerical recipes) solved stepwise in a computer. In theory either method can be used to solve many particular quantitative problems, and sometimes this is done. In practice there are severe restrictions to the class of mathematical problems that can be solved analytically, often requiring a series of sometimes unrealistic assumptions to put the problem into a mathematically tractable format and the mathematical training required to undertake such analytic procedures precludes its use by many. Fortunately if one learns computer programming or even become really good with a spreadsheet, you can solve complex, multiple equations about quantitative relations that the best earlier mathematicians could not.

The use of analytic mathematics was especially important in the development of physics in the early part of the past century, and the creation of the atom bomb was tangible evidence to many of the power of pure mathematics combined with practical application. Even so, the complex fluid dynamics equations required to build the bomb could not possibly be solved by analytic means. As many as half of the Nation's mathematicians spent the summer of 1944 in Los Alamos, New Mexico, many solving the fluid dynamics equations numerically with hand-cranked calculators, something that a single good undergraduate computer student could solve now in an afternoon [8]! The success in physics of mathematics, both analytic and numeric, led many practitioners in other disciplines, including ecology and economics, to attempt to emulate, and at least give the appearance of, the mathematical rigor and sophistication of the physicists. This in turn led ecologist Mary Willson to decry many of their efforts which she said were undertaken for what she has called "physics envy" (Freudian pun intended) [9]. Nevertheless, even Einstein preferred to solve his problems without mathematics when that was possible. Other sciences in which mathematical models have been especially important include astronomy, some aspects of chemistry, and some aspects of biology including demography and in some cases epidemiology. The importance of mathematics for most of biology is a little harder to pin down. Certainly the most important discovery in biology was that of Charles Darwin, who used essentially no mathematics in the development of the theory of natural selection beyond the concept of the potential of organisms for exponential growth. Likewise mathematics by itself had little to do with the development of the cell theory, the structure and nature of DNA, and most modern molecular biology. On the other hand, genetics, from Mendel to contemporary population genetics, and epidemiology has been heavily influenced by, and sometimes tends to lend itself well to, mathematics.

But there have also been many areas in biology where mathematics has had a rather more spotty result. Population biology is intrinsically a quantitative science, and where data is sufficient (as in American actuarial tables for insurance companies), really good mathematical projections are possible. Quite good mathematical projections in laboratory conditions of, for example, one population in a homogenous environment (flour beetles or water fleas) or one predator and one prey (a predator mite and a prey mite) can also be done. The extrapolation of those results to wild populations in nature has been fraught with difficulty and related to more than a little myth making [1]. Basically populations in nature tend to be determined far more from environmental conditions that determine temperature, moisture, food supply, and so on—which cannot be so readily predicted with models—than they are to the simple population equations that make up the basis of population biology. Most of the rest of biology (behavior, physiology, and so on), while certainly quantitative, has resisted the development of any paradigmatic models.

A final problem with models is that there has been frequent confusion between *mathematical* and *scientific* proof. Mathematics can generate real proofs relatively easily because you are working in a defined universe (through the assumptions and the equations used) to which it applies. If you define a straight line as the shortest distance between two points, then you can solve many problems requiring straight lines. But the world handed to us by nature is neither so straight nor so clearly defined, and we must constantly struggle to represent it with our equations. Hence a mathematical proof becomes a scientific proof only in the relatively rare circumstances when the equations do indeed capture the essence of the problem. We all have been seeking to follow in Newton's footsteps, but Newton may have skimmed the cream from what nature has to offer!

16.3.3 So, Then, Why Is Economics, Which Is so Complex, so Analytical?

Nevertheless there remains within academia a great deal of physics envy, the desire to emulate the power and prestige of successful applications of simple analytical equations in physics. Mathematical rigor is sometimes extremely useful to solve an equation but at least equally important for impressing colleagues and deans whether the analysis has a secure connection with reality or not. In some few cases it has led to the most brilliant and important advances in all of human knowledge, such as the equations of Newton or Maxwell. However, mathematical rigor, while useful in its own right and in some applications, is hardly by itself a criteria of acceptable science, although it is often promoted as such. Thus the advanced economist is often reduced to simplifying quite complex economic questions into a format that is analytically tractable—that is, one that can be solved using analytic means. It is a lovely idea, requires enormous skills and concentration, and sometimes generates very useful results. Very often, however, we believe it generates results that represent only the mathematics and not the real system. At the extreme, Krugman [7] has said that the main reason for the financial meltdown of 2008 was that Wall Street turned its analyses from people with financial acumen over to other people with extremely strong mathematical skills. We give some examples throughout this book but explicitly in ▶ Chap. 20.

16.3.4 Now We will Appear to Contradict Ourselves

Despite all of the many problems of modeling, we do not understand how one can use the scientific method, that is, generate and test hypotheses, on complex issues *without* the use of formal modeling. This is as true for management and policy-related issues as for theoretical ones. In our view, quantitative (or occasionally non-quantitative) models are necessary in the complex world of economics (and of environmental sciences) because it allows one to apply the scientific method to complex real systems of nature and of humans and nature. *But it is critical that the right kind of models be used.* And the way to do that is quite simple: try to represent the real system that you are dealing with rather than some abstraction that happens to be analytically tractable. Quite simply most real problems require computer modeling, not analytic modeling. The power of models is to make our assumptions explicit, generally quantitative and hence testable. We shall give some examples of models that we think are pretty good later in this book.

Probably many readers at this point are confused or unhappy about our rather ambiguous treatment of mathematics in this chapter. Well, that is because the world of mathematics as used to understand real-world natural and economic problems is as ambiguous and confused as is represented in this chapter. We hope our quick tour has helped any of you who might be or become mathematical practitioners to separate the wheat from the chaff. Additionally we have given tools to all to help to be able to see through poor ideas that are dressed up in fancy mathematics.

Questions

1. What is the difference between mathematical and quantitative analysis?
2. Under what circumstances is scientific rigor the same as mathematical rigor?
3. Under what circumstances is analytical mathematics most useful?
4. What is the difference between constants and variables?
5. What does "is a function of" mean?
6. What does linear mean? Can you give an example of something that is linear?
7. Give three examples of nonlinear functions or relations.
8. What is an algorithm?
9. How is a correlation different from a function?
10. Define econometrics.
11. Define calculus in terms of something familiar in your everyday life.
12. How does "finite difference" relate to calculus?
13. What does validation mean? Sensitivity analysis?
14. Distinguish between analytical and numeric approaches to solving mathematical equations.
15. Analytical techniques are best suited to what kind of scientific problems?
16. If the equations of economics are often complex, why are they frequently described using analytical approaches?

References

1. Hall, C.A.S. 1988. An assessment of several of the historically most influential theoretical models used in ecology and of the data provided in their support. *Ecological Modeling* 43: 5–31.
2. Hall, C.A.S., and J.W. Day, eds. 1977. *Ecosystem modeling in theory and practice. An introduction with case histories*. New York: Wiley Interscience.

3. Hall, C. 1991. An idiosyncratic assessment of the role of mathematical models in environmental sciences. *Environment International* 17: 507–517.
4. Gowdy, J. 2004. The revolution in welfare economics and its implications for environmental valuation. *Land Economics* 80: 239–257.
5. Kaufmann, R.K. 1992. A biophysical analysis of the energy/real GDP ratio: implications for substitution and technical change. *Ecological Economics* 6: 35–56.
6. Kaufmann, R.K., and D.I. Stern. 1997. Evidence of human influence on climate from hemispheric temperature relations. *Nature* 388: 40–44.
7. Krugman, P. 2009. How did economists get it so wrong? *Editorial New York Times Sept*. 2: 2009.
8. John G. Kemeny. BASIC and DTSS: Everyone a programmer (Obituary). http://www.columbia.edu/~jrh29/kemeny.html
9. Mary Williston Ecological Bulletin, Personal Communication.

17

Is Economics a Science? Social or Biophysical?

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_17

17.1 Introduction: Much Early Economics was Biophysical

We start with a review of the basic ideas of some important economists as given in ▶ Chap. 2. Economies exist independently of how we perceive or choose to study them. For more or less accidental reasons, we have chosen over the past 140 years to consider and study economics as a *social science*. The present social science focus, however, was not particularly the case with earlier economists, before, say, 1880, who were more likely to ask "where does wealth come from?" than are most mainstream economists today. In general, these earlier economists started their economic analysis with the natural biophysical world, probably simply because they had common sense but also because they deemed inadequate the perspective of earlier mercantilists who had emphasized sources of wealth as "treasure" (e.g., precious metals) derived from mining or trade or plunder. In the first formal school of economics, the French *Physiocrats* (e.g., Quesnay 1758; see Christensen [1, 2]) focused on land as the basis for generating wealth.

The biophysical perspective continued with Thomas Malthus' famous *Essays on Population* (there were six of them), which assumed that human populations would grow exponentially—because it seemed unlikely that anyone would control the "passion between the sexes"—unless somehow "checked" by factors that either reduced the birth rate or increased the death rate. Since Malthus had little faith in the "moral restraint" of the working classes and believed that birth control was "vice," he recommended a rather Draconian social policy to increase the death rates among the poor. In Malthus' view, the agricultural production needed to feed this exponentially increasing human population could grow only linearly, i.e., less rapidly than the number of humans. He also opposed the importation of cheaper continental grains to England, as a limited food supply assured increasing rents for his patrons, the landed aristocracy, and squeezed the profits of the rival capitalists. It was this view—that the human prospect was limited by inadequate food supplies and that class conflict was inevitable—that led the Victorian philosopher Thomas Carlisle to give economics the label of "the dismal science." This was because, in the mind of Malthus and other classical political economists, the limited amount of fertile land (a fixed factor of production) ensured that wages would tend towards a meagerly subsistence level. Adam Smith and other *classical* economists focused on both land and especially labor as means of transforming the resources generated by the natural world into materials that we perceive as having wealth. Later, David Ricardo made important observations about the general need to use land of increasingly poor quality as populations (and hence total agricultural production) expanded. Even Marx, while focused firmly on labor, was keenly interested in the long-term adverse effects of large-scale agriculture on soil quality and wrote a great deal about the degradation of the soil in his chapters on ground rent in the third volume of *Capital*. He firmly believed that capitalism exploits the land in the same way it does labor. Marx was an avid follower of the breakthroughs in agricultural chemistry and was especially impressed by the work of Justus von Liebig, who considered English "High Agriculture" to be a generalized system of robbery. Unlike traditional agriculture where crop wastes were returned to the soil, early industrialized agriculture shipped the food to urban areas where the food waste became pollution instead of fertilizer. The important thing is that all the important early economists were all explicitly biophysically based, at least as much as they focused on the social or human aspects of the economies they were trying to understand.

But in the 1870s, these at least partly biophysically based perspectives in economics were displaced by the *marginalist revolution* of William Stanley Jevons, Karl Menger, and Leon Walras. Their perspective was based on abstractions such as "subjective utility" that ignored, essentially for the first time in economics, measurable physical inputs and outputs of material or energy. This novel approach to economics was called *neoclassical*, and the ideas of the marginal revolutionists still dominate today. In the words of the early marginalist Frederic Bastiat: "exchange is political economy." Hence production, a biophysical perspective requiring a knowledge of the natural sciences, became a less important, even nonexistent issue to economists compared to market-based human preferences, and the commonsense biophysical basis for economic analysis was snuffed out intellectually, although of course not in real economies. By the early twentieth century, land (representing all of nature) was simply omitted, along with energy (which had never been considered), from neoclassical production functions. Generations of economists subsequently have

17

been trained from a perspective that is divorced from biophysical reality except, occasionally, as it affects prices, within a worldview that is often extremely mathematical, theoretical, and even doctrinaire. On the other hand, one might say that neoclassical economics does a good job of reflecting the human characteristic of a desire for more of whatever and the reality that much of what happens within economics does indeed occur within what we may call markets. But the overall movement was away from economics being based on material reality and hence amenable to the tools of natural sciences, to one focused on the human or social perspective; in other words, the intellectual basis of economics changed from one that is quite comfortable with the natural sciences to one that is viewed and studied only as a social science.

Although conceptual economics divorced itself from biophysical reality, this was not the case, at least in theory, in one respect, which is with respect to the development of the underlying mathematical theory. At the turn of the last century, economists chose physics (and, more explicitly, the analytic mathematical format of classical mechanics) as a model for capturing the essence of their discipline. This is reflected in the familiar graphs and equations of commodity value and cost vs. quantity, with price determined as the intersection of downward trending demand curves (derived from utility curves) and upward trending cost of supply curves. Although physics served as the model and its intellectual popularity as the motivation, the resulting economic model was physically unrealistic because it represented a dynamic, irreversible process with a static and reversible set of equations as the conservation principles that constrained the equations of physics, were incompatible with capital accumulation and, indeed, growth or even production in the economists' model [3]. Thus in an irony that has escaped most if not all economists, the attempt of economists to add rigor and respectability to their endeavor by emulating physics in fact violated the second law of thermodynamics, something that would disqualify it immediately within physics.

17.1.1 Economics as a Social *Science?*

So far we have focused on whether we should use the word "social" (vs. "natural") in our consideration of the words "social science" as used as a descriptor of economics. Now we want to focus on the use of the word "science" in that descriptor. Banker DeLisle Worrall has said recently (and we agree) [3]:

"There are no laws in economics. A law in the physical sciences, as Beinhocker reminds us is a *universal regularity with no known exceptions*. There is nothing in economics which meets that standard. What we have are theories: explanations for why regularities exist and explanations of how they work. We need to desist from writing papers that "prove" theories; they always turn out to be mathematical exercises of no practical relevance, yielding no insight about how the economy really works. In our empirical work we must accept the reality that the limitations of model specification, measurement error, choice of proxy variable, etc. are so formidable that we can never "prove" anything in economics by appealing to the numbers."

So if we are to take this position, and we do, we have to ask why, then, is economics called a social science, or indeed any kind of a science, if it has no ability to generate laws that we can count on? Why do so many important Wall Street financial institutions turn over their analyses to highly mathematical (but barely financially literate) "quants" when they universally led their institutions and their investors off the cliff? [4].

This reintroduces the most basic message in our book. Should economics be principally about the social sciences, about human wants and desires and the ability of markets to fulfill these optimally, as most economics textbooks would say, or should it be about the biological and physical (i.e., biophysical) conditions that are behind the generation and even distribution of wealth. We believe that of course it should be some mixture, but we also believe that by focusing almost entirely on the social science aspects of economics while essentially ignoring, and even discounting, the biophysical aspects, conventional economics has failed in many ways to understand the processes that are in fact the essence of economics. Consequently mainstream, exchange-based economics is completely inadequate to deal with the new realities imposed upon the world by peak oil and the many issues associated with the end of what has been essentially a "resource free for all on a relatively uncrowded planet." But the planet now is very, very crowded, and depletion is increasingly important for many, probably most, of our resources. Economic theories and concepts can make only a small impact on mitigating these basic problems. Hence we need a whole new approach to economics, one that not only recognizes but is based on the resources themselves, not the prices they command. We call this new approach biophysical economics, and this book is its first synthesis.

Economics as presently perceived may be the most widely, consistently, and incoherently taught course in American higher education, and the same is likely to be true in most other countries. By *widely* we mean that there may be more young people taking an introductory economics course than nearly any other single course in college except perhaps biology, or college algebra, or college composition. By *consistently* we mean that in preparation for writing this book, we reviewed about two dozen basic economics books and found that they are mind-numbingly similar, and all build up a system of economics consistent with the basic neoclassical framework. This consists of a caricature of real economies as that of simply firms and households interacting through markets, with a focus on humans, their wants and needs, and their independence in deciding what is good for them through their individual decisions in markets. In other words there is a consistent body of theory, known as neoclassical economics, that is, accepted or promulgated by essentially all economists, at least as represented in their fundamental textbooks. We assume that the readers of this book have at least passing familiarity with this conventional economics.

By *incoherently* we mean that many of the assumptions that conventional economists must make to generate their world of theoretical economics, the associated equations, and their applications defy logic to one trained in the natural sciences or perhaps even common sense. There are three ways in which conventional neoclassic economics fails these tests: behavioral, biophysical, and moral. Although these concepts have been presented previously, we review each below.

17

17.1.2 Behavioral

The canonical assumptions of *Homo economicus* (non-satiation, self-regarding behavior, strictly rational decision-making) are assumed to predict accurately how people make economic decisions. Thus the basic neoclassical model assumes that people are "rational," meaning selfish or at least self-regarding, so that they make market decisions based on their own self-interest. In fact, as summarized in ▶ Chap. 3, there has been a great increase in the degree to which basic human economic behavior has been tested using the scientific method and in very clever experiments in behavioral economics. The results have tended to show that the *Homo economicus* view is false or at least very poorly predictive. For example, Henrich et al. [5] after examining the results of behavioral experiments in 15 small-scale societies ranging from hunter-gatherers in Tanzania and Paraguay to nomadic herders in Mongolia conclude: "[T]he canonical model [i.e. *Homo economicus*] is not supported in any society studied."

17.1.3 Biophysical

Hall et al. [6] summarized the main ways that the basic neoclassical model failed even the most minimal standards for veracity in natural science: the basic model violated the laws of thermodynamics, had incorrect boundaries, and did not generate its premises by generating and testing hypotheses but rather as logical givens. Most basic models are *not* consistent with the laws of thermodynamics, nor do most economists even think about such laws [7]. This alone would be enough to disqualify any model in the natural sciences, but it has not seemed to bother economists. Gowdy et al. [8] provide many more ways that basic science is violated with the basic neoclassical model. The ability to predict is a crucial criterion for any economic model that is to be used to influence policy and hence the lives of many people.

One can certainly find some hypothesis generation and testing in learned economics journal. For example, Hall [9] examined some 127 articles in the leading economics journal *American Economic Review* and found that for this subset of papers, about 10 percent did test explicit hypotheses, which is good. Only 3 percent, however, could be construed as testing fundamental economic theories. These papers found more often than not that the basic economic theories tested in specific applications were more likely *not* to be supported than the converse. So we might say based on this study that economics is a good science because ideas were being subject to the scientific method or perhaps bad because such results have no impact on the center of gravity of conventional economics, as is clearly stated by leading economists themselves (e.g., Krugman 2008 [4]).

A core belief of many economists is that good models make good predictions and that this is more important than whether or not the model is consistent with known mechanisms [10]. But in fact we find that the core models used by economists (economic man and perfect competition) consistently fail the "good prediction" test. For example, essentially all economists failed to predict the market crash of 2008.

17.1.4 Moral

Most of our students, possibly more idealistic than the average, are also very much put off, for both scientific and moral reasons, by the essential selfishness that is accepted by and even celebrated in the basic economic theory found in introductory economics textbooks. This perspective was made to us even more strongly by our colleague, Donald Adolphson a very popular and thoughtful professor of economics and finance at Brigham Young University in Utah. He said to us:

> "The students at BYU are virtually all practicing Mormons (Mormon is a Christian denomination also known as "The Church of Latter Day Saints", which is very strong especially in Utah and adjacent states). They are trained at home to think of their relation to God and then family first, community second and then the world community. Most travel to a foreign country as a late teenager as part of their preparation for life. When they take Introductory Economics, they are told in their textbooks that the basic neoclassical … starts with the assumption that humans are "rational", rational meaning entirely selfish, or at least self-serving, and principally materialistic. This just strikes them as wrong, and they reject their basic economics textbooks."

Well it strikes us as wrong too. It also strikes most of our own students in upstate New York as wrong morally and with respect to their own motivation. In particular it seems wrong to the majority of our students because they have a high sense of idealism toward nature and toward other people, neither of which they wish to see sacrificed for mere self-serving and often superfluous economic goods and services. This is especially the case when they view the world around them as full of hyper affluence bought at enormous expense to the environment and the enormous discrepancies between rich and poor. They want something else, and they have found it, to some large degree, in the biophysical ways we teach our own economics courses. But there has not been a rallying point, a central synthesis of the broadness of literature needed to understand real economies or a source of synthesized information needed to really understand economics and the basic economic relations of humans to our world. We try to do that in this book.

We cannot accept economics as presently practiced as any kind of science because it does not follow the rules of science as we summarized them in ► Chap. 15. This is true both for the behavioral aspects of humans (i.e., how they in fact interact with others vs. how the basic neoclassical model assumes that they do) and for the degree to which the model is inconsistent with the laws of nature as summarized in ► Chaps. 3 and 15.

17.1.5 Other Economists Agree with Us

Most knowledgeable economists, when pressed, will acknowledge at least some of this, yet economics as a discipline rumbles onward year after year with little real change in the way that our young people are inculcated into this august company. This point of view is not simply ours but was apparent to most of our students (especially those with a focus on, or at least reasonable experience with natural science). While our students can indeed learn the principles of economics in their first course in that subject and can pass and even do well on the tests, they generally do not, or barely, believe the concepts that they are taught there. Because many of the principles seem unrealistic to them, they are often deeply bored. They sometimes use very harsh words to describe their disbelief on what they are being taught. Well, we agree with them and believe that collectively we have been teaching something like one million young people a year in the United States alone something that might reasonably be considered, at worst, complete fabrications or at best a very simplistic and incomplete perspective on the reality and richness of thought that can be brought to bear on economic issues and problems.

As some support for that point of view, we note that, as of 2006, six of the last eight most recent recipients of the Nobel Prize in economics were people whose works challenged, in various very fundamental ways, the basic existing neoclassical paradigm.

We find that there are many, many other scientists and economists who basically believe the same things as we do: that neoclassical economics is intellectually corrupt at its core. Economics has largely isolated itself from harder sciences, trapping itself as entirely a social science by relying on the laws of physics as they were known in the nineteenth century. This perspective is sometimes called "hermetic," in that economics is completely self-enclosed within its own narrow world.

As stated in Gowdy [8]:

> "The distinguished historian of economic thought, Mark Blaug, has remarked that economics has increasingly become an intellectual game played for its own sake (Blaug 1998, pp. 11, 34). A survey of graduate students in economics in the 1980s by David Colander and Argo Kamer (1990) found an astonishing lack of interest in learning about current economic issues or about the literature of economics. Colander and Kamer surmise that, sadly in their view, this may be rational behavior on the part of graduate students in economics. The quickest way to success as an academic economist is to concentrate on mathematics, rather than learning about how actual economies work. Alan Blinder, a former member of President Clinton's Council of Economic Advisors, has characterized training in economics as "increasingly aloof and self-referential"."

Other modern critics include McMurtry [11], Cox [12], Talab [13], Johansson [14], Sutter and Pesky [15], Hall et al. [16], Mirowski [17], and especially Easterly [18] and Piketty [19]. Many of these publications stress the physical harm that a belief in the abstractions of neoclassical economics causes to people, especially poorer people in the developing world.

It is arguable whether at present, the economic field is undergoing a fundamental shift to establish a more scientific foundation. Possibly, just as biology emerged as a true science in the twentieth century, so too will economics in the 21st. But in the meantime, it seems that economists are mostly circling their wagons, defending their assumptions against all attacks from the outside, or more normally simply ignoring them, retreating into their carefully constructed fantasy world of assumptions and impenetrable equations.

17.1.6 Is Economics a Science?

So our answer to the question posed by the title of this chapter is that, no, economics at this time is not a science. Its basic models violate too many scientific principles including the first principles that are necessary for any real model: laws of thermodynamics, the law of the conservation of matter, the ways that people actually do behave according to empirical studies, and so on. In addition even when economics appears to be "borrowing" equations from physics, it is doing so incorrectly, even in violation of the physics it is trying to emulate. Instead of following these principles, principles that all natural science follows or risks rejection or humiliation from peers, neoclassical economics has generated its own world, a world that reflects the real world in only the most basic and contrived ways. While in theory there is a model of physics behind, the equilibrium model is just a copying of the equation form without any understanding of the actual physics—in fact it violates the second law of thermodynamics [6, 17]. Additionally the assumptions of "rational actors" required to make this model work are inconsistent with how humans in fact interact with each other. The generation of theory based on a market concept of perfect information and equal power of interacting buyers and sellers that has not existed since agrarian England, if indeed they ever existed, combined with failure to make and test hypotheses, makes an acceptance of the basic neoclassical model an article of faith, not rationality. Curiously the ascendance and the power of the ideas of the advocates of market theory and self-interest have spilled over to our public and political life. This has destroyed many economies in the less developed world [16, 18], while completely changing the political perspective of many Americans from community, civic responsibility, and fairness in distribution of wealth and care for others to one of unbridled greed and self-focus, while turning, to some degree, universities from learning communities where highly trained and caring professors held students up to their own high standards to commodities where students buy their education and expect high grades with little work. This has also given a green light to those who have enormous financial power to buy and to manipulate our political system, while convincing many that "big government," their only defense against big money corporations, is something to be avoided. The net effect has been an assault on our public institutions, the only entity with enough power to stand up to ever larger and more powerful corporations and their ultra-wealthy directors [20]. It is a very large impact of a theory on reality which is scientifically indefensible at its heart. We conclude that a new, biophysically based economics is critically needed [21].

Questions

1. Why is economics usually considered a social rather than a biophysical science? What is your view?
2. Do you agree, from your own experience, that humans are essentially selfish or at least self-regarding? Or does it depend upon the circumstances?
3. With respect to the previous question, is this pattern of basic selfishness found in all cultures around the world?
4. What are the characteristics of an endeavor that qualify it as a science? Do you think conventional economics qualifies as a science? Why or why not, or where and where not?
5. In the world of conventional economics, what does rational mean? What does it mean to you?
6. Conventional economics is usually classified as a social science. In your opinion does economics qualify as a science? Why or why not?

References

1. Christensen, P. 1994. Fire, motion and productivity: The proto-energetics of nature and economy in Francois Quesnay. In *Natural images in economic thought*, ed. P. Mirowski, 249–288. Cambridge: Cambridge University Press.
2. ———. 1984. Hobbes and the physiological origins of economic science. *History of Political Economy* 21: 4.
3. Worrell, Delisle. 2010. Governor of the Central Bank of Barbados, in an address to the Barbados Economic Society (BES) AGM, Bridgetown, 30 June 2010.
4. Krugman, P. 2009. How did economists get it so wrong? *Editorial New York Times* 2: 2009.
5. Henrich, J., et al. 2001. Cooperation, reciprocity and punishment in fifteen small-scale societies. *American Economic Review* 91: 73–78.
6. Hall, C., D. Lindenberger, R. Kummel, T. Kroeger, and W. Eichhorn. 2001. The need to reintegrate the natural sciences with economics. *Bioscience* 51 (6): 663–673.
7. Hall, C., and J. Gowdy. 2007. Does the emperor have any clothes? An overview of the scientific critiques of neoclassical economics. In *Making world development work: Scientific alternatives to neoclassical economics*, ed. G. LeClerc and C.A.S. Hall, pp 3–12. Albuquerque: University of New Mexico Press.
8. Gowdy, J. 2001. "Economics Interactions with Other Disciplines," theme article Encyclopedia of Life Support Systems, UNESCO, Paris.
9. Hall, C. 1991. An idiosyncratic assessment of the role of mathematical models in environmental sciences. *Environment International* 17: 507–517.
10. Friedman, M. 1955. *Essays in positive economics*. Chicago: University of Chicago Press.
11. John McMurtry, J. 1999. *The cancer stage of capitalism*. London: Pluto Books.
12. Cox, Adam. 2010. Blame Nobel for crisis, says author of 'Black Swan', Reuters (2010-09-28).
13. Taleb, N. N. 2007. The pseudo-science hurting markets (PDF).
14. Johansson, D. 2004. Economics without entrepreneurship or institutions: A vocabulary analysis of graduate textbooks (PDF). *Economic Journal Watch* 1 (3): 515–538.
15. Sutter, D., and R. Pjesky. 2007. Where would Adam Smith publish today? The near absence of math-free research in top journals. *Economic Journal Watch* 4 (2): 230–240.
16. Hall, C.A.S., P.D. Matossian, C. Ghersa, J. Calvo, and C. Olmeda. 2001. Is the Argentine National Economy being destroyed by the department of economics of the University of Chicago? In *Advances in energy studies*, ed. S. Ulgaldi, M. Giampietro, R.A. Herendeen, and K. Mayumi, pp 483–498. Padova: Argentine Economy.
17. Mirowski, P. 1984. *More heat than light*. Cambridge: Cambridge University Press.
18. Easterly, W. 2001. *The elusive quest for growth: Economists' adventures and misadventures in the tropics*. Cambridge: The MIT Press.
19. Piketty, T. 2013. Capital in the twenty first century. (English edition, 2014, Cambridge, MA: Harvard University Press).
20. Sekera, J. 2016. *The public economy in crisis: A call for a new public economics*. New York: Springer.
21. Hall, C.A.S., and K. Klitgaard. 2006. The need for a new, biophysical-based paradigm in economics for the second half of the age of oil. *Journal of Transdisciplinary Research* 1 (1): 4–22.

The Science Behind How Real Economies Really Work

Our book so far has reviewed how economics, as a science, and economies, both historically and today, can be understood much better through an appreciation of the role that energy plays. We also examined how the discipline of economics has viewed economies historically, and developed our perspective on the extreme limitations of the approach used by economists. We then developed the basic science needed to understand economics, a kind of training that is missing from the education of most economists. Now in this section, we apply the concepts of science developed earlier to understand some new and important ways in how real economies have operated and are likely to operate in the future. We focus on the biophysical concepts of peak oil, energy return on investment (EROI), and the role of models, both conceptual and mathematical.

Contents

Energy Return on Investment

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_18

18.1 Introduction [1]

Many important earlier writers, including sociologists Leslie White and Fred Cottrell, and ecologist Howard Odum, have emphasized the importance of net energy and energy surplus as a determinant of human culture [2–4]. Human farmers or other food gatherers must have an energy profit to survive and a significant return for there to be specialists, military campaigns, and cities and substantially more for there to be today's art, culture and other amenities. *Net energy analysis* is a general term for the examination of how much energy is left over from an energy-gaining process after correcting for how much of that energy (or its equivalent from some other source) is required to generate (extract, grow, or whatever) a unit of the energy in question. Net energy analysis is sometimes called, depending upon the specific procedures used, the assessment of energy surplus, energy balance, or, as we prefer, energy return on investment or EROI. To perform this analysis, we start with the more familiar monetary assessment and then develop how this relates to the energy behind economic processes. A somewhat more technical analysis is available in [5] and in our cited papers.

18.1.1 Economic Cost of Energy

In actual economies, energy comes from many sources—from the sun to run ecosystems and from imported and domestic sources of oil, coal, and natural gas, as well as hydropower and nuclear, and from renewable energy—most of that as firewood and hydroelectric but increasingly from wind and photovoltaics. Many of these energy sources are cheaper per unit energy delivered than oil, and some are considerably more expensive.

18

We have said many times that energy became cheap in the last century. How cheap? Such an analysis is not available for the United States, but it has been done for England going back to 1300 by Carey King. He found that the cost of the energy to run the economy (at that time meaning food for humans, fodder for animals, and wood for various things) was typically from 30 to 40 percent of GDP (Fig. 4.4). Another way of looking at his data is that in past centuries about a third of all economic activity was to get the energy to run the rest of the economy. When coal was introduced to the economy, that was reduced to about 15–20 percent and with oil to about 5–10 percent. In other words, as fossil fuels were added to the economy, far less of the total economic activity was needed just to get fuel (and it was more potent fuel too), and this allowed some people to get wealthier, largely because they could appropriate the extra work that most workers could produce by means of fossil-fuel driven machinery.

So let's look at what this real ratio of the cost of energy (from all sources, weighed by their importance) is relative to its benefits:

$$\text{Economic cost of energy} = \frac{\text{Dollars to buy energy}}{\text{GDP}}$$

By this token the relation of the proportional energy cost in dollars is similar, as we shall see, to the proportional energy cost in joules; in 2017 roughly 6 percent (1 trillion dollars) of the US GDP was spent by final demand for all kinds of energy in the US economy to produce the 17 trillion dollars' worth of total GDP (Fig. 18.1).

This percentage increased in the first half of 2008 as the price of oil exceeded $140 a barrel, and then it fell again. This pattern has been seen before. The abrupt rise in the price of energy during the "oil shocks" of the 1970s, the subsequent decline of this value from 1984 through 2000, and the increase again through mid-2008 had large impacts on discretionary spending, that is, the amount of income that people can spend on what they want vs. what they need, because the 5–10 percent change in total energy cost would come mainly out of that 25 percent or so percent of the economy. Thus we believe that changes in energy prices have very large economic impacts, a perspective supported by James Hamilton's analysis [6]. What future energy prices will be is anyone's guess, but even as economies crash in the late part of the first decades of the new millennium and subsequently cease growing or grow only very slowly, there is a great deal of information implying that dollar, and hence presumably energy, costs of fuels are increasing substantially, if irregularly. Our theory is that they will occur in large part due to declining EROI and that they will take a huge economic toll in the future. This chapter develops that argument.

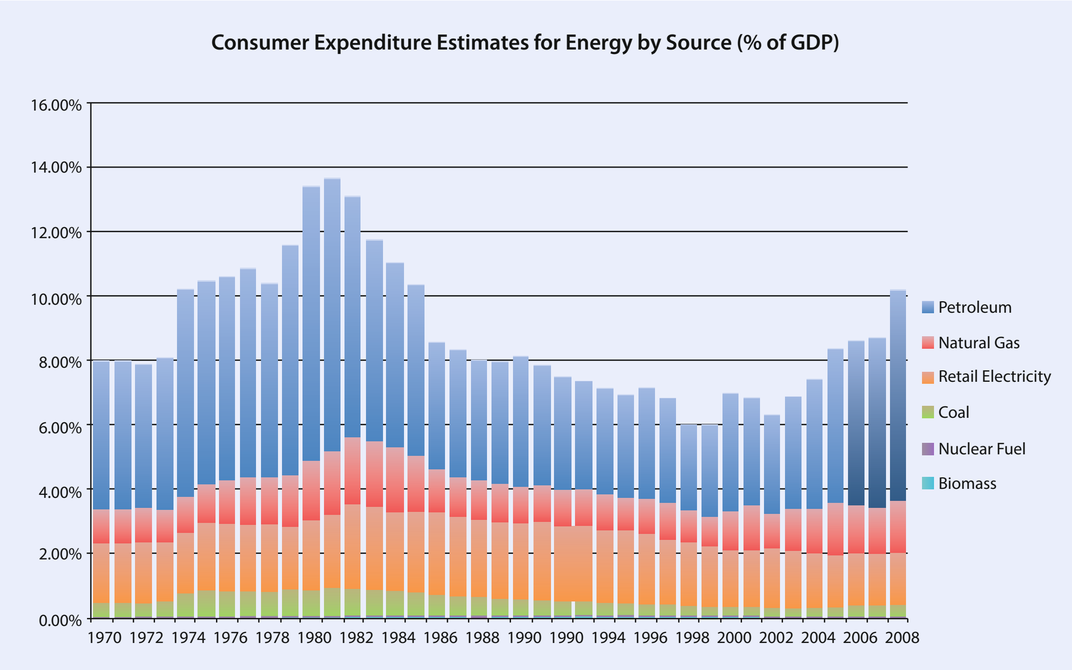


Fig. 18.1 Proportion of GDP for energy

18.2 What Is EROI?

Energy return on investment (EROI or sometimes EROEI, with the second E used to refer to the use of energy in the denominator) is the ratio of energy returned from an energy-gathering activity compared to the energy used in that process. In principle, the idea is to see how much energy society invests to get more energy. Usually the energy is either "already in society" and diverted (such as to make a drill bit) or could readily go to society but is diverted to getting more energy (such as natural gas used to pressurize a field). EROI is calculated from the following simple equation, although the devil is in the details:

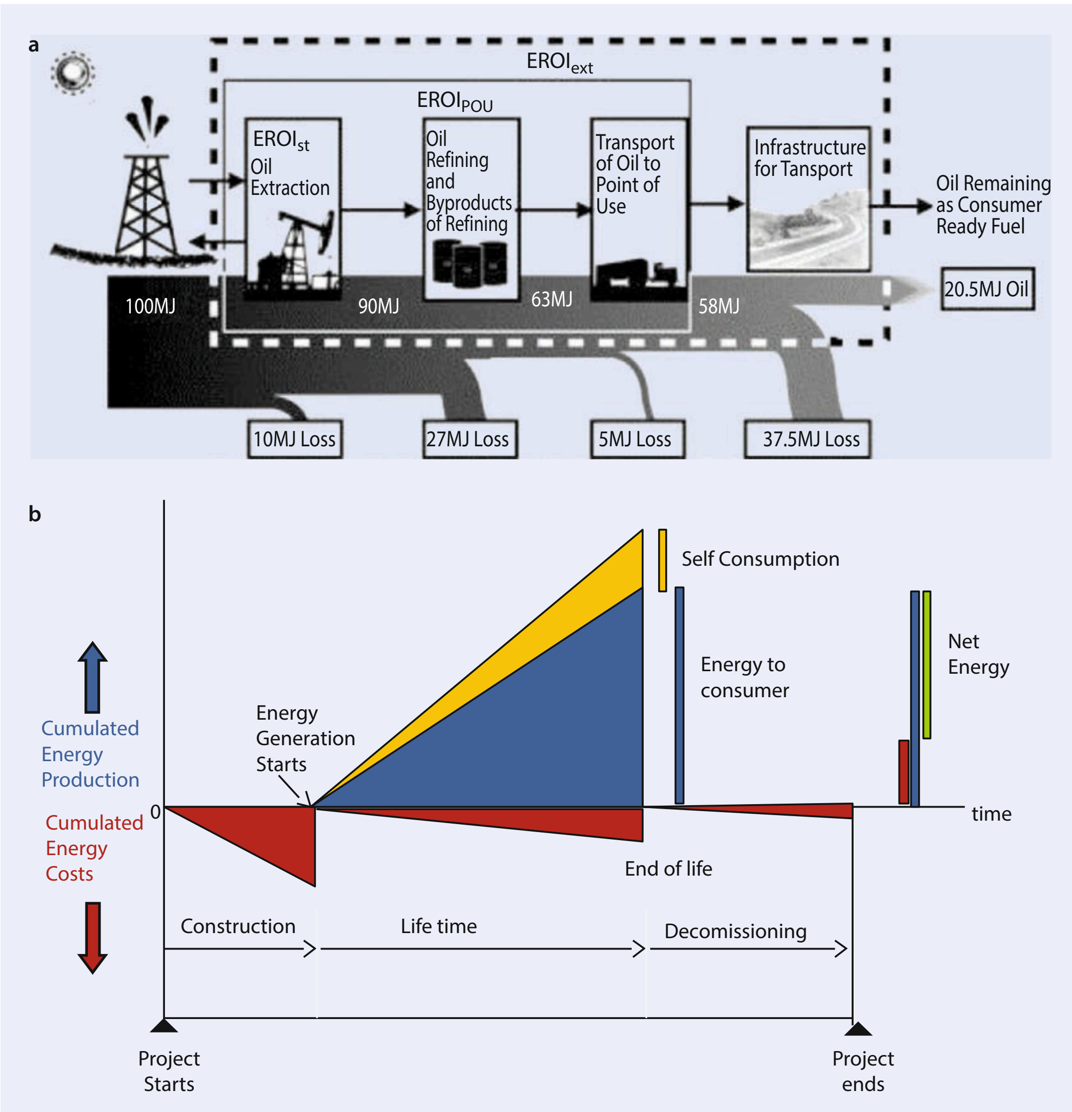
$$\text{EROI} = \frac{\text{Energy returned to society}}{\text{Energy required to get that energy}}$$

Since the numerator and denominator are usually assessed in the same units (an exception we treat later is when quality corrections are made), the ratio so derived is dimensionless, e.g., 20:1 which can be expressed as "twenty to one." This implies that a particular process yields 20 joules on an investment of 1 joule (or Kcal per Kcal or barrels per barrel). EROI is usually and most precisely applied at the mine mouth, wellhead, farm gate, etc. that is at the point that it leaves the production facility. We call this more explicitly EROI_{mm}, and it is not to be confused with *conversion efficiency*, i.e., going from one form of energy to another such as upgrading petroleum in a refinery or converting coal to electricity. More explicit ratios are derived by King [7].

The authors of this book and other advocates of EROI believe that net energy analysis offers the possibility of a very useful approach for looking at the advantages and disadvantages of a given fuel while offering the possibility of looking into the future in a way that markets seem unable to do. Its advocates also believe that in time market prices must approximately reflect comprehensive EROIs, at least if appropriate corrections for quality are made and subsidies removed. Nevertheless we hasten to add that we do not believe that EROI by itself is necessarily a sufficient criterion by which judgments may be made. It is, however, the one we favor the most, especially when it indicates that one fuel has a much higher or lower EROI than others. In addition it is important to consider the present and future potential magnitude of the fuel, and how EROI might change with depletion or if the use of a fuel is expanded. An example of an EROI analysis is given as Fig. 8.7.

18.3 Relation of EROI to Monetary Cost

What is most important about energy to many people is its monetary cost. While the earliest studies of EROI were focused on its physical meaning and whether it was increasing or decreasing for particular fuels, there was always the assumption that high EROI fuels were likely to be cheaper per joule, as less effort had to go into getting them. So coal, which generally had a high EROI, was cheap, and oil much more expensive. This makes sense as it takes much less energy to dig coal out of the ground compared to operating a complex oil well. It would seem that as EROI tended to decrease over time, it was expected that eventually energy would become much more expensive. This is somewhat the case for oil but not for natural gas. King and Hall [8] found that over time the cost of particular energies increased and decreased in opposition to EROI (◘ Fig. 18.2). Prices would tend to increase when EROI was low (due in part to increased drilling rates) and the converse. But sometimes prices decline (as in



◘ **Fig. 18.2** **a** Diagram of energy losses between the wellhead and use for petroleum (from [16]). **b** Schematic for the energy used and gains over time of an energy project. The EROI would be the final value of the blue triangle on top divided by the sum of the final values for the three brown triangles on bottom. These are represented by bars on right

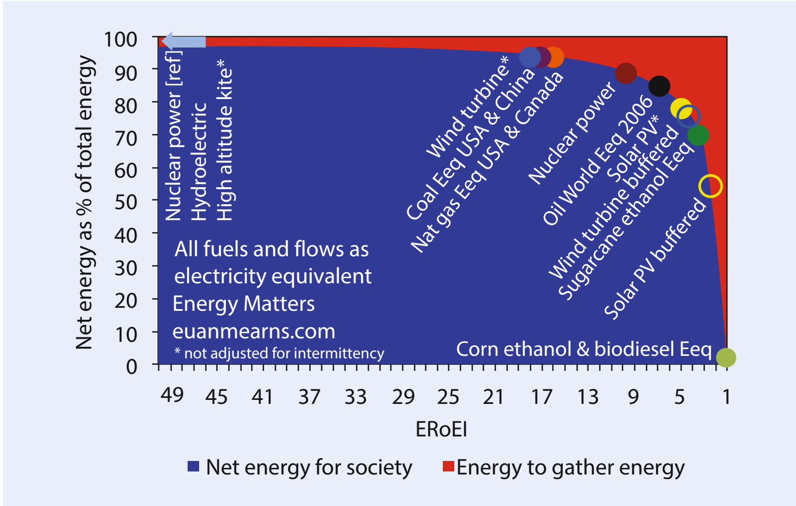


Fig. 18.3 The EROI cliff. "Buffered" means including an estimate of cost of dealing with intermittency. (Courtesy Euan Mearns)

2015–17) due to supply and demand issues unrelated to EROI (▶ Chap. 13). Euan Mearns has developed the concept of the "EROI cliff," where the impact of diminishing EROI is much greater as lower values are approached (Fig. 18.3). The greatest concern is with declining EROI for the premium fuels of the United States and the world, which seems to be occurring although it is increasingly difficult to get the needed data. The large increases in production costs, however, imply a continuing decline in EROI for major fuels. When combined with "peak oil," this might mean very lean times ahead for the production of our most important fuels (see review in Hall [5]).

18.3.1 History

The concept of EROI was derived from Howard Odum's teachings on net energy [2], from earlier work by anthropologists and sociologists as developed in ▶ Chap. 6, and explicitly from Hall's PhD dissertation on the energy costs and gains of migrating fish [9]. The concept was implicit in Hall and Cleveland's 1981 paper [10] on petroleum yield per effort although the term net energy was used there. The first publication using the name EROI was apparently in 1979 [11], and it received much more attention from a paper in science [12]. Somewhat detailed summaries of the literature on EROI then available were put together in a large book [13]. The concept lagged during the "energy lull" 1984–2005 but has picked up post 2005 with increasing energy prices. There are a flurry of new papers now available [e.g., 14–18]. An entire issue on EROI in the online journal *Sustainability* was published in 2011 with many interesting analyses [15]. Some recent controversies are developed in [5] and summarized at the end of this chapter.

There is very little quantitative information about actual EROIs for energy-producing systems for the medium or distant past, which is not surprising because we did not even understand the concept of energy until about 1850. But there is at least one example where Sundberg, mentioned in ▶ Chap. 6, has made a quite detailed assessment of the energy cost of energy in earlier Sweden [18]. From 1560 until 1720, Sweden was the most powerful country in Northern Europe, based mostly on its very productive metal mines but also an aggressive foreign policy backed up by high-quality weapons. The production of these mines required enormous amounts of energy for mining and especially smelting. The source of this energy was wood and especially charcoal (needed to get the high temperatures steel required) cut from Swedish forests. Ulf Sundberg gives a detailed calculation of how a typical forester and his family, self-sufficient on 2 hectares of farmland, 8 hectares of pastures, and 40 hectares of forest (collectively intercepting 1500 terrajoules of sunlight), generated some 760 gigajoules of charcoal in a year to the metal industry. To do that required about half a gigajoules of human energy or 3.5 if we include the draft animal labor. So we might calculate the EROI of the human investment to be as high as 1500:1 or some 250:1 if we include the animals. But that is just the direct energy, as it took 105 GJ to feed, warm, and support the farmer and his family (which includes his replacement) and probably at least that to support the animals. So if we include direct plus indirect energy, the EROI is down to roughly 4:1. The system was sustainable as long as the forests were not overharvested. That was true until the middle of the nineteenth century, but then the forests were severely overharvested, and many Swedes left for America.

Much of the literature on net energy in the first decade of this century tended to be about whether a given project is, or is not, a net surplus, that is, whether there is a gain or a loss in energy from, e.g., making ethanol from corn [19]. The criteria used in much of the current debate are focused on the "energy breakeven" issue, that is, whether the energy returned as fuel is greater than the energy invested in growing or otherwise obtaining it, i.e., if the EROI is greater than 1:1. The general argument seems to be that if the energy returned is greater than the energy invested, then the fuel or project "should be done," and if not, then it should not.

Several of the participants in the current debate about corn-derived ethanol (summarized in reference [20]) argue that ethanol from corn is a clear energy surplus, with from 1.2 to 1.6 units of energy delivered for each unit invested. Further aspects of this argument center around the boundaries of the numerator, i.e., whether one should include some energy credit for nonfuel coproducts (such as residual animal feed—i.e., soybean husks or dry distiller's grains), the quality of the fuels used and produced (e.g., liquid—presumably more valuable—vs. solid and gaseous), and the boundaries of the denominator (i.e., whether or not to include the energy required to compensate for environmental impacts in the future, e.g., for the fertilizer needed to restore soil fertility for the significant soil erosion occasioned by corn production). Such arguments are likely to be much more important in the future as other relatively low-quality fuels (e.g., oil sands or lower-quality shale oil) are increasingly considered or developed to replace conventional oil and gas, both of which are likely to be more expensive and probably less available in the not so distant future. If, of course, the alternatives require much oil and/or gas for their production, which is often the case, then an increase in the price of petroleum will not necessarily make the alternatives cheaper and more available as a fuel. We believe that for most fuels, especially alternative fuels, the energy gains are reasonably well understood, but the boundaries of the denominator, especially with respect to environmental issues, are poorly understood and even more poorly quantified. Thus we think that most calculated and published EROIs, including those we consider here, are higher (i.e., more favorable) than they would be if we had complete information. One study analyzed the reasons for differences in different studies and concluded that they were much less than usually considered [19]. The general tendency has been for recent studies to conclude that the EROI for corn-based ethanol is too close to 1:1 to conclude that they are a significant net energy source, especially if the geography of growing corn is considered [21]. Nevertheless by law 10 percent of gasoline in the United States is composed of ethanol, since the program is very popular in politically powerful corn states.

More recently there have been very strong arguments about what the EROI of solar PV is, with advocates claiming at least 10:1 and critics closer to 3:1 [22]. At this point the issue has not been resolved, although there is agreement among the analysts for procedures to bring their results much closer together [23].

18.4 Seeking an Acceptable EROI Protocol

Given the rather different quantitative responses sometimes derived from different analyses (such as the corn-based ethanol example given above), we need some good and consistent way of thinking about the meaning of the magnitude of the various EROIs of various fuels. It is our opinion that many of the EROI arguments so far are simplistic, or at least incomplete, because the "energy breakeven" point, while usually sufficient to discredit a candidate fuel, should not be the only criteria used. In addition it seems to us that many of the EROI analyses "out there" are generated from the perspective of defeating or defending a particular fuel rather than objectively assessing various potential alternatives. Perhaps we need some way to understand the magnitude, and the meaning, of the overall EROI we might eventually derive for all of a nation or society's fuels collectively by summing all gains from fuels and all costs from obtaining them (i.e., *societal* EROI):

$$\mathrm{EROI}_{soc} = \frac{\text{Summation of the energy content of all fuels delivered}}{\text{Summation of all the energy costs of getting those fuels}}$$

This has been undertaken for England [7] (◘ Fig. 4.7 and 8.7c). This too shows a great increase in the energy return on energy investment over a long period as coal and then oil became the chief supply of energy, with a possible recent uptick that may imply much lower energy returns as England's coal mines and North Sea oil resources are increasingly depleted.

We need to ascertain EROI in a straightforward and universally accepted approach even while accommodating different approaches or philosophies. Of greatest concern are the boundaries of the analysis: should coproducts (such as hulls left from generating biodiesel from sunflower seeds that can be fed to animals, reducing energy needed to make the animal feed) or should we include the costs of the energy to support a laborer's paycheck? Another important issue, which accounts for much of the divergence in the estimates of EROI for PV, is whether the numerator should be multiplied by three to account for the fossil energy that would have needed to be burned to generate that amount of electricity? Since there are no clear and unambiguous answers to those questions, Murphy et al. [24] have advocated a basic EROI approach—using simple standardized energy output divided by the direct (i.e., on site) and indirect (i.e., energy used off site to make the steel used on site)—to generate a *standard* EROI, EROI_{st}. This approach allows the comparison of different fuels even when the analysts do not agree on the methodology that should be used. Murphy et al. advocate the use of other supplementary EROIs, including new approaches that allow for special consideration of other aspects of that EROI at the author's discretion. We believe this allows for both standardization and flexibility:

$$\mathrm{EROI}_{st} = \frac{\text{Energy returned to society}}{\text{Direct and indirect energy required to get that energy}}$$

18.5 The Best Analyses of the Energy Costs

Determining the energy content of the numerator of the EROI equation is usually pretty easy: multiply the quantity produced by the energy content per unit. Determining the energy content of the denominator is usually considerably more difficult. Usually one includes the energy used directly, that is, on site, and this includes the energy used to rotate the drilling bit and pressurize the field, operate the farm tractor, and so on. One usually also includes the energy used indirectly, that is, to make the drilling bit and associated materials, the tractor, and so on. Unfortunately companies generally do not keep track of their energy expenditures, but only their dollar expenditures. Forty years ago a remarkable group at the University of Illinois, including Bullard, Hannon, Herendeen [25], and Costanza [26], undertook such calculations for every sector of the US economy. Using input output analysis (who buys how much of what from other sectors) and very comprehensive energy use information, they were able to generate very detailed determinations of how much energy it took to make all products of the US economy. These allowed us at the time to gain very detailed assessments of where and how energy was used in the US economy and also the energy costs of getting energy.

These analyses also showed that (except for energy itself) it does not matter enormously where money is spent *within final demand* due to the complex interdependency of our economy (i.e., the final products that consumers buy are relatively unimportant to overall GDP/energy efficiency because there are so many interdependencies, i.e., each sector purchases from many other sectors within our economy, although this does not apply to the intermediate products purchased by manufacturers). According to Costanza [26], the market selects for generating a similar amount of wealth per unit of energy used within the whole economic "food chain" leading to final demand. While this is not exactly true, it is close enough for our present purposes, and it is certainly true for the average of all economic activity, with the exception of purchases for energy itself. This is because energy purchases include a similar amount of embodied energy per dollar spent as the societal mean but in addition the chemical energy of the fuel. Unfortunately there has been little such analysis of such "sector interdependencies" since these pioneering works, so that it is hard to make such assessments today. The closest assessment that is available (to our knowledge) is the analyses undertaken by various people at Carnegie Mellon University (Green energy and available on their website).

We next show how an EROI analysis can generate some quite interesting results that can help us understand the importance of EROI for running an actual economy.

18.6 How Much Energy Is Needed to "Get the Job Done": Calculating EROI at the Point of Use

The EROI that is needed to undertake some activity, such as to drive a truck, is far more than just what is needed to get the fuel out of the ground. This was assessed in 2008 by Hall *et al.* [1]. In the spirit of flexibility that we introduced in several sections previously, we introduce here new concepts that start with $EROI_{mm,}$ the standard EROI at the mine mouth (or farm gate, etc.), and then take it further along the use "food chain." While technically it is probably better to calculate the EROI at the mine mouth and then look at the efficiency of using it, we can also calculate the total EROI further along its use path. We call this next step EROI at the "point of use" or $EROI_{pou}$:

$$EROI_{pou} = \frac{\text{Energy returned to society at point of use}}{\text{Energy required to get and deliver that energy}}$$

As we extend the energy cost of obtaining a fuel from the wellhead toward the final consumer, the energy delivered goes down, and the energy cost of getting it to that point goes up, both reducing the EROI. This begins the analysis of what might be the minimum EROI required in society. We do this by taking the standard EROI (i.e., $EROI_{mm}$) and then including in addition in the denominator first the energy requirements to get the fuel to the point of use (i.e., $EROI_{pou}$) and then the energy required to use it to generate $EROI_{ext}$, i.e., extended EROI. As stated it might be more accurate to consider it EROI up to the wellhead and then "food chain efficiency" to the point of use, but this has not been done yet.

A more comprehensive analysis of the EROI required to drive a truck, including all the energy used in the "food chain" to do so, was undertaken by Hall, Balogh, and Murphy [1]. The costs (calculated using 2005 prices and ratios) included:

Refinery losses and costs: Oil refineries use roughly 10 percent of the energy of the fuel to refine it to the form that we use. In addition about 17 percent of the material in a barrel of crude oil ends up as other petroleum products, such as lubricants and asphalt, not fuel. So for every 100 barrels coming into a refinery, only about 73 barrels leaves as usable fuel. Natural gas does not need such extensive refining although an unknown amount needs to be used to separate the gas into its various components and a great deal, perhaps as much as 25 percent, is lost through pipeline leaks and to maintain pipeline pressure. Coal is usually burned to make electricity at an average efficiency of 35–40 percent. What this means, however, is that, e.g., oil resources that have an EROI of 1.1 megajoule returned per megajoule invested at the wellhead cannot provide an energy surplus for a society because not only does it cost about 10 percent of the energy obtained to get the energy out of the ground but then only 73 percent of the remaining oil is delivered to society. Thus this situation could not support itself. One needs a higher EROI.

Transportation costs: Oil weighs roughly 0.136 tons per barrel. Transportation by truck uses about 3400 BTU/ton-mile or 3.58 MJ per ton-mile. Transportation by fuel pipeline requires 500 BTU/ton-mile or 0.52 MJ per ton-mile. We assume that the average distance that oil moves from port or oil field to market is about 600 miles. Thus a barrel of oil, with about 6.2 gigajoules of contained chemical energy, requires on average about 600 miles of travel × 0.136 tons per barrel × 3.58 MJ per ton-mile = 292 MJ per barrel spent on transport or about 5% of the total energy content of a barrel of oil to move it to where it is used (Table 18.1). If the oil is moved by pipeline (the more usual case), this percentage becomes about 1%. We assume that coal moves an average of 1500 miles, mostly by train at roughly 1720 BTU per ton-mile or about 1.81 MJ per ton-mile, so that the energy cost to move a ton of bituminous coal with about 32 MJ per kilogram (kg) (32 GJ/Ton) to its average destination is 1500 miles

Table 18.1 The energy cost of transporting oil and coal

	Energy cost (MJ/ton-mile)	Miles traveled	Energy cost(MJ)	Energy cost of energy unit delivered
Oil truck	3.58[2]	600	292	5%
Pipeline	0.52[2]	600	42	1%
Coal train	1.81[2]	1500	2715	8%

Sources: (1) Energy unit delivered: oil = 1 barrel = 6.2 GJ/barrel; Coal = 1 ton = 32 GJ/Ton. (2) Ref. [1].

× 1.81 MJ per ton-mile = 2715 MJ per ton, or 2.715 GJ per ton of coal, which is about 8 percent (Table 18.1). Line losses if shipped as electricity are roughly similar. So adding between 1 and 8 percent of the energy value of fuels for delivery costs does not seem unreasonable. Perhaps 25 percent of the energy in natural gas is used to move the gas down the pipeline, and an unknown but significant amount of energy is required to build and maintain the pipeline. We assume that these costs would decrease all EROIs by a conservative 5 percent to get it to the user; in other words the fuel must have an EROI of at least 1.05:1 to account for delivery of that fuel.

Thus we find that our $EROI_{pou}$ is about 32 percent less (17 percent nonfuel loss plus 10 percent to run the refinery plus about 5 percent transportation loss) than the $EROI_{mm}$ indicating that at least for oil, one needs an EROI at the mine mouth of roughly 1.5:1 (i.e., 1.0/0.68) to get that energy to the point of final use.

18.7 Extended EROI: Calculating EROI at the Point of Use Correcting for the Energy Required for Creating and Maintaining Infrastructure

We must remember that usually what we want is energy services, not energy itself, which usually has little intrinsic economic usefulness, e.g., we want kilometers driven, not just the fuel that does that. That means that we need to count in our equation not just the "upstream" energy cost of finding and producing the fuels themselves but all of the "downstream" energy required to deliver the service (in this case transportation) including that for (1) building and maintaining vehicles, (2) making and maintaining the roads used, (3) incorporating the depreciation of vehicles, (4) incorporating the cost of insurance, (5) etc. All of these things are as necessary to drive that mile as the gasoline itself, at least in modern society. For the same reason, businesses pay some 55 or 60 cents per mile when a personal car is used for business, not just the 10 cents or so per mile that the gasoline costs. So in some sense, the energy required for delivering the service (a mile driven) is some four to five times the direct fuel costs, and this does not include the taxes used to maintain most of the roads and bridges. Now many of these costs, especially insurance, use less energy per dollar spent than fuel itself and also less than that for constructing or repairing automobiles or roads, and certainly this is not the case with the money used to deliver the fuel itself used in these operations.

On the other hand, the energy intensity of one dollar's worth of fuel is some eight times greater than that for one dollar's worth of infrastructural costs. Table 18.3 gives our estimates of the energy cost of creating and maintaining the entire infrastructure necessary to use all of the transportation fuel consumed in the United States. The energy intensities are rough estimates of the energy used to undertake any economic activity derived from the national mean ratio of GDP to energy (about 8.7 MJ/dollar), from the Carnegie-Mellon energy calculator website, and from Robert Herendeen (personal communication). Specifically, Herendeen estimates for 2005 that heavy construction uses about 14 MJ per dollar. Since in the 1970s insurance and other financial services had about half the energy intensities as heavy industry, our estimate of the energy required for infrastructure replacement and maintenance for the entire United States for 2005 is equal to about 38% of the energy used as fuel itself.

Our calculation, then, of adding in the energy costs of getting the fuel to the consumer in a usable form plus the energy cost of the infrastructure necessary to use the fuel is equal to about 0.32 plus 0.375, respectively, or about 0.695 in total. Thus the EROI at the wellhead necessary to provide transportation from crude oil is 3.3 to 1 (1/0.305). Thus to deliver the transportation services associated with one gallon of fuel put in a car or truck requires more than three gallons produced at the wellhead and probably similar proportions for other types of fuels.

Future research has extended our EROI to including the energy of all of the people and economic activity included directly and indirectly to deliver the energy. This is presented in the next chapter. Since, as we have indicated, roughly 10 percent of the economy is associated with getting energy (this includes even those farmers who grow the grain or laborers who build the airplanes used indirectly to feed laborers or to get engineers to the site), we might say that as a nation that part of the denominator for the $EROI_{ext}$ would be 10 percent of all of the energy used in the country.

An important issue here is EROI vs. conversion efficiency. The EROI technically measures just the energy to get it to some point in society, usually the wellhead. But if we then say "to the consumer," we have to include the refinery losses and energy costs and also the costs to deliver the fuel to the final consumer. It may also include the energy costs of maintaining the infrastructure to use that fuel. This is in reality a bleeding off of the energy delivered or a conversion efficiency of delivering one barrel of oil into transportation services. So whether we should say "the minimum EROI is 3.3:1" or, somewhat more accurately, that to deliver one barrel of fuel to the final consumer and to use it requires a little more than three barrels to be extracted from the ground is somewhat arbitrary, although the second way is technically more correct. Thus given that our national goal is to deliver 36 billion gallons of corn-based ethanol to our drivers, then if we were to include all the costs of getting and using that ethanol, something like 100 billion gallons would be required. Thus ethanol use is subsidized by the transportation infrastructure paid for with petroleum and its taxes and subsidies, and so on.

18

Thus by both economic (▪ Fig. 18.1) and energetic (i.e., assuming an EROI of 10:1) measures calculated here, it appears that at present roughly 10 percent of our economy is required to get the energy to run the other 90 percent, so that in some total sense (including the entire refining, conversion, and delivery chain), the mean EROI for our society is very roughly 10:1, and this seems to be true if numerator and denominator are in either dollars or in energy. (Note: Our use of relatively cheap coal and hydroelectricity, both with a relatively high EROI, lifts the actual ratio "at the wellhead" so that the EROI for energy delivered to society, but not the consumer, is roughly 20:1. By the time the energy is delivered to the consumer, that has fallen to roughly 10:1 considering the larger perspective of the entire energy delivery system—see below). The above analysis suggests that by the time the energy is delivered to the consumer the EROI is cut roughly by two-thirds (0.695). Thus by the time an alcohol fuel with an EROI of 3.3:1 is delivered to the consumer, that is, after the energy costs of refinement and blending, transport, and so on are included, it may no longer deliver an energy surplus. The reason for this decrease in end-user EROI is that it is energy services that are desired, not energy itself, and to create these energy services requires energy transformations that carry, at a minimum, large entropic (conversion efficiency) losses.

Thus we get $EROI_{ext}$ "extended EROI," which modifies that equation to include the energy required not only to get but also to use the energy. We define it formally here as:

$$EROI_{ext} = \frac{\text{Energy returned to society}}{\text{Energy required to } \textit{get}, \textit{deliver}, \textit{and use} \text{ that energy}}$$

This concept is summarized in ▪ Fig. 18.2.

18.8 Why Should EROI Change Over Time? Technology Versus Depletion

There are two basic perspectives, with two very different groups of followers, relating to the long-term trend of efficiency in the production of oil (and other nonrenewable) resources. Many resource analysts emphasize the importance of depletion as humans exploit and eventually exhaust higher grades (i.e., more concentrated resources from more accessible deposits) over time. On the other hand, many resource analysts have expressed serious concern about the depletion of fuels and other resources over time. Why should this happen? Well essentially any business, such as a mining enterprise, is interested in making the maximum profit possible. The best first principle states that humans use the highest-quality sources of natural resources first as this would lead to higher profits. This concept was also of great interest to the classical economist David Ricardo. Given a choice, humans will grow crops on the more fertile soils, mine copper that is 10 rather than 1000 ft. deep, harvest timber from forests that are closer to roads and sawmills, fish larger, closer coastal concentrations, and so on. As the high-quality resources are depleted, lower-quality resources are used. This principle is well understood in economics based on work conducted 200 years ago by David Ricardo and is called the principle of diminishing returns.

18.8.1 EROI for US and North American Domestic Resources and Its Implications for the "Minimum EROI"

We start with historical, ecological, and evolutionary considerations, both because they have helped us a great deal to clarify our own perspectives on these issues and because, in the unsubsidized world where evolution operates, there are no bailouts or explicit subsidies, a very different situation from the one in which we operate in human society today.

In the past, Charles Hall worked with Cutler Cleveland and Robert Kaufmann to define and calculate the energy return on investment (EROI) of the most important fuels for the United States' economy [e.g., 12, 13]. Since that time both Cleveland and Hall (and their coworkers) have undertaken additional and updated analyses for the US oil and gas industry [14–17, 27], and Gagnon and Hall [28] have done that for the world average of private companies (no analyst can get the required information from national oil companies such as Saudi Aramco). Our results indicate that there is still a very large energy surplus from fossil fuels—variously estimated as an EROI (i.e., $EROI_{mm}$) from perhaps 80 to 1 (domestic coal and perhaps some gas) to 11–18 to 1 (US) to 20 to 1 (world) for contemporary oil and gas globally (Table 18.2). In other words, for every barrel of oil, or its equivalent, invested globally in seeking and producing more oil, some 10–20 barrels are delivered to society. Thus fossil fuels still provide a very large energy surplus, obviously enough to run and expand the human population and the very large and complex industrial societies around the world. This surplus energy of roughly 10–20 or more units of energy returned per unit invested in getting it, plus the large agricultural yields generated by fossil-fueled agriculture, allows a huge surplus quantity of energy, including food energy, to be delivered to society. This in turn allows most people and capital to be employed somewhere else other than in the energy industry. In other words these huge energy surpluses have allowed the development of all aspects of our civilization—both good and bad.

But the problem with present substitutes to fossil fuels is that, of the alternatives available, none appear to have the desirable traits of fossil fuels. These include (1) sufficient energy density (Table 4.1), (2) transportability, (3) relatively low environmental impact per net unit delivered to society, (4) relatively high EROI, (5) are obtainable on a scale that society presently demands, and (6) have the needed storage qualities. All of these would presumably greatly reduce the EROI of renewable fuels such as solar PV and wind, but the necessary calculations have not yet been undertaken. Wind, certainly, can make important contributions, but can it be used to generate a very large proportion of society's energy with the required backups or storage? There is no good answer yet.

Thus it would seem that society, both the US and the world, is likely to be facing a decline in both the quantity and EROI of its principal fuels. Our next question is "what are the implications of this?"

Table 18.2 Published EROI values for various fuel sources and regions

Resource	Year	Country	EROI (X:1)[a]	Reference
Fossil fuels (oil and gas)				
Oil and gas production	1999	Global	35	Gagnon (2009)
Oil and gas production	2006	Global	18	Gagnon (2009)
Oil and gas (domestic)	1970	US	30	Cleveland et al. [12], Hall et al. [13]
Discoveries	1970	US	8	Cleveland et al. [12], Hall et al. [13]
Production	1970	US	20	Cleveland et al. [12], Hall et al. [13]
Oil and gas (domestic)	2007	US	11	Guilford et al. [27]
Oil and gas (imported)	2007	US	12	Guilford et al. [27]
Oil and gas production	1970	Canada	65	Freise (2011)
Oil and gas production	2010	Canada	15	Freise (2011)
Oil, gas, and tar sand production	2010	Canada	11	Poisson and Hall (2011)
Oil and gas production	2008	Norway	40	Grandell (2011)
Oil production	2008	Norway	21	Grandell (2011)
Oil and gas production	2009	Mexico	45	Ramirez, in preparation
Oil and gas production	2010	China	10	Hu et al. (2013)
Fossil fuels (others)				
Natural gas	2005	US	67	Sell et al. (2011)
Natural gas	1993	Canada	38	Freise (2011)
Natural gas	2000	Canada	26	Freise (2011)
Natural gas	2009	Canada	20	Freise (2011)
Coal (mine mouth)	1950	US	80	Cleveland et al. [12]
Coal (mine mouth)	2000	US	80	Hall and Day (2009)
Coal (mine mouth)	2007	US	60	Balogh et al. unpublished
Coal (mine mouth)	1995	China	35	Hu et al. (2013)
Coal (mine mouth)	2010	China	27	Hu et al. (2013)
Other nonrenewables				
Nuclear	n/a	US	5 to 15	Hall and Day (2009), Lenzen (2008)
Renewables[b]				
Hydropower	n/a	n/a	>100	Cleveland et al. [12]
Wind turbine	n/a	n/a	18	Kubiszewski et al. (2010)
Geothermal	n/a	n/a	n/a	Gupta and Hall (2011)
Wave energy	n/a	n/a	n/a	Gupta and Hall (2011)
Solar collectors[b]				
Flat plate	n/a	n/a	1.9	Cleveland et al. [12]

Table 18.2 (continued)

Resource	Year	Country	EROI (X:1)[a]	Reference
Concentrating collector	n/a	n/a	1.6	Cleveland et al. [12]
Photovoltaic	n/a	n/a	6 to 12	Kubiszewski et al. (2009)
Passive solar	n/a	n/a	n/a	Cleveland et al. [12]
Passive Photovoltaic	2008	Spain	3 to 4	Prieto and Hall (2013)
Passive Photovoltaic	2008	n/a	10 to 15	Raugei et al. (2012)
Corn-based ethanol	n/a	US	0.8 to 1.6	Patzek (2004), Farrell et al. (2006)
Biodiesel	n/a	US	1.3	Pimentel and Patzek (2005)

From [16]
[a]EROI values in excess of 5:1 are rounded to the nearest whole number
[b]EROI values are assumed to vary based on geography and climate and are not attributed to a specific region/country. See [5, 16] for somewhat more comprehensive analyses which continue to evolve

Table 18.3 Breakdown of the downstream energy costs of refining, transporting, and using one barrel of oil

Process	Energy cost (%)	Energy Cost (GJ)[a]
Nonfuel refinery products[b]	17.0	1.11
Energy used in refining[c]	10.0	0.51
Transport to consumer[d]	5.0	0.23
Energy cost of transportation system[e]	37.5	2.36
Final energy delivered to consumer	*30.5*	*1.99*

[a]1 barrel of crude oil starts with 6.2 GJ
[b]EIA accessed 2007 (▶ http://www.eia.doe.gov/bookshelf/brochures/gasoline/index.html)
[c]Szklo and Schaeffer [26]
[d]Mudge et al. [28]
[e]See Table 18.2

18.9 The Surplus Available to Run the Rest of the Economy

We first generate a simplistic view of the economy in every day units to try to develop for the reader an explanation of how an economy obtains the energy needed for its own function and how differences in EROI might affect that. Assume for the moment that the United States' economy runs 100 percent on domestic oil and that energy itself is not what is desired by the final consumer but rather the goods and services derived from the general economy. In 2016 the US gross domestic product was about 17 trillion dollars, and it used about 97.4 quadrillion BTUs (called quads, equal to 10^{15} BTUs), which is the metric equivalent of about 103 exajoules (1 EJ equals 10^{18} joules). Dividing the two we find that we used an average of about 6 megajoules (1 MJ equals 10^{6} joules) to generate on average one dollar's worth of goods and services in 2017. By comparison, gasoline at $2.50 per gallon delivers about 52 MJ per dollar (at 131 MJ per gallon of gasoline), plus roughly another 20 percent to get that gasoline (extraction and refinery cost = 10 MJ), so if you spend one dollar on energy directly vs. one dollar on general economic activity, you would consume about 62/10 or 10 times more energy.

What is the energy "price" of the oil in the above example (1) to the country (either domestic or if it is imported) and (2) to the consumer—relative to the total economic activity of each entity? One can do some simple math. There are about 6.1 GJ in a standard 42 gallon barrel of oil, so the 97.4 exajoules of industrial energy the United States uses to run its economy for a year would require roughly 17 billion barrels of oil. At $50 per barrel, that amount of oil would take 850 billion dollars to purchase (or at $2.50 a gallon $ 1.8 trillion to the consumer), which is either about 5 percent of GDP or one-tenth if we consider it from the perspective of the consumer (the difference between the two estimates going to the oil companies after production for distribution and profits or to refineries, gas station attendants, etc. as inputs, profits, wages, delivery costs, etc.). Thus the price of energy delivered to the consumer is roughly twice that of the wellhead price (or about three times more than that if converted to electricity). But in 2017 we have to conclude that the energy cost of running our society is cheap.

Now assume that the real price of oil, that is, the price of oil relative to other goods and services, increased by three, that is, to $140 a barrel in today's dollars (which it did briefly in 2008), and that the total size of the economy stayed the same—that is, some other components of the economy were diverted to pay for that oil. If that happened, then about (17 billion times $140 = $2.38 trillion/17 trillion) or 15 percent of the economy would be used to buy the oil to run the other 85 percent (i.e., that part not including the energy extraction system itself). If the price of oil increased to $250 per barrel, about one-third of all economic activity would be required to run the other two-thirds, and at $1000 a barrel, then the output of the entire economy, that is, 17 trillion dollars, would be required to generate the money to purchase the energy required to run the economy, i.e., there would be no net output. While in fact in a real economy there would be many adjustments, alternative fuels, and nuances, this analysis does at least give an overview of the relation of gross to net economic activity and the importance of high EROI in energy and economic terms to the profit of the rest of the economy. As the price of fuel increases (i.e., as its EROI declines), there are large impacts on the rest of the economy. These impacts can be especially influential because changes in the price of energy tend to impact discretionary, not baseline, spending. The implications are explored in the next chapter.

Of course most of our energy costs less than oil so that the 50 dollars a barrel we used in the example above translates to—in the real economy—the equivalent of about $35 a barrel equivalent at the source or $70 a barrel by the time the consumer gets the energy; hence we can assume for this scenario that on average about 5 percent of the dollar economy (i.e., $50 times 17 billion barrels or 800 billion out of 17 trillion dollars) is used just to purchase the energy that allows the rest of the economy to function, which produces the end products we want. This 5 percent of our total economic activity means that roughly 5 percent of all workers' time, 5 percent of the energy used in their jobs, and 5 percent of the total materials consumed were used in some sense to get the energy to the final consumer to make the rest of the economy work. According to the official statistics of the US Energy Information Agency in 2017, the cost of energy to the consumer was about 5 percent of total incomes (Fig. 18.1), so our numbers seem about right on average.

18.10 EROI of Obtaining Energy Through Trade

Now let us assume that the economy runs 100 percent on imported oil (this is the case for many small nations). An economy without enough domestic fossil fuels of the type it needs must import the fuels and pay for them with some kind of surplus economic activity. The ability to purchase the critically required energy depends upon what else it can generate to sell to the world as well as the fuel required to grow or produce that material. For example, Costa Rica to a large degree pays for its imported oil with exported bananas and coffee. These are commodities highly valued in the world and hence readily sold. They are also quite energy intensive to produce, however, especially when produced of the quality that sells in the rich countries. For example, bananas require an amount of money equivalent to about half of their purchase price to pay for the fuel and petrochemicals required for their production including their cosmetic quality. So in this case and in other such cases, the EROI for the imported fuel is the relation between the amount of fuel bought with a dollar or euro and the amount of dollar or euro profits gained by selling the goods and services for export. And the quantity of the goods or services exported to attain a barrel of oil depends upon the relative prices of the fuel vs. the exported commodities.

Kauffman [29] estimated from roughly 1950 through the early 1980s the energy cost of generating a dollar's worth of our major US exports, e.g., wheat, commercial jetliners, etc., and also the chemical energy found in one dollar's worth of imported oil. The concept was that the EROI for imported oil depended upon what proportion of an imported dollar's worth of oil did you need to use to generate the commodities and hence the money obtained from overseas sales that you traded, in a net sense, for that oil. He concluded that before the oil price increases of the 1970s, the EROI for imported oil was about 25:1, very favorable for the United States, but that dropped to about 9:1 after the first oil price hike in 1973 and then down to about 3:1 following the second oil price hike in 1979. The ratio has returned to a more favorable level (from the perspective of the United States) since then because the price of exported goods has increased through inflation more rapidly than the price of oil. As oil prices increased again in this decade, however, and as more of the remaining conventional oil is concentrated in fewer and fewer countries and with their own internal use increasing, the future supply of abundant conventional oil in question, estimating the EROI of obtaining energy through trade may be very useful in predicting economic vulnerability in the near future (Fig. 18.4). As of mid-2017, the United States imports nearly half of the oil it uses.

Now let's revisit our previous example and assume that the US economy of 2007 runs entirely on imported oil rather than domestic oil. Neglecting for the moment debt and certain financial transactions such as cost of transport

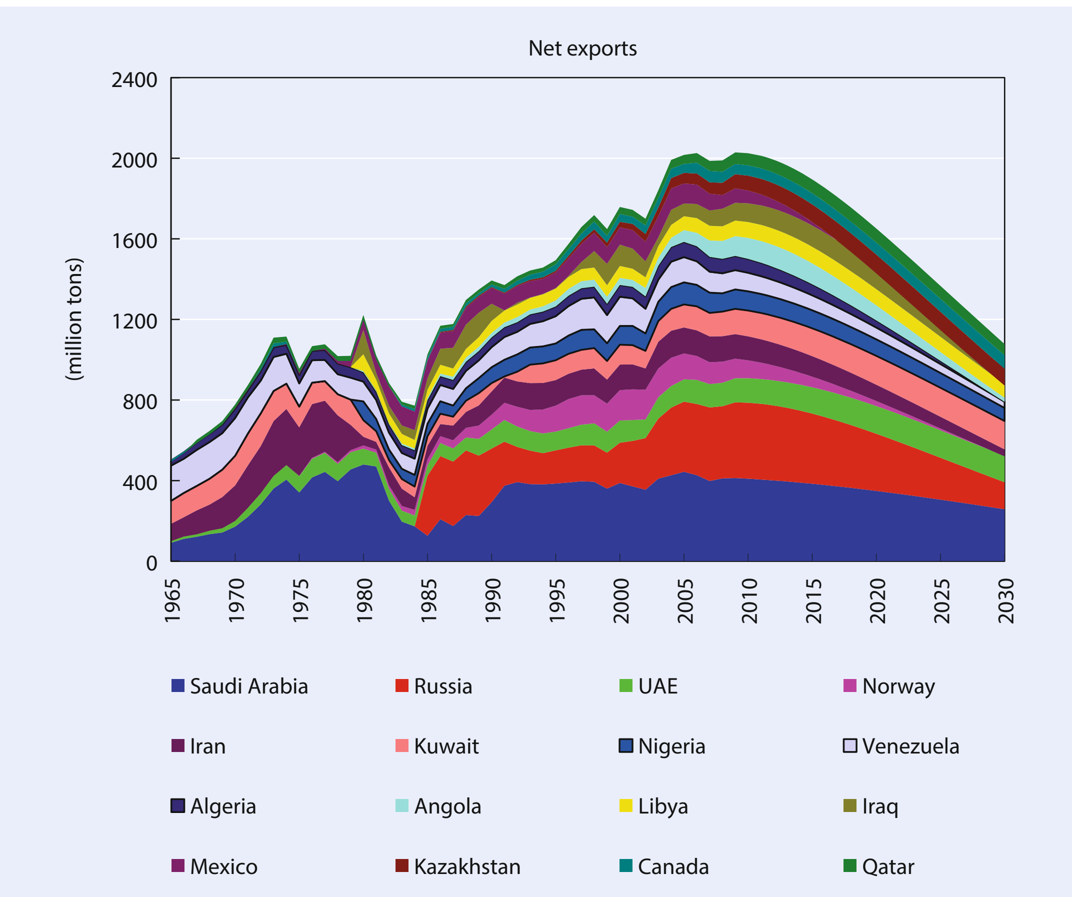


Fig. 18.4 EROI for imported oil for developing nations. (From Lambert et al. [17])

and foreigners investing in our banks, we, in a net sense, take oil, invest it in the economy, sell some of the products abroad to generate foreign exchange, and then use that foreign exchange to purchase oil from someone else—which we then use in the economy to generate more goods and services. To get the 1.2 trillion dollars' worth of oil (17 billion barrels times $70 a barrel) that we would be importing under this scenario, we would have to sell at least 1.2 trillion dollars' worth of our production abroad, which would require $1.2 trillion times 8.7 MJ used per average dollar generated in the economy or 10.4 EJ of our own energy. Thus about one-tenth (10.4 EJ of 105 EJ) of our total energy used and a roughly similar amount of our total economic activity would be required just to get the energy required to run the rest of the economy which produces the goods and services we want. Thus the EROI is about 10:1. This is still a pretty favorable return but only about 40 percent as favorable as it was in 1970 when it was 25:1 or even in 1998. To some degree we have managed to continue to do this through debt (much from Japan and China), which gives us a temporarily higher EROI. Were we to pay off this debt in the future and those who got the dollars wished to turn them into real goods and services (which seems a reasonable assumption), then we will have to take some substantial part of our remaining energy reserves out of the ground and convert it into fish, rice, beef, Fords, and so on that those people would be able to buy from us.

The dollar return on dollars invested is similar: 1.2 trillion of foreign exchange would be required to buy the oil (energy) that allows one to generate in the economy 17 trillion dollars, assuming that we ran only on imported oil. But if the price of oil inflates more rapidly than the prices of goods and services traded for the oil, then the portion of economic activity dedicated to raising foreign exchange to get that oil must increase unless the economy gets more efficient, a complex but probably oversold issue we will avoid here. Cleveland et al. [12] found a very high correlation between quality-corrected energy use and GDP from 1904 to 1984. Since then the economy has increased faster than energy use—although if one uses inflation rates calculated using the pre-Clinton era equation for CPI—such as that provided by ► www.shadowstatistics.com, the GDP declines, and the tight relation between GDP and energy use returns. Nonetheless, we believe that sharp increases (or decreases) in the price of imported oil will probably cause a series of structural changes to our economy that most people will not find particularly desirable. In fact, it is hard to ignore the coincident timing between the increases in the real price of oil culminating in the summer of 2008 and the subsequent financial collapse toward the end of the summer/fall 2008.

18.10.1 The Trade-off between EROI and Total Energy Used in Generating "Civilization"

The basic goods and services that we desire and require to have what we call modern civilization are highly dependent upon the delivery of net energy to society. This is a point made again and again by the authors quoted in the introduction to this chapter. But the total net energy that we have at our disposal, say roughly 90 percent of the 100 or so quads (or 105 EJ), would decrease to 80 if the cost of energy were to double (as what happened in the first part of 2008) or down to 60 if it were to double again and so forth, all of which is very possible. From this perspective, we think it very likely that EROI is likely to become an extremely important issue in defining our future economy and quality of life.

18.11 Conclusion

Our educated guess is that the minimum EROI required for a fuel that will deliver a given service (i.e., miles driven, house heated) to the consumer will be about 3:1 when all of the additional energy required to deliver and use that fuel are properly accounted for. This ratio would increase substantially if the energy cost of supporting labor (generally considered consumption by economists although definitely part of production here) or compensating for environmental destruction were included. While it is possible to imagine that one might use a great deal of fuel with a low EROI, say 2:1, to run an economy, this would mean that half of all economic activity was required to generate the other half, and we would be very, very poor. Even this would not be enough to do anything of economic utility, such as operating a truck. Thus we introduce the concept of "extended EROI" which includes not just the energy of getting the fuel but also that of transporting and using it. This process approximately triples the EROI required to use the fuel once obtained from the ground. Any fuel with an EROI less than the mean for society (about 10 to 1) may in fact be subsidized by the general petroleum economy. For example, fuels such as corn-based ethanol that have marginally positive EROIs (1.3:1) will be subsidized by the infrastructure support (i.e., construction and maintenance of roads and vehicles) undertaken by the main economy which is two-thirds based on oil and gas. These may be more important questions than the exact math for the fuel itself, although all are important.

Finally future analysis might even go so far as to include the money/energy to support and replace the oil worker. We believe this is important as there is little argument about the need to amortize the maintenance and depreciation of the oil derrick, so why not some prorated portion of medical care for the worker or education of his or her children for eventual replacement of the worn-out worker? Mainstream economists have some serious problems with this line of reasoning because they say that, e.g., medical care of workers or their children is consumption, not production. But, as with energy itself, a certain amount of consumption is essential for production and maybe we need to rethink when and how we draw the line between them. Perhaps it is best considered from the perspective of the two paragraphs above: as the EROI of fuels presumably declines into the future, then the rest of us will be supporting more and more workers in the energy industry, and there will be fewer and fewer net dollars and energy delivered to the rest of society. And if we are to support all the infrastructure to train engineers, physicians, and skilled laborers needed by society, we would need a far higher EROI from our primary fuels. This is explored in the next chapter.

Questions

1. Define net energy, energy surplus, and EROI. Are they just different ways to say the same thing?
2. Who were some of the pioneering thinkers about net energy?
3. Why do they think of net energy as a "determinant of human culture"? Do you agree?
4. Define the economic cost of energy.
5. What are some of the precedents that led to the development of the concept of EROI? What was the role of fish?
6. EROI can be calculated at various points, starting at the wellhead or the farm gate. Give some additional places in the energy use "food chain" where it might be useful to calculate EROI?
7. Do you think that including the energy to maintain the infrastructure required to use a fuel should be included in an EROI assessment? Why or why not?
8. What are some typical EROIs for various fuels in the US?
9. Why would environmental considerations change an estimate for EROI?
10. What is the approximate proportion of our economy attributable to energy costs?
11. Explain how a nation with no energy resources invests energy to get energy.
12. What is the relation between EROI and the amount of, e.g., education, medical care, and culture that a society can sustain?

References

1. Modified from Hall, C.A.S., S. Balogh, and D.J.R. Murphy. 2009. What is the minimum EROI that a sustainable society must have? *Energies* 2: 25–47. which contains additional supporting references. SEE Hall, C.A.S. 2017. Energy return on investment (Springer) for a more thorough treatment of EROI.
2. Odum, H.T. 1972. *Environment, power and society*. New York: Wiley-Interscience; Odum, H.T. 1973. Energy, ecology and economics. Royal Swedish Academy of Science. *AMBIO* 2(6):220–227.
3. White, L.A. 1943. Energy and the evolution of culture. *American Anthropologist* 45: 335–356.
4. Cottrell, F. 1955. *Energy and society*. New York: McGraw-Hill.
5. Hall, C.A.S. 2017. *Energy return on investment: A unifying principle for biology, economics and development*. New York: Springer.
6. Hamilton, J.D. 2009. Causes and consequences of the oil shock of 2007–08. Brookings Papers on Economic Activity, 1, Spring, 215–261.; Hamilton, J.D. 2011. Nonlinearities and the Macroeconomic Effects of Oil Prices. *Macroeconomic Dynamics* 15: 472–497.
7. King, C.W. 2015. Comparing world economic and net energy metrics, Part 3: Macroeconomic Historical and Future Perspectives, *Energies* 8: 12997–13020.
8. King, C., and Hall, C.A.S. 2011. *Relating financial and energy return on investment: Sustainability: Special Issue on EROI*, 3:810–1832.
9. Hall, C.A.S. 1972. Migration and metabolism in a temperate stream ecosystem. *Ecology* 53: 585–604.

10. Hall, C.A.S., and C.J. Cleveland. 1981. Petroleum drilling and production in the United States: Yield per effort and net energy analysis. *Science* 211: 576–579.
11. Hall, C.A.S., C. Cleveland, and M. Berger. 1981. Energy return on investment for United States Petroleum, Coal and Uranium. In *Energy and ecological modeling,* ed. W. Mitsch, 715–724. Symp. Proc., Elsevier Publishing Co.; Hall, C.A.S., M. Lavine and J. Sloane. 1979. Efficiency of energy delivery systems: Part I. An economic and energy analysis. Environ. Mgment. 3: 493–504.
12. Cleveland, C.J., R. Costanza, C.A.S. Hall, and R. Kaufmann. 1984. Energy and the United States economy: A biophysical perspective. *Science* 225: 890–897.
13. Hall, C.A.S., C.J. Cleveland, and R. Kaufmann. 1986. *Energy and resource quality: The ecology of the economic process*, 1986. New York: Wiley.
14. Cleveland, C.J. 2005. Net energy from the extraction of oil and gas in the United States. *Energy: The International Journal* 30 (5): 769–782.
15. Hall, C.A.S., and D. Hanson, eds. 2011. *Sustainability: Special Issue on EROI*. Sustainability 3: 1–393
16. Hall, C.A.S., J.G. Lambert, and S.B. Balogh. 2014. EROI of different fuels and the implications for society. *Energy Policy* 64: 141–152.
17. Lambert, J., C.A.S. Hall, S. Balogh, A. Gupta, and M. Arnold. 2014. Energy, EROI and quality of life. *Energy Policy* 64: 153–167.
18. Sundberg, U. 1992. Ecological economics of the Swedish Baltic Empire: An essay on energy and power, 1560–1720. *Ecological Economics.* 5: 51–72.
19. See review by Farrell et al. 2006, [Farrell, A.E., R.J. Plevin, B.T. Turner, A.D. Jones, M. O'Hare, D.M. Kammen. 2006. Ethanol can contribute to energy and environmental goals. *Science* 311: 506–508.] as well as the many responses in the June 23, 2006 issue of Science Magazine for a fairly thorough discussion of this issue.
20. Hall, C.A.S., B. Dale, and D. Pimentel. Seeking to understand the reasons for the different EROIs of biofuels. *Sustainability* 2011, 2011: 2433–2442.
21. Murphy, D.J., C.A.S. Hall, and R. Powers. 2011. New perspectives on the energy return on investment of corn based ethanol. *Environment, Development and Sustainability* 13 (1): 179–202.
22. Prieto, P., and C.A.S. Hall. 2012. *Spain's photovoltaic revolution: The energy return on investment*. New York: Springer. Prieto and Hall; Raugei, M., Pere Fullana-i-Palmer, and Vasilis Fthenakis. 2012. *The Energy Return on Energy Investment (EROI) of photovoltaics: Methodology and comparisons with fossil fuel life cycles. Energy Policy* 45: 576–582.
23. Raugei, M. Personal communication.
24. Murphy, D., Hall, C.A.S., Cleveland, C., P. O'Conner. 2011. *Order from chaos: A preliminary protocol for determining EROI for fuels*. Sustainability: Special Issue on EROI, 1888–1907.
25. Bullard, C.W., B. Hannon, and R.A. Herendeen. 1975. *Energy flow through the US economy*. Urbana: University of Illinois Press. Hannon, B. 1981. *Analysis of the energy cost of economic activities: 1963 2000*. Energy Research Group Doc. No. 316. Urbana: University of Illinois.; Herendeen, R., and C. Bullard. 1975. The energy costs of goods and services. 1963 and 1967. *Energy Policy* 3: 268.
26. Costanza, R. 1980. Embodied energy and economic valuation. Science 210: 1219–1224.
27. Guilford, M., C.A.S., Hall, P., O'Conner, and C.J., Cleveland. 2011. *A new long term assessment of EROI for U.S. oil and gas: Sustainability: Special Issue on EROI*. Pages 3: 1866–1887.; The Oil Drum: search for "Hall EROI".
28. Gagnon, N.C., A.S. Hall, and L. Brinker. 2009. A preliminary investigation of energy return on energy investment for global oil and gas production. *Energies* 2 (3): 490–503.
29. Kauffman. 1986. Energy return on investment for imported petroleum. In *Energy and resource quality: The ecology of the economic process*, ed. C.A.S. Hall, C.J. Cleveland, and R. Kaufmann. New York: Wiley.

18

Peak Oil, EROI, Investments, and Our Financial Future

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_19

19.1 Introduction

The enormous expansion of the human population and the economies of the United States and many other nations in the past 100 years have been facilitated by a commensurate expansion in the use of fossil fuels (Fig. 8.1) [1]. To many energy analysts, that expansion of cheap fuel energy has been far more important than business acumen, economic policy, or ideology, although they too may be important [1–15]. While we are used to thinking about the economy in monetary terms, those of us trained in the natural sciences consider it equally valid to think about the economy and economics from the perspective of the energy required to make it run. When one spends a dollar, we do not think just about the dollar bill leaving our wallet and passing to someone else's. Rather, we think that to enable that transaction, that is, to generate the good or service being purchased, an average of about 5000 kJ of energy (roughly half the amount of oil that would fill a standard coffee cup) must be extracted and turned into roughly a half kilogram of carbon dioxide. Take the money out of the economy and it could continue to function through barter, albeit in an extremely awkward, limited, and inefficient way. Take the energy out and the economy would immediately contract or stop. Cuba found this out in 1991 when the Soviet Union, facing its own oil production and political problems, cut off Cuba's subsidized oil supply. Both Cuba's energy use and its GDP declined immediately by about one-third, groceries disappeared from market shelves within a week, and soon the average Cuban lost 20 pounds [16]. Cuba subsequently learned to live, in some ways well, on about half the oil as previously, but the impacts were enormous. While the United States has become more efficient in using energy in recent decades, most of this is due to using higher-quality fuels, exporting heavy industry, and switching what we call economic activity (e.g., [17]), and many other countries, including efficiency leader Japan, are becoming substantially less efficient [18–20].

19.2 The Age of Petroleum

The economy of the United States and the world is still based principally on "conventional" petroleum, meaning oil, gas, and natural gas liquids (Fig. 19.1). *Conventional* means those fuels derived from geologic deposits, usually found and exploited using drill bit technology. Conventional oil and gas flows to the surface because of its own pressure or with pumping or additional pressure supplied by injecting natural gas, water, or occasionally other substances into the reservoir. *Unconventional* petroleum includes shale oil, oil

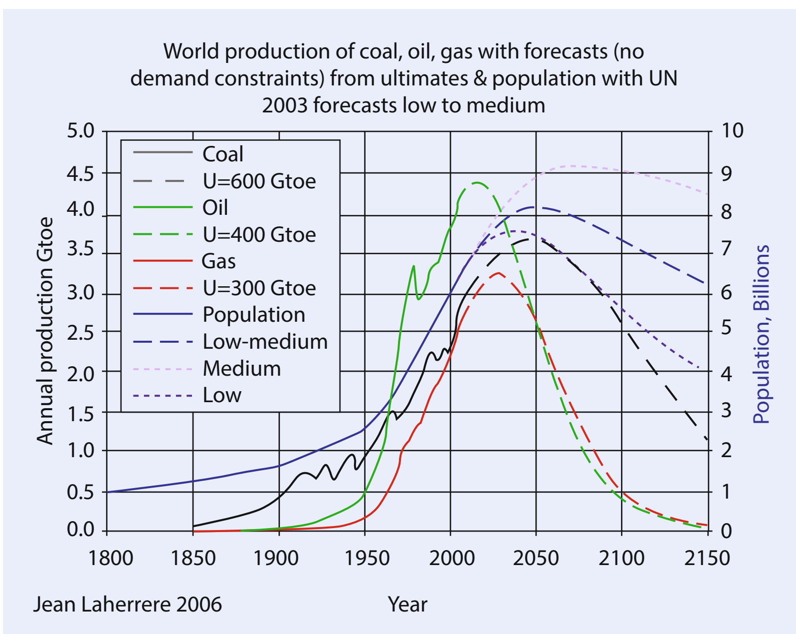


Fig. 19.1 Pattern of past fossil energy use and human population for the world, including projections (Gtoe = Giga tons oil equivalent. Source: Jean Laherrere)

sands, and other bitumens usually mined as solids and converted to liquids and also natural gas from coal beds and/or "tight" deposits where the gas is found in low concentrations in rock. For the economies of both the United States and the world, from half to two-thirds of our energy comes from conventional petroleum, about 30–40% from liquid petroleum, and another 20–25% from gaseous petroleum (Fig. 19.1). Coal, hydroelectric, and nuclear provide most of the rest of the energy that we use. Hydroelectric power and wood together are renewable energies generated from current solar input and provide about 5% of the energy that the United States and world uses, "New renewables" including windmills and photovoltaics generate about 2%. In recent years the annual increase in oil and gas use has been greater than the power coming from the new renewables or indeed their total production so that they are mostly not displacing fossil fuels but just adding to the mix. All of these proportions have not changed very much since the 1970s in the United States or the world. We believe it is most accurate to consider the times that we live in as *the age of petroleum*, for petroleum is the foundation of our economies and our lives. Just look around.

Petroleum is especially important because it has important and unique attributes leading to high economic utility that include very high energy density and transportability [20], massive availability, and relatively low price. Its future supply, however, is worrisome [21–23]. The issue is not the point between supply and potential demand. Barring a massive worldwide recession, demand will continue to increase, perhaps slowly, as human populations, petroleum-based agriculture, and economies (especially Asian) continue to grow. Petroleum supplies have been growing since 1900 at roughly four or five, but recently at two or one, percent per year. While most governments are trying to make their economies grow more rapidly, a trend many observers think is that high growth rates are unlikely to occur again anytime soon [23, 24]. Peak oil refers to the time at which an oil field, a nation, or the entire world reaches its maximum oil production and then declines. It is not some abstract issue debated by theoretical scientists or worried citizens but an actuality that occurred in the United States in 1970 and in some 60 (of 95) other oil-producing nations since [25–28]. Several prominent geologists have suggested that it may have occurred already for the world, although that is not clear yet, in part because the official statistics are including increasingly other liquid hydrocarbons such as natural gas liquids and biofuels under "oil" [29–31]. At some time, presumably, it will not be possible to continue to increase petroleum supplies or even to maintain current levels of supply, regardless of technology or price. At this point we will enter (or have entered) the "second half of the age of oil" [31]. The first half was one of year-by-year growth; the second half will be of year-by-year decline in supply, with possibly an "undulating plateau" around the peak. Natural gas will probably last a decade or two longer than oil as a major fuel source. We are of the opinion that it will not be possible to fill in the growing gap between supply and demand of conventional oil with alternatives on the scale required [32], and even were that possible, the investments in money, energy, and time required would mean that we needed to start some decades ago [33]. When or as the decline in global oil production begins, we will see the "end of cheap oil" and a very different economic climate.

The very large use of fossil fuels in the United States means that each of us has the equivalent of 60–80 hardworking laborers to "hew our wood and haul our water" as well as to grow, transport, and cook our food; make, transport, and import our consumer goods; and provide sophisticated medical and health services. Energy produced by these energy slaves even allows us to visit our relatives and take vacations in far away or even relatively nearby places. Simply to grow our food requires the energy of about a gallon of oil per person per day, and if a North American takes a hot shower in the morning, he or she will have already used far more energy than probably two-thirds of the Earth's human population use in an entire day.

19.3 How Much Oil Will We Be Able to Extract?

So the next important question is how much oil and gas are left in the world? The answer is a lot, although probably not a lot relative to our increasing needs and maybe not a lot of the high-quality stuff that we can afford economically or energetically. Although we will probably always have enough oil to lubricate our bicycle chains, the question is whether we will have anything like the quantity that we use now at the prices that allow the things we are used to having and whether growth is possible. Worldwide we have consumed about 1.3 trillion barrels of oil, mostly in the past 25 years. The current debate is fundamentally about whether there is 1, 2, or even 3.5 trillion barrels of economically extractable oil left. Fundamental to this debate, yet mostly ignored, is an understanding of the capital, operating and environmental costs, in terms of both money and energy, necessary to find, extract, and use whatever new sources of oil remain to be discovered and to generate whatever alternatives we might be able to develop. These investment issues, in terms of both money and energy, will become ever more important.

There are two distinct camps for this issue. One camp, the "technological cornucopians," led principally by economists such as Michael Lynch [34, 35], believes that market forces and technology will continue to supply (at a price) whatever oil we have a need for in the indefinite future. They argue that we now are able to extract only some 35% of the oil from a field, that large areas of the world (deep ocean, Greenland, Antarctica) have not been explored and may have substantial supplies of oil, and that substitutes, such as oil shale and tar sands, abound. They are buoyed by the failure of many earlier predictions of the demise or peak of oil production, two recent and prestigious analyses by the US Geological Survey and the Cambridge Energy Research Associates that tend to suggest that remaining extractable oil is near the high end given above, the recent discovery of the deepwater Jack 2 well in the Gulf of Mexico, and the development of the Alberta tar sands.

A second camp, the "peak oilers," is composed of scientists from diverse fields inspired by the pioneering work of M. King Hubbert [25], a few very knowledgeable politicians such as former U.S. congressman Roscoe Bartlett of Maryland, private citizens from all walks of life, and, increasingly, members of the investment community. Some of them come together once a year under the auspices of the International Society of BioPhysical Economics. All believe that there remains only about one additional trillion barrels of extractable conventional oil and that the global peak—or "bumpy plateau"—will occur soon or, perhaps, has already occurred (▣ Fig. 19.2). The arguments of these people and their organization, the Association for the Study of Peak Oil (ASPO), were spearheaded by the analyses and writings of geologists Colin Campbell and Jean Laherrere. They are supported by the many other geologists who agree with them, the many peaks that have already occurred for many dozens of oil-producing countries, the recent collapse of production from some of our most important oil fields, and that we now extract and use two to four barrels of oil for each new barrel discovered (▣ Fig. 8.3). They also believe that essentially all regions of the Earth favorable for oil production have been well explored for oil, and there are few surprises left except perhaps in regions that will be nearly impossible to exploit.

There are several issues that tend to add confusion to the issue of peak oil. First, some people do, and some do not, include natural gas liquids or condensate (liquid hydrocarbons that condense out of natural gas). These can be refined readily into motor fuel and other uses so that many investigators think they should simply be lumped with oil, which most usually they are. Since a peak in global natural gas production is thought to be likely one or two decades after a peak in global oil, inclusion of natural gas liquids extends the time or duration of whatever oil peak has occurred or may be occurring. The second is what characteristics of the peak will cause the largest economic impact? Is it the peak itself or the ratio between the declining production rate and the potential consumption rate? Both the production and the consumption of oil and also natural gas which had been growing at roughly 4% a year before 1970 declined gradually to 2% by 2005 and 1% or not at all since then. The great expansion of the economies of China and India has recently more than compensated for some reduced use in other parts of the world. Meanwhile the growth rate of the human population has continued so that "per capita peak oil" has probably occurred, perhaps as early as 1978 [36]. What the future holds possibly may have more to do with limiting carbon release than the declining physical production rate. Whenever we start on the inevitable downside of the global Hubbert curve, prices will rise.

The rates of oil and gas production (more accurately extraction) and the onset of peak oil are dependent upon interacting geological, economic, and political factors. The geological restrictions are the most absolute and depend on the number and physical capacity of the world's operating wells. In most fields the oil does not exist in the familiar liquid state but in what is more akin to a complex oil-soaked brick. The rate at which oil can flow through these "aquifers"

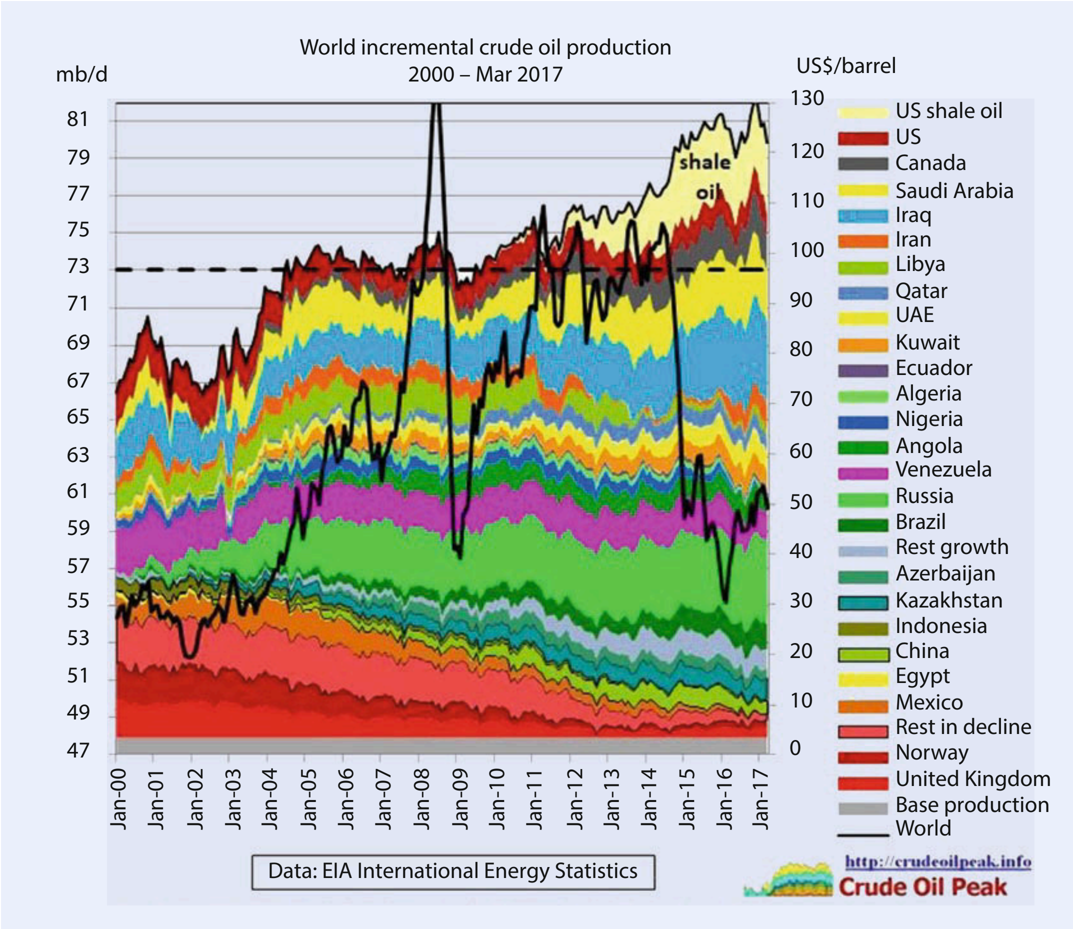


Fig. 19.2 Production of conventional oil in the world. This does not include recent additions of, e.g., natural gas liquids or biofuels to the data on "oil" but only conventional oil (Data source: Matt Mushalik)

depends principally upon the physical properties of the oil itself and of the geological substrate and upon the natural pressure forcing the oil through the substrate to the collecting wells. The natural pressures are increasingly replaced by pumping more gas or water into the structure. Detergents, CO_2, and steam can increase yields, but too-rapid extraction can cause compaction of the "aquifer" or fragmentation of flows which reduce yields.

So our physical capacity to produce oil depends upon our ability to keep finding large oil fields in regions that we can reasonably access, our willingness to invest in exploration and development, and our willingness to not produce too quickly. The usual economic argument is that if supply is reduced relative to demand, then the price will increase which will then signal oil companies to drill more, leading to the discovery of more oil and then additional supply. Although that sounds logical, the empirical record shows that the rate at which oil and gas is found has little to do with the rate of drilling (Fig. 8.4). Recent experience may be changing that for "tight" oil and gas, where smaller amounts (compared to the past) of oil and gas can be obtained by drilling many low yielding wells.

Finally, output can be limited or (at least in the past) enhanced for political reasons—which are even more difficult to predict than the geological restrictions. Certainly the events of the "Arab spring" of 2011 were completely unpredictable. Empirically there is a fair amount of evidence from post-peak countries, such as the United States, that the physical limitations become important when about half of the ultimately recoverable oil has been extracted. But why should that be? In the United States, it certainly was not due to a lack of investment, since most geologists believe that the United States had been over-drilled. We probably will not know until we have much more data, and much of the data are closely guarded industry or state secrets. But whether or not the world has reached peak oil, most individual producing countries have [36, 37]. According to one analyst, if one looks at all of the 60 or so post-peak oil-producing countries, the peak occurs on average when about 54% of the total extractable oil in place has been extracted [37]. Finally oil-producing nations often have high population and economic growth and are using an increasing proportion of their own production, leaving less for export [38].

The United States clearly experienced "peak oil" in 1970 (although this might or might not be followed by a second peak based on unconventional oil about now; see ▶ Chapter 13). As the price of oil increased by a factor of 10, from 3.50 to 35 dollars a barrel during the 1970s, a huge amount of capital was invested in US oil discovery and production efforts. The drilling rate increased from 95 million feet per year in 1970 to 250 million feet in 1985. Nevertheless the production of crude oil decreased during the same period from the peak of 3.52 billion barrels a year in 1970 to 3.27 in 1985 and has continued to decline to 1.89 in 2005 even with the addition of Alaskan production. (There has been an upswing to a possible second peak in 2017). Natural gas production has also peaked and declined, although less dramatically. Thus despite the enormous advancement of petroleum discovery and production technology and despite very significant investment, US production of conventional oil has continued its downward trend nearly every year since 1970, and the US still imports nearly half the oil it uses. When drilling rates are high, apparently poorer prospects, on average, tend to be drilled. The technological optimists are correct in saying that advancing technology is important. But there are two fundamental and contradictory forces operating here, technological advances and depletion. In the US conventional oil industry, it is clear that depletion is trumping technological progress, as oil production is declining and oil is becoming much more expensive to produce. Because oil exploration and development is very energy intensive, it can lead to less net oil being delivered to society. As of 2017 there is a lot of drilling, and a lot of production, taking place, but even at prices high by historical standards, almost none of the oil companies are making a profit.

19

19.4 Decreasing Energy Return on Investment

Energy return on investment (EROI or EROEI) is simply the energy that one obtains from an activity compared to the energy it took to generate that energy. The calculations are generally straightforward, although the data may be difficult to get and the boundaries uncertain (see previous chapter). When the numerator and denominator are derived in the same units, as they should be (the units can be barrels per barrel, kcals per kcal, or MJoules per MJoule), the results are in a unitless ratio. The running average EROI for the *finding* of US conventional oil has dropped from greater than 300 kJ returned per kilojoule invested in 1919 to about five for one today. The EROI for *producing* that oil has declined from 30 to 1 in the 1970s to around ten for one today. This illustrates the decreasing energy returns as oil reservoirs are increasingly depleted and as there are increases in the energy costs as exploration and development are increasingly deeper and offshore [13, 21, 39]. Even that ratio reflects mostly pumping out oil fields that are half a century or more old since we are finding few significant new fields. A new, or newly analyzed, troubling trend is that the EROI for "elephants," (i.e., the largest oil fields that still generate most of our oil), has been declining regularly in addition to their declining production [40]. The increasing energy cost of a marginal barrel of oil or gas is one of the factors behind their increasing dollar cost, although if one corrects for general inflation, the price of oil has increased only a moderate amount since 1970.

The same pattern of declining energy return on energy investment appears to be true for global petroleum production, but getting such information is very difficult. With help from the extensive financial database on "upstream" (i.e., preproduction) maintained by the John H. Herold Company, Gagnon and colleagues [41] were able to generate an approximate value for global EROI for producing new oil and natural gas (considered together). Their results indicate that the EROI for global oil and gas (at least for that which was publicly traded) was roughly 23:1 in 1992, increased to about 33:1 in 1999, and since then has fallen to approximately 18:1 in 2005. The apparent increase in EROI during the late 1990s reflects the effects of reduced drilling effort, as was seen for oil and gas in the United States (e.g., ◘ Figs. 19.3 and 19.4).

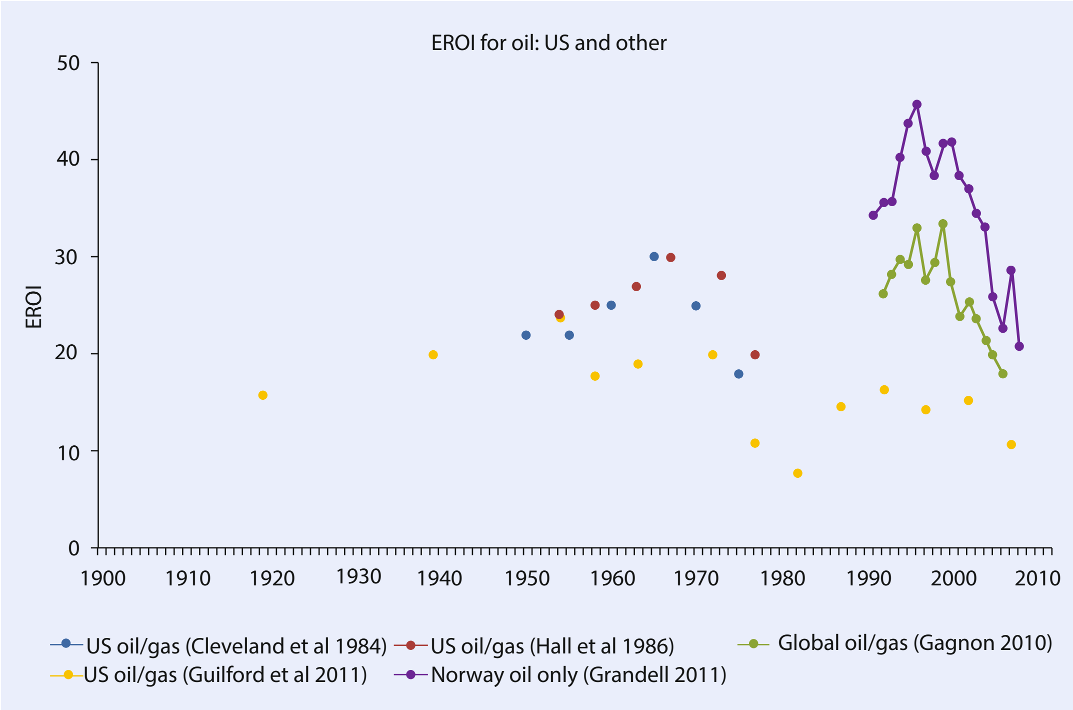


Fig. 19.3 EROI for oil and gas in the United States according to three more or less independent studies but all based on data from the US Bureau of Census (blue, yellow, and red dots: From Guilford et al. 2011). EROI for global publically traded companies (Green dots: From Gagnon et al. [41]. EROI for Norwegian oil: Grandell et al. 2011)

If the rate of decline continues linearly for several decades, eventually it would take the energy in a barrel of oil to get a new barrel of oil. While we do not know whether that extrapolation is accurate, essentially all EROI studies of our principal fossil fuels do indicate that their EROI is declining over time and that EROI declines especially rapidly with increased exploitation (e.g., drilling) rates. This decline appears to be reflected in economic results. In November of 2004, *The New York Times* reported that for the previous 3 years, oil exploration companies worldwide had spent more money in exploration than they had recovered in the dollar value of reserves found. The quantity of oil found in 2016 was only about 10% of the amount we produced and burned [42]. This illustrates that even though the EROI for producing oil and gas globally may still be about 15:1, it is possible that the energy breakeven point has been approached for finding new oil. Whether we have reached this point or not, the concept of EROI declining toward 1:1 makes irrelevant the reports of several oil analysts who believe that we may have substantially more oil left in the world. It simply does not make sense to extract oil, at least for fuel, when it requires more energy for the extraction than is found in the oil extracted.

How we weather this coming storm will depend in large part on how we manage our investments now. There are three general types of investments that we make in society. The first is investments into getting energy itself, the second is investments for maintenance of, and replacing, existing infrastructure, and the third is for discretionary expansion. In other words before we can think about expanding the economy, we must first make the investments into getting the energy necessary to operate the existing economy and also into maintaining the infrastructure that we have to compensate for the entropy-driven degradation of what we already have. The required investments into the second and especially the first category are likely to increasingly limit what is available for the third. The dollar and energy investments needed to get the energy needed to allow the rest of the

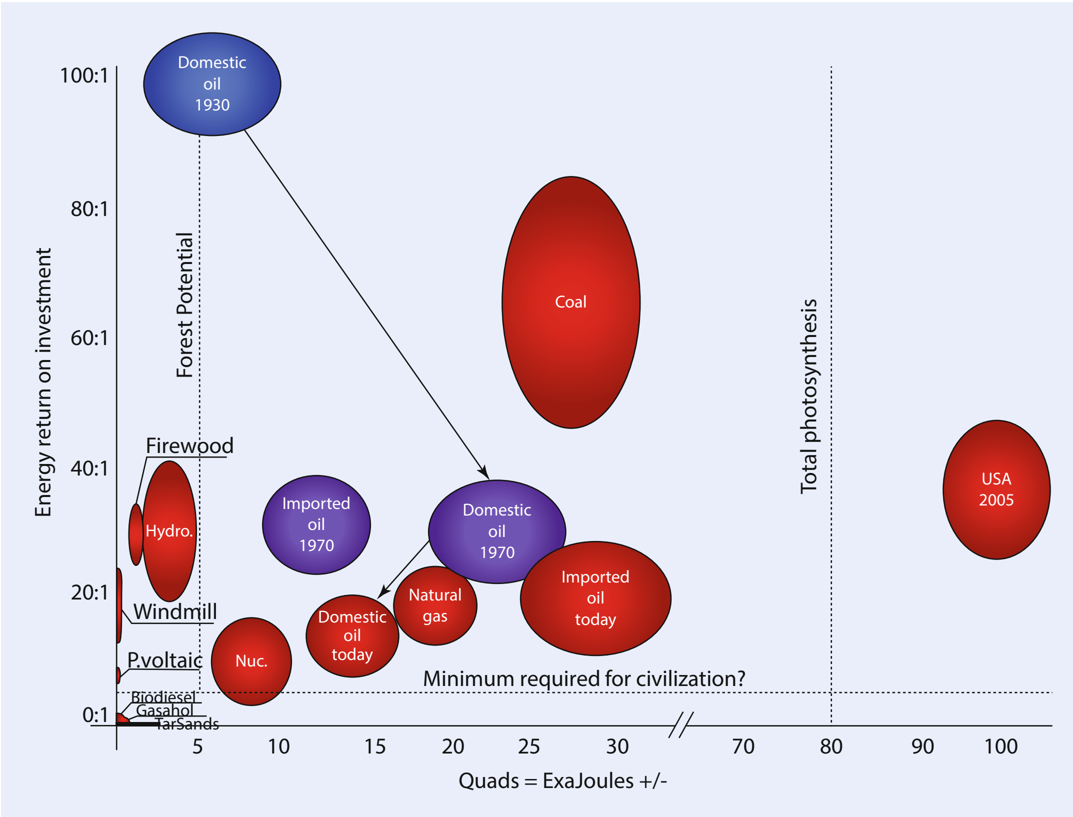


Fig. 19.4 "Balloon graph" representing quality (*y axis*) and quantity (*x axis*) for various fuels at various times in the US economy. *Arrows* connect fuels from various times (i.e., domestic oil in 1930, 1970, 2005), and the size of the "balloon" represents part of the uncertainty associated with EROI estimates (Source: US EIA, Cutler Cleveland and C. Hall's own EROI work). Note added in 2017: the high oil EROI value for 1930 represents the EROI for finding, not producing, oil and is slightly misleading although in a sense accurate. It might be better to use a value of 30:1 for 1970 which was the peak of EROI for production. See Guilford et al. (2011) for an update

economy to operate and grow have been very small historically, but this is likely to change dramatically. This is true whether we seek to continue our reliance on ever-scarcer petroleum or whether we attempt to develop some alternative. Technological improvements, if indeed they are possible, are extremely unlikely to bring back the low investments in energy that we have grown accustomed to.

The main problem that we face is a consequence of the "best first" principle. This is, quite simply, the characteristic of humans to use the highest-quality resources first, be they timber, fish, soil, copper ore, or fossil fuels. The economic incentives are to exploit the highest quality, least cost (both in terms of energy and dollars) resources first (as was noted by economist David Ricardo in 1891 [43]). We have been exploiting fossil fuels for a long time. The peak in finding new oil was in the 1930s for the United States and in the 1960s for the rest of the world. Both have declined enormously since then. An even greater decline has taken place in the efficiency with which we find oil, that is, the amount of energy that we find relative to the energy we invest in seeking and exploiting it.

That pattern of exploiting and depleting the best resources first also is occurring for natural gas. Natural gas was once considered a dangerous waste product of oil development and was flared at the well head. But during the middle years of the last century, large gas pipeline systems were developed in the United States and Europe that enabled gas to be sent to myriad users who appreciated its ease of use and cleanliness, including its relatively low carbon dioxide emissions, at least relative to coal [44]. US natural gas originally came from large fields, often associated with oil fields, in

Louisiana, Texas, and Oklahoma. Its production has moved increasingly to smaller fields distributed throughout Appalachia and the Rockies. A national peak in production occurred in 1973 as the largest fields that traditionally supplied the country peaked and declined. Later as "unconventional" fields were developed, a second, somewhat smaller peak occurred in the 2000s. Gas production had fallen by about 6% from that peak, and some investigators predict a "natural gas cliff" as conventional gas fields are increasingly exhausted and as it is increasingly difficult to bring smaller unconventional fields on line to replace the depleted giants. However, this "cliff" appears unlikely to occur for at least several decades because of the new technologies of horizontal drilling and hydrofracturing, which as of this writing are bringing in new "unconventional" gas at just about the rate that the conventional supplies are declining. It is quite difficult to predict the future of natural gas because of the many economic, environmental and social issues associated with horizontal drilling and hydraulic fracturing.

19.5 The Balloon Graph

All sources of energy used in the economy, except the free solar energy that drives ecosystem processes, have an energy cost, and all of them have different magnitudes of importance to society. The energy cost of obtaining coal or oil or photovoltaic electricity is straightforward even if difficult to calculate, but there are other sources and other ways payment is needed. For example, we pay for imported oil in energy as well as dollars, for it takes energy to grow, manufacture, or harvest what we sell abroad to gain the dollars with which we buy the oil (or we must in the future if we pay with debt today). In 1970 we gained roughly 30 mJ for each megajoule used to make the crops, jet airplanes, and so on that we exported [39]. But as the price of imported oil increased, the EROI of the imported oil declined. By 1974 that ratio had dropped to nine to one and by 1980 to three to one. The subsequent decline in the price of oil, aided by the inflation of the export products traded, eventually returned the energy terms of trade to something like it was in 1970, at least until the price of oil started to increase again after 2000, again lowering the EROI of imported oil. A rough estimate of the quantity used each year and the EROI of various major fuels in the United States, including possible alternatives, is given in ◘ Fig. 19.4. An obvious aspect of that graph is that qualitatively and quantitatively alternatives to fossil fuel have a very long way to go to fill the roles of fossil fuels. This is especially true when one considers the additional qualities of oil and gas, including energy density, ease of transport, and ease of use. The alternatives to oil available to us today are characterized by even lower EROIs, limiting their economic effectiveness. It is critical for CEOs and government officials to understand that the best oil and gas are simply gone, and there is no easy replacement.

If we are to supply into the future petroleum at the rate that the United States consumed in recent decades, let alone an increase, it will require enormous investments in either additional unconventional sources or payments to foreign suppliers. That will mean a diversion of the output of our economy from other uses into getting the same amount of energy just to run the existing economy. In other words, from a national perspective, investments will be needed increasingly just to run what we have, not to generate new real growth. If we do not make these investments, our energy supplies will falter, and if we do, the returns may be small to the nation, although the returns to the individual investor may be large. Further, if this issue is as important as we believe it is, then we must pay much more attention to the quality of the data we are getting about the energy costs of all things we do—including getting energy. Finally the failure of increased drilling to return more fuel (◘ Fig. 8.4) calls into question the basic economic assumption that scarcity-generated higher prices will resolve that scarcity by encouraging more production. Indeed scarcity encourages more exploration and development activity, but that activity does not necessarily generate more resources. Oil scarcity will also encourage the development of alternative liquid fuels, but their EROIs are generally very low.

19.6 Economic Impacts of Peak Oil and Decreasing EROI

Whether global peak oil has occurred already or will not occur for some years or, conceivably, decades, its economic implications will be enormous because we have no possible substitute on the scale required and at the EROI that is needed. Any alternatives will require enormous investments in money and energy when both are likely to be in short supply. Despite the projected impact on our economic and business life within relatively few years, neither government nor the business community is in any way prepared to deal with either the impacts of these changes or the new thinking needed for investment strategies. There are many reasons for this, but they include the role of economists in downplaying the importance of resources in the economy, the disinterest of the media, the failure of government to fund good analytic work on the various energy options, the erosion of good energy record keeping at the Departments of Commerce and Energy, and the focus of the media on trivial "silver bullets" despite the inability of any one of them (except economic contraction and in some few cases conservation) to contribute anything like 1% to the total energy mix.

Of perhaps greater concern is that none of the top ten or so energy analysts that we are familiar with are supported by government or, generally any, funding. There are not even targeted programs in the National Science Foundation or the Department of Energy where one might apply if one wishes to undertake good objective, peer-reviewed EROI analyses to see what options might actually be able to contribute significantly. Consequently much of what is written about energy is woefully misinformed or simply advocacy funded by various groups that hope to look good or profit from various perceived alternatives. Issues pertaining to the end of cheap petroleum will be the most important challenge that Western society has ever faced, especially when considered within the context of our need to deal simultaneously with climate change and other environmental issues related to energy. Any business or political leaders who do not understand the inevitability, seriousness, and implications of the end of cheap oil or who make poor decisions in an attempt to alleviate its impact are likely to be tremendously and negatively impacted as a result—and the rest of us with them. At the same time, the investment decisions we will make in the next decade or two will determine whether civilization is to make it through the transition away from petroleum or not.

What would be the impacts of a large increase in the energy and dollar cost of getting our petroleum or of any restriction in its availability? While it is extremely difficult to make any hard predictions, we do have the record of the impacts of the large oil price increases of the 1970s as a possible guide. These supply restrictions or "oil shocks" had very serious impacts on our economy which we have examined empirically in past publications [10]. At the time many economists did not think that even large increases in the price of energy would affect the economy dramatically because energy costs were but 3–6% of GDP. But by 1980, following the two "oil price shocks" of the 1970s, energy costs had increased dramatically until they were 14% of GDP. Actual shortages had additional impacts, when sufficient petroleum to run our industries or businesses were not available at any price. Other impacts included an exacerbation of our trade imbalances as more income was diverted overseas, adding to the foreign holdings of our debt and a decrease in discretionary disposable income as more money was diverted to access energy, whether via higher prices for imports, more petroleum exploration, or the development of low EROI alternative fuels. As EROI inevitably declines in the future, more and more of the economy's output will have to be diverted into getting the energy to run the economy. This in turn will affect those sectors of the economy that are not essential. Consumer discretionary spending will probably fall dramatically, greatly affecting nonessential businesses such as tourism and the economy more generally.

19.7 The "Cheese Slicer" Model

We have attempted to put together a conceptual and computer model to help us understand what might be the most basic implications of changing EROI on the economic activity of the United States. The model was conceptualized when we examined how the US economy responded to the "oil shocks" of the 1970s. The underlying foundation is the reality that the economy as a whole requires energy (and other natural resources derived from nature) to run, and without these most basic components, it will cease to function. The other premise of this model is that the economy as a whole is faced with choices in how to allocate its output in order to maintain itself and to do other things. Essentially the economy (and the collective decision-makers in that economy) has opportunity costs associated with each decision it makes. Figure 19.5 shows our basic

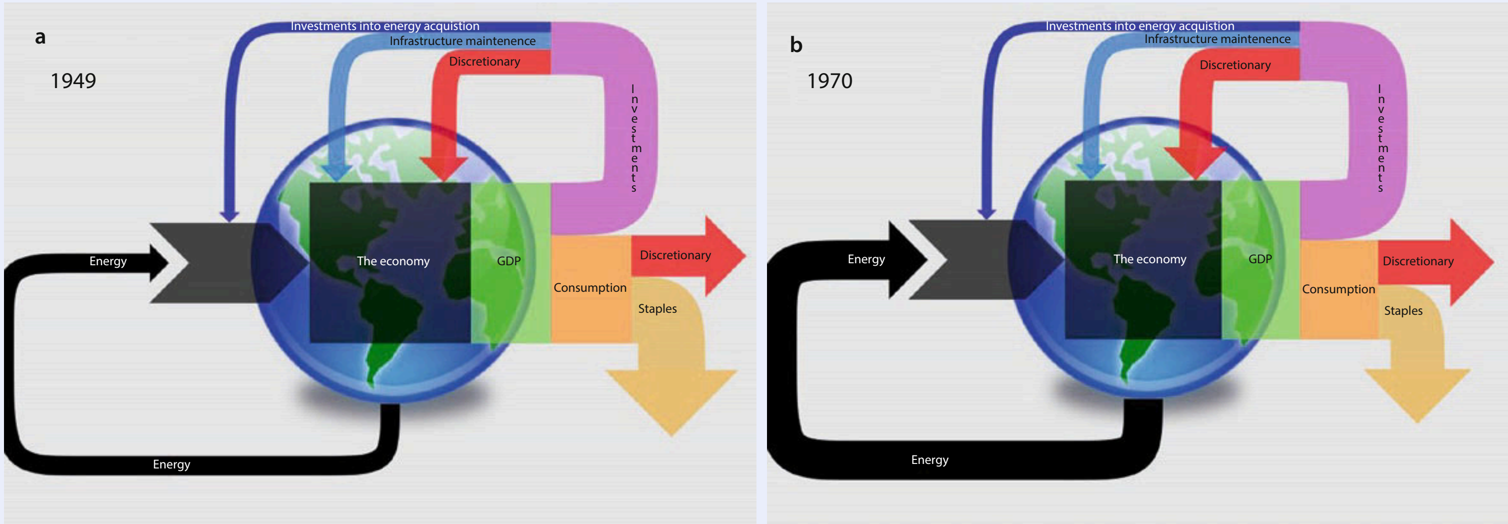


Fig. 19.5 The "cheese slicer" diagrammatic model (45), which is a basic representation of the fate of the output of the US economy (Source: U.S Department of Commerce & Andrea Bassi). **a** 1949 and **b** 1970. The *box* represents the US economy, the *input arrow* from the left represents the energy needed to run the economy, and the *large arrow* on the left of the box represents the output of the model (i.e., GDP) which is then subdivided as represented by the *output arrow* going to the right—i.e., first into investments (into getting energy, maintenance, and then discretionary) and then into consumption (either the basic required for minimal food, shelter, and clothing or discretionary). In other words the economic output is "sliced" into different uses according to the requirements and desires of that economy/society. **c** Same for **a** and **b** but for 1981, following large increases in the price of oil. Note change in discretionary investments. **d** Same for **a** and **b** but for 1990, following large increases in the price of oil. Note change in discretionary investments. **e** Same for **a** and **b** but for 2007, following large decreases then small increases in the price of oil. Note change in discretionary investments. **f** Same for **a** and **b** but projected for 2030, with a projection into the future with the assumption that the EROI declines from 20:1 (on average) to 10:1. **g** Same for **a** and **b** but projected for 2050, but a projection into the future with the assumption that the EROI declines to 5:1

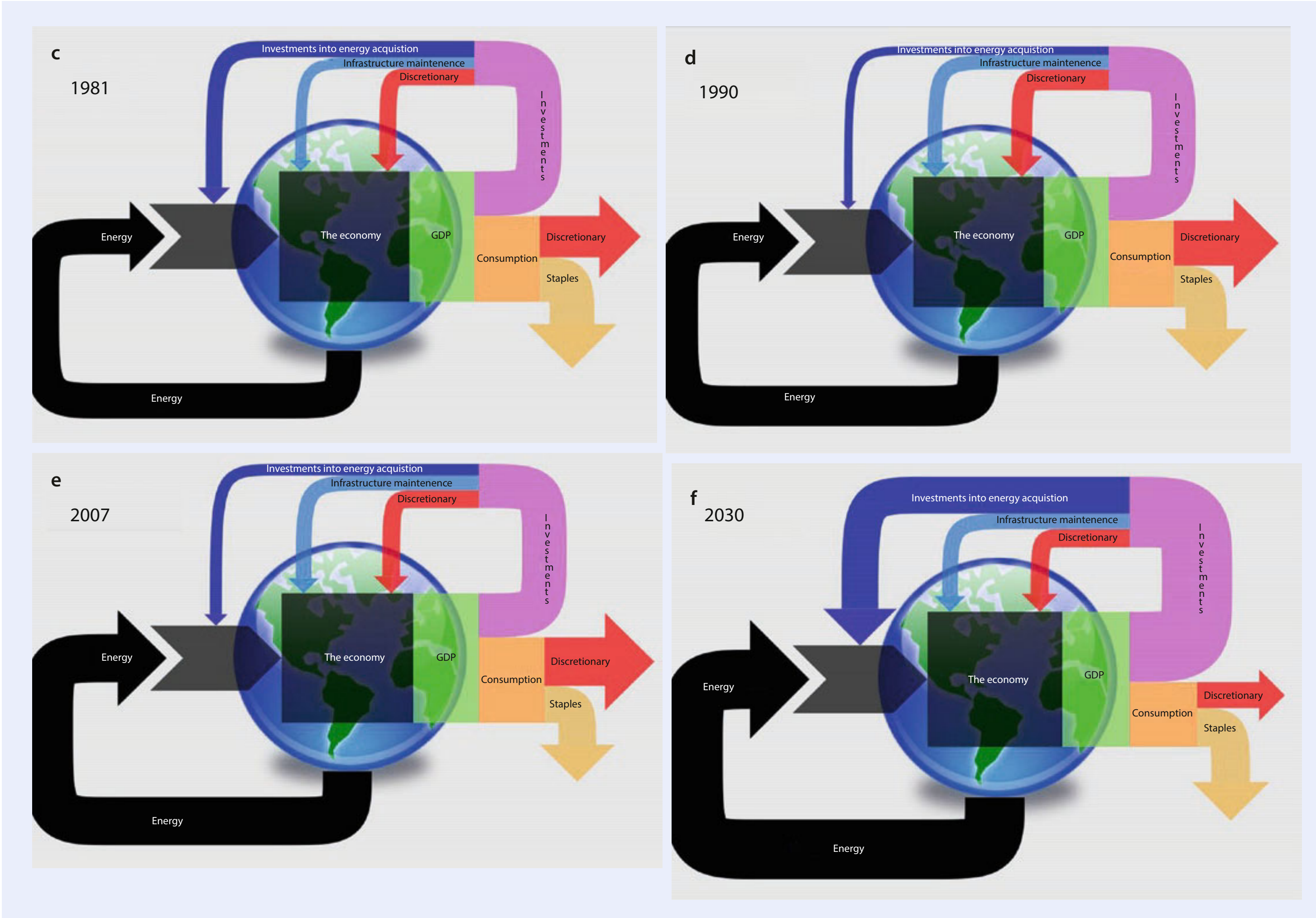
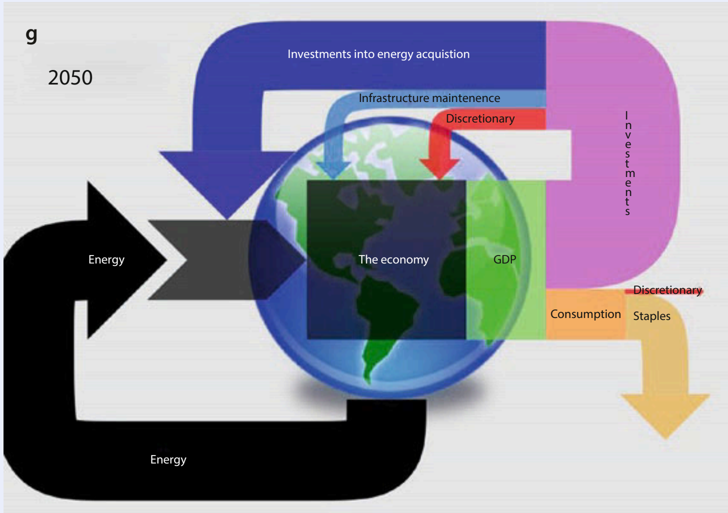


Fig. 19.5 (continued)

Fig. 19.5 (continued)



conceptual model parameterized for 1949 and 1970, before the oil shocks of that decade. The large square represents the structure of the economy as a whole, which we put inside a symbol of the Earth biosphere/geosphere to reflect the fact that the economy must operate within the biosphere [45]. In addition, of course, the economy must get energy and raw materials from *outside* the economy, that is, from nature (the biosphere/geosphere). The output of the economy, measured as GDP, is represented by the large arrow coming out of the right side, where the depth of the arrow represents 100% of GDP. For the sake of developing our concept, we think of the economy, for the moment, as an enormous dairy industry and cheese as the product coming out of the right-hand side, moving toward the right. This output (i.e., the entire arrow) could be represented as either money or embodied energy. We use money in this analysis but the results are probably not terribly different from using energy. So, our most important question is "how do we slice the cheese," that is, how do we, and how will we, divide up the output of the economy with the least objectionable opportunity cost. Most mainstream economists might answer "according to what the market decides," that is, according to consumer tastes and buying habits. But we want to think about it a little differently because we think things might be profoundly different in the future [43].

Most generally the output of the model (and the economy) has two destinations: *investment* or *consumption*. *Required* expenditures (without which the economy would cease to function) include (1) top line in blue are the investments into, or payments for, energy (i.e., the amount of economic output that is used to secure and purchase the domestic and imported energy needed for the economy), (2) investments in maintaining societal infrastructure (i.e., countering depreciation: repairing and rebuilding bridges, roads, machines, factories, vehicles—represented by the middle top arrow feeding back from output of the economy back to the economy itself), (3) some kind of minimal food, shelter, and clothing for the population (represented by the bottom rightward pointing arrow) required to maintain all individuals in society at the level of the federal minimum standard of living. This energy is absolutely critical for the economy to operate and must be paid for through proper payments and investments—which we consider together as investments to get energy. No investment in energy, no economic output. This "energy investment" feedback is represented by the topmost arrow from the output of the economy back upstream to the "workgate" symbol [44]. The width of this line represents the investment of energy into getting more energy. Of critical interest here is that as the EROI of our economy's total combined fuel source declines, then more and more of the output of the economy *must* be shunted back to getting the energy required to run the economy if the economy is to remain the same size.

Once these necessities are taken care of, what is left is considered the *discretionary* output of the economy. This can be either discretionary consumption (a vacation or a fancier meal, car, or house than needed, represented by the upper right pointing arrow in the diagrams) or discretionary investment (i.e., building a new tourist destination in Florida or the Caribbean, represented as the lower of the arrows feeding back into the economy). During the last 100 years, the vast wealth generated by the United States economy has meant that we have had an enormous amount of discretionary income. This is in large part because the expenditures for energy represented in ◘ Fig. 19.5 have been relatively small in the past.

The information needed to construct the above division of the economy is reasonably easy to come by for the US economy, at least if we are willing to make a few major assumptions and accept a fairly large margin of error. Inflation-corrected GDP, i.e., the size of the output of the economy, is published routinely by the US Bureau of Commerce. The total investments for maintenance in the US economy are available as "depreciation of fixed capital" (US Department of Commerce, various years). The minimum needed for food, shelter, and clothing is available as "personal consumption expenditures" (or the minimum of that required to be above poverty) which we selected from the US Department of Commerce for various years. The investment into energy acquisition is the sum of all of the capital costs in all of the energy-producing sectors of the United States plus expenditures for purchased foreign fuel. Empirical values for these components of the economy are plotted in ◘ Fig. 19.5. When these three requirements for maintaining the economy, investments and payments for energy, maintenance of infrastructure, and maintenance of people, are subtracted from the total GDP, then what is left is discretionary income.

We simulated two basic data streams: the US economy from 1949 to 1970 (representing the growth prior to the "oil crises" of 1973 and 1979) and the impact of the oil crisis and the recovery from that, which had occurred by the mid-1990s. Then we projected this data stream into the future by linearly extrapolating the data used prior to 2005 along with the assumption that the EROI for society declined from an average of roughly 20:1 in 2005 to 5:1 in 2050. This is an arbitrary scenario but may represent what we have in store for us as we enter the "second half of the age of oil," a time of declining availability and rising price when more and more of society's output needs to be diverted into the top arrow of ◘ Fig. 19.5.

19.8 Results of Simulation

The results of our simulation suggest that discretionary income, including both discretionary investments and discretionary consumption, will move from the present 50 or so percent in 2005 to about 10% by 2050 or whenever (or if) the composite EROI of all of our fuels reaches about 5:1 (◘ Figs. 19.5e and f).

19.9 Discussion

Individual businesses would be affected by increasing fuel costs and, for many, a reduction in demand for their products as people's income go increasingly for energy. This simultaneous inflation and recession happened in the 1970s and is projected to happen into the future as EROI for primary fuels declines. According to the economic theory called the Phillips Curve, the "stagflation" that occurred in the 1970s was not supposed to happen. According to Keynesian economics, inflation occurred only when the economy's aggregate demand exceeded its ability to produce. Unemployment was the result of too little aggregate demand. The simultaneous occurrence of inflation and unemployment rocked the very foundations of Keynesian analysis. But an energy-based explanation is easy [46]. As more money was diverted to getting the energy necessary to run the rest of the economy, disposable income and hence demand for many nonessentials declined, leading to economic stagnation. Meanwhile the increased cost for energy led to inflation, as no additional production occurred from higher prices. Unemployment increased during the 1970s but not as much as demand decreased, for at the margin labor became relatively useful compared to increasingly expensive energy. Individual sectors might be much more impacted as what happened in 2005, for example, with many Louisiana petrochemical companies that were forced to close or move overseas when the price of natural gas increased. On the other hand, alternate energy businesses, from forestry operations and woodcutting to solar devices, might do very well.

When the price of oil increases, it does not seem to be in national or corporate interest to invest in more energy-intensive consumption, as Ford Motor Company found out in 2008 with its former large emphasis on large SUVs and pickup trucks. (Although when energy prices declined again in 2016, big trucks came back!) When oil was cheap, we over-invested in remote second homes, cruise ships, and Caribbean semi-luxury hotels, so that we had a massive loss of the value of real estate. This was called the "Cancun effect"—such hotels require the existence of large amounts of disposable income from the US middle class and cheap energy. If EROI declines, that disposable energy may have to be shifted into the energy sector with an opportunity cost to the economy as a whole. Investors who understand the changing rules of the investment game are likely to do much better in the long run, but the consequence of having the "rug" of cheap oil pulled out from the economy will impact us for a long time.

So what can the scientist say to the investor? The options are not easy. As noted above worldwide investments in seeking oil have had very low returns in recent years. Investments in many alternatives have not fared much better. Ethanol from corn projects may be financially profitable to individual investors because they are highly subsidized by the government, but they are a very poor investment for the nation. It is not clear that ethanol makes much of an energy profit, with an EROI of 1.6 at best and less than one for one at worst, depending upon the study used for analysis [32, 47]. Biodiesel may have an EROI of about three to one. Is that a good investment? Clearly it is not relative to remaining petroleum. However real fuels must have EROIs of 5 or 10 or more returned on one invested to not be subsidized by petroleum or coal in many ways, such as the construction of the vehicles and roads that use them. Other biomass, such as wood, can have good EROIs when used as solid fuel but face real difficulties when converted to liquid fuels, and the technology is barely developed. The scale of the problem can be seen by the fact that we presently use several times more fossil energy in the United States than is fixed by all green plant production, including all of our croplands and all of our forests (Pimentel, D. Personal communication). Biomass fuels may make more sense in nations where biomass is very plentiful and, more importantly, where present use of petroleum is much less than in the United States. Alternatively, one might argue that if we could bring the use of liquid fuels in the United States down to, say, 20% of the present, then liquid fuels from biomass could fill in a substantial portion of that demand. We should remember that historically we in the United States have used energy to produce food and fiber, not the converse, because we have valued food and fiber more highly. Is this about to change?

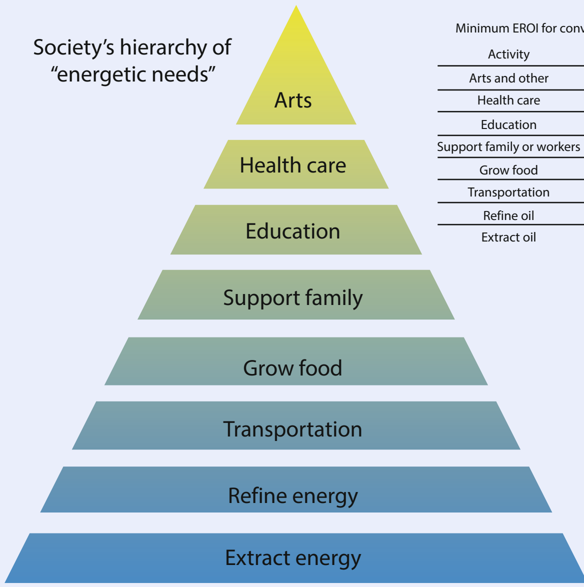
Energy return on investment from coal and possibly gas is presently quite large compared to alternatives (ranging from perhaps 50:1 to 100:1 at the point of extraction), but there is a large energy premium, perhaps enough to halve the EROI by the time they are delivered to society in a form that society finds acceptable. The environmental costs may be unacceptable, as may be the case for global warming and pollutants derived from coal burning. Injecting carbon dioxide into some underground reservoir seems unfeasible for all the coal plants we might build, but it is being pushed hard by many who promote coal. Nuclear has a debatable moderate energy return on investment (5–15:1, some unpublished studies say more). Newer analyses need to be made. Nuclear has a relatively small impact on the atmosphere, but there are large problems with public acceptance and perhaps safety in our increasingly difficult political world.

Wind turbines have an EROI of at least 15–20 returns on one invested, but this does not include the energy cost of backup or electricity "storage" for periods when the wind is not blowing. They make sense if they can be associated with nearby hydroelectric dams that can store water when the wind is blowing and release water when it is not, but the intermittent release of water can cause environmental problems. Photovoltaics are expensive in dollars and energy relative to their return, but the technology of both PV and storage is improving. One must be careful about accepting all claims for efficiency improvements because many require very expensive "rare-earth" doping materials, and some may become prohibitively expensive if their use expands greatly because of material shortages, even for copper [48, 49]. According to one savvy contractor, the efficiency in energy returned per square foot of collector has been increasing, but the energy returned per dollar invested has been constant as the price of the high-end units has increased. Additionally while photovoltaics have caught the public's eye, the return on dollar investment is about double for solar hot water installations. Wind turbines, photovoltaics, and some other forms of solar do seem to be a good choice if we are to protect the atmosphere, but the investment costs up front will be enormous compared to fossil fuels and the backup issue will be immense. Meanwhile the use of fossil fuel in the past decade has increased enormously relative to all of the solar.

Energy and money are not the only critical aspects of development of energy alternatives. Recent work by Hirsch and colleagues [33] has focused on the investments in time that might be needed to generate some kind of replacement for oil. They examined what they thought might be the leading alternatives to provide the United States with liquid fuel or lower liquid fuel use alternatives, including tar sands, oil shales, deepwater petroleum, biodiesel, high MPG automobiles and trucks, and so on. They assumed that these technologies would work (a bold assumption) and that an amount of investment capital equal to "many Manhattan projects" (the enormous project that built the first atomic bomb) would be available. They found that the critical resource was time. Once we decided to make up for the decline in oil availability, these projects would need to be started one or preferably two decades in advance of the peak to avoid severe dislocations to the US economy. Given our current petroleum dependence, the rather unattractive aspects of many of the available alternatives, and the long lead time required to change our energy strategy, the investment options are not obvious. This, we believe, may be the most important issue facing the United States at this time: where should we invest our remaining high-quality petroleum (and coal) with an eye toward insuring that we can meet the energy needs of the future. We do not believe that markets can solve this problem alone or perhaps at all. Research money for good energy analysis unconnected to this or that "solution" is simply not available.

Human history has been about the progressive development and use of ever higher-quality fuels, from human muscle power to draft animals to water power to coal to petroleum. Nuclear at one time seemed to be a continuation of that trend, but that is a hard argument to make today. Perhaps our major question is whether petroleum represents but one step in this continuing process of higher-quality fuel sources or rather is the highest-quality fuel we will ever have on a large scale. There are many possible candidates for the next main fuel, but few are both quantitatively and qualitatively attractive. In our view we cannot leave these decisions up to the market if we are to solve our future climate or peak oil problems. One possible way to look at the problem, probably not a very popular one with investors or governments, is to pass legislation that would limit energy investments to only "carbon-neutral" ones, remove subsidies from low EROI fuels such as corn-based ethanol, and then perhaps allow the market to sort from those possibilities that remain. Or should we generate a massive scientific effort, as objectively as possible, to evaluate all fuels and make recommendations?

A difficult decision would be whether we should subsidize certain "green" fuels. At the moment alcohol from corn is subsidized four times: in the natural gas for fertilizers, the corn itself through the Department of Agriculture's 100 or so billion dollar general program of farm subsidies, the additional 50 cents per liter subsidy for the alcohol itself, and a 50 cents per gallon tariff on imported alcohol. It seems pretty clear that the corn-based alcohol would not make it economically without these subsidies as it has only a marginal (if that) energy return. Are we in effect simply subsidizing the depletion of oil and natural gas (and soil) to generate an approximately equal amount of energy in the alcohol? We think so. Wind energy appears to have a relatively high EROI, enough to make it a reasonable candidate, although there are additional energy costs relative to backup technologies for when the wind is not blowing that have not been well calculated. So should wind be subsidized or allowed to compete with other "zero emission" energy sources? A question might be the degree to which the eventual market price would be determined by, or at least be consistent with, the EROI, as all the energy inputs (including that to support labor's paychecks) must be part of the costs. Otherwise that energy is being subsidized by the dominant fuels used by society.



Minimum EROI for conventional sweet crude oil

Activity	Minimum EROI required
Arts and other	14:1
Health care	12:1
Education	9 or 10:1
Support family or workers	7 or 8:1
Grow food	5:1
Transportation	3:1
Refine oil	1.2:1
Extract oil	1.1:1

Fig. 19.6 "EROI pyramid" of increasing abilities to support economic activities as a function of the mean EROI of a society. The values run from 1.1:1 to extract energy to 3:1 to provide transportation, etc. to perhaps 12 or 15:1 to provide for the complex amenities of civilization. Values up through transportation are based on Hall et al. 2008 and are fairly solid; higher values are increasingly speculative quantitatively. Graph from Lambert et al. (2014) as inspired by Maslow's pyramid of human needs

19.10 What Level EROI Do We Need?

We have stated that the criteria used by some investigators for an "acceptable" EROI has been only that it is positive, i.e., above one return for one invested. But in fact, as we developed in the last chapter, we need at least 3:1 to drive a truck if we include the cost of getting the fuel to the truck and pay for the depreciation of the infrastructure to use it. But we need also to pay for the "depreciation" of the workers as well, meaning the energy required to educate his or her children, provide for health care, and in general support the family, not to mention the various cultural amenities that make life good. We have developed this concept in some detail elsewhere [50] but provide a summary as Fig. 19.6.

19.11 Conclusion

It seems obvious to us that the US economy is very vulnerable to a decreasing EROI for its principle fuels. Increasing impacts will come from an increase in expenditures overseas as the price of imported oil increases more rapidly than that of the things that we trade for it, from increased costs for domestic oil and gas as reserves are exhausted and new reservoirs become increasingly difficult to find and as we turn to lower EROI alternatives such as biodiesel and/or photovoltaics. Our "cheese slicer" model suggests that as economic requirements for getting energy increase, a principal effect will be a decline in discretionary income as a proportion of GDP. Since more fuel will be required to run the same amount of economic activity, the potential for increased environmental impacts is very strong. On the other hand, protecting the environment, which we support strongly, may mean turning away from some higher EROI fuels to some lower ones. We think all of these issues are very important yet are hardly discussed objectively in our society or even in economic or scientific circles.

Questions

1. What was the experience of Cuba that allows us to understand better the role of energy in an economy?
2. What is meant by the phrase "the second half of the age of oil"?
3. Argue for or against the following question: the important issue is "when will we run out of petroleum."
4. How much oil do we discover for each barrel that we burn?
5. What happens to pressure as an oil field matures? Why?
6. What is the "cheese slicer" model?

7. Explain the difference between investment and consumption?
8. What is discretionary consumption?
9. What is the "Cancun effect"?
10. What resource does Hirsch and his colleagues think is especially important to adjust to a post-peak oil society?

Acknowledgments We thank our great teacher, Howard Odum; many students over the years; colleagues and friends including Andrea Bassi, John Gowdy, Andy Groat, Jean Laherrere, and many others who have helped me to try to understand these issues. Jessica Lambert created ◘ Figs. 19.4 and 19.6. Nate Hagens made many useful comments. The Santa Barbara Family Foundation, ASPO-USA, the Interfaith Center on Corporate Responsibility, and several individuals who wish not to be named provided much appreciated financial help.

References

1. Hall, Charles A.S., Robert C. Powers, and William Schoenberg. 2008. Peak oil, EROI, investments and the economy in an uncertain future. In *Renewable energy systems: Environmental and energetic issues*, ed. D. Pimentel, 113–136. London: Elsevier.
2. Soddy, F. 1926. *Wealth, virtual wealth and debt*. New York: E.P. Dutton and Co.
3. Tryon, F.G. 1927. An index of consumption of fuels and water power. *Journal of the American Statistical Association* 22: 271–282.
4. Cottrell, F. 1955. *Energy and society*. Dutton: Greenwood Press.
5. Georgescu-Roegen, N. 1971. *The entropy law and the economic process*. Cambridge: Harvard University Press.
6. Odum, H.T. 1972. *Environment, power and society*. New York: Wiley-Interscience.
7. Kümmel, R. 1982. The impact of energy on industrial growth. *Energy – The International Journal* 7: 189–203.
8. ———. 1989. Energy as a factor of production and entropy as a pollution indicator in macroeconomic modelling. *Ecological Economics* 1: 161–180.
9. Jorgenson, D.W. 1984. The role of energy in productivity growth. *The American Economic Review* 74 (2): 26–30.
10. ———. 1988. Productivity and economic growth in Japan and the United States. *The American Economic Review* 78: 217–222.
11. Daly, H.E. 1977. *Steady-state economics*. San Francisco: W. H. Freeman.
12. Dung, T.H. 1992. Consumption, production and technological progress: A unified entropic approach. *Ecological Economics* 6: 195–210.
13. Hall, C.A.S., C.J. Cleveland, and R.K. Kaufmann. 1986. *Energy and resource quality: The ecology of the economic process*. New York: Wiley-Interscience.
14. Ayres, R.U. 1996. Limits to the growth paradigm. *Ecological Economics* 19: 117–134.
15. Ayres, Robert U., and Benjamin Warr. 2009. *The economic growth engine: How energy and work drive material prosperity*. Northhampton Mass: Edward Elger.
16. Quinn, M. 2006. The power of community: How Cuba survived peak oil. Text and film. Published on 25 Feb 2006 by Permaculture Activist. Archived on 25 Feb 2006. Can be reached at megan@communitysolution.org.
17. Kaufmann, R. 2004. The mechanisms for autonomous energy efficiency increases: A cointegration analysis of the US energy/GDP ratio. *The Energy Journal* 25: 63–86.
18. Hall, C.A.S., and J.Y. Ko. 2006. The myth of efficiency through market economics: A biophysical analysis of tropical economies, especially with respect to energy, forests and water. In *Forests, water and people in the humid tropics: Past, present and future hydrological research for integrated land and water management*, ed. M. Bonnell and L.A. Bruijnzeel, 40–58. Cambridge University Press: UNESCO.
19. LeClerc, Grégoire, and C.A.S. Hall, eds. 2006. *Making world development work: Scientific alternatives to neoclassical economic theory*. Albuquerque: University of New Mexico Press.
20. Smil, V. 2007. Light behind the fall: Japan's electricity consumption, the environment, and economic growth. *Japan Focus*, April 2; EIA. (2009). U.S. Energy Information Agency website. Accessed June 2009.
21. Cleveland, C.J. 2005. Net energy from the extraction of oil and gas in the United States. *Energy*. *The International Journal* 30 (5): 769–782.
22. Campbell, C., and J. Laherrere. 1998. The end of cheap oil. *Scientific American* (March): 78–83.
23. Heinberg, R. 2003. *The Party's over: Oil, war and the fate of industrial societies*. Gabriella Island, B.C. Canada: New Society Publishers.
24. Galbraith, J. K. 2014. The end of normal. Simon and Shuster, N.Y.
25. Hubbert, M.K. 1969. Energy resources. In *Resources and man*. National Academy of Sciences, 157–242. San Francisco: W.H. Freeman.
26. Strahan, D. 2007. *The last oil shock: A survival guide to the imminent extinction of petroleum man*. London: Hachette Publisher.
27. Energyfiles.com. Accessed August 2007. www.energyfiles.com.
28. Deffeyes, K. 2005. *Beyond oil: The view from Hubbert's peak*. New York: Farrar, Straus and Giroux.
29. EIA. 2007. U.S. Energy Information Agency website. Accessed June 2007.
30. IEA. 2007. European Energy Agency, web page. Accessed August 2007.
31. Campbell, C. 2005. *The 2nd half of the age of oil*. Paper presented at the 5th ASPO Conference, Lisbon Portugal. See also. In *The first half of the age of oil: An exploration of the work of Colin Campbell and Jean Laherrere*, ed. C.A.S. Hall and R.C. Pascualli. New York: Springer. 2012.

32. Murphy, D.J., C.A.S. Hall, and Bobby Powers. 2011. New perspectives on the energy return on investment of corn based ethanol. *Environment, Development and Sustainability* 13 (1): 179–202.
33. Hirsch, R., Bezdec, R. and Wending, R. 2005. Peaking of world oil production: Impacts, mitigation and risk management. U.S. Department of Energy. National Energy Technology Laboratory. Unpublished Report.
34. Lynch, M.C. 1996. The analysis and forecasting of petroleum supply: Sources of error and bias. In *Energy watchers VII*, ed. D.H.E. Mallakh. Boulder: International Research Center for Energy and Economic Development.
35. Adelman, M.A., and M.C. Lynch. 1997. Fixed view of resource limits creates undue pessimism. *Oil and Gas Journal* 95: 56–60.
36. Duncan, R.C. 2000. *Peak oil production and the road to the Olduvai Gorge*. Keynote paper presented at the Pardee Keynote Symposia. Geological Society of America, Summit 2000.
37. Brandt, A.R. 2007. Testing Hubbert. *Energy Policy* 35: 3074–3088.
38. Hallock, J., P. Tharkan, C. Hall, M. Jefferson, and W. Wu. 2004. Forecasting the limits to the availability and diversity of global conventional oil supplies. *Energy* 29: 1673–1696; Hallock, J., Jr., W. Wu, C.A.S. Hall, and M. Jefferson. 2014. Forecasting the limits to the availability and diversity of global conventional oil supply: Validation. *Energy* 64: 130–153.
39. Cleveland, C.J., R. Costanza, C.A.S. Hall, and R.K. Kaufmann. 1984. Energy and the US economy: A biophysical perspective. *Science* 225: 890–897.
40. Musnadi, M., and A. Brandt. 2017. Energetic productivity dynamics of global super-giant oilfields. *Energy and Environmental Science* 10: 1493–1504.
41. Gagnon, N., C.A.S. Hall, and L. Brinker. 2009. A preliminary investigation of energy return on energy investment for global oil and gas production. *Energies* 2 (3): 490–503.
42. Holter, M. 2016. Oil discoveries at a 70 year low signal supply shortfall ahead. Bloomberg The year ahead. August 30.
43. Ricardo, David. 1891. *The principles of political economy and taxation*. London: G. Bell and Sons. (Reprint of 3rd edition, originally pub 1821).
44. Hughes, D. 2011. Will natural gas fuel America in the 21st century? Post Carbon Institute. There are a number of other very good reports on oil and gas resources by Hughes available from the Post Carbon Institute.
45. Odum, H.T. 1994. *Ecological and general systems: An introduction to systems ecology*. Niwot: University Press of Colorado. Millennium Institute. 2007. Data principally from the U.S. Department of Commerce. Extrapolations via the Millennium Institute's T-21 model courtesy of Andrea Bassi.
46. Hall, C., D. Lindenberger, R. Kummel, T. Kroeger, and W. Eichhorn. 2001. The need to reintegrate the natural sciences with economics. *Bioscience* 51: 663–673; Hall, C.A.S. 1992. Economic development or developing economics? In *Ecosystem rehabilitation in theory and practice, Vol I. Policy issues,* ed. M. Wali, 101–126. The Hague, Netherlands: SPB Publishing.
47. Farrell, A.E., R.J. Plevin, B.T. Turner, A.D. Jones, M. O'Hare, and D.M. Kammen. 2006. Ethanol can contribute to energy and environmental goals. *Science* 311 (5760): 506–508. and also the many letters on that article in Science Magazine, June 23, 2006.
48. Andersson, B.A., C. Azar, J. Holmerg, and S. Karlsson. 1998. Material constraints for thin-film solar cells. *Energy* 23: 407–411.
49. Gupta, A.J. in press. *Materials: Abundance, purification, and the energy cost associated with the manufacture of Si, CdTe, and CIGS PV*. Amsterdam: Elsevier.
50. Lambert, J., C.A.S. Hall, S. Balogh, A. Gupta, and M. Arnold. 2014. Energy, EROI and quality of life. *Energy Policy* 64: 153–167.

The Role of Models for Good and Evil

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_20

The words "model" and "modeling" are found increasingly in economics and indeed in science in general. Therefore, it is important that we consider here some of the most important characteristics of these words and concepts and introduce the reader to how they are used in energy studies and in economics. Most generally the word "model" means a simplification. For example, we all work with models of human behavior that a person will act in a particular way because he or she is a teenager, a man, a woman, rich, poor, black, white, or whatever. Of course one of the important aspects of growing up is to realize that such models are wrong so often that they are essentially useless and that we need to meet and judge people one person at a time. So perhaps one of the most important things you can learn about models of any kind is that they are often wrong. But often they are correct, and usually, even when incorrect, they can be useful by allowing us to generate and test hypotheses formally. So armed with a healthy skepticism we can go on to see what models really are. This chapter is in some respects a continuation of ► Chap. 16 on mathematical tools and our earlier work in ecology [1].

20.1 Definitions: Models and Analytic vs. Simulation Models

There are many definitions of the word "model." One, given above, is a purposive simplification, such as a model airplane. A second is "a device for predicting a complex whole from the operation of parts that are thought to be known." The definition that we like the most, from Hall and Day [1], is "a formalization of our assumptions about a system." Whether we formalize them or not, we use models constantly: models of scientific outcome, models of economic decisions, models of your own behavior, or that of others. In our view, quantitative (or occasionally non-quantitative) models are necessary in the complex world of ecology and economics because they allow one to apply the scientific method to complex real systems. Despite the many problems of modeling, we do not understand how one can use the scientific method (i.e., to generate and test hypotheses) for any reasonably complicated system without the use of formal modeling. This is as true for management and policy-related issues as for theoretical ones.

The power of models is to make our assumptions explicit, and hence testable. The power of mathematics (in its broad sense) and of mathematical models is to make the results of the prediction *quantitatively* explicit and hence quantitatively predictable. Generally we are seeking a *solution*, that is, a quantitative prediction for the value of some variable at some different place or time. The process of examining whether a model is correct or at least adequate is called *validation*. The examination of the degree to which uncertainty in model formulation or (more generally) parameterization allows you to trust your results, or reach certain conclusions, is called *sensitivity analysis*. It is through validation and sensitivity analysis that models generate their (occasional) tremendous power in resolving truth, such as that is accessible to the human mind.

What then is the role of mathematics in this process? First of all, it is necessary to distinguish *mathematical* from *quantitative*. Quantitative means simply using appropriate numbers in your analysis: 3 salmon vs 7 salmon. This does not necessarily require any particular mathematical skills (although getting accurate numbers may require enormous skills of a different kind). Mathematical means using the complex tools of quantitative analysis to manipulate those numbers, often to make a prediction, that is, an educated guess of the value of something modeled (such as oil production or GDP). These tools include algebra, geometry, calculus, simulation, and so on. There are two principle means of manipulating or solving numbers: *analytic* (or closed form) and *numeric* (or simulation). The first, generally using paper and pencil, gives usually explicit and exact solutions to relatively simple equations for a particular point in time or set of conditions, generally using rather complex and difficult equations (e.g., see ► Chap. 16). The second gives approximate answers to a broader set of possible equations using (generally) simpler equations put together in complex patterns and solved stepwise *in a computer*. In theory, either method can be used to solve many particular quantitative problems, and sometimes this is done. In practice the mathematical training required to undertake analytic approaches precludes its use by many. In addition there are severe restrictions to the class of mathematical problems that can be solved analytically, often requiring a series of sometimes unrealistic assumptions to put the problem into a mathematically tractable format. On the other hand, simulation allows one to solve very complex problems using relatively simple mathematics. Hence, there is a curious paradox: the most complex mathematics actually requires the simplest basic equations for starters.

A final problem is that there has been frequent confusion between *mathematical* and *scientific rigor* or *proof*. Mathematics can generate real proofs relatively easily because you are working in a defined universe (through the assumptions and the equations used) to which it applies. If you define a straight line as the shortest distance between two points, then you can solve many problems requiring straight lines. But the world handed to us by nature is not so defined, and we must constantly struggle to represent it with our equations. Hence, a mathematical proof becomes a scientific proof only in the relatively rare circumstances when the equations do indeed capture the essence of the problem.

20.2 Models and Reality

The truth of the matter is that humans have often found that models of reality are much easier to deal with conceptually and operationally than reality itself, which tends to be very messy. There is a very long history of models getting in the way of truth, and this continues today. Perhaps the clearest and oldest example of both the strength and the potential fallacies of models are those associated with our understanding of astronomy. Most educated ancients were extremely interested in astronomy because of their belief that the movements of heavenly bodies had great importance to their day-to-day affairs (this lives on today as astrology). Sometimes the reasons were clear and scientific. As agriculture became more and more important, it became obvious that an understanding of the movement of the sun North and South with the seasons was a much more reliable index of when to plant than was temperature, which could not be measured anyway and which varied much more than the stately daily progression of the sun. Thus, the ancients built entire buildings and even cities to help measure the movement of the sun and other heavenly bodies, as beautifully told by archeoastronomers such as Anthony Aveni [2]. These ancient astronomers needed very large instruments in those times before the invention of brass instruments so that the relation of inaccuracies in construction was not too large compared to the size of the instruments. Consequently, they built entire cities that would track the movement of the sun and other heavenly bodies through the seasons. Those who planted according to the schedules of astronomer-priests tended to get rewarded with larger and more reliable crops, and political power flowed to the priests accordingly. Stonehenge, the pyramids of Egypt and Mexico, as well as many lesser known ancient cities are built, at least in part, as giant celestial observatories.

Most of these ancient astronomers thought that the sun, the moon, and the planets went around the Earth (after all it was obvious) and that all heavenly bodies traveled in perfect circles, since that would reflect the perfection of God—as well as God putting humans at the center of all things. Probably the greatest of these ancient astronomers was the Greek-Egyptian Ptolemy, and today we must understand him as a person with tremendous mathematical and modeling skills. Ptolemy could predict the seasons and even the movement of the planets with great precision and even predict when the Nile would flood even though the rain that caused this to occur was thousands of miles to the South. They were able to do this with relatively simple mathematics—with one exception. In order to explain the observations of the interior planets (Venus and Mercury), Ptolemy and his colleagues had to come up with a series of circular “epicycles” in which these planets circled the sun, which was circling the Earth. This was a remarkably successful approach to astronomy and could explain the observed data to within a few percent.

We now know that Ptolemy, who was a remarkably intelligent person and gifted mathematician, was dead wrong. It took more than a thousand years for the Polish astronomer Copernicus to come along and, with the extremely accurate sightings of Tycho Braye, show that not only did the Earth revolve around the sun but that also the Earth did not follow a circular orbit, but rather an ellipse. The search for circular perfection over a thousand years had got in the way of understanding reality – that the earth went around the sun in an elliptical orbit. Or we might say putting too much faith in religious perfection got in the way of science. But for a long time, the incorrectly developed model of Ptolemy was a better predictor than the correctly constructed model of Copernicus, because Ptolemy was such a good mathematician. Today there are many scientists and economists who do not wish to abandon mathematically "perfect" solutions for a more accurate but less elegant explanation.

Seeking perfect and simplified models in other disciplines has also interfered with understanding truth and reality. In fact it is probably more often the rule than the exception. The creation model from the bible of the origin of species story of the creation probably made as much sense as any other explanation until Charles Darwin came along and gave us a model that was much more consistent with our observations and the fossil record. God may certainly exist (that question is well outside the aegis of Science) but so does evolution. Just ask any hospital administrator or agricultural pest manager who has to deal with the routine evolution of hospital and agricultural pathogens. That is pretty hard to explain by the actions of a benevolent God. Meanwhile, apparently more Americans believe in Angels than in evolution.

For another example, fisheries science lost decades of understanding and ultimately contributed to the destruction of many of the world's most important fisheries because fisheries scientists, hampered by the overwhelming complexity of interannual variations in numbers of a fish species, chose to believe a model, the Ricker curve, rather than to look more carefully at their own data. In fisheries science, for a long time, it was believed that the number of fish in a population (i.e., sockeye salmon in British Columbia) depended principally upon the number of parents, with there being the possibility of both too few and too many parents for maximum production of young. This idea "allowed" managers to let fishermen take large numbers of salmon that had been considered "excess." More recent work has shown that the main determinants of the salmon populations are climate and other environmental factors, although number of parents may also be important although not necessarily in the way initially proposed [3, 4]. More generally in population biology, it has become clear that the simple, elegant mathematical models that once dominated thinking about populations were almost certainly misleadingly incomplete when not entirely wrong [5]. But we have learned, and now it is usually the case that we use ongoing data from the fisheries itself to set the seasons and otherwise manage the fish through "adaptive management" (where politics or too much greed does not get in the way). Likewise, ecologists and game managers believed for too long in the simplistic, "perfect," and almost always wrong logistic and Lotka-Volterra mathematical models of population dynamics, rather than to concentrate on the environmental factors that were generally far more powerful predictors of actual populations. The point is that in all of these issues, there is a huge tendency for people to want to believe in models of perfection rather than in the messy reality that surrounds us, although usually that too can be understood, even if painfully slowly, by a proper use of science.

20.3 The Importance of Paradigms

Models are far more than complex mathematical entities that live in mathematics books or in computers. More generally models are *conceptual*—that is, mental pictures of the structure and/or function of a system or of how something operates. Such conceptual models are often called *paradigms* when they become expansive and general. Nearly all disciplines, including economics, work from what is usually called a paradigm or a set of paradigms. Paradigms (sometimes called "preanalytical visions") are conceptual constructs that synthesize the main ideas of a discipline, explain a wide set of observations, and allow for the positioning of new ideas into the existing intellectual structure. Examples include evolution in biology and plate tectonics in geology. Before Charles Darwin's synthesis in 1859, there were many observations of nature that simply did not make sense or at least were not related to each other. These observations include the fact that organisms tended to have far more offspring than was needed to replace themselves, that animal breeders were able to change very much the characteristics of their animals and these characteristics were passed on to their offspring, and that there existed a vast record of past life in rocks that in some cases showed a regular progression of change from one stratum to another. At that time, the principal idea as to where life had come from for most Europeans was the story of creation in the bible. Darwin was himself religious and initially believed like other educated people of his time that the biblical explanation for creation was all that one needed to know. But Darwin also knew from the work of the earlier geologists Hutton and Lyell that the earth was very old and that processes that had shaped the earth in the past were often still occurring now. Finally he knew that these processes (such as erosion of landscapes) could be very powerful even though they were very slow because they played out over such a large amount of time. Darwin brilliantly synthesized all of these different observations, and many more, in his book *The Origin of Species*. His concept of evolution through natural selection has become a paradigm for all of the biological world since then. His particular genius was to come up with the mechanism, *natural selection*, that could explain the process, which was evolution. As we gain new information, we have made additions and revisions to his basic idea, but the idea itself has withstood the test of time very well. For example, in the past several decades, we have made astonishing progress in understanding the nature of DNA and the many ways it works at the cellular and the molecular level. Nevertheless, all of this exceedingly detailed and powerful new information has not changed the basic way that we understand how evolution works and in fact adds considerable additional insight and support. Evolution is, essentially, *the* paradigm for biology.

Similarly in the 1950s, geology was a rather sleepy science which had a whole series of unrelated observations about the earth: that volcanoes appear in specific regions and that earthquakes were associated with these regions as were mountain chains. As probably every schoolchild has thought about when staring at a map of the world during a boring class, the shape of the African West Coast snuggles up very nicely against South American and so on. Additionally they knew that biologists had found that a particular type of tree, the southern beech (the genus *Nothofagus*), was found in very similar, but not exactly the same, forms in Southern South America, in Australia and New Zealand, and in South Africa. Although biologists such as Philip Darlington had considered for a long time that the continents must have moved, geologists were not buying it, or more usually not even thinking about it, because they had no idea of a *mechanism* to move the continents. Remember *theory* is the device that explains the mechanism that underlies our observations. The continents were just too large, there was no concept of where the energy might come from to do that much work, and the concept was too weird. But in the 1950s, a group of geologists, many of them at Princeton University, began to connect the dots [6]. The most important knowledge was coming from, surprisingly, the bottom of the oceans. Oceanographers had begun to map the bottom of the ocean with powerful new sonar, and they found a very surprising thing—the middle of the Atlantic (and other) oceans had a series of underwater volcanoes that stretched from Iceland in the North (Iceland is itself a series of volcanoes) to below the tip of South America. Further studies showed that some of these volcanoes were actually active, spewing forth lava and heat under water, and that the sea bottom on either side of the volcanoes was spreading away from each other. Here was the needed mechanism to explain continental drift! It was energy from deep inside the earth, moving up in these oceanic rift zones, that was pushing the continents apart! Soon geology was abuzz with excitement and many new concepts tumbled out, all aided by this continental drift paradigm. For example, we could now see and even measure with lasers that the Red Sea was hinging apart, and the beautiful rift lakes of East Africa could be seen as the first stage in land masses splitting apart. In time lakes such as Tanganyika and Malawi will split apart entirely, and the sea will pour into what is now the middle of Africa—as it has with the Red Sea and the area between Africa and Madagascar.

20.4 Paradigms and Models

In these, and other, examples, scientists are usually most satisfied when they can formalize their paradigm, or some derivative of it, as a *model*. As we stated earlier, the definition for a model that we like best is "a formalization of our assumptions about a system." The beauty of models from that perspective is that it says essentially that a model is a working hypothesis about how the world works, and as such it can be tested explicitly. It allows one to put reasonably complex issues such as continental drift into a format where they can be tested quantitatively. While most of you who have thought much about models probably think of them as some kind of mathematical or computer entity, in fact there are five major types or classes of models: *conceptual*, *physical*, *diagrammatic (or graphical)*, *mathematical*, and *computer*. Each of them in some way attempts to capture the essence of a problem or situation, in a formalized although simplified way. Of course models can be good or bad, correct or incorrect, and complete or incomplete. But they should be consistent with the general principles of science outlined at the start of ► Chap. 15, they should contain appropriate mechanisms, and they should explain considerable empirical observations. A good paradigm meets all those criteria and can be considered a sort of super model that cements knowledge in an entire discipline. Both of the paradigms given above, natural selection and continental drift, meet those criteria. But many other models, and even some paradigms, have been found to be sadly lacking.

Clearly we have to build our models and our paradigms very carefully. The beauty of science and the scientific method is that it allows one to construct tentative models of how the world might work. Subsequent testing may find through empirical (i.e., related to observation and data) observation and testing that the assumptions used to construct that model were good or poor. Then the model can be adjusted or abandoned. There is no disgrace in constructing a model that turns out to be incorrect. That is how science moves forward. In science when one model or paradigm is shown to be false, there is often another to take its place, or sometimes we have to conclude that a good model just is not possible yet or maybe ever.

Models are great devices for bringing problems that are otherwise too large or too small into a scale humans can understand and conceptualize. Trying to imagine something as large as the atmosphere or the world economy, or as small as a hydrogen atom, can be a pretty daunting task. But by using models, we can write them all down on a piece of notebook paper or carry them around on a data stick. Most people are visual creatures. From an evolutionary perspective sight dominates most of our other senses. Think about it, would you tolerate the amount of pollution you now do if you primarily sensed the world through the chemistry of smell, as does a salamander?

A model must be simpler than the world it is attempting to explain. If it were not, it would be a description and not a model. What we mean by simplification is that the model contains fewer variables than the world we are trying to explain by means of a model. In addition, the independent variables should be as independent as possible. If they are not, it becomes very difficult to separate cause from effect. Finally, if the models start out with its variables arrayed linearly, the model is simpler. We would like to include a word of warning for the student who is not yet accomplished in modeling. Please do not confuse simple with easy. Simple means there are but a few independent variables which are linear and which do not strongly interact. Simple does NOT mean immediately apparent by casual observation. Even simple models often require a great deal of work to get the relations with the real world good.

But here is the rub. Just because a model is simple, or even just because it might make considerable intuitive sense, does not mean that the model has correctly captured the essence of the system or the essence required for the question being asked at the time. It is amazing how infrequently this question has been asked. In our extensive and very different experience with modeling, we do not believe that 10% of the models that we have seen are in any way sufficiently well constructed to be appropriate for the questions they are being put to. Is the Hubbert model of oil production sufficient to predict the future? If so what data do we need to parameterize it? Certainly the simple firms and households model (◘ Fig. 3.1) is completely inappropriate to resolve questions about national debts, pollution, climate change, or a host of other issues that it or its manifestations have been used for.

Why is this so? Is it because economists have no other place to turn to? That economists need a model to get there work done whether or not they have a good one to work with? One gets that impression from the review of development models by LeClerc [7]. Some people are poor, others are wealthy and want to help, money is thrown at aid, and economists are supposed to come up with a good development scheme and a model to justify it. Sometimes, of course, that works, more generally it seems to not work. Or maybe it would work if the population were not growing at the same time, eating up whatever new wealth is produced.

20.5 So, then, Why Is Economics, Which Is So Complex, So Analytical?

Nevertheless, there remains within academia a great deal of what we might call "physics envy," that is, a desire to emulate the power and prestige of successful applications of simple equations in physics.

Mathematical rigor is often very important for impressing colleagues and deans whether the analysis has a secure connection with reality or not. In some few cases, it has led to the most brilliant and important advances in all of human knowledge. Mathematical rigor, however, while useful in its own right and in some applications, is hardly by itself a criteria of acceptable science, although it is often promoted as such. Thus the advanced economist is often reduced to simplifying quite complex economic questions into a format that is analytically tractable, that is, can be solved using analytic means and sometimes using essentially ideological concepts such as "free markets generate the optimal use of resources." It is a lovely idea, requires enormous skills and concentration, and sometimes generates very useful results. Very often, however, we believe it generates results that represent only the mathematics and not the real system. We give some examples throughout this book but especially in ▸ Chaps. 4 and 11. Of course we use models too, so the reader should ask "have they validated their models." We think so and give as an example the work of Hallock et al., where we predicted the oil production for 46 countries and then went back 10 years later to see how we did [8](◘ Fig. 8.8). The answer is "pretty well for most countries, but miserably for a few."

What you need to make *analytical* mathematical models work is really very simple systems, often described as the *two body* system. Real atmospheric or real economic systems are not so simple, and pushing real systems kicking and screaming into a small enough box (i.e., few enough equations) to be analytically tractable is not science. In our opinion, there are very few real problems in economics that can be adequately represented by such simple relations, and much of the economics that is done by complex analytic analysis is deriving mathematical and not economic results. But the use of analytical mathematics does have one major benefit. Through the manipulation of equations, you can transform a cause and effect relation that is obtuse into a way in which you can sometimes see, derive and test patterns.

20.6 How Have Models in Fact Been Used in Economics

We believe that models have rarely been used in economics in their proper role, that is, as a formalization of our assumptions that would allow the testing of the hypotheses that are represented by the equations therein. Rather models have been used mostly as conceptual shortcuts that take the very complex biophysical and social entities that real economies are composed of and represent them as caricatures that demand acceptance (or dismissal), but not testing. While of course any model needs to be in some sense a simplification, the important issue is that that simplification must represent the basic reality modeled. But in fact we have shown in ▸ Chap. 3 that the most important models in economics, such as the firm-household model, do not represent the essential biophysical reality that constitutes real economies. Why should this absurdly simplistic model that is not even true as a first approximation be allowed to be represented over and over in introductory economics text books with so few economists speaking up (other than Leontief) that the king has no clothes?

It is true that within economics there are complex empirically-based models. An example is the University of Pennsylvania Wharton model, a huge, data-rich computer simulation of linked economic transactions throughout the economy. It gives very detailed predictions about each section of the economy, although it failed to predict the 2008 market crash [9]. As such it is a useful predictive device, and as such it *could* be used to generate and test hypotheses. But it was not generated upon a series of hypotheses about how the economy works, nor was it asked, to our knowledge, to test the basic hypotheses of economics. Instead the structure of the economy is specified (given), and then a massive amount of information is fed into the calibration phase of the model. The computer cleverly fits all of the actual data collectively to all of the equations in a process known as parameterization. The net effect is that the model can predict well small changes, say from 1 year to the next, because of the "can't fail structure" of the model, which is in some ways a tautology. But in no way that we are aware of does this model test the underlying conceptual base of the neoclassical model of economic reality. We love the article by Krugman in the *New York Times* "How did economists get it so wrong: Mistaking beauty for truth" [10] where Krugman found no economists predicted the market and housing crash despite their extremely sophisticated "quant" mathematical models. This reminded us of Ptolemy's model of the solar system: it worked very well when you are recreating the known, but if the mechanism is incorrect, you have no chance with changing forcing functions.

20.7 Some Problems with the Standard Neoclassical Model

Thus, while there are some good attributes to the basic neoclassical supply-demand-market model, there are also some extreme problems, as the reader probably has guessed by now. The first problem, well understood by any economist, is that of *externalities*. Externalities refer to a gain, or (more generally) a loss, associated with a market transformation (often a third party) that is not expressed in the market price. Classic examples are pollution impacting a downstream fishery, or worker's compensation. In, say 1850, when an industrial worker making, say, chairs loses an arm in a piece of machinery such as a mechanical saw or lathe it was catastrophic financially as well as physically, for if he had lost his arm he was no longer able to do his job and so was let go, generally leaving his family with no means of support. If he was lucky, the mill owner might take pity on him and allow him to push a broom at a reduced salary, but this was hardly guaranteed. Nevertheless that loss of limb and of income was in fact part of the cost of producing that chair. Worker's compensation, which was mandated by federal law, recognized that if the cost of some kind of insurance could be entered into the price of the good being produced, that would form a fund called "workman's compensation" which was held in reserve for the occasional catastrophe. The injured worker was given money to pay for medical treatment as well as a pension to cover wages lost. This fund was paid for by the manufacturer, who passed it along in the price of his or her products. Then we could say that the cost (losing a limb and also income) had been *internalized* into the price of the product. The market alone could not do that, nor should it be expected that it would. It requires intervention by governments (usually at the state level), and although the idea was at first fought by early manufacturers, it has been pretty well accepted by all now and in fact is also responsible in part for making working conditions safer, for example, by the simple and very effective practice of putting covers over belts and gears to avoid the possibility of a worker slipping and getting chewed by the machinery. From [11]: "Externalities are ad hoc corrections introduced as needed to save appearances, like the epicycles of Ptolemaic astronomy…As long as externalities involve minor details, this is perhaps a reasonable procedure. But when vital issues (e.g. the capacity of the earth to support life) have to be classified as externalities, it is time to restructure the basic concepts and start with a different set of abstractions that can embrace what was previously external."

Many advocates of the use of "pure" neoclassical economics have called for a reduction in government intervention in the economy. A classic example is their response to the Glass-Steagall Act, passed in 1933, which had as its main objective the keeping of investment banks (think Goldman Sachs) separate from commercial banks, of the sort that exists on main streets and may be principally concerned with e.g., making housing loans. It was thought that this was part of the poor banking processes that resulted in the great market crash of 1929. The idea was to protect those who had their money in the commercial banks (which was presumably protecting their money) from the possible impacts of an investment bank making too risky loans with main street cash. So in fact after a dozen attempts, the large Wall Street agencies in 1999 were able to repeal the Glass-Steagall Act, allowing financial agencies much greater freedom in what they did. Many blame this repeal and the repeal of other such financial controls on the great market crash of 2008.

20.8 If the Basic Neoclassical Model Is Unrealistic, Why Do Economists Continue to Use It?

As suggested in ► Chap. 6, all cultures live at least partly by myths, a set of deeply held and sometimes true beliefs that validate the everyday experiences and propagate patterns of behavior thought compatible with the social and economic well-being of at least some of its members. When today we examine ancient cultures we often marvel at what we perceive to be the strange and foolish (to us) myths that guided their activities. Ancient Mayans apparently believed that sacrificing virgins would bring rains and prosperity; the people of Easter Island apparently thought (or their priests and leaders bamboozled them into thinking) that constructing huge statues would insure the continuance of their early economic well-being and later compensate for their reduction in quality of life occasioned by their overexploitation of birds, forests, and soils; ancient Egyptians thought that the worship of Ra would make the Nile flood properly; and Medieval Europeans thought that the plague was caused by their sins, as do contemporary social conservatives when they attribute the AIDS pandemic to "immoral lifestyle choices." Some of these, and other, ways that humans have hooked their myths to their economic well-being and lack thereof are marvelously developed in Jared Diamond's important book *Collapse*. (The story on Greenland is particularly poignant). In many cases what we call today the "myths of old" were, and sometimes still are, very serious religious issues to the people who follow them. After all, a myth to one group is generally someone else's religion or cultural values. Most of these old myths appear today as more or less harmless ways to try to understand or control a world before science gave us more powerful tools, but some, like sacrificing virgins or the letting of blood from a sick person to "drain their bad humors," were extremely destructive by today's standards, and, apparently, they sometimes led to the destruction of their cultures.

Contemporary Western society also operates according to a number of sometimes contradictory myths embodied in various established conventions, religious tenets, folk wisdoms, and, we think, economic "truths." For example, the future might tell us that the basic tenets of market capitalism, including the primacy and/or virtue of individual initiative, survival of the economic fittest, the need for economic growth, the indefinite possibilities of exploitation of particular resources, material consumption as the road to happiness, unlimited substitution, that technology will solve any economic shortages, that nature is there to be exploited as we wish, and so on, are just as much myths as whatever the Easter Island statues relate to. Or they may not be. Or they may be perceived as an extremely effective way by which people at one time can live very well, but at the expense of their descendants. The application of conventional economics, whether that is a series of myths or a pipeline to reality (or some mixture) have given the residents of the wealthy North an unprecedented material standard of living and tremendous technological achievements. Yet we do believe that these myths (or realities) now threaten to undermine the affluent society they helped build without necessarily generating the unalloyed happiness it was supposed to and while, clearly, generating enormous misery (and also sometimes happiness) to many in the global South.

We believe that what separates myths from reality is the judicious use of the scientific method. Of course science itself has been, and remains, hardly immune to the need for and use of myths. In the natural sciences, we are familiar with large-scale "paradigm shifts," where fundamental scientific ideas that have been widely accepted and well developed are suddenly found to be quite wrong, leading to the replacement of the entire conceptual basis of a discipline. Some examples include: One of our freshmen asked us "Do you mean that capitalism actually encourages its own destruction by encouraging the destruction of its resource base as rapidly as possible?" "Well, we responded, that hypothesis is consistent with the data." We could not say that his statement was not true, but did say that it is a long-term process in the wealthy parts of the world. We believe that the sixteenth-century replacement of the Earth-centered Ptolemaic theory of the solar system with the sun-centered Copernican view, the ongoing replacement of population-intrinsic concepts for the population dynamics of commercial fisheries with a more complex ecosystem view, and only 50 years ago the replacement of a static view of continents with that of dynamic plate tectonics were all good paradigm shifts. The main questions are: do we need, and are we ready for, such a paradigm shift in economic models? And if we are prepared intellectually for that task can we possibly implement it given the enormous intellectual and financial investment in neoclassical economics as it is applied around the world? Our answer to the first question is that it is probably too late for most of the older economists steeped in neoclassical theory, but there are many younger economists, economists to be, and certainly a vast army of environmental, geological, and physical scientists ready to learn and help create a new economics more consistent with their own empirically based view of the world. The answer to the second is that it would be an extremely difficult and demanding task to actually implement a new policy approach to economics, even if it could be agreed as to what that should be. The other side of the coin is, however, that it might be much worse not to do so, especially if the plight of the third world degrades substantially, which we perceive as not unlikely as oil and other critical materials become increasingly scarce over the coming decades.

And there exists procedures by which we can do this and it is called the scientific method as it is applied in the biophysical sciences. Within this framework, one is able to come up with whatever hypothesis you might want as to what is the truth. But if you are inconsistent with known reality you are likely to get shot down by your colleagues. For example, when in the early 1950s several scientists were closing in on the structure of DNA (where the "A" stands for acid), the great biochemist Linus Pauling proposed a chemical model that he thought represented DNA. But Watson and Crick, who later came up with the correct structure, noticed that Pauling's structure was not an acid and so immediately shot down Pauling's model—it was not consistent with known science. Likewise, all kinds of mechanical devices for generating energy have been shown false because they are thermodynamically incorrect. We think that there has to be a lot more of this kind of analysis applied to all economic models.

We expect a great deal of resistance from established economists to what we develop here. Past criticisms of neoclassical welfare economics are almost invariably dismissed by economists as attacks on a "straw man." This response by economists is so prevalent it is worth addressing in some detail. In one sense, economists are correct to point out that the theory of many present-day economists has gone far beyond e.g., the restrictive and unscientific assumptions of *homo economicus* and perfect competition. A growing number of economists, particularly the most respected theorists in the field, have already abandoned the models of human behavior we criticize in ▶ Chap. 3. The applied work and policy recommendations of most economists, however, remain grounded in these models. Most economists still believe that contemporary work in behavioral economics and game theory can be integrated into the standard welfare model. This is wishful thinking. If the restrictive assumptions of *Homo economicus* are relaxed to incorporate current knowledge about actual human behavior, the conditions for efficient resource allocation by markets (Pareto efficiency) cannot be met.

Another tactic by economists in responding to criticism is to claim that welfare theory is based on very general, reasonable assumptions—ignoring their unsupportable interpretations of those assumptions. For example, according to economist Herbert Gintis, a definition of the "rational actor model" is that it: "holds that individual choice can be modeled as maximization of an objective function subject to informational and material constraints." In other words, people try to do the best they can with the limited means at their disposal. Their objective is said to be "utility" or "well-being" broadly defined. These seem to be reasonable and harmless assumptions. But in economic texts and applied work, "well-being" is equated *only* with the consumption of market goods chosen in a manner that conforms to the mathematical requirements of constrained optimization. We ask the reader to think what are the most important factors in your own life. For most of us family, friends, health, justice, fairness, a clean, non-degraded and uncrowded environment, spiritual issues, and good associates are all ahead of issues that could be bought or sold in the market. But every leading text in economic theory follows the pattern based on consumption. For example, the respected economists Pyndyck and Rubenfeld [12] write:

> In everyday language, the word *utility* has rather broad connotations, meaning, roughly, "benefit" or "well-being." Indeed, people obtain "utility" by getting things that give them pleasure and by avoiding things that give them pain. In the language of economics, the concept of **utility** refers to *the numerical score representing the satisfaction a consumer gets from a market basket.*

Thus, the complex issue of individual utility becomes reduced to only the consumption of collections of market goods. The analysis of market choice proceeds by making the three basic assumptions of completeness, transitivity, and that more is always preferred to less. These are the kinds of assumptions economists have refused to empirically test until recently. Without these assumptions Walrasian (neoclassical) analysis cannot work. As a leading microeconomic text points out regarding just one of these assumptions: "…substantial portions of economic theory would not survive if economic agents could not be assumed to have transitive preferences." Therefore, we believe that attempting to "fix" the NCE model through, for example, internalizing externalities (such as by adding a dollar value for essential properties or services of nature into the existing market-based evaluative scheme) is missing the point of what should be (in our opinion) our major undertaking, which is to start our economic conceptualization from scratch in a way that represents what actually occurs in a real economy. In essence we must put our conceptual economic models *inside* nature where it *must* exist (e.g., Fig. 20.2), rather than attempt through internalizing externalities to put nature inside the economic framework (Fig. 20.1). We believe that this is necessary for two reasons. First, we believe that the basic structure of neoclassical economics is so flawed as to be impossible to jury-rig back to credibility and, second, that for this and other reasons, the practical consequences of the application of NCE result in actions that are immoral and self-defeating. While we recognize that probably most older economists will not agree with our assessment, we do think it is about time to flush this question into the open so that we can have a much more substantive discussion of what kind of economics we should be constructing.

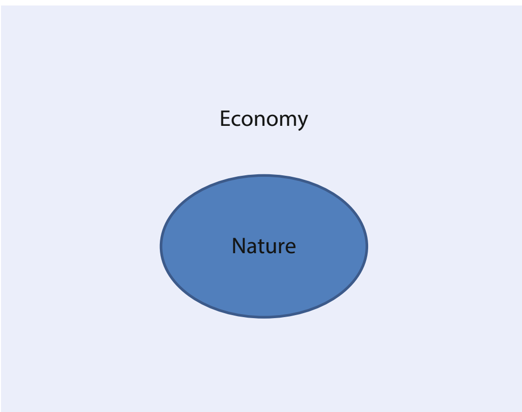


Fig. 20.1 Too often in ecological economics nature is placed "within" the economy where functions of nature are given monetary values that were originally evaluated in the economy

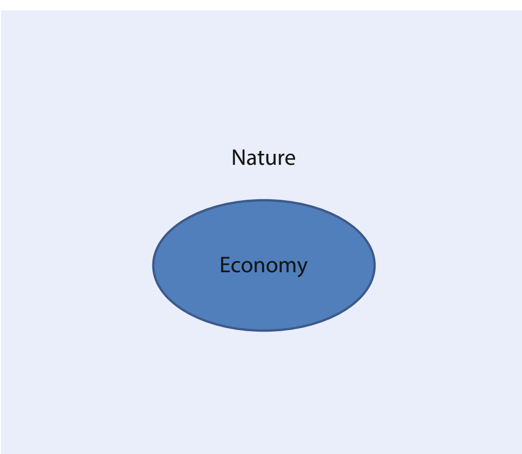


Fig. 20.2 The economy must exist within nature for it cannot exist any other way

20.9 A Final Thought on the Proper Use of Mathematics

Part of what defines science as science in most peoples' minds (including scientists themselves) is the use of mathematics, and mathematical models, to define and resolve problems. The power of mathematics (in its broad sense) is to make the results of the prediction quantitatively explicit and hence quantitatively predictable. The process of examining whether your model is a correct or at least adequate is called *validation*. An examination of the degree to which uncertainty in model formulation (how it is structured) or parameterization (what numerical coefficients are assigned) allows one to trust your results or reach certain conclusions is called *sensitivity analysis*. It is through validation and sensitivity analysis that models generate their (occasional) tremendous power in resolving truth, such as that is accessible to the human mind.

The use of mathematics was especially important in the development of physics in the early part of the past century, and the creation of the atom bomb was tangible evidence to many of the power of pure mathematics combined with practical application. Nevertheless, even Einstein preferred to solve his problems without mathematics when that was possible. Other sciences in which mathematical models have been especially important include astronomy, some aspects of chemistry and some aspects of biology such as demography and in some cases epidemiology. The importance of mathematics for most of biology is a little harder to pin down. Certainly the most important discovery in biology was that of Charles Darwin, who used essentially no mathematics in the development of the theory of natural selection beyond the concept of the potential of organisms for exponential growth. Likewise, mathematics by itself had little to do with the development of the cell theory, the structure and nature of DNA, and most modern molecular biology. On the other hand genetics, from Mendel to contemporary population genetics, has been heavily influenced by, and sometimes tends to lend itself well to, mathematics.

A final problem that we repeat from before is that there has been frequent confusion between *mathematical* and *scientific* proof. Mathematics can generate real proofs relatively easily because you are working in a defined universe (through the assumptions and the equations used) to which it applies. If you define a straight line as the shortest distance between two points, then you can solve many problems requiring straight lines. But the world handed to us by nature is neither so straight nor so cleanly defined, and we must constantly struggle to represent it with our equations. Hence, a mathematical proof becomes a scientific proof only in the relatively rare circumstances when the equations do indeed capture the essence of the problem.

20.10 Now We Will Appear to Contradict Ourselves

Despite all of the many problems of modeling, we do not understand how one can use the scientific method, that is, generate and test hypotheses, on complex issues without the use of formal modeling. This is as true for management and policy-related issues as for theoretical ones. The reason is that models are an explicit formalization of our assumptions about a system, and such allow for explicit testing of how you think the world works. In our view quantitative (or occasionally non-quantitative) models of at least sufficient complexity are necessary in the complex world of economics (and of environmental sciences) because it allows one to apply the scientific method to complex real systems of nature and of humans and nature. *But it is critical that the right kind of models be used.* And the way to do that is quite simple: Try to represent the real system that you are dealing with rather than some abstraction that happens to be analytically tractable. Quite simply most real problems require computer modeling, not analytic modeling. The power of models is to make our assumptions explicit, generally quantitative, and hence testable.

20

Questions

1. What is a model? Where do you find models?
2. When speaking of a model, what do we mean by "solution"?
3. What is the difference between something that is mathematical and something that is quantitative? Can something be both?
4. What is the difference between an analytical solution and a numerical one?
5. Explain how some cities were astronomical instruments.
6. Give one or more examples of a model that is conceptually incorrect but that nevertheless gives good predictions.
7. Give one or more examples of models that are very commonly used but that are probably incorrect.
8. What is a paradigm? Give several examples.
9. Can you give five general types of models (Hint: One of them is computer).
10. Can you explain the apparent paradox that one can use complex mathematics only on a rather simple model?
11. What is an externality?
12. How can we separate myth from reality?
13. What are some of the ways in which some economists today have criticized the basic models of contemporary economics?
14. Discuss the conceptual advantage of putting our economic models inside our models of nature vs. the opposite.
15. What is validation? Sensitivity analysis? How would you use them in economics?

References

1. Hall, C.A.S., and J.W. Day, eds. 1977. *Ecosystem modeling in theory and practice. An introduction with case histories*. New York: Wiley Interscience.
2. Aveni, Anthony F., ed. 2008. *People and the sky: Our ancestors and the cosmos*. London: Thames and Hudson.
3. Pyper, B.J., F.J. Mueter, and R.M. Peterman. 2005. Across species comparisons of spatial scales of environmental effects on survival rates of Northeast Pacific salmon. *Transactions of the American Fisheries Society* 134: 86–104.
4. McAllister, M.K., and R.M. Peterman. 1992. Experimental design in the management of fisheries: A review.

North American Journal of Fisheries Management 12: 1–18.
5. Hall, C.A.S. 1988. An assessment of several of the historically most influential theoretical models used in ecology and of the data provided in their support. *Ecological Modeling* 43: 5–31.
6. Hess, H.H. 1962. History of ocean basins. In *Petrologic studies: A volume in honor of A. F. Buddington*, ed. A.E.J. Engel, Harold L. James, and B.F. Leonard, 599–620. New York: Geological Society of America.
7. LeClerc, G. 2008. In *Making development work: A new role for science*, ed. G. LeClerc and Charles Hall, 13–38. Albuquerque: University of New Mexico Press.
8. Hallock, J., P. Tharkan, C. Hall, M. Jefferson, and W. Wu. 2004. Forecasting the limits to the availability and diversity of global conventional oil supplies. *Energy* 29: 1673–1696. Hallock J., Jr., W. Wu, C.A.S. Hall and M. Jefferson. 2014. Forecasting the limits to the availability and diversity of global conventional oil supply: Validation. *Energy* 64: 130–153.
9. http://knowledge.wharton.upenn.edu/article.cfm?articleid=2234
10. Krugman, P. "How did economists get it so wrong: Mistaking beauty for truth". *New York times*. Sept 2, 2009.
11. Daly, H. and J. Cobb 1992. *For the Common Good: Redirecting the economy toward Community, the Environment and a sustainable future*. Beacon Press Boston.
12. Pyndyck, R.S., and D.L. Rubinfeld. 2005. *Microeconomics*. Saddle River: Prentice Hall, Inc.

Applying a Biophysical Economics Approach to Developing Countries

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_21

It seems imperative that we as individuals who care about the human condition and about nature must create a new way to undertake developmental economics and perhaps economics in general. The reasons this is so important have been reviewed in previous chapters and include our dissatisfaction with the intellectual foundations of conventional economic models used in development and of the results that have occurred with their use, the general sense of many development economists themselves that conventional economics has failed, the need to do something that will work, the concern that most knowledgeable people have that the future, and especially the future of most developing nations, will be much more constrained by the "end of cheap oil," and the need to protect whatever nature is left. We try to develop such a model in this chapter, summarizing certain approaches and even successes of the past, and use a biophysical basis to try to generate a synthesis to help the reader. We are not foolish enough to believe that we can in one fell swoop cure all the economic problems that generations of traditional economists have not been able to, but we believe that we do provide a useful basis here for beginning that process and for generating useful results now for field workers.

We undertake this analysis with the full understanding that conventional (e.g., neoclassical) economics, for whatever its limitations, is an extremely well-developed and well-integrated approach where, in general, the players are well entrenched and agree upon the rules. And we acknowledge that their influence is increasing in the applied world, even as many academic economists step back from the pure model. For example, "computable general equilibrium" (CGE) models, which are pure applications of NCE, are increasingly used in world trade organization (WTO) negotiation rounds that affect billions of lives. In addition conventional economics has been developed in such a way (e.g., by emphasizing money rather than energy, demography, and other resources as we do) as to appear to be a logical extension of the day-to-day economics with which we are all familiar. These are significant hurdles to overcome for those of us that believe that a more useful and accurate economics can be developed. Nevertheless we perceive the importance of this to be so great as to require our best efforts to do so. A point in our favor is that we know that we are not alone in challenging NCE, and our best allies may be some of the economists themselves, especially those who spend their time in the realities of the developing world.

We have spent considerable time in the past developing a biophysical assessment for the country of Costa Rica, and much of what follows is based on our experience in that assessment [1]. That book has 26 chapters with detailed assessments of essentially all important aspects of the Costa Rican economy. It has in addition (on a CD bundled with the book) a comprehensive and user-friendly visualization and model that we think is extremely important in communicating biophysical information and assessments to both other professionals and also to lay people. The basic idea of the main visualization and model is that there is a central image—that of the country of Costa Rica, shown with the mountains visible in a three-dimensional representation—with ten small graphs around the edge with lots of different information that is plotted over time as you watch the rather amazing deforestation unfold in the central image and the green, forested country turns to agriculture and pastures represented in yellow, while the numbers of humans, cows, hectares of used and degraded lands, and so on grow nearly exponentially on the graphs around the margin.

One characteristic of these analyses—which may be good or bad depending upon your point of view—is that there is (usually) no attempt to reduce the various different information sets to a single scalar (such as is usually the objective in, e.g., money-based economic cost-benefit analyses). Rather the idea is to put all of the dynamic information, including land use, demographic, environmental, economic, and so on, on the screen simultaneously and then let the user or decision-maker (or the people effected) decide whether they prefer the existing path of development (by whatever criteria they choose) or might rather have something else. This approach can be particularly effective when integrated with historical patterns of, e.g., land use. Most people living in Costa Rica today are too young to understand how much their country has changed in one human lifetime, but they can see that clearly –and are often amazed—when they see this as an n-dimensional visualization. So most of the rest of this chapter is a discussion of what kind of information you might want to include in such a visualization or perhaps in some simpler analytic structure such as a spreadsheet. The model allows for the implementation of policy and the observation of how that would effect the many parameters. When shown this model, Oscar Arias, the former President of Costa Rica and Nobel Prize winner, said to Hall "I like it. It forces the decision-maker to see the consequences of his decisions". Would we have such a model for the present United States and a president who would pay attention to it!

A rough guess as to the cost of developing the kind of overall biophysical analysis for a small- to medium-sized developing country is on the order of one to ten million dollars, assuming that you are undertaking this analysis with competent and not greedy investigators and that the biophysical and economic database is well developed, as was the case for Costa Rica. Our very thorough assessment of Costa Rica was done on a small fraction of that, although much of the work was subsidized with sabbatical pay from Hall's university, essentially free graduate student help, other projects that had already funded Leclerc, and the data, interest, skills, and good will of numerous Costa Ricans. Most of the examples we give here are aimed at such a national level, although the biophysical approach that we are advocating is in theory applicable at any regional level that the investigator might choose. The most important scale issue is that much of the data is generally most readily available at the national level.

21.1 Other Somewhat Related Biophysical Approaches

Before we give our own approach, we think it is useful to review a number of other biophysical approaches that have been developed either to evaluate/assess specific environmental impacts of economic activity or for some other explicit reason. While these approaches do not give the full and comprehensive environmental and economic analysis we advocate, we think it is important to review them as they can be very useful supplements to the analysis that we give below.

We would also like to emphasize that our attempts to build a biophysical assessment are only marginally related to most of what is being done under the aegis of "environmental economics" or even the bulk of the activity in "ecological economics." Although the goal of environmental economics (and a substantial part of ecological economics) is to integrate the environment into economic analyses, in fact it has been mostly about putting a dollar price tag on all kinds of environmental objects and services, and while we applaud such analyses, that is not at all our objective here. One basic reason for this is that we believe that the dollar or other monetary unit is basically defined in market situations for nonessentials, the demand for which hardly represents real human wants and needs because it is often tremendously influenced by advertising. In addition dollar values often give extremely poor information about basic resources: for example, as wild salmon increasingly are disappearing and are hence of less and less value to our society, their price goes up indicating they are becoming more valuable than when they were cheap and abundant!

Hence we believe that giving a dollar value to many things is often a rather poor estimate of the value of our most prized things, including our relations to those people close to us, justice before the law, the maintenance of natural environments, and the milieu of the Earth that allows us to exist here in the first place. In fact all of these are under assault by dollar-based aspects of our economy, and hence in our opinion, dollar-based criteria are not appropriate for making assessments of the value of nature or our most essential resources. That said we of course realize that we live in a monetary-based world where many things must be valued in monetary units for routine day-to-day transactions. So we try to walk an appropriate tightrope between using and not using monetary estimates.

The first assessment procedure we review is to examine the environmental requirements for a given region (for our purposes a social and economic unit such as a country or city) in terms of the quantity of land required to support the activities on that area considered. The most comprehensive and thorough such analysis is called the ecological footprint which is run by Mathis Wackernagel [2]. For example, they found that the land area required to support the needs of the city of Vancouver, Canada, was about 18 times the land area of the city itself. This included land areas needed for growing crops and producing cows, fish, and other animals consumed, growing timber, mining minerals, and so on (about half the area required) as well as assimilating the sewage, toxins, CO_2, and other wastes produced (the other half). Such assessments always show that the areas actually in use supporting people are much greater than the areas the humans actually occupy and give lie to those who say that the Earth can support much larger human populations (or even the present level) indefinitely. They conclude that about three Earths are needed for today's population and level of affluence if we are to live on income rather than by running down capital. Over time the authors have developed and refined their methodology impressively and made its use on their website very straightforward and easy. Because they trace back virtually all the major material substances used by different groups of people, their complete list of material used constitutes a ready-made list of the biophysical materials required to support an economy. What they have not done yet is to relate the materials required to the level of monetary activity or ask these questions of developing countries. Once this is done, we will have one rather good biophysical assessment at our fingertips.

The second approach is to undertake energy analysis, which in its many variants means essentially how much energy does it take to undertake various economic activities. These methods were developed most importantly at the University of Illinois in the 1970s by Bruce Hannon, Clark Bullard, and Robert Herendeen and were applied to most aspects of our economy including agriculture, manufacturing, provision of services, and so on [3–5]. A feature of these studies was that they calculated not only the direct energy used (such as the energy used in a tractor factory to make the tractor) but the indirect energies as well (i.e., the energy to mine and refine the iron, plastics, and so on used by the tractor factory). As a rough estimate, about half the energy used to make some product sold in "final demand" occurs in obtaining and refining the raw materials. Summaries of the results of such studies are given in the above publications and Hall et al. [6] and Cleveland [7]. An important aspect of this research is that the numbers are old, as there has been little Federal funding of such energy research for decades as energy analysis has fallen into political disfavor or, more accurately, indifference, because in the minds of many (but not us), the market has resolved the energy issues of the 1970s. However a recent study by Carnegie Mellon has updated these analyses to 2002 (by methods that seem pretty defensible according to Robert Herendeen), and these estimates are readily available on their website [8]. Sergio Ulgaldi and his students at the University of Naples are putting together a web-based system for calculating the material costs for many different commodities (e.g., a new building) including the associated environmental costs.

Howard Odum, Mark Brown, and others have argued that, while the above energy analysis is useful, it is incomplete because it does not take into account either the environmental energies required to manufacture something or correct for the fact that different types of energies have different qualities. For example, a kilojoule of electricity has value to society beyond its ability to simply heat water and hence more value than a kilojoule of coal, because of its special properties and because it takes about three heat units of coal in a power plant to produce one unit of electricity, the rest more or less of necessity being released into the air and water. Likewise a kilojoule of sugar fixed by a plant has more value than a kilojoule of the sunlight that made it and so on. Odum has generated the idea of embodied energy or more explicitly emergy (with an m, as in energy memory, a concept analogous to the embodied labor, or total energy required to make, in a manufactured item) as a term to reflect the various qualities of energy. Odum and his student Mark Brown have developed an extensive accounting scheme to measure this and to compute the quantities of emergy required to make, or cause to happen, many things [9–12]. Transformity is a word used to evaluate the different qualities of different types of energy. An advantage of this approach is that it is obvious that if we want to account, e.g., the oil used to manufacture something, we are missing all together the large quantities of environmental energies that are just as much needed to make it. These energies include, for example, the energy used to distill freshwater from the sea and lift it to mountain tops which allows it to form rivers and hence become available to plants and to humans. Likewise the sun runs photosynthesis and everything that derives from that even though we do not pay Mother Nature for either the water or many of the products of photosynthesis. In addition it includes in the analysis an emergy assessment of the environmental services foregone because of the activity in question. While the idea is tremendously appealing to us, and the comprehensiveness essential in our view, the difficulty in estimating transformities makes its use less desirable to some.

It may be that all of these techniques are measuring something quite similar and that their utility may converge. Their use has not been compared often. Hall, Brown, and Wackernagel compared the carrying capacity of Costa Rica for humans using a comprehensive economic approach that went well beyond market costs, as well as two biophysical assessments: ecological footprints and emergy analysis [13]. The results of the three approaches were very similar, giving hope that we are approaching a true cost using both biophysical and comprehensive economic analysis. However, although each of these procedures is helpful in assessing a biophysical economic analysis, we still feel that it is useful to generate a more explicit summary as to how we can undertake biophysical economics. We do this below; however we look forward to the day when scientists and policy makers agree on a set of assessment procedures to be integrated in one useful package. We look forward to the time in the not too distant future when as part of the biophysical analysis of any item or activity all that would be necessary would be to go to one website, maintained by skilled professionals, and type in the quantity (in tons or dollars of a particular year) to get all of the material, energy, emergy, footprint, environmental degradation, and so on associated with that economic activity. A step in that direction is the triple bottom-line approach (economic, energy, and environmental) of Barney Foran, with free software available to help with the assessment [15]. Later this can be done also for different countries or international corporate entities to give more explicit values. Perhaps someday there will be a label on your breakfast cereal that gives, in addition to calories and sodium per serving, an assessment of the fuel and solar energy required to make it as well as the soil and biodiversity loss, maybe all summarized in terms of energy.

21.2 Explicit Procedures for Creating a Biophysical Economic Analysis for a Country or Region

While we wait for this future web-based synthesis, there is a great deal of quantitative analysis we can do and in fact that can help provide the basis for this web synthesis. We base what follows on our earlier work related to preparing our previous book *Quantifying Sustainable Development: The Future of Tropical Economies* [1]. This assessment included extensive discussion of our (and others) biophysical approaches with contributors and our extensive previous experience with assessing land use change [15, 16]. We also base our assessments on simply living for much of our lives in the developing tropics (especially LeClerc, who has done everything to escape his native Canadian winters) and reading a large number of newspapers and scientific papers there. Hall [1] represents the most serious attempt to date to develop a complete biophysical economic model of a national economy which we summarize and extend in this chapter.

We will be the first to recognize that this is a very imperfect activity, that we are just learning how to undertake such analyses, and that there are many changes that will be developed over time. Nevertheless we have found that this approach in part or in full has served us and our colleagues and students well for analyzing many basic characteristics of a country or a region.

We have come to the conclusion that there is a way to undertake routine biophysical economic analysis, including a rapid assessment of development, and to use this process to help construct better development schemes. We propose a methodology that unfolds in five steps that can be put simply as:

Step 1 State your objectives (with the right people including your critics).

Step 2 Assemble a time series database of critical biophysical parameters.

Step 3 Make an assessment of critical economic parameters with as much data as possible from the past.

Step 4 Construct a comprehensive simulation of the future.

Step 5 Make the right decisions.

We assume that after these steps are taken into account for devising a development scheme, money will flow in the right directions; schools will be built, equipped, and populated; and institutions will improve. Nevertheless we are also quite aware of the potential for, e.g., corruption of leaders to undermine our efforts. Does the use of explicit and open science make corruption less possible? We think so but do not really know! Part of what must be done is the professionalization of all government institutions and personnel, including accountants.

21.3 Step 1: State your Objectives (with the Right People)

It is not possible to undertake a journey, no matter how sophisticated your vehicle, if you do not know where you are going (unless of course your objective is simply the activity itself). So the first thing to do in undertaking a biophysical assessment is to ponder, discuss, and then state explicitly your objectives. Often people confound problems and objectives. An objective should not be a series of problem-solving activities; it should be seen as a long-term desired future condition. For the Costa Rica study [1], the main objective was to determine to what degree, and in what ways, the country was or could become sustainable. This led logically to the next set of objectives which was then to determine what we meant by sustainability, which in turn led to some interesting literature that showed that very different people had very different perspectives on what sustainability meant, most of which were antithetical to each other!

A second part of this analysis is to examine what objectives people had in the past for related issues and how well these were achieved. In other words, a review of pertinent literature both for the region being analyzed and also of past public and private development projects, their objectives, procedures, successes, and failures. Many of these analyses use (or should use) time series data of, e.g., economic, agricultural, or other data. It simply is not possible to understand whether whatever plan you are undertaking is successful or not unless you have a yardstick of the past trends in time to compare it to. An important issue is to state objectives as hypotheses which then can be tested, something that is rarely done. While it is often difficult to test hypotheses, one can often restate policy objectives as hypotheses and then see if ensuing data are consistent with that hypothesis or not [17].

Very often the objectives will be stated in social, economic, or environmental terms. Given that we agree with that perspective, the reader might be curious as to why we then focus so much on the biophysical aspects of analysis. The answer is simple: we believe that social, economic, and environmental issues must be addressed and, where possible, resolved within the context of the biophysical systems within which they must take place. It is very easy to list the various things that you would like to have: higher incomes with greater equitability, less pollution, greater welfare, and so on. Given that for the developing world these and other objectives are very often not met means that there are serious constraints. Some of course are social, and we include here especially corruption and the very unequal distribution of whatever wealth is available. But much of what gets in the way of achieving one's social or economic objectives is biophysical, including resource availability, climatic constraints, and biophysical mismanagement including, for example, overfishing, soil erosion, fuel limitations, ability to generate foreign exchange, and so on. It is important to understand what these are or might be.

And it is especially the biophysical aspects of development that have been neglected during decades of neoclassical economic policies. Therefore the biophysical context must be restored in mainstream thinking, possibly as the framework within which the social and economic possibilities are considered, hence our biophysical emphasis, although we in no way wish to diminish the importance of the social, political, and economic elements. In fact we believe that the reader will find that most of our papers try to integrate the biophysical and the social sciences toward attempting to meet their objectives.

If we are interested not only in the progress of science but also in its impact in the development of the country studied, then we have to find the right people to develop the models with. These people will help at many levels: to clarify the objective, to obtain the data (not easy in many developing countries), to provide key insights to interpreting the data and for prospective analysis, and to make the connection with policy so that we can extend its use beyond the scientific paper. If we are all involved from the start in developing an analytic model (i.e., "companion modeling") [19], there is a good chance that we learn from each other and end up with a model (or a family of models) that is not only more relevant but one that will continue to be used for policy making. Allan and Holland, and Beaulieu in [17] give several hints about how to identify who you should work with and how to connect to a development process. A good starting point is to do a stakeholders' analysis and work, with the right people, on a shared vision for the country or region. This is where genuine objectives will appear more clearly to all and when the collective learning process will begin.

21.4 Step 2: Assemble a Database of Critical Biophysical Parameters

The first step in undertaking biophysical analyses (once past time trends of pertinent data have been prepared) is normally to determine the physical characteristics of the country or region being analyzed. Such analyses are far easier than in the past due to the increased availability of good digital summaries compared to 20 years ago. An example of how such a database has been developed is given in Barreteau et al. [19]. The best way to do this is to generate an assessment of the physical resources of the region in question.

An essential requirement is a summary of energy resources including any known oil, gas, and coal deposits; assessments of what might be found in the future; developed and potential hydroelectric, solar, and wind potential (for which you need meteorological information); biomass possibilities; and so on. In all of these assessments, it is important to realize that in general the better resources were developed first, such that increased exploitation may be more energy and monetarily expensive. For all of these generate a time series of their use.

But different types of energy have different properties or qualities, and often it is useful to take that into account. Generally the data available will be in the form of heat units (i.e., therms, BTUs, kilowatt-hours, kcal, or the most commonly accepted units used today which is joules). By heat units we mean that the energy is measured by its ability to heat water, for example, 1 kilocalorie (kcal) is the energy required to heat 1 kilogram (about 2.2 pounds) of water 1 degree centigrade. These units are all intraconvertible and there is no real difference among them. When fossil fuels are compared to electricity generated from hydro or nuclear power, it is generally best to multiply them by a factor of about 2.6 to account for the difference in their ability to do work and also their opportunity (or conversion) cost if they were made from fossil fuels. Additionally we need to undertake an assessment of the various environmental energies that must be supplied for the economy to work properly. As stated above this can be done most comprehensively using an energy analysis.

Similar assessments are required for natural resources that are not energy sources, such as:

1. Nonfuel mineral resources, such as metal ores. The important components of this are the size of the reserve (in tons), the quality (i.e., percent metal in ore, both at present and as exploitation proceeds), the depth and ease of extraction, the energy cost of extraction of different amounts, and so on. Since in general the best grades were used first in the past, the remaining resources may not be as cheaply or profitably exploited as was once the case. Since often the exploitation of minerals occasions significant pollution, any such impacts, and a social and monetary estimate of that damage, must be made before the project begins. These issues must be considered in addition to expected market prices and other routine economic factors.
2. Water resources, both quantitatively and qualitatively, first in overview and then spatially. Some of the information that needs to be generated or summarized includes rainfall and flow of major rivers (both as a mean and for drought and wet years), ground water resources and their vulnerability to depletion/salinization, evapotranspiration and soil moisture over space and time, water bodies that are significantly polluted, and so on.
3. Land resources for examining agricultural (and other) potential, i.e.:
   - A soil map, ideally with the soil units related to crop productivity, including where possible potential and actual erosion
   - A digital elevation map
   - A land use map

21.4.1 Taking Demography into Account

We believe that fundamental to what one is trying to achieve with almost any biophysical model is a proper representation of human demography. Fortunately, excellent datasets exist for less developed countries (LDSs), from nationwide census data every 5 or 10 years to yearly estimates based on samples in between. (Note that because NCE is based on the behavior of individual firms, it is insensitive to demography!)

For prospective analysis it is necessary to generate a demographic model based on actual demographic data. One simple model is:

$$P_t = P_0 e^{rt}$$

where P is the population level (normally in millions), P_t is the population at time t years into the future, P_0 is the population at some initial time t, the natural log of e is approximately 2.718 and r is the "intrinsic rate of growth," the rate at which the population is growing or, better, is expected to grow. The value r (in units of proportion of the existing population per year) is the birth rate (b) minus the death rate (d). Hence the term ert is a number that will usually be greater than 1.0 and will be the factor by which the population is larger (relative to the initial population) over time. The doubling time of a population can be calculated by dividing the number 70 by the growth rate expressed as a percentage, so, for example, a population with a 2 percent per year growth rate will double in 35 years. This simple model is often reasonably accurate, at least within the restrictions of knowing the value of r, for a few decades.

But there have been many who believe that to continue to use an exponentially growing model is seriously flawed, as populations cannot grow exponentially indefinitely as they would run out of food, resources, and/or space (i.e., carrying capacity). Some models, attempting to represent that fact, will assume or simulate some sort of empirical plateau, (in other words, r diminishes) or saturation of growth. A logistic or S-shaped curve is used often to simulate that saturation effect. Although the logistic equation is simple and has some perhaps good logic behind it, in fact few populations in nature follow that pattern, and attempts to use that model to predict human populations in the past failed miserably. The debate between "implosionists" and "explosionists" is still alive (because the data support either view equally well), and while the S-curve is still the most widely used distribution for making human population projections in less developed countries (LDCs) (see ▶ www.prb.org), the beginning of the plateau could be put at any time after 2050.

Both the exponential and the logistic model have a number of liabilities, including that they are not sensitive to changing values of r over time and are insensitive to the more detailed demographics such as the number of pre-reproductive vs. post-reproductive females, and of course they are for only one geographical unit. More complex and accurate, or at least sensitive, models can be made using what is known as a Leslie matrix, which is usually solved in a spreadsheet or a computer program. A simple example in FORTRAN is given in ◘ Table 21.1. Data for all of the world's countries can be obtained from FAO or the CIA database. Sometimes the growth and death rates are given for 5-year intervals when annual values are needed. To use this data, it is necessary to enter the data into a spreadsheet such as Excel and fit, e.g., a second- or third-order polynomial to the data to get a relation from which you can generate values for each year as well as predictions into the future.

Additional demographic information can be developed including poverty assessments, health, and labor productivity.

Additional geographical information needs to be developed on the location and extent of built infrastructure including cities, villages, transportation, industries, ports, airports, protected areas, land tenure (private and public), and so on. These can be built into additional geographical information systems (GIS) data layers as is well understood from conventional GIS analyses. This information is useful in understanding the accessibility of resources to populations and as drivers for predicting land use change. Often our overall objective is to simulate how future land use, economic, and food security scenarios might be as influenced by demography, erosion, policy, climate change, and so on.

21.5 Step 3: Make an Assessment of Critical Economic Parameters over Time

The first step is to undertake an assessment of the current economy and its recent history. There are a number of locations to find *empirical*

Table 21.1 A simple Leslie matrix in FORTRAN

```
PROGRAM LESMATRIX
!************************************************************************
! Dictionary:
!************************************************************************
! ACLS                  = Age class of the human population.  1 equals all people before
!                         their first birthday, 2 = all people between their first and
!                         second birthday and so on.
! PopNum(YR,ACLS )      = Population number for each age class for each year
!                         This state variable is updated each year.
!  DRate(ACLS)          = Age-specific death rate
!  Births (ACLS)        = Number of births per year per female by age class (this may be
!                         known only on average)
!************************************************************************
!************************************************************************
! Define variable type:
!************************************************************************
INTEGER PopNum(100,100) , YR, ACLS
REAL DRate (100), BRate(100)
!************************************************************************
!************************************************************************
! Open read and write files        :
!************************************************************************
 OPEN (1,FILENAME = "LeslieMat.DAT", Status = "OLD")
 OPEN (2,FILENAME = "LeslieMat.OUT", Status = "UNKNOWN")
! Read in initial population numbers (in thousands or millions) & age-specific death rates
!************************************************************************
READ (1,900) (PopNum(1,ACL), ACL= 1,80)
READ (1,900) (DRate      (ACL), ACL = 1,80)
READ (1,901)  BRate      (ACL),ACL = 1,80)
! Write output headers:
!************************************************************************
WRITE (2,902) "Table 1,  Population levels by age class"
WRITE (2,903) "Year  Age Class >" , (ACLS(I), I = 1,80)
! Solve equations annually for 50 years starting in year 2000
!************************************************************************
DO YR = 1, 50
        Ryr = 2000 + YR              ! Real Year
        PopNum(Yr,1) = BirTot          ! Births from end of last year considered age class one

       ! Do for 80 year classes (assume 80 is oldest year people live or at least reproduce

    DO ACLS = 2,80                     ! New members of first age class already added in as births
        Births  = RepPop * BRate (ACLS)  ! Sum up number of potentially reproducing females
                                          ! (here age 15 to 50)
                                          ! Move each year class forward, reduced by their
                                          ! death rate
        PopNum(YR,ACLS) =  PopNum(YR-1,ACLS-1) –(1.0 * DRate(ACLS)
        IF (ACLS.GT.15.AND.ACLS.LT.50) RepPop = RepPop + Pop(YR,ACLS)
        BirTot = BirTot + Births
    END DO
  WRITE (1,904) YR, (PopNum(YR,ACLS), ACLS = 1,80)
END DO
!************************************************************************
!Format:
!************************************************************************
900 FORMAT (80I6)
901 FORMAT (F8.2)
902 FORMAT (A20)
903 FORMAT (A15,80I6)
904 FORMAT (15X,80I6)
!************************************************************************
END PROGRAM LESMATRIX
```

*Source: Charles A.S. Hall, with the assistance of Athena Palmer

21

information for this, but probably the easiest is to get the data off the web, generally by using Google or another search engine. Good sources are the large multilateral organizations (United Nations Food and Agricultural Organization, United Nations Development Programme, World Trade Organization, Non Governmental Organization, World Resources Institute), and the unavoidable World Bank. Several organizations provide country fact sheets, the US Central Intelligence Agency Fact Book (▶ http://www.cia.gov/cia/publications/factbook/index.html/) and The Economist (▶ www.economist.com/countries/), and as the digital divide gets narrower, there are more and more data from LDC government sites available. These government sites often contain key documents on policies, feasibility studies, law texts, economic summaries, etc. Travel books are quite useful to have a grasp of country's idiosyncrasies. A problem with many sites is that there is no time series data which makes the FAO (Food and Agricultural Organization of the United Nations) data probably the most generally useful, as they have data back to 1961.

From this information a time series of economic activity can be derived. Some data we suggest might be considered include a time series of basic monetary economic information, including GDP over time.

While any analysis of any raw GDP data almost always shows a rapid increase over time, this is very misleading as much of the increase is due to inflation. So the first thing to do is to correct the data for inflation, normally by expressing all data in terms of monetary units for 1 year, for example, "2000 dollars" or "2004 Pesos." This is done by using "implicit price deflators" (the easiest ones can be found in the "Statistical Abstracts of the United States"). This is especially useful when dealing in US dollars, although it is more accurate to use corrections implicit for the country in question. In the United States and many other countries, there are also more specific correctors for different sectors of the economy, for example, for energy and for food.

A second step is sometimes required, which is to make an additional correction for purchasing price parity (PPP). If a nation's GDP is corrected for inflation relative to the US dollar, as it often is, it is also necessary to correct for the fact that the increase in prices expressed in dollars does not reflect the fact that there is often far less inflation for local products such as food than, e.g., imported computers or fuel paid for in dollars. On the other hand, if you are interested in the issue of how much, it costs for e.g. imported oil (which must be paid for in dollars or euros), then correcting for PPP is not useful. Since for many developing countries the inflation rate applied to dollars is considerably greater than the rate applied to local items, this can be an important issue.

To express the meaning of the GDP changes (corrected as appropriate as given above) in terms of how it effects the average person's ability to purchase goods and services, the total national GDP, corrected as above, needs to be corrected to per capita values. The total national GDP tells you little about how well individuals in that country are doing in terms of their own economic welfare or purchasing power. Dividing the total wealth production by the number of people gives you per capita wealth, which is roughly proportional to at least some important aspects of the average person's material well-being. To do this one simply divides the total GDP (corrected as above) by the number of people in the country for that year to get the per capita GDP. This then results in a decrease in the effect of GDP increases and in many cases where the population increases more rapidly than the GDP people, on average, get poorer.

Even per capita changes do not tell the whole story, for most of the GDP may go to only a relatively few people. One way to examine this issue is to use or compute the "Gini index," named after Italian economist Corrado Gini. This measures the degree of inequality in a society. If there were perfect equality, the Gini coefficient would = 0. If nobody except the richest individual had any money, the Gini coefficient would = 1. Therefore, the larger the Gini coefficient, the greater the degree of inequality. In 1968 the Gini coefficient for the United States was 0.388, by 2015, it had risen to 0.480, indicating a substantial rise in the degree of inequality.

An extremely important aspect of sustainability is whether a nation is able to do whatever economic activity it does without going into international debt, which tends to be a killer aspect of development that leads many otherwise excellent development schemes into failure. Since the desire for foreign products, both those essentials, e.g. for the development of food production but also luxury items, requires payment in foreign exchange, that is, dollars or euros, it is essential for a country to export enough to pay for these items. The alternative is foreign debt, which in many countries is more or less the largest problem in making an economy that works. Costa Rica, for example, needs to use about 15 percent of the foreign exchange it generates through the sales of bananas, coffee, and tourist services simply to pay for interest on its foreign debt. It uses perhaps another 20 percent of the foreign exchange it earns to pay for the generation of the exports, i.e., for the fertilizers, plastics, and fuels required to make bananas. Since there are enormous demands in Costa Rica for imported items (from cars, buses, and trucks to fuel to run them, to computers, to apples) and a rather limited international demand (or more properly a huge oversupply) of bananas and coffee, then it is real tough for countries like Costa Rica not to get into debt. On top of this, governments often borrow from external banks to, e.g., make payrolls or provide health services. While Costa Rica has done much better than many countries (including the United States) in not running up external debt, it is a very difficult issue. Hence it is useful to plot imports, exports, and their difference, as well as debt and its accumulation or decrease over time.

Another issue that contributes to a large difference between imports and exports is that developing countries tend to be desperate for development capital and that capital is rarely available internally. So, for example, Costa Rica needs more electric power as its economy grows, and that can be supplied by developing more hydropower. But the Costa Rican government does not have the investment capital for that. So Japanese power companies are more than happy to build the hydropower plants that are needed because they are happy to collect the revenue from those plants. The problem is solved, sort of, but there is a new revenue flow out of the country. The point is that development projects need to be examined not only from the perspective of their promised gains but also their costs including, of course, their costs and gains to whom.

21.6 Undertaking a Biophysical Assessment of the Current Economy

The next major step is to look at the biophysical resources needed to make the economy do what it does and, presumably, to do more of the same in the future. Since we also have developed time series of economic activity and also time series of energy used, we can quite easily develop the energy intensity, which is the energy used per unit of economic activity, either for the economy as a whole or for some aspect of interest. This is the first step required to understand the biophysical resources needed for the operation of the economy. A similar concept (actually the inverse) is assessing the *efficiency* of an economy. In general efficiency is the output of a process divided by the input. *Efficacy*, a similar sounding but very different term, is the effectiveness of some activity regardless of the efficiency; in other words it is getting the job done. For example, we might say that the US economy is very efficacious, that is, it produces a great deal of goods and services. But its efficiency, that is, the total dollar value of its output compared to the quantity of energy used to generate that wealth, is rather low compared to many other nations. One straightforward measure of efficiency that we might want to calculate then is the output of the economy divided by its energy input, which if we have the information derived above we can do very easily in a spreadsheet or computer program. The efficiency of the economy can be seen by the ratio of the two and the changing efficiency by the changing slope of that line.

A critical issue is that most developing countries are dependent upon imported petroleum, which is unlikely to be indefinitely cheap or even available [19]. Thus contingency energy supplies and their potential cost need to be considered. Increases in energy prices tend to raise havoc with LCDs. For oil-producing countries that have become dependent upon revenues from petroleum, peak oil, which is inevitable, tends to cause political chaos [20].

Depending upon the objectives of the study, other indices can be used, such as imported vs. domestic energy or GNP per unit of water, or agricultural production per unit of energy or fertilizer used, or GNP per unit foreign exchange gained or lost or many other objectives. When we have done these analyses in the past, we have often found that GDP increases more or less in step with energy, water, fertilizer use, and so on, so that efficiency does not change much over time. This has important implications for the economic aspect of efficiency for if efficiency is not increasing that implies that the only way to generate wealth is through the further exploitation of resources, something that has ultimately serious environmental and supply implications. Much more detailed analyses can be undertaken through the use of input-output analyses.

An important aspect of a biophysical (or any) assessment is that there are often not clear ways to achieve several goals at once, and one is left with trade-offs. Several of the chapters in this book are focused on that issue. Finally, development projects that were once very good often crash over time, as is classically illustrated with wild fisheries and aquaculture. These crashes are often, but not always, predicted through fishery science but not ever, to our knowledge, through market assessments alone.

21.7 Predicting the Future Energy Needs of a Society

Presumably any such biophysical analysis will show that the economy of the region is moderately to very energy intensive and that any expansion of the economy is likely to be even more so. Most development is presently based on oil. Thus future expansion of the economy presupposes the physical and economic availability of oil or at least some other equally useful form of energy, if that exists (which we doubt). At present there are about 38 oil-exporting nations. The economies of most of the smaller- and medium-sized exporters are becoming themselves much more energy intensive over time, and most will become net importers themselves within decades as their own domestic use intercepts their production [19, 20]. Thus it is important now to consider how, if economies are to be expanded, that might be done in a way that makes them dependent upon perhaps unreliable or at least very expensive future oil supplies. This is an issue not normally considered within conventional economics as the present market price of oil makes it a seemingly attractive choice. But we feel it important to go beyond that mentality. As of this writing (July 2017), there have been both large price increases and declines recently in the price of oil, although correcting for inflation it is often still higher than in the past decades. One of our colleagues in Great Britain said that he felt he was standing on the shore of the North Sea and although the storm had not hit yet the first large waves were starting to roll in. In other words the price increases that we have observed recently are only a small sign of what lies ahead as the world truly approaches the end of cheap oil. What this will mean for the world can only be guessed at, but for the non-oil-producing nations of the developing world, the impact is likely to be enormous as populations and economies that had expanded based on cheap oil have the rug pulled out from under them. It is unlikely to be a pretty sight.

21.8 Predicting Land Use Change

An important part of many assessments of the future capacity of a nation or a region for providing economic or environmental services is an assessment of how much land is available in different categories (this is loosely related to the concept of ecological footprint). The principle tools for doing this are several computer models that start with one map of land use for a given year and then make assessments of what the land use might be in the future based on rates and patterns of development. Both rates and patterns tend to be derived from existing patterns that can be extracted digitally from one or more existing maps of land use. One of our favorite models for doing this, not surprisingly, is one that we derived ourselves. This model, called GEOMOD, is bundled with the most recent version of IDRISI, a commercial software package with powerful modules for assessing and predicting patterns of land use [1, 16, 21, 22].

One might start with, for example, a map of the forested vs. non-forested region of Costa Rica, as we did in our original analysis. It had been our experience based on looking out airplane windows while flying over many regions of the tropics, especially the hilly or mountainous tropics, that development tended to start along rivers, often at lower elevations, and then work progressively up stream and upslope over time, with the development usually proceeding from one already developed place to an adjacent forested one. This is consistent with the idea that farmers will develop land in a way that represents the least effort or energy investment on their part (hence adjacent properties on flatter land) with the highest potential for agricultural production (usually soils near a river on flatter land). Our first assessments used a DEM (digital elevation model) to represent topography, with originally the land represented as a checkerboard of 1 kilometer by 1 kilometer cells, each of which was assigned a one for forest and a two for deforested land. We would provide GEOMOD with an initial or start-up map, with the areas developed or deforested for a particular region represented by one and the original forested area represented by two (or with more categories, another number, or color when displayed). We then used a search window to search row by row and column by column for cells that had already been developed, meanwhile examining the nine (or sometimes more) cells around each developed cell using a process called "adjacency" as one criteria for which cells are likely to be developed. If there was a non-developed (forested) cell next to a developed cell, then we had an "edge," and that forested cell was a candidate for development. This was done for the entire map, meanwhile keeping track of the elevation and/or slope for each candidate cell. Then enough of the lowest elevation (and/or flattest) cells were developed to meet the proscribed rate of development expected for that time step (usually 1 year). Over time this process will result in the spread of development upstream and upslope, simulating a basic pattern by which humans use land. The final project will be maps of human use of land into the future.

It is worth mentioning that it is imperative to reexamine, on a regular basis, our assumptions on farmer's decision-making rules. This typically involves interviews and surveys in the field. Often we find that what we thought initially was wrong, even if it seemed perfectly logical. For example, the main cabbage production areas in Nepal are rocky high altitude slopes, classified as "not suitable for agriculture" by western planners.

Over time variants of GEOMOD have been developed (see website of Gil Pontius at Clark University) that can use many different properties of the environment (e.g., distance from roads or cities, soil types, and so on) that give the option of undertaking much more sophisticated assessments and predictions of land use change. There are a number of good chapters in [1, 16], and that use GIS and related spatial analysis techniques to examine geographical aspects of development and development possibilities, often while paying especial attention to scale issues. All of these chapters show the incredible role that geographical analyses linked with computers now play in virtually every aspect of examining development issues.

21.9 Predicting Net Economic Output as a Function of Land Type

All land does not have the same capacity for economic production, and this is especially true when specific uses are examined. For example, only about 19 percent of the total land area of Costa Rica was flat and fertile enough to be utilized for any use, including specifically row crop agriculture, which would be likely to cause irreparable damage if applied to other land categories (in other words if the land was too steep, then erosion would destroy the potential for production in a relatively short time). Another 9 percent of the land was suitable for pastures and another 16 percent for tree crops such as coffee, which causes less erosion because of its continuous cover. The rest of the country, more than 56 percent, should have no human use at all except for forestry that would maintain tree cover. In fact as of about 1990, far more than 56 percent of the country has been developed for agriculture, pastures, or urban areas. More recently much of this steep land has been reverting to forest as the futility of its economic development is increasingly clear.

Farmers and many other humans are well aware of what land is best to use for various purposes and tend to use the best land first, as is represented by the farmer's choices given in the above example for GEOMOD. Thus over time the land available for development tended to be of poorer and poorer quality, as represented, for example, in the pioneering work of David Ricardo. What this means for development is that average values of, e.g., crop production cannot be used to project what the yields might be for some development project. For example, coffee can be grown anywhere in Costa Rica. But high-quality coffee, of which Costa Rica has some of the best, requires very explicit environmental conditions (e.g., precipitation, temperature, soil, and so on) to get high yields, which tend also to mean best quality coffee beans. We found that for Costa Rica as of 1990 nearly all of the land that was best for growing coffee already had it growing there (or was covered by urbanized areas) and that if there were to be increased coffee production yields would probably be less or else more energy intensive than the average of what was occurring already. This is an example of what has been called, variously, diminishing returns to investment or declining (energy or other) return on energy investment as the best resources are used We found, quite remarkably, that for most crops any increase in area of land cropped would produce an instantaneous reduction in yield per hectare, as the land being used for production would be, on average, of lower quality. In any land use model, we have to make sure, however, that the decision rules that we put in the model are rules that farmers actually follow. This generally implies to run interviews and surveys in the field. One of the best ways to challenge and test our hypothesis is to go in the field and talk to farmers.

21.10 Assessing the Energy and Other Cost of a Development Scheme

If there is an economic plan for development, then the next step is to assess the energy, material, and other resource requirements for such a project. While this can be an extremely difficult and comprehensive issue and there is not yet a clear-cut formula for how to undertake it as we discussed above, a recent computer program derived to examine the material costs of any development project, one that we think is very good, is being developed in Italy by Sergio Ulgaldi and others at the University of Siena, Italy. Thus if we have a list of, e.g., materials required for a development project, then we can assess the most important aspects of their use rather straightforwardly. The user simply puts in dollar amounts to be spent for different development categories according to the spreadsheet provided, and the results are then printed out, a far cry from the old days of undertaking such calculations by hand as we used to do.

21.11 Include Social Assessments

As we stated in step one, many of the issues that are most important to people interested in development are of course social and economic in nature. There is no easy formula for integrating the biophysical and the socio-economic approaches, although much can be undertaken with an open mind, a willingness to work outside of one's own discipline, and, perhaps most useful, an ability to find and work with others from other disciplines. It is worth noting persistent attempts by economists to put a dollar value to "social capital," just like they do with the environment (an enterprise that we believe is seriously flawed and doomed to fail). Many of the chapters in [1, 16] are especially good at attempting to integrate biophysical and socio-economic approaches, and it is almost impossible to list specific chapters as most do in fact integrate both sciences.

21.12 Construct a Comprehensive Simulation of the Future and Make the Right Decisions!

A final step in undertaking a thorough assessment of the biophysical possibilities and constraints of a region is to examine alternative future environments in which one's decisions might be played out. Prospective analysis plays a fundamental role in shaping the development of a country. However it is poorly done at best, policy makers having to juggle with too many parameters and being forced to use shortcuts, which opens the door to misconceptions and prejudice, wrong interpretation of the data, and shortsighted emergency measures. In *The Art of the Long View*, Swartz [23] describes the critical role of scenario analysis for positioning ourselves properly into the future. Scenarios are not predictions or forecasting: they are "vehicles for helping people to learn, alternative images of the future, to change the managerial view of reality."

At the core of prospective analysis, one can easily imagine an environment to run and discuss comprehensive simulations of the future, e.g., based on the previous three steps. It can contain some or all of the entities included above plus whatever other elements the user feels appropriate, including elements of neoclassical economic analysis, and the results can be compared or by the right person even integrated! Again our example of this approach is given in the CD that is included in [1]. We believe especially in the development of good graphics and real-time simulations for communication to stakeholders and hold up the above CD as an example. Although many people are extremely suspicious of any such simulation models, we think that formalizing one's knowledge and assumptions through modeling is a critical approach that needs to be undertaken much more with the decision-makers of the developing countries in the future.

We must also face the fact that whatever good we might be able to do with the approach that we advocate can be undermined, like anything else, by the corruption and unresponsiveness of government in much of the developing (and developed) world. We have no magic solution to this either, although we are confident in the positive impact of a neutral and transparent scientific approach. But the main problem that we scientists face is that we are not very good at communicating our results to the public and therefore we have limited influence on the decisions that affect our society. This is where good computer graphics showing to the general populace the past and projected future aspects of their economies and environments as a function of whatever policies are implemented can be key. In fact we believe that if well done political debates about the future might be carried out with the aid of good computer simulations and visualizations shown on national television! We often think while watching political debates that it would be very interesting if the promises of the candidates were subject to modeling reality checks (i.e., testing of politicians hypotheses) to see what was in fact possible and at what cost! Beaulieu in [16] gives one example of a fairly successful application of science to politics.

21.13 Make the Right Decisions

Most people who are involved with such a comprehensive analysis are interested in implementing the results in what is normally called policy. Of course that can be an extremely difficult process, but if you have worked with the right people from the start, it will be possible to actually make better decisions. So it is important to involve decision-makers from the beginning. From them (and ideally from the general effected populace as well) the scientist or economist can get a much clearer idea of desired ends (which might be quite different from what the scientist or economist assumes). In turn the decision-maker can learn to have a systemic, longer-term perspective for their country.

"Hybrid" forums where scientist and citizens meet and exchange views are ideal for social-technical debates and the education of each. Again the use of dynamic graphs that can convey to the user possible futures as a function of policy today can be very useful. Finally with the new insights gained from the entire process given above, reexamine if and where conventional economics has failed and propose amendments to neoclassical economics-based policy or develop an entire new perspective based on the analyses we have given above. It is a big charge to develop an entirely new economics, but we think it critical, and what we have here is a formal start. And of course throughout the entire process of undertaking biophysical economic assessments and plans, the scientific method must be used, theories need to be advanced in a way consistent with first principles, hypotheses need to be generated and tested, and so on. The final arbitrator of the correctness of our analyses is not whether this or that theory is the basis for our efforts but whether our predictions and policy prescriptions come to pass. This closes the loop on what is our basic wish: to bring the scientific method to our development economics.

21

? Questions

1. Explain some virtues of the process of visualization of model output (as was done, e.g., for Costa Rica).
2. Distinguish among "environmental economics," "ecological economics," and "biophysical economics."
3. What is an ecological footprint? How does that relate to biophysical economics?
4. What is emergy analysis? How does it differ from energy analysis?
5. Give one example where biophysical economic, footprint, and energy analyses give substantially the same answer.
6. Give five steps that can be followed in developing a biophysical analysis.
7. How can social, political, and economic elements be incorporated into a biophysical analysis?
8. What kinds of issues might one want to gather data on in a biophysical assessment?
9. What is a simple way of translating a simple growth rate into a doubling time? For example, the United States had 300 million people growing at 1 percent a

year in about the year 2000. If this 1 percent a year growth rate continues, when would the United States have 600 million people? How old would you be then if you were still alive?
10. What are time series data? How do they help us to understand biophysical economics?
11. What kind of corrections need to be made for raw economic data (e.g., GDP) when examining data over time?
12. What is the Gini index? How does that help to put a more nuanced perspective on, for example, GDP data?
13. What are a few important considerations in how imports, exports, and their difference might influence our economic policies?
14. How does a prediction of land use change understand possible economic possibilities? How does that relate to land quality?
15. What are some of the pitfalls that await even the best possible plan that one might develop? How can citizen involvement assist in that process?

References

1. Hall, C.A.S. (editor, Gregoire Leclerc and Carlos Leon, associate Editers). 2000. *Quantifying sustainable development: The future of topical Economies*. Academic Press, San Diego.
2. We refer the reader to the following web sites as good examples of other definitions of ecological economics and how ecological economics can actually be done. http://www.wordiq.com/definition/Ecological_economics; http://www.fs.fed.us/eco/s21pre.htm; http://www.anzsee.org/ANZSEE8.html; https://www.reference.com/science/can-test-ecological-footprint-bf9a91a2c19a45bd?aq=footprint+calculator&qo=cdpArticles
3. World-wide Fund for Nature International (WWF). 2004. *Living Planet report 2004*. Global Footprint Network, UNEP World Conservation Monitoring Centre, WWF, Gland Switzerland.
4. Herendeen, R., and C. Bullard. 1975. The energy costs of goods and services. 1963 and 1967. *Energy Policy* 3: 268.
5. Bullard, C.W., B. Hannon, and R.A. Herendeen. 1975. *Energy flow through the US economy*. Urbana: University of Illinois Press.
6. Hannon, B. 1981. *Analysis of the energy cost of economic activities: 1963 2000. Energy research group doc. No. 316*. Urbana: University of Illinois.
7. Hall, C.A.S., C.J. Cleveland, and R. Kaufmann. 1984. *Energy and resource quality: The ecology of the economic process*. New York: Wiley Interscience. 577 pp. (Second Edition. University Press of Colorado).
8. Cleveland, C.J. 2004. The encyclopedia of energy. Elsevier (www.Carnegie-Mellon).
9. Odum, H.T. 1996. *Environmental accounting emergy and environmental decision making*. New York: John Wiley & Sons.
10. Brown, M. 2004. Energy quality, emergy, transformity: The contributions of H.T. Odum to quantifying and understanding systems. In *Through the macroscope: The legacy of H.T*, eds. M. Brown and C.A. S. Hall. Odum. Ecological Modelling, Elsevier Amsterdam, special issue: Vol. 178.
11. Brown, M., and R. Herendeen. 1996. Embodied energy analysis and emergy analysis: A review. *Ecological Economics: A Comparative Review* 19: 219–236.
12. Herendeen, R. 2004. Energy analysis and emergy anlysis – a comparison. In *Through the macroscope: The legacy of H.T,* Odum, M. Brown and C.A.S. Hall eds. Ecological Modelling, special issue: Volume 178: 227–238.
13. Brown, M., M. Wackernagel, and C.A.S. Hall. 2000. Comparable estimates of sustainability: Economic, resource base, ecological footprint and Emergy. In *Quantifying sustainable development: The future of topical economies*, ed. C.A.S. Hall, 695–714. San Diego: Academic Press.
14. Foran, B., M. Lenzen, C. Dey, and M. Bilek. 2005. Integrating sustainable chain management with triple bottom line accounting I Integrating sustainable chain management with triple bottom line accounting. *Ecological Economics* 52: 143–157. 5.
15. Detwiler, P., and C.A.S. Hall. 1988. Tropical forests and the global carbon cycle. *Science* 239: 42–47.
16. LeClerc, G., and Charles Hall, eds. 2008. *Making development work: A new role for science*. Albuquerque: University of New Mexico Press.
17. Kroeger, T., and D. Montanye. 2000. An assessment of the effectiveness of structural adjustment policies in Costa Rica. Chpt. 24. In *Quantifying sustainable development: The future of topical Economies*, ed. C.A.S. Hall. San Diego: Academic Press.
18. Barreteau, O, M. Antona, P. d'Aquino, S. Aubert, S. Boissau, F. Bousquet, W. Daré, M. Etienne, C. Le Page, R. Mathevet, G. Trébuil, J. Weber. 2003. *Journal of Artificial Societies and Social Simulation* 6(1) (http://jasss.soc.surrey.ac.uk/6/2/1.html).
19. Hallock, J., P. Tharakan, C. Hall, M. Jefferson, and W. Wu. 2004. Forecasting the availability and diversity of the geography of oil supplies. *Energy* 30: 2017–2201; Hallock Jr., J. L., W. Wu, C. A.S. Hall, M. Jefferson. 2014. Forecasting the limits to the availability and diversity of global conventional oil supply: Validation. Energy 64: 130–153.
20. Ahmed, N. 2017. *Failing states, collapsing systems: Biophysical triggers of political violence*. Springer, New York.
21. Hall, C.A.S., H. Tian, Y. Qi, G. Pontius, and J. Cornell. 1995. Modeling spatial and temporal patterns of tropical land use change. *Journal of Biogeography* 22: 753–757.
22. Pontius, R.G., Jr., J. Cornell, and C.A.S. Hall. 1995. Modeling the spatial pattern of land-use change with GEOMOD2: Application and validation for Costa Rica. *Agriculture, Ecosystems & Environment* 85: 191–203.
23. Swartz, P. 1996. *The art of the long view: Planning for the future in an uncertain world*, 272p. New York: Doubleday.

Understanding How Real-World Economies Work

Much of what traditional economics believes "works" because of clever technology, substitutions, and intelligent investments, in fact, do so only because we have had huge amounts of cheap energy to throw at the problem. But if indeed we are at or approaching "the end of cheap oil" and "the second half of the age of oil," not to mention serious climate disruption, we need a new way to think about how we do economics. Economic growth, which solved many issues in the past, now is declining and even coming to a halt in much of the world, associated with a similar slow down in the availability and use of energy. These concepts also apply to a very much broader suite of the basic resources and environmental conditions required to fuel our economy. While many people are taught and believe that technology has made natural resources increasingly irrelevant, this book contains a great deal of evidence to show the contrary. Our national and global society is becoming more, not less, dependent upon natural resources, as fossil fuels, for example, underlie essentially everything we do economically, including building their "renewable" replacements. Additionally, many of the things that are treated as *externalities* in conventional economics, that is as supposedly secondary issues not properly included in prices, are instead what we believe to be often the *main* issues of economics. Depletion of highest quality fuels is one such issue. More generally, understanding and protecting the basic systems of the Earth, such as the atmosphere, far from being a luxury or an "externality" as is indicated in conventional economic analysis, are *the* critical issues for economics. This section gives a number of applications of biophysical economics to these important contemporary issues.

Contents

Peak Oil, Secular Stagnation, and the Quest for Sustainability

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_22

22.1 Introduction [1]

As we write and rewrite these final chapters, in 2010 and then 2017, the US national economy continues to struggle. There has been little inflation-corrected growth of the economy for more than a decade, wages for most Americans remain stubbornly low. The stock market crashed during the Great Financial Crisis, beginning in the summer of 2007. By 2009, the Dow Jones Industrial Average was down from its then-historic high of 14, 198 points to a low as 8000, barely half the peak of the preceding fall. However, the stock indexes have subsequently recovered, with the post-crisis period surpassing all previous records. As of November 10, 2017, the Dow Jones Industrial Average stood at 23,422.21, an all-time high. This boom in stock values, coupled with wage stagnation, has been a primary driver of the inequality mentioned in Section III. In 2014, economists Emmanuel Saez and Thomas Piketty found that during the recovery, from 2009 to 2012, that the top 1% of income earners received 95% of all the gains in income [2]. More than half of this increase was from capital income and capital gains. Since stock prices have increased another 58%, there is no reason to believe that the share of the top 1% of income earners has decreased. Many of our states are facing severe budgetary problems, schools and colleges everywhere are facing severe budget shortfalls. Political promises of the left and the right are increasingly viewed with suspicion or hostility. The rich get richer, and the poor get poorer.

The most abrupt change in our economy began in the summer of 2008 with the highest oil prices ever (almost $150 a barrel) and historically high prices for other energy and most raw materials. The Dow Jones Industrial average was down from its then historic high of 14,198 to as low as 8000, barely half the peak the preceding fall. Each week the stock market lost 5% or 10% of its value. A series of disasters struck the financial markets, with many of the largest, most prestigious, and seemingly impervious companies declaring bankruptcy, by the end of November. Many investors lost from one-third to one-half of the value of their stocks. Since then the financial markets have recovered, but the growth of the real economy has been tepid, at best. Europe and Japan have continued to grow very slowly, if at all, a situation called "secular stagnation." Few understand the role of energy in either secular stagnation or as a driving force in the financial explosion. In earlier years, periods of financial excess would occur at the end of boom periods of economic growth. However, since the 1970s, financial speculation showed marked increases even in times of slow growth or recession. Mainstream economics tends to view the rise of the financial sector and speculation as a drain on the economy, as investment in the real economy (factories, mines, oil wells) is displaced by purely financial investments in paper claims on real assets. Yet, if one believes, as do we, that the normal state of a monopolized economy is towards slow growth or stagnation, then profit expectations in the real economy decline with growing excess capacity. Money channeled into finance would not necessarily be invested in the real economy. It may not be invested at all, but held as cash in corporate coffers. Perhaps financial speculation is one of the few things that is keeping the economy growing at a tepid 2%, rather than experiencing permanent recession or depression. After the financial crash, the nation's central bank, or the Federal Reserve, flooded the economy with liquidity to avoid another great depression. Most of this money flowed into the financial sector, propping up stock prices [3].

Fewer still understand the underlying role of energy. The North Sea, once the source of enormous amounts of oil for the United Kingdom and Norway, has declined greatly. Europe is again beholden to Russia and the Middle East for its economic lifeblood. The summer of 2017 also saw the twelfth year in a row in which the global production of conventional oil essentially did not rise (although there was a modest increase in "all liquids", often reported as "oil", driven by an increase in natural gas liquids) leading some to say that the long predicted "peak oil," the time of maximum global oil production, had indeed arrived. Total energy use in the United States had not increased for almost a decade. The use of oil went down by about eight percent since its peak in early 2008. World conventional oil production has essentially been flat. It is not quite clear whether this is a good sign of decreasing use of CO_2-emitting fossil fuels or a sign that our economies are beginning to be in real trouble. Meanwhile, populations and their aspirations continue to increase relentlessly in much of the "less-developed" world, especially India and China.

From the point of view of mainstream economics and business executives, economic growth is the most important of all goals. Most policies are justified to the degree that their proponents

22

say they will produce economic growth. But if the energy needed to drive economic growth is in decline, and the concentrated economy produces slow growth on its own terms, then perhaps few, if any, policies can produce economic growth. Slow growth is likely to become "the new normal," for both biophysical and internal economic reasons. A group to which we belong believes that the world has entered a new mode, one that was predicted in many ways in the 1960s and 1970s by some geologists, ecologists, and economists. This is a world of limits, one in which our once-trusted tools of conventional economics are no longer sufficient by themselves, if indeed they ever were, of righting economic wrongs and allowing us all to maximize our material well-being. While there is no question that under the auspices of conventional economics, many parts of the Western world, and increasingly Asia, have done very well in increasing human material well-being, the perspective that we raise is whether our growth in wealth has been due to really understanding our economies or, as we believe, more to simply our increased ability to pull more cheap oil, gas, and coal out of the ground to allow the increased economic work that is the basis of our wealth. To some extent, any set of theories about economics in the past was bound to be at least partly correct because with more and more energy it was possible to generate more and more wealth, whatever one's theoretical premises!

22.2 What Is the Source of the Crash of 2008?

Many factors merged to cause the financial crash of 2008—the subprime mortgage crisis, high foreclosure rates, and Wall Street's sale of opaque financial products known as derivatives. Behind these are many aspects of greed, corruption, and malfeasance, not to mention the moral hazard caused by lax political oversight. It is not the intent of this book to focus on the personalities and moral shortcomings behind these issues, but we believe one good and detailed summary of much of this can be found at ▶ http://www.informationclearinghouse.info/article28189.htm. While we do not wish to downplay these "moral" issues, we also believe that the root cause of the current downturn and our difficulty in climbing out of the recession was the same one that sparked four out of the last five world recessions: the high price of oil [4]. Why did most economists and financial analysts (and models like the Wharton model) not see this coming? One hypothesis, advanced by Nobel laureate Paul Krugman is that the economics profession "went astray because economists, as a group, mistook beauty, clad in impressive-looking mathematics, for truth" [5]. We agree. As the market debacle has shown, mathematical elegance in economics is not a substitute for scientific rigor, something we have discussed in many previous papers [6, 7], and in chapter 20. If physical quantity of energy and its effect on energy prices are crucial functions impacting the economy, and they are not in our models, then of what utility are the models?

As of this writing, global production of conventional oil has been nearly flat since 2005, so that peak oil, or at least a cessation of reliable growth at the former rate of two to four percent per year, appears to have occurred—with the remaining debate only about whether there may be a subsequent peak and how soon we begin a slide down the other side, even given the temporary respite from hydraulic fracturing. If we have passed the global peak in oil production, then indeed the end of cheap oil will soon be upon us, and our ability to grow or even maintain economies is likely to decrease. Because of the critical importance of liquid and gaseous petroleum for essentially everything we do, we have serious reservations as to whether conventional economics and business or governmental policies can guide us again to growth or indeed to manage an economy where growth is no longer possible (e.g., ◘ Fig. 4.10). Thus the question becomes: "Can we improve upon our ability to do economics and financial analysis by using procedures that focus more on the energy available (or not) to undertake the activity in question?" In other words, are finances beholden to the laws of physics?

We think yes. Thus the question becomes: can we supplement or improve upon our ability to do economics? Resource scientists have predicted such a financial crash, or more accurately cessation of growth, for a long time [7–11]. Any good physical or biological scientist knows that all activity in nature—or anywhere—is associated with energy use. Consequently, many in the scientific community were not the slightest bit surprised by the financial crash or its timing. Colin Campbell, a former oil geologist and cofounder of the Association for the Study of Peak Oil, predicted in 2006 that we are likely to see an end of year after year economic growth and a movement to an "undulating plateau" in oil production, prices, and economic activity, with periodic high prices in oil-generating financial stress and a cessation or even reduction of growth. These financial strains would, in turn, cause a decrease in oil use and hence a price decline, with lower oil prices then leading to new economic growth and new increases in oil use and, eventually, oil prices. In other words, he foresaw very large impacts of restrictions in oil availability, and consequent price increases, on the market. According to Campbell, "Every single company on the stock market is overvalued from the perspective of what the cost of running that company will be after peak. Value is determined by performance which has been based on cheap oil." This approach has been used to develop a model by Murphy and Hall [9] which seems to be a pretty good predictor of the present situation.

Many other analysts have remarked upon, and even predicted, the probable impact of peak oil, or at least oil price increases, on the financial status of the United States and the world. A thoughtful, chilling, and ultimately correct view of the implications of peak oil on the American economy was presented by Gail Tverberg in January 2008 on the energy blog site "The Oil Drum" [10]. Her predictions, which we thought impossibly pessimistic at the time, have been vindicated in great detail. Many analysts foresaw these issues as early as the 1960s, including the authors of the famous but cavalierly dismissed "Limits to Growth" study of 1972, ecologists Garrett Hardin and Howard Odum, economists Kenneth Boulding, Paul Baran, Paul Sweezy, Nicholas Georgescu-Roegen, John Bellamy Foster and others. But for those who bothered to read and think about what these authors were saying, the future is clear. Charles Hall made his retirement decisions in 1970 based on the assumption that peak oil and a crash of stocks would occur in about 2008 [11]. The reason is that all of these people understood that—of necessity—real growth is based on growth in real resources, and that there are limits to those resources. The case for peak oil was clearly laid out almost 60 years ago by Hubbert [12, 13] who predicted, in 1955 that the US peak in oil production would occur in 1970, which it did. The United States has struggled to exceed the 1970 value in the intervening half century but has not done so as of November 2017 (see Fig. 7.5) and still imports nearly half the oil it uses.

While many economists place a great deal of faith in increasing technology, in fact technology does not operate on a static playing field but continually competes with declining resource quality. There is little or no evidence that technology is winning this game over time because the energy return on investment keeps falling [14–17]. It is important to understand that, at least so far, the Limits to Growth model is an almost perfect predictor of our current situation [18]. Resource-based analysts understand and appreciate that the recent turmoil in much of our financial structure has many plausible causes. But they also know energy underlies all of these issues. The fundamental dilemma is this: if oil, the most important energy source to fuel the economy, goes through the inevitable path of growth, plateau, and eventual decline (i.e., peak oil) while the financial market is built on the assumption of unfettered growth, then something has to give. Eventually the aspirations and assumptions of indefinite growth in assets, production, and consumption must collide with the reality of an ever-constricted source of the energy that fuels real growth.

Part of the financial stress is attributable to cheap oil that then becomes dear. Starting in the early 1990s, relatively inexpensive oil, declining interest rates, and globalization all contributed to economic growth and to declines in risk premiums for virtually all asset classes. Capital went further out on the risk curve to make up for reduced returns and increased leverage (that is, a reduction in "money in the vault" relative to what was loaned) became the new norm. As volatility seemed to disappear, even more leverage was piled on to the system. Along with the changing landscape in global credit markets came cheap financing for US home buyers. The low price of energy also greatly increased discretionary income which further encouraged people to take advantage of this cheap financing, adding to massive residential development. According to financial analyst George Soros this created a self-reinforcing "reflexive" system, where increasing home values increased collateral, which encouraged further borrowing in the household sector and in lines of credit for consumption and so on [19]. The system had been built on the premise that large amounts of discretionary spending would always be available and the notion that everyone was entitled to a McMansion, a "lawyer foyer," and a home theater. Since the construction of homes far outpaced population growth, most of the growth was due to the perceived demand for these larger houses. To get the area needed, we had to build out from the cities. The largest growth in real estate had been in the exurban areas, which were most vulnerable to gas price spikes.

Discretionary wealth—that which is available for nonessential investments and purchases—is extremely sensitive to volatile energy prices [21]. Since most oil use is not discretionary but needed for getting to or undertaking work, it is relatively price inelastic, that is the response of consumers is not particularly sensitive to changes in price. Consequently, discretionary income dropped substantially when gasoline and other energy prices, which had been creeping up from a very low level in 1998, increased sharply in 2007–2008. The United States reached a "tipping point" in 2006–2008 [20] as the price of oil rose temporarily to nearly $150 a barrel. The assumption that the suburban lifestyle would be sustainable became a question in many potential owner's mind. This perception appeared to be an important initiator of a decline in aggregate demand, particularly for exurban real estate. It also may have initiated the massive de-leveraging initiated we are now experiencing globally. (There is a good summary of the various analyses by Rubin, Hamilton, and others who argue that oil price increases were behind these, and past, recessions [22–23]). Massive household debt could not be supported when the value of the underlying collateral declined: a decline triggered by the spike in energy prices. As the collateral disappeared, huge derivative positions that had been built in the previous decade experienced margin calls. A spiral of forced selling pressured all asset classes further, and forced the banking sector essentially to freeze in September of 2008. Will this faltering of the suburban model be a preview of our ultimate response to peak oil? Perhaps. Examining the general pattern of oil price increases and probability of them continuing can help us understand these things better in the longer term.

22.3 Energy Price Shocks and the Economy

At the start of 1973, oil was cheap at $3.50 a barrel. The United States was still the world's largest producer. Peak oil had just occurred in the United States in 1970, but no one noticed. The economy kept growing, fueled by increasing oil imports. As domestic oil production in the United States declined from 1970 to 1973, foreign suppliers gained leverage. In late 1973, both political events that precipitated the Arab Oil Embargo and an accident that severed an export oil pipe in the Middle East caused the price of oil to jump from 3 to 12 dollars a barrel. In a matter of months, these events created the largest recession since the Great Depression. The price spike had at least four immediate effects upon and within the economy: (1) oil consumption declined, (2) a large proportion of capital stocks and existing technology became too expensive to use, (3) the marginal cost of production increased for nearly every manufactured good, and (4) the cost of transportation fuels increased.

By 1979, the price of oil had increased by a factor of 10, to $35 a barrel. The proportion of gross domestic product that went to buying energy increased from 6% to 8% to 14%, restricting discretionary spending while causing previously unseen "stagflation". The prices of other energies, and commodities more generally, increased at nearly the same rate, driven in part by the price increase of the oil that was behind all economic activities. Then, in the 1980s, all around the world, oil that had been found but not developed (as it had not been worth much previously) suddenly became profitable, and it was developed and overdeveloped. By the 1990s, the world was awash in oil, and the real price fell to nearly what it was in 1973. The energy portion of GDP fell to about 6%, essentially giving everyone an extra 8% of their incomes to play with. The impact on discretionary income, perhaps a quarter of the total, was enormous. Many invested in the stock market, but then found themselves victims of the "tech bubble" of 2000, as excess capacity began to build in the technology sector. Real estate was considered a "safe" bet, so many invested in what was really surplus square footage. Speculation became rampant as real estate became valued for its financial returns rather than as a place to live. For a while, it seemed as if investment in real estate was a sure path to wealth. As we now recognize, most of that increase in wealth was illusory. With energy price increases from 2000 to the summer of 2008, an extra 5% to 10% "tax" from increased energy prices was added to our economy as it had been in the 1970s, and much of the surplus wealth disappeared. Speculation in real estate was no longer desirable or possible as consumers tightened their belts because of higher energy costs. Then the housing market crashed.

While this energy perspective is not a sufficient explanation for all that has happened, the similar economic patterns in response to the energy price increases of both the 1970s and of the last decade give the "energy trigger" considerable credibility. In systems theory language, the endogenous aspects of the economy that the economists focus on (Fed rates, money supply, etc.) became beholden to the exogenous forcing functions of oil supply and pricing that are not part of economists' usual framework.

22.4 The Relation of Oil and Energy More Generally, to Our Economy

While economics is overwhelmingly taught as a social science, in fact, our economy is completely dependent upon the physical supply and flow of resources. Specifically, our economy is overwhelmingly dependent upon oil, which supplied about 40% of US energy use in the 2000s, followed by natural gas and coal at about 25% each, and nuclear at a little less than 5%. Hydropower and firewood supply no more than 4% each. Wind turbines, photovoltaics, and other "new solar" technologies together account for less than 2% (although that percentage may be increasing). Global percentages are similar. Our economy has been based on increased use of fossil fuels for most of its growth. Until 2008, we added much more new capacity with fossil fuels than with new solar, which has added a bit to the total use rather than displaced fossil fuels. Since 2008, growth in both energy and the economy has been very slow (Fig. 13.5), and the remaining economic activity is still based on about the same energy mix, although in the US gas is displacing coal.

Because of the enormous interdependency of our economy, there is not a huge difference in the energy requirements for the various goods and services that we produce. A dollar spent for most final demand goods and services uses very roughly the same amount of energy no matter what the good or service is. An exception is money spent for energy itself, which includes the chemical energy plus another ten or so percent which is the energy needed to get it (i.e., the embodied energy). For 2017 an average dollar spent in the economy required about 5 megajoules for that activity. Money spent for chemicals such as paint might use 12, but for most final demand goods and services the number is nearer to the mean. For heavy construction in the petroleum industry, the estimate is about 11 MJs per dollar and for very heavy industry such as obtaining oil and gas about 16 MJs per dollar. Year by year less energy is used per dollar, due mostly to inflation but also increasing efficiency, especially as the economy turned from goods to services and manufacturing moved overseas. There continues to be decreasing energy return on energy invested (EROI) for our major fuels as we go after ever more difficult resources [15–17].

22.5 Energy and the Stock Market

We include here some preliminary analyses that we think show the importance of energy to Wall Street and the economy more generally. First, Wall Street prices reflect not only something about the real operation of the economy but also a large psychological factor often called "confidence." Our hypothesis is that the energy used by the economy is in some sense a proxy for the amount of real work done. Thus over time, the inflation-corrected Dow Jones Industrial Average (DJIA), an index of financial speculation about the potential future profits of top industrial corporations, should have the same basic slope as the use of energy in society. It should also "snake" around the real amount of work done, reflecting issues of confidence, speculation, and so on. Over sufficient time, however, the DJIA must return approximately to the real energy use line. To test this hypothesis, we plotted the DJIA from 1915 until 2008 along with the actual use of energy by the US economy. Our hypothesis would be supported if the slope of these two lines are similar over the longer time period. In fact from 1915 until 2010, the DJIA had the same basic slope as the use of energy, and it has greater variability, consistent with our hypothesis (Fig. 7.8). We hypothesize that the Dow Jones will, over the long run, continue to snake about the total energy use in response to periods of irrational exuberance and the converse. If US total energy use continues to stagnate or decrease, as it has for the last decade, this hypothesis implies no sustained real growth for the Dow Jones. Investors and analysts should question whether any speculative boom can continue indefinitely. Failure to assess critically this possibility was a factor in both the financial panic that preceded the Great Depression, and the in Great Financial Crisis of 2008-09.

In the past, we also hypothesize that the amount of wealth generated by the US economy should be closely related to fuel energy use. Cleveland et al. found that the gross national product of the United States was highly correlated with quality-corrected energy use from 1904 to 1984 ($R^2 = 0.94$) [24]. This high correlation appeared to be much poorer for the period 1984 until 2008, a period during which inflation-corrected GDP doubled while energy increased by only a third. It is possible that the divergence is due not to increasing efficiency but rather an increasing proclivity of governments to underreport on inflation (see the online group ► shadowstatistics.com). Correcting for this, if indeed that is needed, would make the relation of energy use and GDP growth much tighter through the 1990s and 2000s. Also, it is very clear that much of U.S. heavy industry has been moved overseas, although we still import the products.

22.6 A Financial Analyst Concurs

Jeff Rubin, Chief Economist at CIBC World Markets, wrote in a recent book that defaulting mortgages are only one symptom of the high oil prices [22]. Higher oil prices caused Japan and the European Nations to enter into a recession even before the most recent financial problems hit. According to Rubin: oil shocks create global recessions by transferring billions of dollars of income from economies where consumers spend every cent they have, and then some, to economies that sport the highest savings rates in the world. While those petro-dollars may get recycled back to Wall Street by sovereign wealth fund investments, they don't all get recycled back into world demand. The leakage, as income is transferred to countries with savings rates as high as 50%, is what makes this income transfer far from demand neutral. By any benchmark, the economic cost of the recent rise in oil prices is nothing short of staggering. The oil impact is much more staggering than the impact of plunging housing prices on housing starts and construction jobs, which, according to the press, has been the most obvious brake on economic growth from the housing market crash. And those energy costs, unlike the massive asset write downs associated with the housing market crash, were borne largely by Main Street, not Wall Street, in both America and throughout the world. This big increase in oil prices has caused the annual fuel bill of OECD countries to increase by more than $700 billion a year, with $400 billion of this going to OPEC countries. Rubin asks: "Transfers a fraction of today's size caused world recessions in the past. Why shouldn't they today?" We and others believe that there is ample evidence that our economy is beholden to energy supplies and prices, and that good investors and good economists need to learn a great deal more about energy. This is one reason why we are attempting to tackle this problem head on through the development biophysical economics. But getting the economists to rethink their intellectual training will be a tough job, no matter how much that is needed [23].

22.7 Is Growth Still Possible?

There was little inflation-corrected growth of the US economy or in its use of energy from 2004 through about 2015. Is this just part of the normal business cycle or something new? Numerous mainstream theories have been posited over the past century that have attempted to explain business cycles. Each offers a unique explanation for the causes of—and solutions to—recessions, including Keynesian Theory, the Monetarist Model, the Rational Expectations Model, Real Business Cycle Models, NeoKeynesian models, etc. Yet, for all the differences among these theories, they all share one implicit assumption: a return to a growing economy is both desirable and possible, i.e., GDP can grow indefinitely. Historically, the US economy has grown at rather slow average rate of 1.9% per year since the Civil War. Some decades, such as the 1890s or the

1930s, showed profound declines. However, the decades after the Second World War until the 1970s showed sustained growth. Economists began to take the unique postwar phenomenon as normal. Of course, economic growth in the era preceding the use of fossil fuels was less than 1%. But if we are entering the era of peak oil, then for the first time in history we may be asked to grow the economy while simultaneously decreasing oil consumption, something that has yet to occur in the United States for 100 years. Oil more than any other energy source is vital to today's economies because of its ubiquitous application as transportation fuel, as a portable and flexible energy carrier and as feedstocks for manufacturing and industrial production. Historically, spikes in the price of oil have been the proximate cause of most recessions [4]. On the other hand, expansionary periods tend to be associated with the opposite oil signature: prolonged periods of relatively low oil prices that increase aggregate demand and lower marginal production costs, all leading to, or at least associated with, economic growth. This has happened (modestly) from 2008 to 2017.

By extension, for the economy to sustain real growth there must be an increase in the flow of net energy (and materials). Quite simply economic production is a work process and work requires energy. Thus to increase production over time, i.e., to grow the economy, we must either increase the energy supply or increase the efficiency with which we use our source energy. This is called the energy-based theory of economic growth. This logic is an extension of the laws of thermodynamics, which state that: (1) energy cannot be created nor destroyed, and (2) energy is degraded during any work process so that the initial inventory of energy can do less work as time passes. As Daly and Farley [26] describe, the first law places a theoretical limit on the supply of goods and services that the economy can provide, and the second law sets a limit on the practical availability of matter and energy. In other words, to produce goods and services energy must be used, and once this energy is used it is degraded to a point where it can no longer be reused to power the same process again.

22.8 An Energy-based Theory of Economic Growth

This energy-based theory of economic growth is supported by data: the consumption of every major energy source has increased with GDP since the mid-1800s at essentially the same rate that the economy has expanded (Figs. 22.1 and 22.2). Throughout this growth period, however, there have been numerous oscillations between periods of growth and recessions. Recessions are defined by the Bureau of Economic Research as "a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales" [27]. From 1970 until

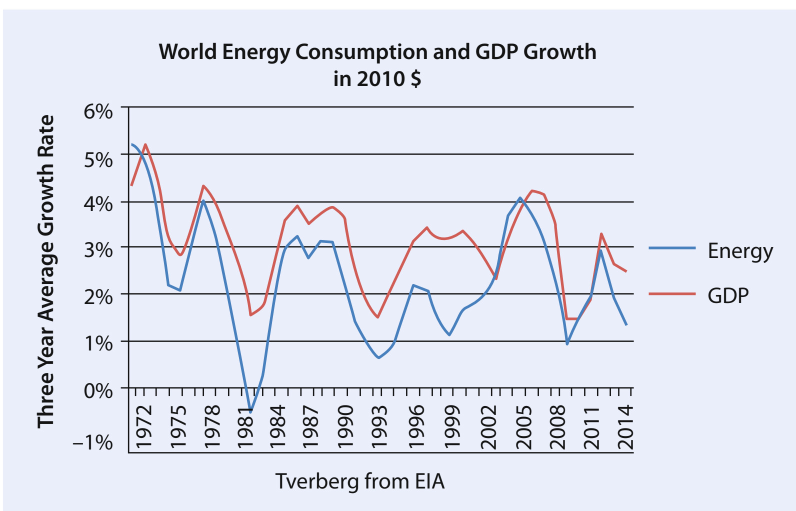


Fig. 22.1 Correlation of year on year (YoY) changes in oil consumption with YoY changes in real GDP, for the United States from 1970 through 2008 (Source: Gail Tverberg. Oil consumption data from the BP Statistical Review and real GDP data from the St. Louis Federal Reserve)

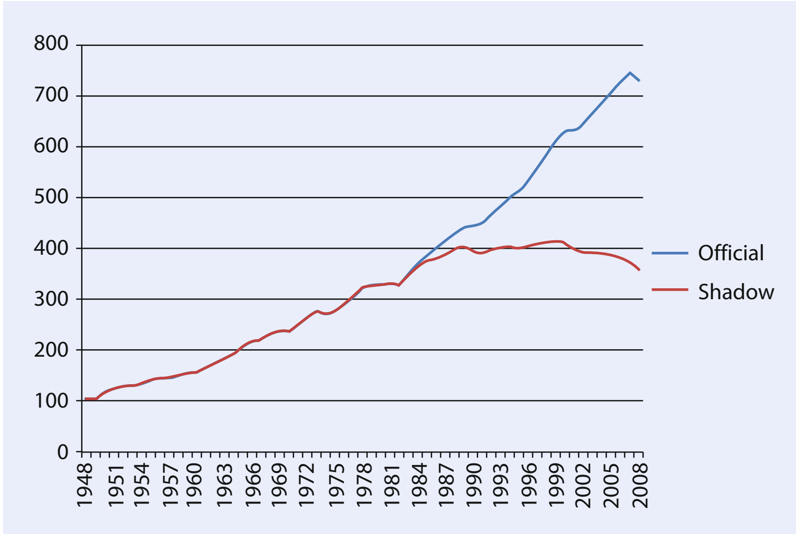


Fig. 22.2 One attempt to correct the GDP for the "deflated" inflation factor by using the inflation corrections year by year since 1984 supplied by the group shadowstatistics. If larger inflation estimates are used, the economy has grown very little since 1984, and there may have been no improvement in efficiency which is how energy is changed to GDP (Source: Hannes Kunz) (see also: ▶ http://www.leap2020.eu/the-true-us-gdp-is-30-lower-than-official-figures_a5732.html)

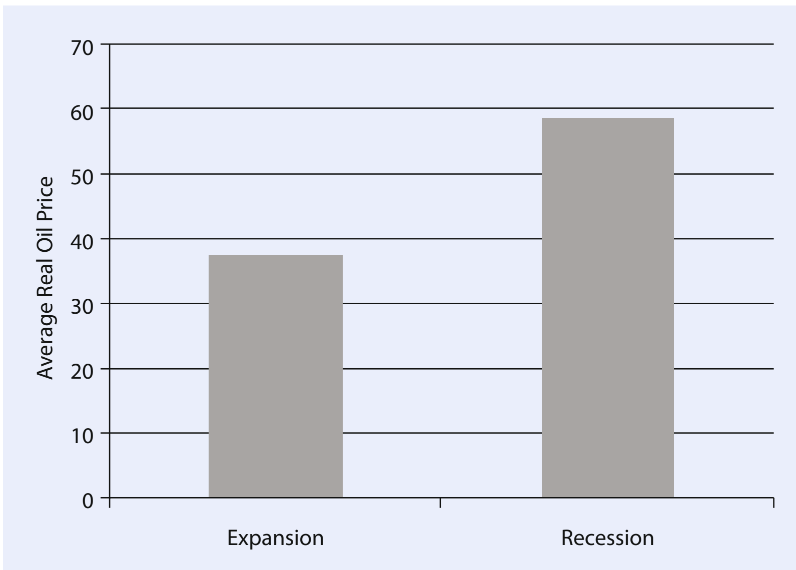


Fig. 22.3 Real oil prices averaged over expansionary and recessionary periods from 1970 through 2008

2007, there have been five recessions in the United States. Examining these recessions from an energy perspective elucidates a common mechanism underlying each recession: oil prices are lower and oil consumption increases during periods of economic expansion while oil consumption decreases and oil prices are higher during recessions (Fig. 22.3). Oil price increases precede essentially all recent recessions.

Plotting the year on year (YoY) growth rates of oil consumption and real GDP provides a more explicit illustration of the relation between economic growth and oil consumption (Fig. 22.1). But correlation is not causation, and an important question is whether increasing oil consumption causes economic growth, or conversely, whether economic growth causes increases in oil consumption [28]. Cleveland et al. [29] analyzed the impact of these two factors on the causal relation between energy consumption and economic growth. Their results indicated that increases in energy consumption caused economic growth, especially when they adjusted the data for quality and accounted for substitution. Other subsequent analyses that adjusted for energy quality support the hypothesis that energy consumption causes economic growth, not the converse [30]. In sum, our analysis indicates that about 50% of the changes in economic growth over the past 40 years are explained, at least in the statistical sense, by the changes in oil consumption alone. In addition, the work by Cleveland et al. [29] indicates that changes in oil consumption cause changes in economic growth. These two points support the idea that energy consumption, and oil consumption in particular, is of the utmost importance for economic growth.

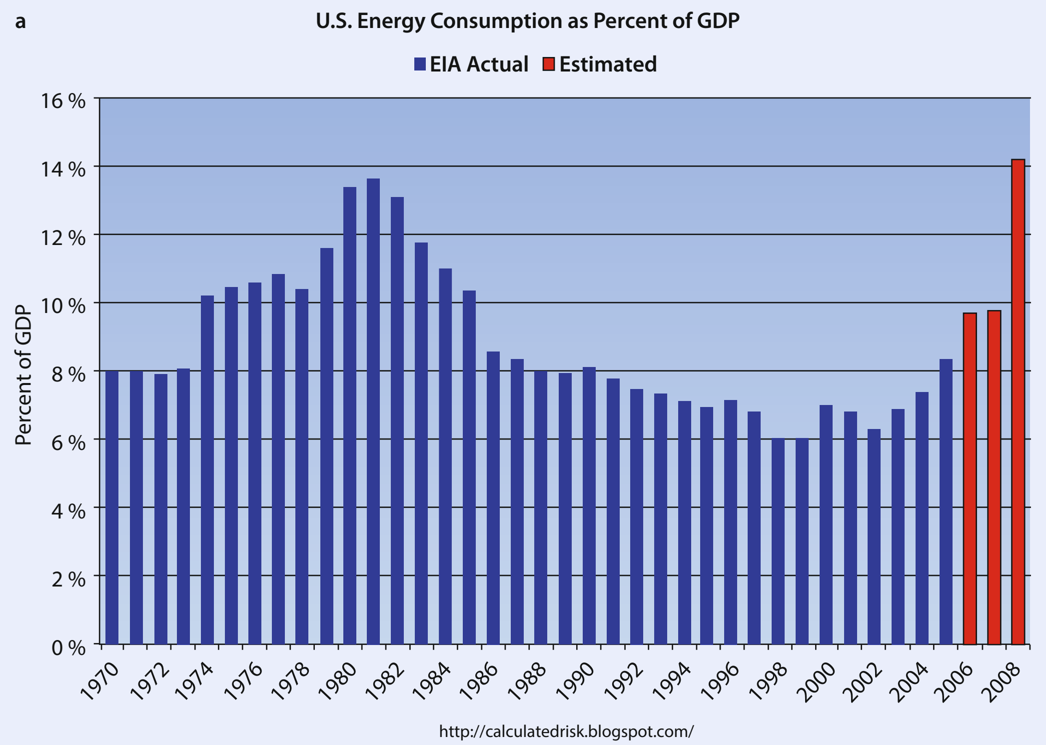
Yet changes in oil or energy consumption are rarely used by neoclassical economists as a means of explaining economic growth. For example, Knoop [31] describes the 1973 recession in terms of high oil prices, high unemployment, and inflation yet omits mentioning that oil consumption declined four percent during the first year and two percent during the second year. Later in the same description, Knoop claims that the emergence from this recession in 1975 was due to a decrease in both the price of oil and inflation, and an increase in money supply. To be sure, these factors contributed to the economic expansion in 1975, but what is omitted, again, is the simple fact that lower oil prices led to increased oil consumption and hence greater physical economic output. Oil is treated by economists as a commodity, but in fact it is a more fundamental factor of production than either capital or labor. Thus we again present the hypothesis that higher oil prices and lower oil consumption are both precursors to, and indicative of, recessions. Likewise, economic growth requires lower oil prices and simultaneously an increasing oil supply. The data support these hypotheses: the inflation-adjusted price of oil averaged across all expansionary years from 1970 to 2008 was $37 per barrel compared to $58 per barrel averaged across recessionary years, whereas oil consumption grew by two percent per year on average during expansionary years compared to decreasing by three percent per year during recessionary years (Figs. 22.1 and 22.3).

Although this analysis of recessions and expansions may seem like simple economics, i.e., high prices lead to low demand and low prices lead to high demand, the exact mechanism connecting energy, economic growth, and business cycles is rather more complicated. Hall et al. [21] and Murphy and Hall [9, 32] report that when energy prices increase, expenditures are reallocated from areas that had previously added to GDP, mainly discretionary consumption, toward simply paying for the more expensive energy. In this way, higher energy prices lead to recessions by diverting money from the general economy toward energy only. The data show that recessions occur when oil expenditures as a percent of GDP climb above a threshold of roughly 5.5%, or, stated somewhat differently, when all energy becomes more than 12 percent of the economy (Fig. 22.4).

22.9 Predicting Future Economic Expansion

Each time the US economy emerged from a recession over the past 40 years there was an increase in the use of oil even while a low oil price was maintained. Unfortunately oil is a finite resource. What are the implications for future economic growth if following a recession: (1) oil supplies are unable to increase with demand or (2) oil supplies increase but at an increased price? To undertake this inquiry, we must examine first the current and probable future status of oil supply; then we can make inferences about what the future of oil supply and price may mean for economic growth.

Since oil consumption causes change in economic growth, understanding how both peak oil and net energy will impact oil supply and price is important to understanding the ability of our economy to grow in the future. To that end, we review both the theory and current status of peak oil and net energy as they pertain to oil supply, and then discuss how both of these may influence oil price. Optimists about future oil availability usually start with the correct observation that there is a great deal of oil left in the Earth, probably three to ten times what we have extracted, and, usually, with the assumption that future technology driven by market signals will get much of that oil out. There are at least two problems with that view. The first is that of "peak oil." It is clear we have, or soon will, reach a physical limit in our ability to pump more oil out of the ground. For a long time, oil production grew at three to four percent a year. Now there has been little or no growth in global oil production since 2004. The second problem is that the oil left in the ground will require an increasing quantity of energy to



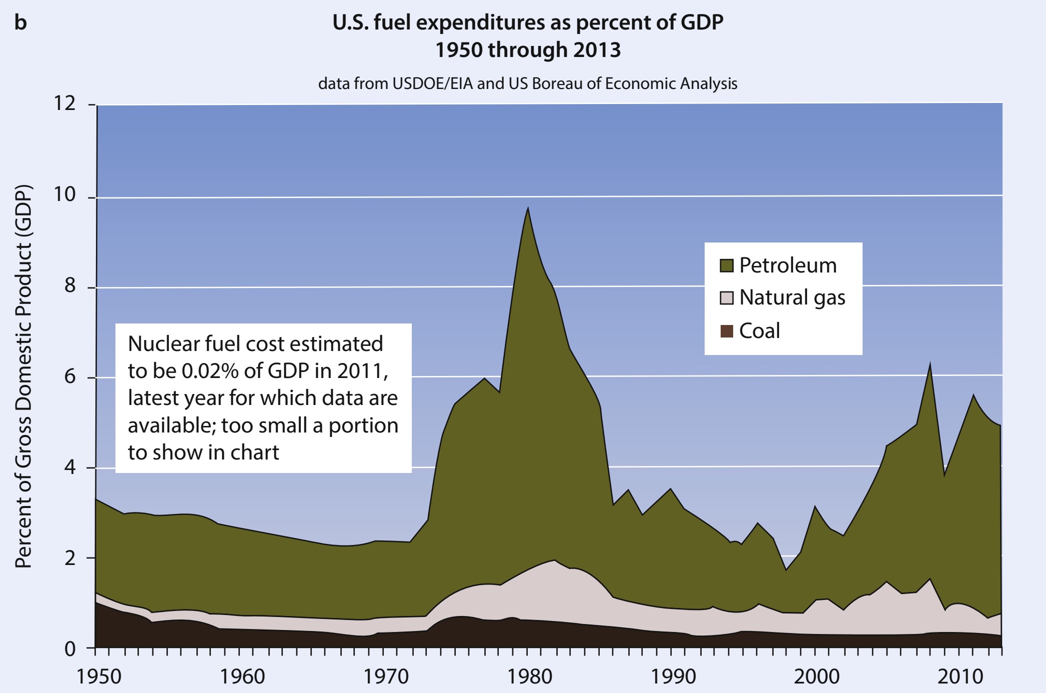


Fig. 22.4 Two estimates of the cost of fuel as a percent of U.S. GDP. **a** The threshold above which the economy moves toward recessions is about 10 to 12 percent. **b** A second, more conservative estimate

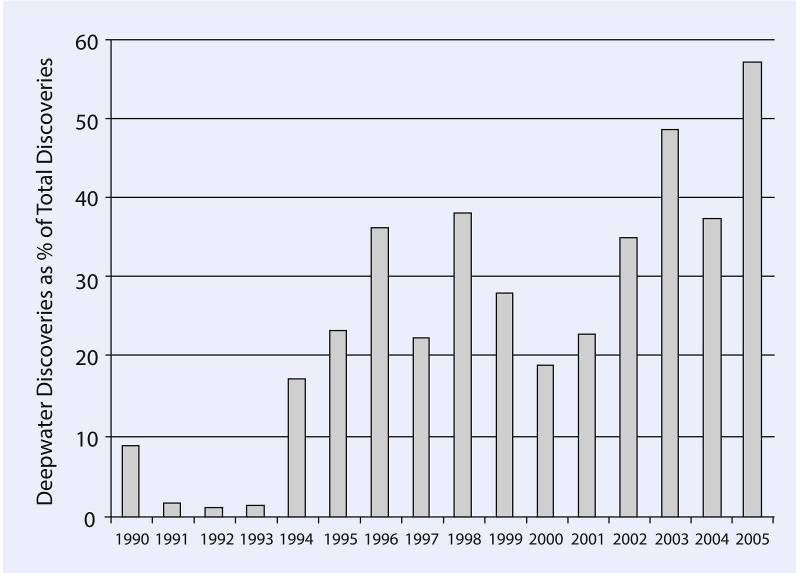


Fig. 22.5 Deepwater oil discoveries as a percent of total discoveries from 1990 through 2005 (Source: Jackson 2009)

extract, at some point as much as is in the oil. There is a clear trend that the EROI of oil production is declining in each region for which data are available. This shows that depletion is more important than technical advances. Gagnon et al. [16] report that the EROI for global oil extraction declined from about 36:1 in the 1990s to 18:1 in 2008. This downward trend results from at least two factors: first, increasingly supplies of oil must come from sources that are inherently more energy intensive to produce, simply because firms have developed cheaper resources before expensive ones. For example, in 1990 only two percent of discoveries were located in ultra-deepwater locations, but by 2005 this number was 60 percent, (Fig. 22.5). Second, enhanced oil recovery techniques, such as the injection of steam or gases are being implemented increasingly. For example, nitrogen injection was initiated in the once super-giant Cantarell field in Mexico in 2000, which boosted production for 4 years, but since 2004 production from the field has declined precipitously. Although enhanced oil recovery techniques increase production in the short term, they also increase significantly the energy inputs to production, offsetting much of the energy gain for society. Thus it seems that additional oil is unlikely to be available, and if so it will have a low EROI and hence high price.

Forecasting the price of oil, however, is a difficult endeavor as oil price depends, in theory, on the demand as well as the supply of oil. Following the economic "crash" of 2008 most OECD economies around the world have been contracting or at least not growing. Thus the flat rate of oil production since 2004 did not cause a huge sustained increase in the price of oil. One thing we can do with some accuracy is to examine the cost of production of various sources of oil to calculate the price at which different types of oil resources become economical (Fig. 22.6). We can then estimate how much oil would be available at a given price. If the price of oil is below the cost of production, then most producers of that oil will cease operation. If we examine the cost of production in the areas in which we are currently discovering oil, hence the areas that will provide the future supplies of oil, we can calculate a theoretical floor price below which an increase in oil supply is unlikely.

Roughly 60% of the oil discoveries in 2005 were in deepwater locations (Fig. 22.5). Based on estimates from Cambridge Energy Research Associates [33], the cost of developing that oil is between \$60 and \$85 per barrel, depending on the specific deepwater province. Therefore, oil prices must exceed roughly \$60 to \$90 per barrel to support the development of even the best deepwater resources. These data indicates that an expensive oil future is necessary if we are to expand our total use of oil, that is, to grow economically. But these prices will discourage that very growth (Fig. 22.6). Indeed, it may be difficult in the future even to produce the remaining oil resources at prices the economy can afford. As

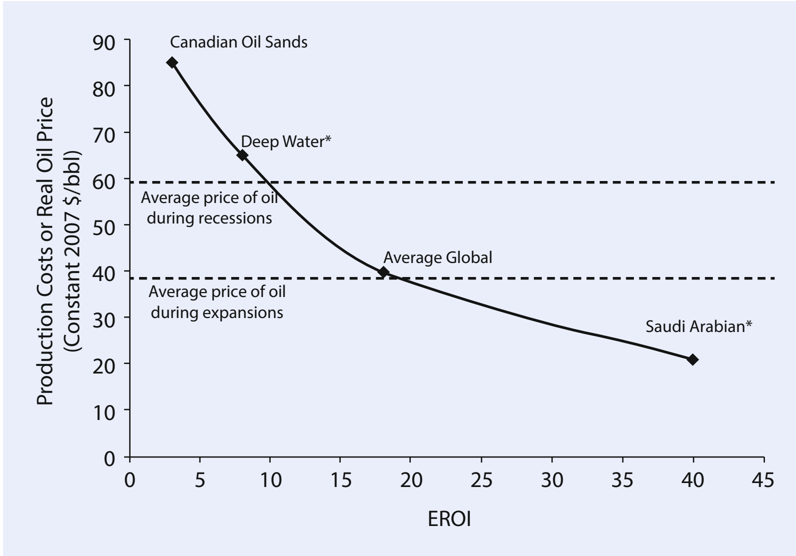


Fig. 22.6 Oil production costs from various sources as a function of the EROI of those sources. The dotted lines represent the real oil price averaged over both recessions and expansions during the period from 1970 through 2008. Data on EROI from Murphy and Hall [32], Gagnon et al. [24], and the data on the cost of production comes from CERA [33], *asterisks* are educated guesses

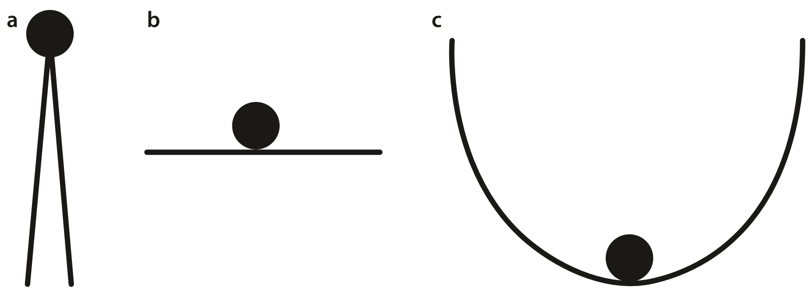


Fig. 22.7 Three types of equilibrium: unstable **a**, neutral **b**, and stable **c**. The third situation seems to represent what we face in the world today

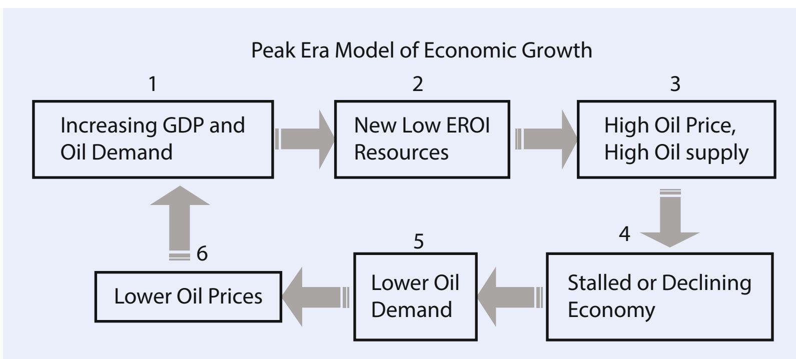


Fig. 22.8 Peak oil era model of the economy. Cycle of relation of economic grown (or recession) and oil prices

a consequence, the economic growth witnessed by the United States and globe over the past 40 years may be a thing of the past.

One way to think about this situation is to borrow a concept from systems theory. A very general concept is that many systems seek an equilibrium point because there are dynamic forces that resist change. An example is a marble in a bowl (Figs. 22.7 and 22.8). The marble seeks its equilibrium position at the bottom of the bowl. One can push the marble up the side with your finger, but the marble easily slips off your finger and goes back to the equilibrium position. This might represent the situation our economy is in now, kept at a more or less constant GDP by growth being discouraged by rapidly increasing oil prices at levels of consumption barely above where we are now, but maintained from further shrinking by decreased oil prices with contraction—indeed this is a recipe for a steady state economy.

22.10 EROI and Prices of Fuels

Since EROI is a measure of the efficiency with which we use energy to extract energy resources from the environment, it can be used as a proxy to estimate generally whether the cost of production of a particular resource will be high or low, or perhaps even energy costs themselves [34]. For example, production from Canadian oil sands have an EROI of roughly 4:1, whereas the production of conventional crude oil has an average EROI of about 10–20:1 and Saudi crude much higher. The production costs for oil sands are roughly $85 per barrel compared to roughly 60 dollars for average US oil and $20 per barrel for Saudi Arabian conventional crude. Thus there is an inverse relation between EROI and price, indicating that high EROI resources are generally relatively inexpensive to develop and low EROI resources are generally more expensive to develop (Fig. 22.6). As oil production continues, we can expect to move further toward the upper left of that picture. We see no evidence that technology has lowered EROI even as it extends our resources. In summary, relatively low EROI appears to translate directly into higher oil prices, so that if we have to move to lower EROI oil in the future the price is likely to be higher, restricting economic activity and growth [35]. At the time of this writing it is not known whether renewable energies such as photovoltaic or wind turbine electricity can replace a substantial portion of fossil fuels.

22.11 Summary

The main conclusions to draw from this discussion are:

1. Over the past 40 years, economic growth has required increasing oil consumption;
2. The supply of high EROI oil cannot increase beyond current levels for any prolonged period of time;
3. The average global EROI of oil production will almost certainly continue to decline as we search for new sources of oil in the only places we have left: deep water, arctic, and other hostile environments;
4. We have globally no more than 20–30 years of conventional oil remaining at anything like current rates of consumption and anything like current EROIs, and less if oil consumption increases and/or EROI decreases;
5. Increasing oil supply in the future will require a higher oil price because mostly only low EROI, high-cost resources remain to be discovered or exploited;
6. Developing these higher-cost resources is likely to cause economic contraction as oil costs exceed five and total energy costs exceed ten percent of GDP;
7. Using oil-based economic growth as a solution to recessions is untenable in the long-term, as both the gross and net supplies of oil have, or will soon, begin, at some point, an irreversible decline.

A similar assessment could be developed for other energy resources.

This growth paradox leads to a highly volatile economy that oscillates frequently between expansion and contraction periods, and as a result, there may be numerous peaks in economic activity and in oil production but little trend. In terms of business cycles, the main difference between the pre- and peak oil era is that business cycles appear as oscillations around an increasing trend in the pre-peak era but as oscillations around a flat trend following the peak. For the economy of the United States and most other growth-based economies, the prospects for future, oil-based economic growth are bleak, and we do not have another model that would allow for growth. It seems clear that the economic growth of the past 40 years will not continue for the next 40. A resolution to these problems can occur if economic growth was no longer the goal. Society must begin to emphasize energy conservation over growth and adjust our population numbers, jobs, living patterns, and aspirations accordingly.

Questions

1. What events of 2008–2011 might be construed as indicating some limits to the three or four percent per year growth that the United States and much of the world had previously expected? These new limits may or may not be related to biophysical limitations. How would you assess this situation?

2. Have events since the publication of this book in late 2011 changed your answer to the previous question?
3. What was the main reason that Nobel Prize economist Paul Krugman put forth for the market crash in 2008?
4. Do you think that finances are beholden to the laws of physics? Why or why not?
5. What is the relation of an "undulating plateau" to peak oil?
6. Can you discuss financial "leverage" with respect to energy and other resources?
7. Discuss some of the financial issues that were related to the "oil crises" of the 1970s.
8. Do you think that price gives signals as to the future availability of energy? Why or why not?
9. If energy supplies are indeed restricted is economic growth still possible? What would be the requirements for that?
10. What is the relation historically between the price of energy and discretionary spending?
11. What has been the relation between the amount of oil that is consumed in a given year and the price of that oil? What might be a reason for that?
12. As the EROI for a given source of oil declines how does that relate to its price?
13. How might we best respond to a future of limited oil supplies should it occur, which seems likely?
14. Due to the depletion of high EROI oil the economic model for the peak era, i.e., roughly 1970–2020, is much different when viewed as net rather than gross energy from oil. Why is that?

References

1. Derived in part from: Hall, C.A.S., and A. Groat. 2010. Energy price increases and the 2008 financial crash: A practice run for what's to come? *The Corporate Examiner* 37: 19–30. and Hall, C.A.S., and D.J. Murphy. Adjusting to the new energy realities in the second half of the age of oil. *Ecological Modeling* 223: 67–71.
2. Saez, E. 2014. Income and wealth inequality: Evidence and policy implications. Neubauer Collegium Lecture. University of Chicago.
3. Magdoff, H., and P. Sweezy. 1987. *Stagnation and the financial explosion*. New York: Monthly Review Press.
4. Hamilton, J. 2009. Causes and consequences of the oil shock of 2007–08. *Brookings Papers on Economic Activity, Spring* 2009: 215–225.
5. Krugman, P. 2009. How did economists get it so wrong? *New York Times* Sept 2.
6. Hall, 2001, C.A.S. Hall, D. Lindenberger, R. Kummel, T. Kroeger, and W. Eichhorn. 2001. The need to reintegrate the natural sciences with economics. *Bioscience* 51: 663–673.
7. Gowdy, J., C. Hall, K. Klitgaard, and L. Krall. 2010. The End of faith based economics. *The Corporate Examiner* 37: 5–11. Odum, H.T. 1973. Energy, ecology, and economics. *Ambio* 2: 220–227; Huang, J., R. Masulis, and H. Stoll. 1996. Energy shocks and financial markets. *Journal of Futures Markets* 16: 1–27.
8. Deffeyes, K. 2001. *Hubbert's peak: The impending world oil shortage*. Princeton: Princeton University Press. Campbell, C., and J. Laherrere. 1998. The end of cheap oil. *Scientific American* 278: 78–83.
9. Murphy, D.J., and C.A.S. Hall. 2011. Energy return on investment, peak oil, and the end of economic growth. Annals of the New York Academy of Sciences. *Special Issue on Ecological Economics* 1219: 52–72.
10. Tverberg, G. 2008. Peak oil and the financial markets. A forecast for 2008. The oil Drum. http://www.theoildrum.com/node/3382
11. Hall, C. 2004. The Myth of sustainable development: Personal reflections on energy, its relation to neoclassical economics, and Stanley Jevons. *Journal of Energy Resources Technology* 126: 86–89.
12. Hubbert, M.K. 1969. *Energy resources*. In the National Academy of Sciences–National Research Council, Committee on Resources and Man: A Study and Recommendations. W. H. Freeman. San Francisco.
13. ———. 1974. *U.S. Energy Resources: A review as of 1972*. Background paper prepared for U.S. Senate Subcommittee on Interior and Insular affairs, 93d Congress, 2nd Session. Serial 93–94 (92–75). U.S. Government Printing Office, Washington.
14. Hall, C.A.S., and C.J. Cleveland. 1981. Petroleum drilling and production in the United States: Yield per effort and net energy analysis. *Science* 211: 576–579.
15. Cleveland, C.J. 2005. Net energy from the extraction of oil and gas in the United States. *Energy: The International Journal* 30 (5): 769–782.
16. Gagnon, N.C., A.S. Hall, and L. Brinker. 2009. A preliminary investigation of energy return on energy investment for global oil and gas production. *Energies* 2 (3): 490–503.
17. Guilford, M., C.A.S. Hall, P. O'Conner, and C.J. Cleveland. 2011. A new long term assessment of EROI for U.S. oil and gas: *Sustainability*: Special Issue on EROI. Pages 1866–1887.
18. Hall, C.A.S., and J.W. Day Jr. 2008. Revisiting the limits to growth after peak oil. *American Scientist* 97: 230–237.
19. Soros, G. 1987. *The Alchemy Furnace. Reading the mind of the market*. New York: Wiley.
20. Gladwell, M. 2000. The tipping point: How little things can make a big difference. New York: Little Brown Co.
21. Hall, C.A.S., R. Powers and W. Schoenberg. 2008. Peak Oil, EROI, Investments and the Economy in an Uncertain

Future. In *Biofuels, Solar and Wind as Renewable Energy Systems: Benefits and Risks*, ed. D. Pimentel. Springer. The Netherlands.

22. Rubin, J. 2008. Just how big is Cleveland. CIBC World markets. http://research.wibcwm.com/economic_public/download/soct08.pdf
23. Sauter, R. and S. Awerbuch. 2003. Oil price volatility and economic activity: A survey and literature review. IEA Research Paper IEA, Paris August 2003.
24. Cleveland, C.J., R. Costanza, C.A.S. Hall, and R. Kaufmann. 1984. Energy and the United States economy: A biophysical perspective. *Science* 225: 890–897.
25. Rubin, J. 2012. The end of growth. Random House Canada Toronto; Aucott, M. and Hall C.A.S. 2014. Does a Change in Price of Fuel Effect Affect GDP Growth? An Examination of the U.S. Data from 1950–2013. Energies 7: 6558-6570.
26. Daley, H., and J. Farley. 2004. *Ecological economics: Principles and applications*. Washington, DC: Island Press.
27. NBER. 2010. *US business cycle expansions and contractions*. Washington, DC: National Bureau of Economic Research.
28. Karanfil, F. 2009. How many times again will we examine the energy-income nexus using a limited range of traditional econometric tools? *Energy Policy* 37: 1191–1194.
29. Cleveland, C., R. Kaufmann, and D. Stern. 2000. Aggregation and the role of energy in the economy. *Ecological Economics* 32: 301–317.
30. Stern, D. 2000. A multivariate cointegration analysis of the role of energy in the US macroeconomy. *Energy Economics* 22: 267–283.
31. Knoop, T.A. 2010. *Recessions and depressions: Understanding business cycles*. New York: Praeger.
32. Murphy, D.J., and C.A.S. Hall. 2010. Year in review – EROI or energy return on (energy) invested. *New York Annals of Science* 1185: 102–118.
33. CERA. 2008. *Ratcheting down: oil and the global credit crisis*. Cambridge, Massachusetts: Cambridge Energy Research Associates.
34. King, C. and C. Hall. (2011). *Relating financial and energy return on investment*. Sustainability: Special Issue on EROI. Pages 1810–1832.
35. Jacobson, M. Z., Delucchi, M. A., Cameron, M. A. & Frew, B. A. Proc. Natl Acad. Sci. USA 112, 15060–15065 (2015). See also Clack, C. et al. Proc. Natl Acad. Sci. USA 114: 6722–6727 (2017). And J. Sole, A.García-Olivares, A.Turiel, J. Ballabrera-Poy (2018). Renewable transitions and the net energy from oil liquids: A scenarios study. *Renewable Energy* 116: 258-271.

Fossil Fuels, Planetary Boundaries, and the Earth System

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_23

23.1 Introduction

The decline in resource quality, measured in part by falling energy returns on investment, is a serious problem for our futures in and of itself. Access to fossil fuels has given many of us a far more comfortable life. Cellular phones are a ubiquitous possession, even among the world's poor. Fossil-derived energy has lifted the burden of heavy manual labor. Most Americans now work in offices and access electronic media at will. These feats would be impossible in the absence of electricity. Many people complain vociferously about how difficult commercial airline travel is these days with late flights, extra baggage charges, no food, and cramped seating. But imagine crossing the country in a Conestoga wagon. As an exercise, try thinking about the energy that is embodied in your day-to-day consumption patterns. Not surprisingly, people are reluctant to do without the goods they have acquired and become used to. When environmental educator Ray Bowdish, speaking at a recent symposium, asked his students what they could simply not give up, a frequent answer was "my truck!"

Fossil fuels have a direct effect on the economy, as we have seen in prior chapters. Gross domestic product has increased exponentially with a similar increase in fossil fuel use (Fig. 4.1). Since the peak of domestic conventional oil production in 1970, every spike in oil prices has been followed by a recession in the US economy. Murphy and Hall [1] enunciated the growth paradox. Maintaining business-as-usual economic growth requires new sources of oil or other very high quality energy. Since the only remaining sources require higher prices to bring them forth, this helps reduce the potential for economic growth. As a result, the economy exhibits a high degree of volatility. Petroleum geologist Colin Campbell called this an undulating plateau. As we saw in the last chapter, declining energy returns on investment in the long term are a factor in long-term slow growth or secular stagnation, as well as in cyclical variations.

This portends a future in which we should not assume automatically that our children and grandchildren will be better off than are we. But if this were our only problem, we could manage, for a while, by turning to coal that is still abundant and possesses a high energy return on investment relative to many alternative fuels. Perhaps we will be able to use the remaining fossil fuels to power the transition to a solar economy that provides nearly the same amount of energy as we have access to now. Or maybe we can't do this. It is a distinct possibility that petroleum and coal are one-time gifts of nature that simply have no substitutes, much to the ignorance and eventual chagrin of mainstream economists. Maybe we will have to give up our trucks, as difficult as that may be.

Unfortunately, this is *not the only problem* we face regarding our energy future. Growth of the human economy and its social systems, driven by fossil fuel consumption, is overwhelming the proper functioning of the Earth systems. We are reaching the limits of what Herman Daly called the full world, as we have reached at least three of our planetary boundaries and are hurtling precipitously toward others. In many ways the human economy appears already too big for its supporting biophysical systems, and a system in overshoot cannot grow its way into sustainability [2].

This represents a challenge for our economy or at least as it presently exists. We have, clearly, a capitalist economic system. As we showed in ► Chap. 5, capitalism must grow. Periods when the economy does not grow are called stagnation, recession, and even depression and come complete with rising unemployment, shuttered businesses, poverty, and declining opportunities. Is this a "new normal" we must learn to live with? Everyone from the smallest entrepreneur to the chair of the largest multinational corporation will tell you about the growth imperative. Companies that do not meet their growth projections see their stock values decline. People who work for them see their salaries stagnate or see one another at the unemployment office and often seek some kind of a resurrection from extremists who promise, without evidence, a return to "the good old days." Yet if we are in overshoot already, then we must shrink in order to live within nature's change limits. Degrowth means getting smaller, and a steady-state economy means staying smaller permanently. A question we must ask is "how can biophysical economics provide the insights to help us, and generations to come, cope with or adapt to nature's new normal?"

23.2 A Systems Approach

Let us return momentarily to the biophysical economics model presented in ▶ Chap. 3 (◘ Fig. 3.3) that showed the economic process from the acquisition of solar energy to extraction, production, and consumption. At the very bottom, right-hand corner was a pile of fetid waste. This is the part of the system we must now consider. In addition, Herman Daly's model of the embedded economy (◘ Fig. 3.2) showed the sun's energy creating sources of economic potential by means of photosynthesis, as well as showing human access to the ancient products of photosynthesis known as fossil fuels. On the right-hand side of the embedded economy model were the planet's sinks. The sinks, including the land, the oceans, and the atmosphere, are where the wastes of the human enterprise accumulate. If the growth of the open economic system confined within a finite and nongrowing ecosystem continues, we use more resources and create more effluents. If we put more waste into our sinks than they are able to assimilate, we suffer from myriad environmental problems, from litter to pollution to a climate change. But how do we know how much is too much? These are questions that can be answered by biophysical science. Only on the basis of solid science can we develop an economic policy to cope with our new system-level constraints.

Evidence continues to pile up regarding the overuse of our sinks. Landfills, which are basically holes in the ground, eventually fill up and close. Stories of garbage barges traveling the world looking for a place to dump their trash and frequently being rejected are the stuff of newspaper headlines. In the Finger Lakes area of New York, home to both the State University of New York College of Environmental Science and Forestry and Wells College, acrimonious local politics heat up periodically as to whether to keep the local landfill, affectionately known as Seneca Meadows, open to the trash of New York City for the revenues or to close it for issues of health, traffic congestion, and water quality. *The Washington Post* recently reported that one-third of plastics escape collection systems. In 2015 over eight million metric tons, or five full garbage bags per foot of coastline, ended up in the world's oceans. *The Post* cited a World Economic Forum study that predicted the mass of plastic would outweigh that of fish by 2050 [3]. Much of this plastic is concentrated into ocean vortexes known as gyres. The plastic floating in the North Pacific Gyre, to the northeast of Hawaii, is twice the size of Texas. The plastics break down into tiny parts where they contain large quantities of PCBs and DDT. As larger fish dine on the smaller ones who have ingested the plastic, the toxins bioaccumulate in the process made famous by Rachel Carson in *Silent Spring* [4]. In addition, the North Pacific garbage patch is but one of five. The first mate of a research vessel put it bluntly, saying that it was "just a reminder that there's nowhere that isn't affected by humanity" [5].

23.3 Planetary Boundaries

In 2009 a team headed by Johan Rockström of the Stockholm Resilience Centre [6] published a paper in the prestigious journal *Nature*, entitled "A Safe Operating Space for Humanity." The team, which included Nobel Prize-winning atmospheric chemist Paul Crutzen and climatologist James Hansen, identified nine "planetary boundaries" that are necessary to remain within, to assure the proper functioning of the Earth's biophysical systems. The list includes climate change, ocean acidification, biodiversity loss, stratospheric ozone depletion, and disruption of nitrogen and phosphorous cycles, among others (Table 23.1).

Rockström and his team calculated meticulously the preindustrial, or pre-fossil fuel era, baselines and proposed thresholds where the kinds of tipping points could occur that would lead to irreversible changes that could affect the entire Earth system. Before the age when coal was used to propel machinery by means of the steam engine in the late 1700s, the atmospheric concentration of carbon dioxide was 280 parts per million volume. Their proposed threshold to avoid a tipping point was 350 parts per million. Current concentrations exceed 400 parts per million. Regarding the climate, the team also measured radiative forcing. If you recall, the planet receives about 1400 Watts per every square meter that is in the sun, although about 30% bounces off the atmosphere. If you have ever seen the picture of "Earthrise" taken by the Apollo 11 astronauts,

Table 23.1 Planetary boundaries, preindustrial baselines, and current levels

Earth system process	Parameters	Proposed boundary	Current status	Preindustrial value
Climate change	1. Atmospheric carbon dioxide concentration (parts per million by volume)	350	387	280
	2. Change in radiative forcing (watts per meter squared)	1	1.5	0
Rate of biodiversity loss	Extinction rate (number of species per million species per year)	10	>100	0.1–1
Nitrogen cycle (part of a boundary with the phosphorus cycle)	Amount of N_2 removed from the atmosphere for human use (million of tonnes per year)	35	121	0
Phosphorus cycle (part of a boundary with the nitrogen cycle)	Quality of P flowing into the oceans (million of tonnes per year)	11	8.5–9.5	−1
Stratospheric ozone depletion	Concentration of ozone (Dobson unit)	276	283	290
Ocean acidification	Global mean saturation state of aragonite in surface sea water	2.75	2.90	3.44
Global freshwater use	Consumption of freshwater by humans (km^3 per year)	4000	2600	415
Change in land use	Percentage of global land cover converted to cropland	15	11.7	Low
Atmospheric aerosol loading	Overall particulate concentration in the atmosphere, on a regional basis	To be determined		
Chemical pollution	For example, amount emitted to, or concentration of persistent organic pollutants, plastics, endocrine disrupters, heavy metals, and nuclear waste in, the global environment or the effects on the ecosystem and functioning of Earth system thereof	To be determined		

what you saw was the light bouncing off the atmosphere. This is also known as albedo. The Earth's atmosphere is sensitive to very small changes. A slight human-induced change in the amount of radiation that the Earth receives can have a large effect on the planet's ability to trap heat. In pre-fossil fuel times, there was no human-induced radiative forcing. The proposed boundary is 1 Watts per square meter. The current threshold is 1.5. They measured biodiversity loss by the rate of extinction. The baseline figure was the loss of between 0.1 and 1 species per every million. The boundary is 10, while the current rate exceeds 100 species per 1 million. In other words, we are currently seeing the greatest mass extinction since the end of the age of the dinosaurs, also known as the Cretaceous-Tertiary boundary.

Human, fossil fuel-based activity is also disrupting our biogeochemical cycles. Liebig's law of the minimum states that growth of a system is limited, not by total resources, but by the most limited resource. For centuries, the resource limiting agricultural yields was nitrogen and phosphorous fertilizer. A shortage of manure led the British to scour the battlefields of the Napoleonic Wars to obtain the phosphorous-leaching bones of the fallen soldiers. In the 1900s the quest for nitrogen turned to the shores of South America for phosphorous-rich bird guano. But the removal of the ancient supplies outstripped the new droppings and the resource peaked. In the early twentieth century, German chemist Fritz Haber and chemical engineer Karl Bosch figured out how to make "bread from air" in the process described in ► Chap. 15. Since then the runoff from nitrogen and phosphorous fertilizers has been accumulating in hypoxic "dead zones" at the mouths of our major rivers, including the Mississippi. The extra nutrients cause algal blooms which consume all available oxygen upon their death, leaving the water body unable to support other forms of life. Before the invention of the Haber-Bosch process, people removed almost no nitrogen for human use from the atmosphere. The proposed safe threshold is 35 million metric tons per year. We currently remove 121 million metric tons per year. This creates a dilemma. Agricultural scientist Tim Wise of the Global Development and Environment Institute put the matter bluntly. "Ask any third-world farmer what sustainable agriculture is and they will tell you, more fertilizer." Yet the current level of nitrogen removal exceeds the proposed safe, or sustainable, level of by a factor of 3 and one-half times.

Human use of carbon, other species, and nitrogen has gone beyond the safe operating space of the planetary boundary already. Other categories are close to the edge. In preindustrial times, more phosphorous was returned to the ocean than taken out in fish harvests. The Rockström team calculates that a safe threshold would be 11 million metric tons of phosphorous flowing into the ocean. The current level is between 8.5 and 9.5 tonnes. The ocean is becoming more acidic as the carbon released into the atmosphere by the burning of fossil fuels is eventually sequestered in the ocean. Before the age of fossil fuel enabled industrial agriculture and suburban housing, humanity drew about 415 cubic kilometers of water from our aquifers. The Stockholm team estimates the safe level to be 4000 km^3, while we currently extract about 2600. Changes in the conversion of wild lands into cropland were miniscule in the days before industrialization. We now convert a little less than 12%. The proposed boundary is 15%, so we are about 80% of the way to exceeding the safe levels of land use. The impact of deforestation to feed the needs for food production, potential medicines, and forest products and beef for the developed world is continually pressuring the world's remaining tropical forests, and the debt-ridden governments of the often-poor nations in which the forests are located have a difficult time resisting the conversion of forest to land to resettle a burgeoning urban population and the need for foreign exchange.

The team has yet to calculate thresholds for chemical pollution and atmospheric aerosols. Aerosols are difficult to calculate, yet are an important component of climate science. Black aerosols, such as those produced by diesel engines, absorb sunlight and heat the planet. White aerosols, including sulfur dioxide that are emitted by coal-burning power plants, especially in areas of the world that lack stringent pollution-control laws, and volcanic eruptions, as well as nitrogen oxides from automobile tailpipes, reflect electromagnetic radiation and cool the planet. When Mount Pinatubo erupted in 1991, the oceans cooled for several subsequent years. If we underestimate the amount of reflective aerosols actually in the atmosphere, they may be masking the actual degree of climate forcing. Getting the estimate correct will take time and effort, but the effort will be well worth it for the scientific understanding, if not for the fate of humanity (◘ Fig. 23.1).

Perhaps the most interesting case is that of stratospheric ozone depletion, for it shows humans are capable of taking collective action to reverse environmental damage. The upper atmosphere contains a relatively small amount of three molecules of oxygen bonded together called ozone (O_3). Ozone is very rare, with only three molecules of ozone for every ten million molecules of oxygen, and it is measured in Dobson units, with one DU equaling only 0.01 mm in thickness. Ozone is created by a complex interaction with ultraviolet radiation and absorbs the shortest, and most harmful, wavelengths of ultraviolet light, UV-B and UV-C. These are the most powerful component of sunlight that can cause skin cancers and reductions in crop yields. In 1971 Dutch atmospheric chemist Paul Crutzen made the connection between nitrogen oxides and ozone depletion. In the same year, British scientist James Lovelock found molecules of chlorinated fluorocarbons (CFCs) ubiquitously mixed into the entire atmosphere. Sherwood Rowland and his postdoctoral fellow, Mario Molina, found that the stable CFC molecules interacted with stratospheric ozone, splitting the ozone molecule into oxygen (O_2) and chlorine monoxide (ClO). The ozone layer was thinning, especially in

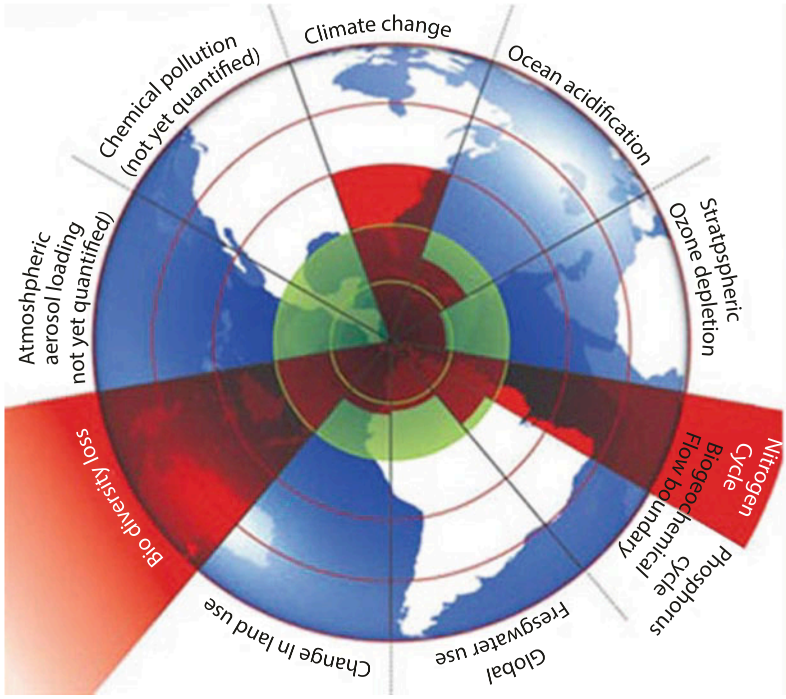


Fig. 23.1 The Green represents a safe operating space, while the distance that the red segment is from the origin represents the proportion of that category relative to the proposed maximum that the Earth can sustain

the southern latitudes, bringing with it the capacity for more UV light to reach the Earth's surface, and cause considerable environmental harm, such as damage to the retinas of vertebrate eyes. For their efforts Crutzen, Rowland, and Molina won the 1995 Nobel Prize in Chemistry.

But where did all the CFCs come from? It is a classic story of the unintended consequences of industrial production. It would be difficult to argue that human life has not been improved by refrigeration. Modern humans in electrified societies are far less likely to sicken or die from foodborne pathogens. Early refrigerators used toxic materials such as ammonia, methyl chloride, and even liquid sodium (which explodes on contact with oxygen) as refrigerants, and the use of them was limited. But, during the 1920s, the number of US homes served by electricity increased from 25% to 80%, and the possibility for mass marketing of a safe refrigerator was on the verge of possibility. General Motors, owners of Frigidaire, commissioned chemist Thomas Midgley to create a safe refrigerant. His invention of a chlorinated fluorocarbon, with the trade name of Freon, seemed like the answer. It was odorless, tasteless, nontoxic, long lasting, and cheap. CFCs found their way into myriad propellants, from whipped cream to deodorants to hairspray. In the 1950s some 20,000 tons per year found their way into the stratosphere. By 1970 the figure stood at 750,000 tons [7]. But the unintended consequences of the miracle invention were being mixed into the upper atmosphere and participating in the ozone-reducing reactions that were now threatening life on the planet. Could humanity respond?

In 1985 the nations of the world met in Montreal and ratified the Montreal Protocol on ozone-depleting substances. We found that we could live without spray deodorant and without Freon without too much sacrifice. Moreover, complying with the treaty was made easier by the simple technological change of adding hydrogen to the chlorine and fluorine. A chlorinated hydrofluorocarbon does not have the ozone-depleting potential of a CFC. Unfortunately, unintended consequences still remain, as these CFHCs are powerful greenhouse gases.

23.4 Climate Change

Scientists have known about the connection between atmospheric composition and temperature for a long time. In the 1820s Jean Baptiste Fourier hypothesized that the thickness of the atmosphere and the conditions of the planet's surface determined the Earth's average temperature. In 1859 John Tyndall concluded that the atmosphere and its trace gases (primarily carbon dioxide), along with water vapor, were transparent to visible light but "opaque" to the less energetic wavelengths of infrared radiation. In other words, carbon dioxide allowed high-energy photons to pass through the atmosphere but trapped some of the resultant heat from escaping into space, much like the glass in a greenhouse. Heat-trapping gases such as carbon dioxide, methane, sulfur hexafluoride, and CFHCs are today known as "greenhouse gases." Swedish chemist Svante Arrhenius confirmed Tyndall's hypothesis and expressed concern that we are warming the planet by "emptying our coal mines into the sky" [8].

Today there is a broad scientific consensus that the observed increases in the temperature of the Earth are linked to the level of carbon dioxide and other heat-trapping gases, although no respectable climate scientist dismisses other causes or thinks that atmospheric dynamics are simple and straightforward. For example, increases in temperature usually precede increases in carbon dioxide. Moreover, the most powerful greenhouse gas is not CO_2 but water vapor. The "debates" or outright denial of the connection between carbon and climate are generally found among politicians, business executives, and workers with jobs to lose instead of scientists. Climate change is a difficult issue to conceptualize, even for many atmospheric scientists. There are many variables and the theory is often far ahead of the data. To begin with weather is not climate. Weather is what is happening here, today. Climate is long-term averages. There is a big difference. Today in Massachusetts, USA, the temperature increased by 20 °F over the course of the day. It was cool and rainy in the morning and hot and muggy in the afternoon. But average temperature across the world is different. Climate scientists John Anderson and Alice Bows conclude that we must keep the increase in average temperature to less than 2 °C (3.6 °F). For Anderson and Bows, 2° is not the threshold between safe and dangerous; it is the threshold between dangerous and extremely dangerous. In the developed world, a positive feedback loop has developed. As the temperature warms, more people purchase and use air conditioning. This uses more electricity and puts more carbon into the atmosphere. The planet warms. People use more air conditioning. The planet warms………The ubiquitous use of air conditioning is a fairly recent phenomenon. When one of your authors (Klitgaard) grew up in the hot and arid Southwest, nobody he knew had an air conditioner. Now they are part of the "middle class life" in most parts of the country. And carbon dioxide emissions continue to increase.

In 1992 the United Nations held its Framework Convention on Climate Change. Since that time diplomats have been meeting in regular Conferences of Parties (COPs) to try to reach agreement on limiting the emission of greenhouse gases, with little to show for the effort. Poor nations saw the industrialized part of the world grow rich on the power provided by coal and other fossil fuels and ask why they are now precluded from doing the same. Rich nations do not want to lose their competitive advantage to newly industrializing nations like China and India, with low per capita incomes but high total emissions. But finally, in 2016, the nations of the world signed the Paris Accord, committing themselves to enact policies to stay within the 2° threshold. Whether the agreement will be successful is now questionable, as new US President Donald Trump has vowed to pull the United States out of the Accord because it gives too much competitive advantage to China. Business aside, what are the scientific concerns, and what is the evidence?

In 1957 Roger Revelle and Charles Keeling began to measure carbon dioxide concentrations in the Northern Hemisphere at an observatory on Mauna Loa in Hawaii, and the taking of atmospheric samples continues to this day. In 1957, they measured concentrations of 315 parts per million volume (ppmv). The latest readings are nearly 409 ppmv. Look at the graph in Fig. 23.2.

You should notice two crucial details, a sawtooth pattern and a trend. The sawtooth pattern is

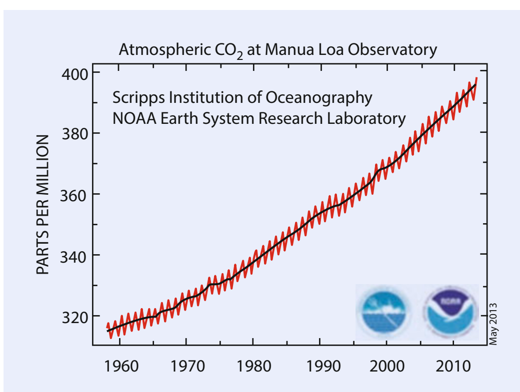


Fig. 23.2 Keeling curve (Courtesy NOAA)

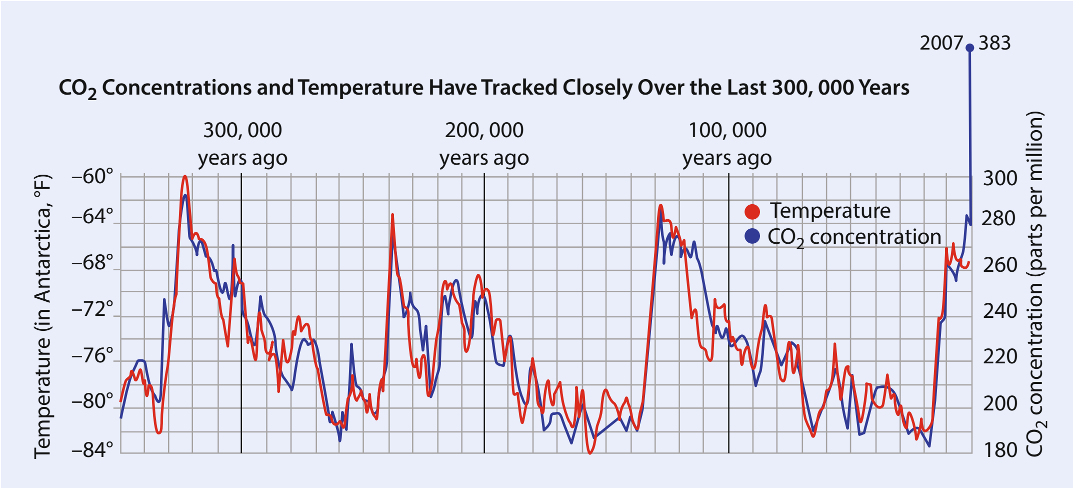


Fig. 23.3 Vostok ice core data

the Earth breathing [11]. In the Northern Hemisphere, leaves of deciduous trees fall in the autumn. Oxygen production ceases as photosynthesis stops and carbon dioxide concentrations increase. In the spring new leaves form, oxygen production begins again, and carbon dioxide concentrations fall as new leaves are formed. But the disturbing feature is the trend of growth. Can this be attributed to humans and their burning of fossil fuels, or is it just a natural variation? Atmospheric scientists have collected data back to more than 350,000 years ago by taking ice core sample in Antarctica, where ice rarely melts. Since dirt accumulates in thin layers between winter snowfall, a clear year-by-year record can be found by drilling deep into the ice. Ice contains air bubbles, and one can, using sophisticated machinery, test for the composition of ancient air. A main sampling station is at Vostok Station in Antarctica. Figure 23.3 displays the Vostok ice core data. As one can see, temperature and carbon dioxide concentrations are correlated closely. As CO_2 concentrations rise, after a lag period, temperature follows. When carbon dioxide falls, so does temperature, and certain patterns repeat through history. Temperature rises rapidly and cools more slowly. But look at the very right hand of the graph, and you will see something unusual: climate stability. The epoch in which humans evolved, known as the Holocene, is marked by unusual climate stability and warmth, which is critical to humans—at least until recently. The study of the ancient climate is known as paleoclimatology. Will it continue as carbon dioxide levels increase to a level greater than anything observed in the past 350,000 years? Will we, as a species, be able to adapt? What are the problems we might anticipate?

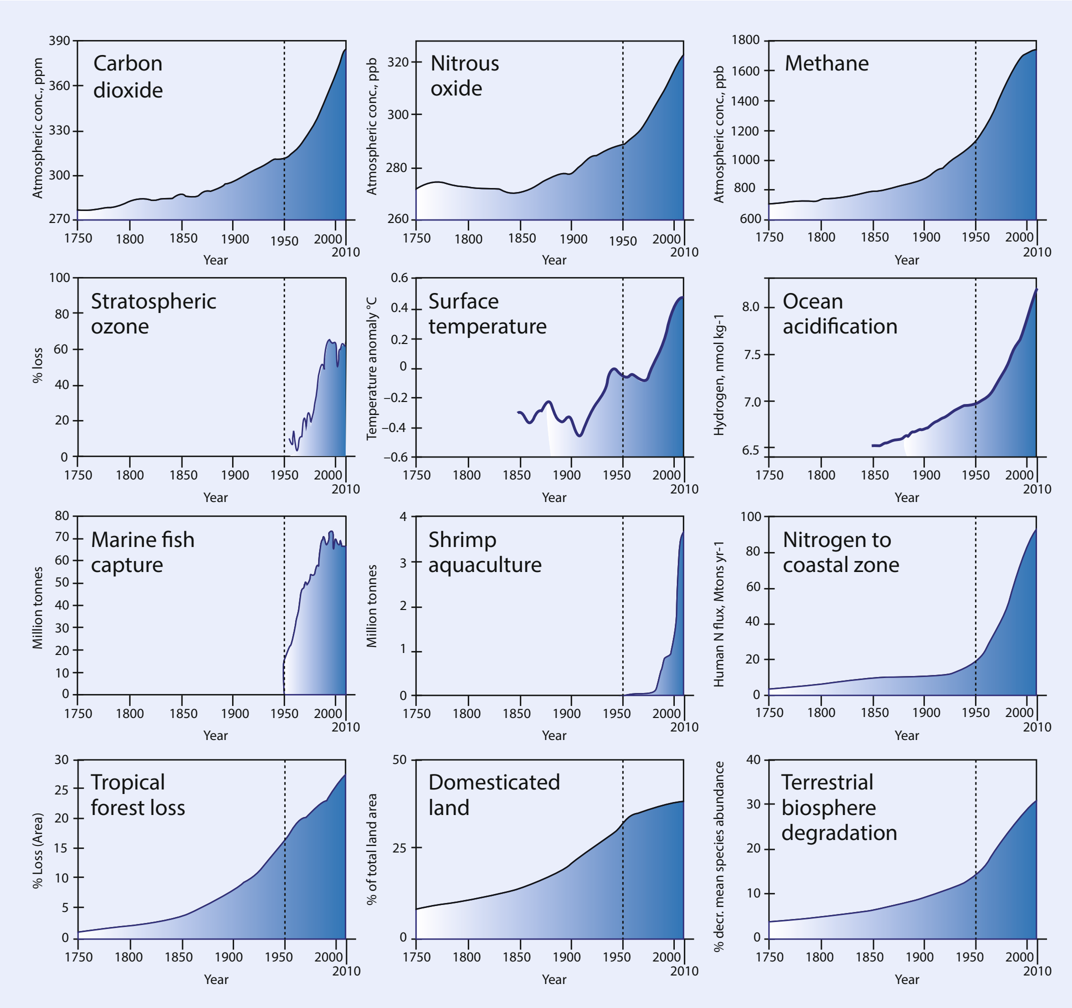
One problem is increased volatility of the weather. Computer simulations predict more frequent and stronger storms, as tropical cyclones feed on warmer water, and more severe thunderstorms and tornadoes are born from the clash of dry and humid air masses. Evapotranspiration increases exponentially with temperature. As the temperature warms and crosses the arid West, the air becomes desiccated and seeks out all available moisture from the ground. As the same warming air masses cross the humid Gulf of Mexico, they pick up more moisture. The part of the country where the cool dry air masses flowing eastward from the Rocky Mountains meet the warm humid air of the Gulf is known as tornado alley. Increased storm damage is now a fact of life in states such as Oklahoma, Texas, and Arkansas. Those on the coast fear storm surge from more powerful oceanic storms. The National Oceanic and Atmospheric Administration has now taken to naming winter storms as they name hurricanes.

Another environmental change attributable to global warming is sea level rise. Water expands as it gets warmer, so part of sea level rise comes from thermal expansion. Part of the rise comes from a positive feedback mechanism called the albedo effect. *Albedo* measures reflectivity, and a reduction in reflectivity can lead to the absorption of more of the sun's radiation. Newly fallen snow reflects about 99% of the radiation that hits it, and blacktop absorbs and reradiates as heat nearly all of the radiation that strikes it. Physicists call an object that absorbs and reradiates 100% of the radiation that strikes it a "perfect black body." As the planet warms, the ice shelves, which act as "speed bumps", which inhibit the flow of the glaciers towards the oceans to keep the ice sheets in place, begin to melt. This exposes additional dark ocean water which absorbs more solar radiation. This raises the temperature and melts more ice. The process continues as long as the ocean continues to warm. If the ocean would become warm enough to melt the ice surrounding frozen methane in the Arctic tundra and in the oceans, a 6 °C warming could be a distinct possibility. The sea levels will also rise because the ice shelves have melted and the moving ice sheets add their mass to the ocean. If you recall from ► Chap. 6, humans were likely to have migrated from Asia to the Americas during an ice age. Enough ocean water was taken up in ice to lower the sea levels and create a "land bridge" upon which our ancestors could walk. If carbon dioxide concentrations continue to rise, the opposite will occur. Nearly the entire state of Florida, along with nations such as Bangladesh, is likely to be flooded, along with many coastal cities. Moreover, the drinking water source of billions of Asians can be found in a handful of Himalayan glaciers which are the headwaters of the Ganges, Brahmaputra, Mekong, Irawati, and Yangtze Rivers. The combination of sea level rise and reduced water supply could create a climate refugee problem of epoch proportions. It is likely that these events will occur in the same time frame as running short of petroleum. Will the people of the developed world, deprived of their sources of comfort and convenience and perhaps facing economic dislocation or even collapse, open their arms and welcome billions of climate refugees?

Climate change also has biological effects. According to climatologist James Hansen, more than a thousand studies have shown an average migration rate for various species toward the poles of about four miles per decade. However, the lines of equal temperature called isotherms have been moving poleward at a rate of 35 miles per decade. If carbon emissions continue at the present rate, the isotherm movement will double to 70 miles per decade by the end of the century [9]. Polar and Alpine flora and fauna are simply being pushed off the planet.

Will we remain in the Holocene, or are we entering a new geological epoch dominated by human action called the Anthropocene? Geologists are still debating the issue. A proponent of the idea that we are now in a new geological epoch is Will Steffan, director of the International Geosphere-Biosphere Program (IGBP). Steffan and colleagues published their analyses in the 2004 *Global Change and the Earth System* [10]. In it they recorded the trajectory of the human enterprise from 1750, the humble beginnings of the fossil fuel age, to 2000. They presented a series of 24 graphs, including both the Earth system and the socioeconomic system. The results were shocking. Nearly every series they looked at was escalating exponentially, with a sharp increase around 1950. They dubbed the period "The Great Acceleration." On the Earth systems side, carbon dioxide emissions, tropical forest loss, ocean acidification, and coastal nitrogen pollution, among others, were all rising exponentially. In the socioeconomic realm world and urban populations, real gross domestic product, primary energy use, and foreign direct investment showed similar exponential patters. The series are presented in ◘ Figs. 23.4 and 23.5.

People should ask themselves at least two questions. Can the exponential growth of both Earth systems and socioeconomic trends be compatible with the Holocene stability with which our species evolved and thrived, or will the acceleration of carbon dioxide emissions that have not been seen for at least 400,000 years push us into a period of instability or chaos? How will humans adapt and react? We will take up these and other pressing questions in our final chapter (◘ Fig. 23.6).



■ **Fig. 23.4** Earth system trends

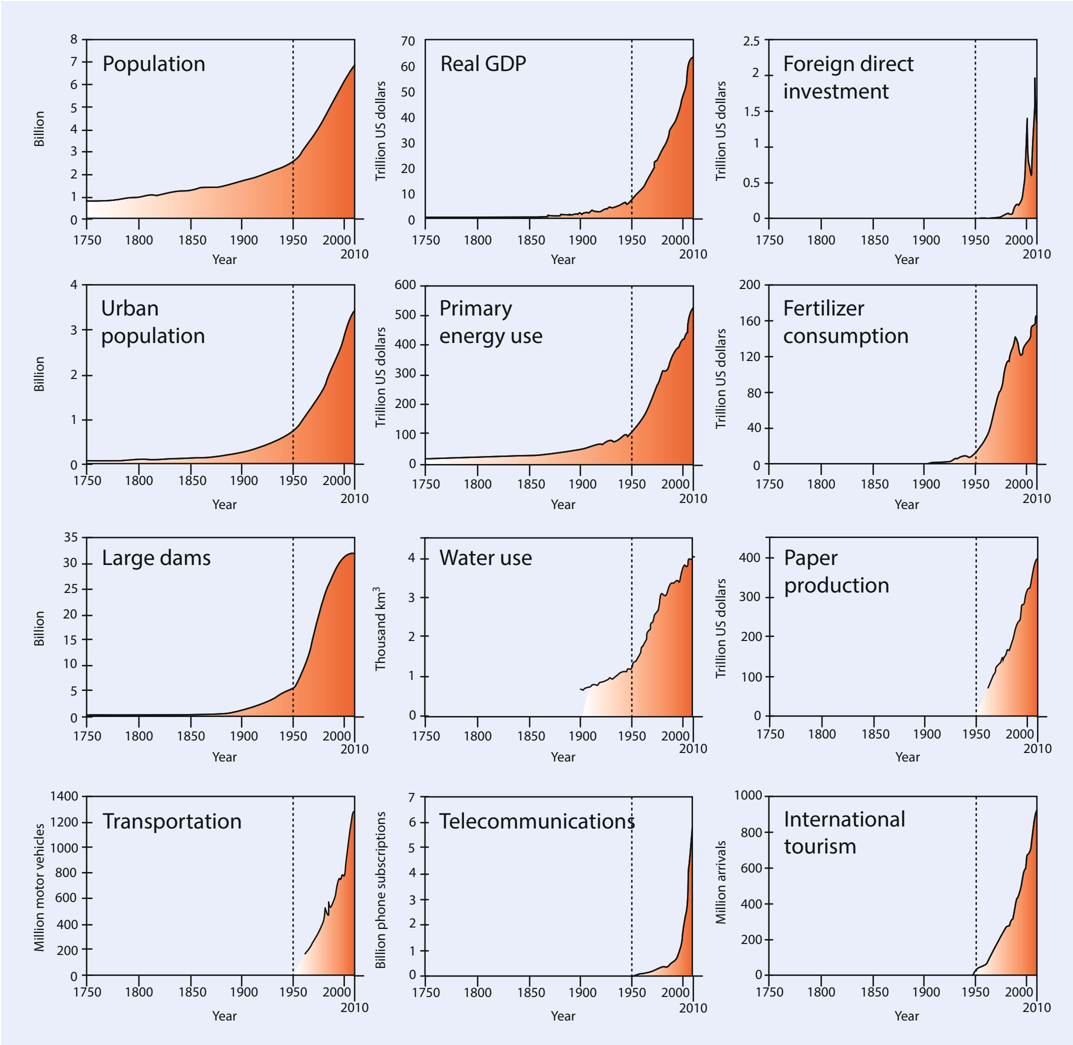


Fig. 23.5 Socioeconomic trends

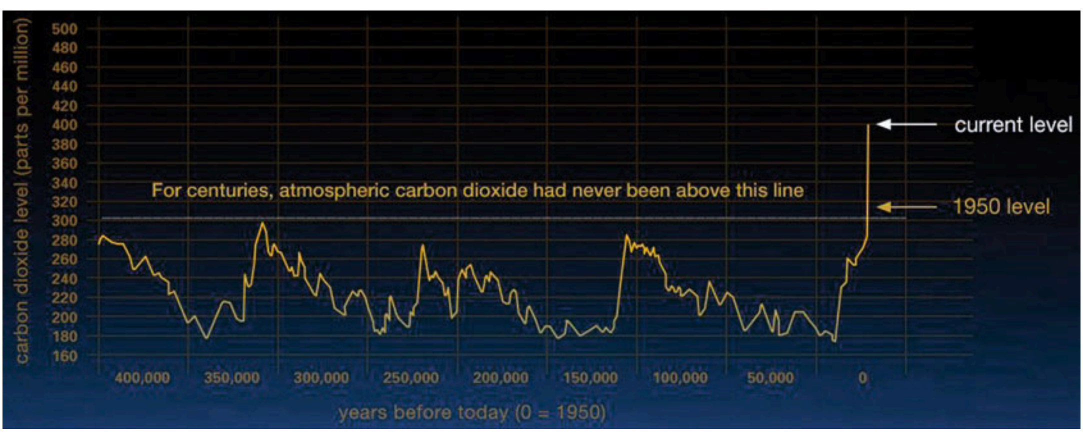


Fig. 23.6 Carbon emissions in historical perspective

23

References

1. Murphy, D., and C.A.S. Hall. 2011. Energy return on investment, peak oil, and the end of economic growth. In *Ecological economic reviews*, Annals of the New York Academy of Sciences 1219, ed. R. Costanza, K. Limburg, and I. Kubiszewski, 52–72. Boston: Blackwell Publishing on behalf of the New York Academy of Sciences.
2. Klitgaard, K. 2013. Heterodox political economy and the degrowth perspective. *Sustainability*. 5: 276–297.
3. Kaplan, S. 2016. By 2050 there will be more plastic than fish in the world's oceans. *Washington Post*. January 20.
4. Carson, R. 1962. *Silent spring*. Boston: Houghton-Mifflin.
5. Hoshaw, L. 2009. Afloat in the oceans, expanding islands of trash. *New York Times*. November 9.
6. Rockström, J., W. Steffan, K. Noone, A. Persson, S. Chapin III, E. Lambin, T.M. Lentin, M. Scheffer, C. Folke, H. Joachim, H.J. Schellnhuber, B. Nykvist, C.A. De Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P.K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R.W. Corell, V.J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J.A. Foley. 2009. *Nature* 461 (24): 472–475.
7. Angus, A. 2016. *Facing the anthropocene*. New York: Monthly Review Press.
8. Foster, J.B., B. Clark, and R. York. 2012. *The ecological rift*. New York: Monthly Review Press.
9. Hansen, J. 2009. *Storms of my grandchildren*. New York: Bloomsbury.
10. Steffan, W., A. Anderson, P.D. Tyson, J. Jäger, P.A. Matson, B. Moore III, F. Oldfield, K. Richarson, H.J. Schellnhuber, R.L. Turner II, and R.J. Wasson. 2004. *Global change and the earth system*. Berlin: Springer.
11. Hall, C.A.S., C. Ekdahl, and D. Wartenberg. 1975. A fifteen-year record of biotic metabolism in the Northern hemisphere. Nature 255: 136–138.

Is Living the Good Life Possible in a Lower EROI Future?

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0_24

In order to assess the prospect of living the good life in an energy-short, climate-compromised future, we need to address the issue of what constitutes the good life. Neoclassical economics would have us believe, or assume, that increased consumption of goods and services indefinitely improves our welfare, as does increased choice. Yet, if one views the empirical evidence (see Fig. 24.5) or reads the advances in behavioral psychology, the strength of the relation is rather weak. According to psychologist Tim Kasser, those who are the most acquisitive and materialistic have the highest levels of clinical depression and the most tenuous personal relationships [1]. But even if there were a strong positive connection between mass consumption and human happiness, the prospect of continuing the acquisitive lifestyle, which has long been limited to a small share of the population, will likely be limited in the future, both by limitations of the energy and materials required and by the environmental consequences, as summarized in the previous chapter. If future growth is so constrained, where will we seek happiness? Can a life of lower material and energy consumption be a happy one?

Traditionally, economists have thought not. In 1996, John Kenneth Galbraith produced a short book entitled *The Good Society* [2]. For Galbraith, the good society provided for stability of employment, where no one was denied the income to allow them access to the basic requirements of nutrition, shelter, and safety. It was a society that minimized an entrenched bureaucracy and an imperialistic military, and one that welcomed immigrants and improved the lot of the planet's poor while protecting the environment. It is a vision your authors share. But 20 years ago, few economists recognized the biophysical limits to human activity. Consequently, Galbraith declared:

"Very specifically, the good society must have substantial and reliable economic growth—a substantial and reliable increase in production and employment from year to year."

However if biophysical limits and the inability to absorb economic surplus lead to slow, and sometimes declining, rates of economic growth, the question becomes one of how do we live the good life in the absence of economic growth? While we may not need another brand of toothpaste or underarm deodorant, how do we provide sufficient employment in the absence of economic growth? We certainly do not have all the answers, but we know enough to realize we have to raise the questions, and that we must do so within a biophysical context.

We do not see this automatically as a bad future, depending on how we deal with it. As boys, we both had a wonderful childhood on opposite coasts in the 1950s and 1960s during a period when the US energy use was only 20% of what it is now. We could go fishing and surfing (respectively) on our bicycles and had no need for soccer moms driving us around in an SUV. We played sports all the time with neighborhood friends and went camping and hiking to our heart's content. Nature was abundant, everywhere, exciting, and fascinating. Even today's perspective that there are dangerous people out there and children must be driven everywhere for protection was not valid—and even today youngsters are considerably more likely to die or be hurt in an automobile accident than be kidnapped!

For the record, we, deeply involved in all this stuff as professional ecologists, economists, and energy analysts for the last four to five decades, are neither optimists (which is our nature) nor pessimists about our energy and economic future because we really have no way to predict the future beyond some easy and very coarse extensions of present trends (for demographics, probably oil, possibly gas, conceivably coal). The hardest things to predict would be human behavior—will we go quietly into declining affluence (as we are sort of doing now)? Will the unemployed or never to be employed cause riots or become terrorists? Will people vote in an authoritarian government who promises to bring back better days? Will we be able, in some way that we do not yet know, to do things with human hands we do now with fossil fuel? Will we be able to make some kind of transition to a new energy source? If the economic pie must shrink, will the rich respond by attempting to keep their absolute amount constant—while the poor get a smaller part of a shrinking pie? Or what? Will society get behind the proposal to spend the falling share of income that does not go to acquiring energy on perpetual war over the scraps that remain? For the record, we, deeply involved in all this as professionals and modelers since the 1960s and 1970s, can be neither optimists nor pessimists because we cannot predict these things and do not trust anyone who says we can. We think we have to go into the future with the following model and something like the following probabilities (you can choose your own percentages): we will go off the cliff, energetically, economically, or environmentally (25%), we will make a transition to a new energy source that will benevolently replace oil and gas (25%), or we will muddle along, gradually getting materially poorer but adjusting to that (50%). The point is that we do not think anyone knows those percentages, and so we must go into the future with a huge breadth of possibilities. That in itself might be pretty difficult. Some would trust the market to adjust, others might not, and others might think we can all move to communes and grow our food or have other mechanisms of adjustments. Many people who think about these things retreat to a bunker mentality and are stocking their country houses with food and ammunition. Paul Raskin, of the Tellus Institute, offers three possible scenarios: Conventional Worlds, Barbarization, and Great Transitions. The *Conventional Worlds* scenario can take the form of Market Forces or Political Reform. In the first approach, all one must do is trust that the market will produce solutions that are not only efficient and equitable, but also sustainable. In the second, incremental reforms, without challenging the inner workings of the system, will lead to a sustainable future. A less optimistic scenario is *Barbarization*. In this scenario, the privileged of the world retreat to a Fortress World, militarizing their borders, and keep what resources remain to themselves. It is a very repressive world that is likely to represent Breakdown, or the second variant of Barbarization, which is complete chaos, or Thomas Hobbes' "war of all against all." *Great Transitions* is the transformation of the present system into a more humane and livable future. These include Eco-communalism, where groups of like-minded people live collectively and pastorally in small-scale communities in touch with nature and culture, or as a New Sustainability Paradigm, where sustainability is defined as progressive global social evolution, and is extended to the world's poor, not just to those affluent enough to live peacefully in rural communes. Related to this perspective are calls for "the end of growth," "degrowth," and the steady-state economy [3].

24.1 What Are the Main Issues for Transitioning to the Future?

The main problem that we face is that we (the United States, the world, wherever) will require massive new investments in whatever might be the next energy source at a time when most citizens will be experiencing a decline in their own purchasing power. For example, if gasoline costs $10 a gallon (and this is just to extract that gallon from an aging, energy-requiring field), who will want to pay an additional 5 dollars a gallon as an investment in whatever fuel or other technology will replace that gallon? The answer is probably few, if any, and that implies that we just continue on the path of using ever-lower-grade, more expensive conventional resources, slowly grinding into ever-greater poverty. Will the rising price of fossil fuels make renewable energy technologies more competitive? Alternatively, will their intermittency and inflexibility limit their ultimate use [4]? It depends on the structure of markets and the power of the largest corporations, but also the EROI and flexibility of the alternatives. Moreover, in a world of expensive energy and excess capacity, *Conventional Worlds* scenarios such as a reduction in corporate tax rates are highly unlikely to produce economic growth, as promised by its proponents.

If one accepts the importance of a biophysical basis for economics, then there are some important implications of our analysis for economics and for society. The first issue pertains to the economic pie and how we will cut it. As we developed in some detail in ► Chap. 7, "the American dream" gave for perhaps the first time in human history the hope of significant and ever-increasing prosperity to a broad swath of people through a number of generations and for an entire nation. As we believe, this book makes clear it is not clear at all that this prosperity is continuing or will or can continue, and in fact, there is a great deal of evidence that we have reached the end of any increase in affluence: The GDP and the average take-home pay for workers in of the United States have barely budged for decades, there is increasing evidence that such growth as took place from the mid-1990s until 2017 was based in large part on debt or speculation. Many state governments are broke or are cutting back on such former entitlements such as good university education for all or many. Underemployment remains stubbornly high, colleges and universities are having increasing difficulties balancing their budget, many people's retirement plans have lost a great deal of their net worth, and housing prices are again inflated. Of course, none of this is new, for the United States has gone through depressions and recessions often enough before, and many Conventional Worlds thinkers believe that if we just wait, we will come out of the present period of meager growth. As of the publication of this book, we are still waiting, and we believe we will wait for a long time, as there is ample evidence to show that a monopolized economy tends to produce stagnation instead of rapid growth.

What if the recent recession is not part of a cycle but is the new reality, one in which new and unprecedented energy constraints exacerbate the already existing tendency of a concentrated economy to grow very slowly, if at all? What if energy restrictions, such as David Murphy's concept, developed in ◘ Fig. 4.8 that any increase in growth sets into motion its own demise because of the need to use much more expensive oil? In other words, what if the national (and global) economic pie can no longer grow?

Traditionally, as we developed in ▶ Chap. 7, the concept of the American dream, that is the continually growing pie, had previously resolved or defused many contentious issues in the United States for some time: labor had made more in their salaries (at least until the late 1990s), while management has made much more, large portions of total wealth were "skimmed off" by Wall Street and other entities and it was hardly noticed. Government could be corrupt or inefficient and still the roads got fixed and public universities expanded. Each generation still had the sense that they were better off than their parents, and so on. There were few complaints because everyone made more, at least a little more. But that seems no longer to be the case. So if any one group does better now, it has to be at the expense of some other group or some other use of the money—in other words, the question is if the pie is no longer getting larger, indeed if because of energy constraints it can no longer get larger, how will we slice it? This is forcing some ugly debates back into the public vision, and provides fodder for both responsible politicians and demagogues. And indeed, if total energy availability and economic largess is actually shrinking, then we will need to ask some very hard questions about how we should share and spend what is left.

Probably, this will force individuals and our nation to focus on what is most important. One way to think about this, a commonly used perspective, is "Maslow's hierarchy of human needs." This theory, proposed by Abraham Maslow in his 1943 paper "A theory of human motivation" [5], proposes that humans will attempt to meet their needs in more or less the following order: First, they will meet their physiological needs which are the literal requirements for human survival, including breathing, nutrition, water, sleep, homeostasis, excretion, and sexual activity. These require clean air and water, food, clothing, and shelter. Second, once physiological needs are satisfied, an individual will attempt to meet safety needs in an attempt to derive a predictable, orderly world in which perceived unfairness and inconsistency are under control, the familiar frequent and the unfamiliar rare. These include personal security, financial security, health and well-being, a safety net against accidents/illness, and their adverse impacts. For example, in the world of work, these safety needs manifest themselves in such things as a preference for job security, grievance procedures for protecting the individual from unilateral authority, savings accounts, insurance policies, reasonable disability accommodations, and the like. Third, once the above needs are met, humans seek love and belonging, i.e., emotionally based relationships in general, such as friendship, intimacy, and family. These include large social groups, such as clubs, office culture, religious groups, professional organizations, sports teams, gangs, or small social connections (family members, intimate partners, mentors, close colleagues, confidants). They need to love and be loved by others. Fourth, again once the above have been met, humans seek esteem, to be respected and to have self-esteem and self-respect and also the esteem of others. Also known as the belonging need, esteem presents the normal human desire to be accepted and valued by others through a sense of contribution in, for example, a profession or hobby. Finally, according to Maslow, people seek self-actualization, the need to understand what a person's full potential is and to realize that potential, to become everything that one is capable of becoming, for example, an ideal parent, athlete, scholar, painter, or inventor.

Maslow's theory has been criticized from a number of angles including the lack of evidence that humans in fact follow that hierarchy, or indeed any such hierarchy, and from the perspective that his pyramid may be more representative of people from an individualist vs. socialist society. Nevertheless, his theory is broadly accepted in psychology and even marketing.

Our own research on the implications of declining net energy, while not consciously based on Maslow's theories, is consistent with them in that we have the sense that discretionary spending will be increasingly abandoned as humans attempt to meet their needs for food, shelter, and clothing, what we call "staples" (see ◘ Fig. 19.7). Presumably if and as the amount of net energy declines in our society due to having gone through peak oil and declining EROI, humans will increasingly give up categories higher on the pyramids (fifth above) and concentrate increasingly on the more basic requirements including food, shelter, and clothing. What this may mean in modern society is that expensive vacations, then education, and then health care would be abandoned if and as the economy is increasingly restricted. On the other hand, the first author's mother said that during the depression, people would give up a lot of basics to go to the movies, which were an escape from the often grim daily reality.

24.1.1 Labor

During the last four decades under the pressure of cost minimization, the economies of the United States, Japan, and Germany have been driven into substituting powerful, cheap energy and increasingly automated capital for weak, expensive labor. The low price of fossil fuels relative to their productive power has tended to generate large profits. In other words, labor productivity, the amount of value added per hour that the laborer works, has been greatly increased by subsidizing the efforts of a laborer with more fossil energy, for example, a larger tractor for a farmer. This substitution has not occurred to the degree that it might for various reasons [6] but nevertheless has contributed enormously to unemployment. Will robots put more people, such as truck or taxi drivers, out of work? Heterodox labor economists have known for a long time that an increase in productivity without a subsequent growth in spending manifests itself as unemployment and excess capacity. New resource and environmental constraints may further preclude growth to a degree unimaginable to mainstream economists who do not include energy in their employment models. Will an increase in the price of energy relative to labor increase substantially the amount of labor employed? If labor can again be more valuable in production, real wages would have to fall because goods and services would become more expensive relative to real purchasing power of salaries (otherwise, the labor would not become relatively cheaper). Jean Laherrere has shown an uncanny relation between oil price and unemployment (Fig. 24.1) which may be something to worry about.

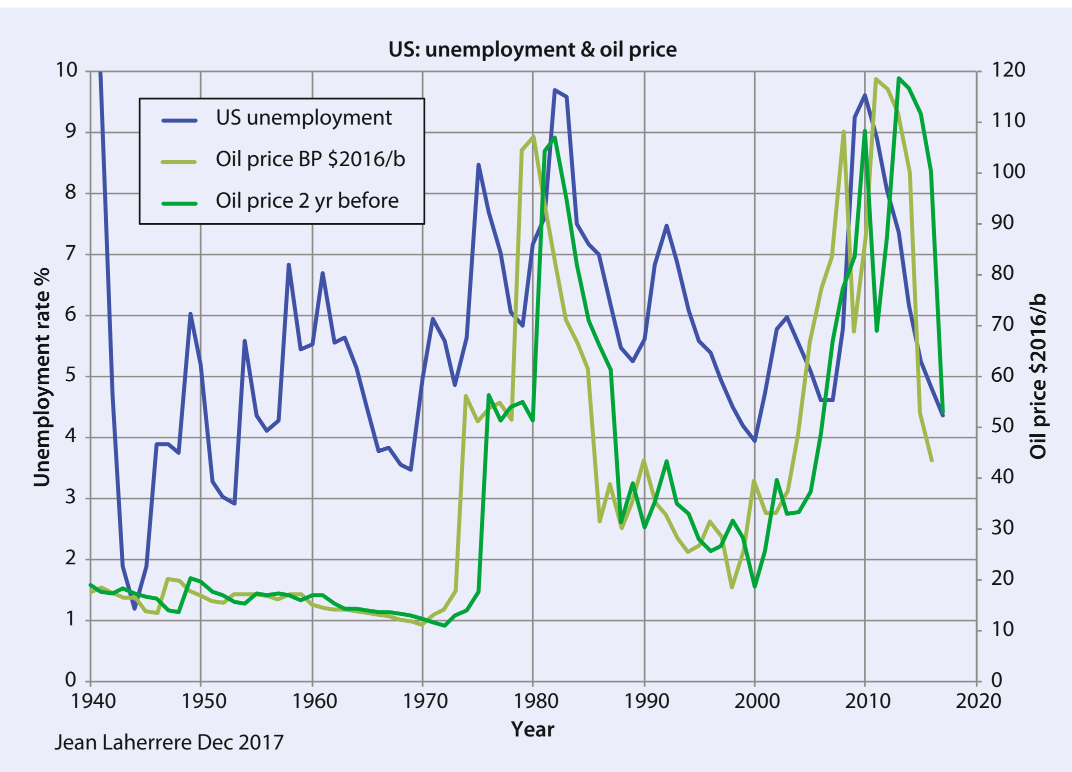


Fig. 24.1 The relation between oil price and unemployment the following year for the United States

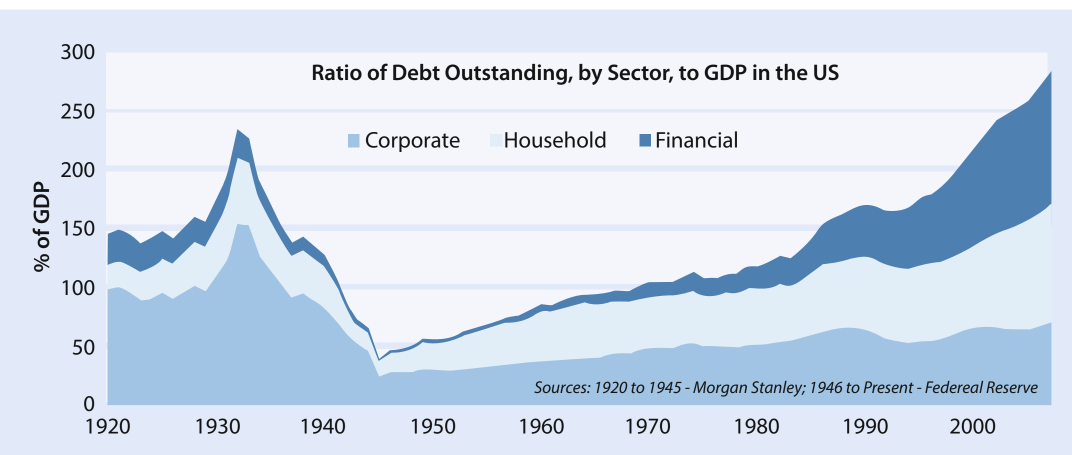
24.1.2 Debt

An enormous, perhaps overwhelming, aspect of our future will be debt. The concept and importance of debt to the American dream were presented in ▶ Chap. 7, and the connection of debt to energy in ▶ Chaps. 4 and 7. Where once we could grow our way out of the importance of debt, this looks more and more difficult if growth becomes a thing of the past. Debt has become an enormous political football with some very curious political dimensions because nominally fiscal conservatives in the past generated the largest part of our debt, at least until the current situation. Few know that the Reagan administration generated far more debt, even corrected for inflation, than Franklin Roosevelt! Given our belief that debt is a lien against future energy use (i.e., if debts are to be honored, then some portion of a nation's future energy use must be diverted to nonproductive, nonconsumptive payment of interest or principle on debt), then it is worrisome to consider that in the future when we will need large amounts of our energy resource to invest in new energy technologies (including conservation), then we must consider that some significant portion of whatever energy is available will be just dissipated on meeting debt loads. Of course, our huge debt load (◘ Fig. 24.2) may never be paid unless we greatly reduce the energy/dollar relation through massive inflation.

24.1.3 International

This book focuses on the United States, and it must seem clear that we have problems enough with energy. But it is worse for many other countries. For example, the United States imports about one-quarter of its energy, while Europe and Asia import two-thirds, making the countries there far more vulnerable to whatever the future energy situation becomes. Europe had a momentary respite with the enormous North Sea oil fields, from which nearly 50 billion barrels of oil have been extracted and another 10–30 billion might yet be from smaller fields. This oil bonanza allowed Britain to have a few decades of tremendous affluence and led many to believe that Margaret Thatcher's political policies had somehow saved the day. But now the oil and gas reserves of the British portion of the North Sea are nearly gone. Britain is struggling with the fact that the oil was essentially spent in a wild binge, and a new cold hard reality is upon her as civil servants and students explode with the drastic cuts in government largess. Norway, on the other hand, has developed its oil and gas at a more measured pace and placed much of the revenues into a trust fund to help all future Norwegians, one of the relatively few examples we have of a mineral bonanza being used to help all citizens, although that too has suffered from falling returns on investments [7].

At the extreme, many tropical developing countries are especially vulnerable because of their



◘ **Fig. 24.2** Total American debt (Sources: 1920 to 1945 – Morgan Stanley; 1946 to Present – Federal Reserve) increasing reliance on oil for their increasing populations, increasing use of fertilizer and other inputs required for agriculture, and importance of tourism. The first author has a great deal of experience attempting to understand the relation of energy to what is normally called "development" in the tropics. Many tropical countries are poor, or at least not affluent, and essentially all wish to become more wealthy. Hall was initially attracted to the country of Costa Rica which was promoting itself as a "laboratory for green, sustainable development." Unfortunately, his experience from years of living there and studying quantitatively all major aspects of its economy, detailed in two large books on the subject [8], was that Costa Rica, no matter how lovely and how well developed the ecotourism and solar energy (mostly hydroelectric) industries, was at least as dependent upon petroleum as any place else, was far from sustainable for at least 18 reasons, and had no real plan as to how to continue its moderate standard of living without oil. This is for a country that is relatively well-off with respect to its sustainability and its government! Thus, unfortunately, I think peak oil is likely to hit the developing world especially hard. Likewise, even the modest increases in the price of oil have already impacted fully developed, but highly oil-dependent, Puerto Rico (a "dependent" of the United States) especially hard as it has recently declared bankruptcy and an inability to service its debts. These regions, whose economy once depended almost entirely on agricultural production unsubsidized by fossil fuel, cannot possibly feed their swollen populations now from indigenous agriculture. They have no contingency plans for peak oil. This is all the more true in the aftermath of Hurricanes Irma and Maria. The island, before the hurricanes, was completely dependent upon electricity produced by means of diesel generators, whose very maintenance was crippled by an unelected financial control board. Now the vast majority of the island is without electricity. Agriculture and tourism have ground to a halt, and the infrastructure lies in ruins. Will development mean a reconstruction of the fossil economy or will the new electricity system be built upon renewable energy? It probably depends more upon the vested interests of fossil fuel industries and neoliberal politicians than upon the technology of alternative forms of energy. How much net energy is required to deal with the inevitable storms and floods, which appear to be increasing?

Likewise, agricultural production for the world more generally may be very susceptible to peak oil and gas (which would limit the production of nitrogen fertilizer) and peak other requirements. The site ironically called "Sustainable Phosphorus Futures" suggests global peak phosphorus by 2030 [9, 10]. Irrigation, used on perhaps 15% of US crops, is often dependent upon deep groundwater that requires more energy over time as it is pumped from deeper and deeper depths as the fossil water is depleted. More generally around the world, agriculture has shifted to procedures that are energy intensive in many ways, and we expect all to be impacted in various ways by peak oil. Since the growth of the global population is not too different from the growth of fossil energy, we would not be surprised to see those curves to continue to be related on the downslope of the energy curve. As the physicist Albert Bartlett states [11], there is little doubt that populations will decline, what we have is a choice about whether it is due to procedures that we might like (i.e., reproductive control) or the things we like much less, such as starvation, disease, pestilence, and war.

24.2 Choosing a Better Future

To the best of the authors' imperfect ability to predict, it appears very unlikely that there is a "supply" approach out of the circumstances that peak oil will leave us with. Every analysis that we respect as realistic shows a future with peak oil either about now, an "undulating plateau" for not many additional years at best, and then declining oil into the future. Coal and natural gas may be able to fill in part of the gap (but with enormous difficulty for liquid fuels) for some additional decades, but growth or probably even a steady-state energy economy seems unlikely after a decade. To us, it seems that the die is inevitably cast because we simply are not finding oil as rapidly as we are using it (Fig. 8.3). Globally 80% of our oil comes from some 400 large oil fields discovered before 1970, and at least a quarter of these are presently subject to declines in production and EROI, and more will join that group soon. Thus, whatever new oil we find, and we will find a lot, will have to make up for some of that decline and is almost ensured not to add to any increase in oil supplies worldwide (Fig. 8.9). There are indeed enormous quantities of low-grade fossil fuels left in the ground, but their low EROI and huge investments required make it unlikely that they can replace the role of oil or the forthcoming shutdown of 100 nuclear plants. Other low-grade types of oil such as tar sands are making a small difference but are almost inconsequential on a global basis. All new oil supplies are likely to be much more expensive than the existing oil production. Natural gas may not peak for several decades but is unlikely to more than compensate for declining oil at best.

Coal is harder to predict. There is a lot of talk about "peak coal" (e.g., Patzek [12]; Mohr et al. [13]) much of which is based on the difficulty in mining increasingly thin seams, and the estimate of the total size of the resource is much smaller than was the case a decade or two ago [14]. Peak coal has already come to the world's largest coal user, China (Qi, Tsinghua University. Personal communication) see also [15], but clearly in the United States, Russia, and some few other regions, coal remains extremely abundant. Alaska alone has huge resources of exploitable high-quality coal. US production in 2009 was about one billion tons, and the Powder River formation in Wyoming alone contains some 40 billion recoverable tons. The total recoverable coal base estimated by the US EIA is about 500 billion tons. But Rutledge [14] gives a much smaller number based on what is recoverable. Thus, it seems that if we are willing to make the investment and suffer the environmental consequences, coal can be as abundant as we wish it to be in the United States for a century at least. But coal is currently being displaced by natural gas, so it is a bit hard to predict the future patterns of consumption.

Few if any alternatives, including conservation, appear to be able to fill in for the anticipated decline of oil and then gas. The most recent estimates for all fossil fuels show a projected peak in all fuels by 2025 or 2050, earlier than previously anticipated [13, 16] (Fig. 24.3). Replacing them, if possible, will take an enormous investment in money, energy, and time to be viable. Replacing oil for trucking will be especially difficult [17]. There are some very ambitious plans for replacing all or most use of fossil fuels with solar renewables (e.g., Jacobson [15]), but the actual ability to do that is hard to predict and has been severely criticized by Clack et al. among others [4]. Biomass (other than traditional solid forms such as firewood) can make a certain gross contribution but unless things change considerably little net difference. New solar technologies (including wind turbines and photovoltaics) are a great hope for the future but to date contribute no more than about 2 % (Fig. 4.1) and the pace of development has slowed recently. All of these alternatives would have a much lower EROI than what we are used to if provisions for intermittency are included. Thus, we do not necessarily foresee a future United States without energy but rather substantial problems in providing or substituting for the liquid and gaseous hydrocarbons that have been our lifeblood and the engines of rapid economic growth.

Unless we as a country decide to increase our coal use enormously, which would be difficult but certainly not impossible, given present environmental concerns and infrastructure limitations, it seems likely that the future will be one of an increasingly constricted energy supply. This implies, as developed again and again in this book, the end of economic growth and some extremely large adjustments of our citizens to a new steady-state or declining economic condition. If we pay off our huge international debts, this implies an even more constricted economic situation. While for many this will seem like a very gloomy future, for us this is not necessarily the case. Given the environmental destruction we have observed in our lives due to rampant development, we will not miss its continuation, should that be. It depends upon how we adjust, including the fairness of dividing what is left. While others have written better or at least more comprehensively on this issue, we do wish to summarize some few aspects of this issue.

24.3 What We Need to Do: A Biophysical Plan for a Sustainable Future

The conventional wisdom, consistent with Raskin's *Conventional Worlds* scenario, suggests that we can reach sustainability without fundamentally changing ourselves or our institutions. Those in the wealthy, industrialized, world can maintain their energy-intensive and elevated levels of consumption merely by means of technological change. For example, we can

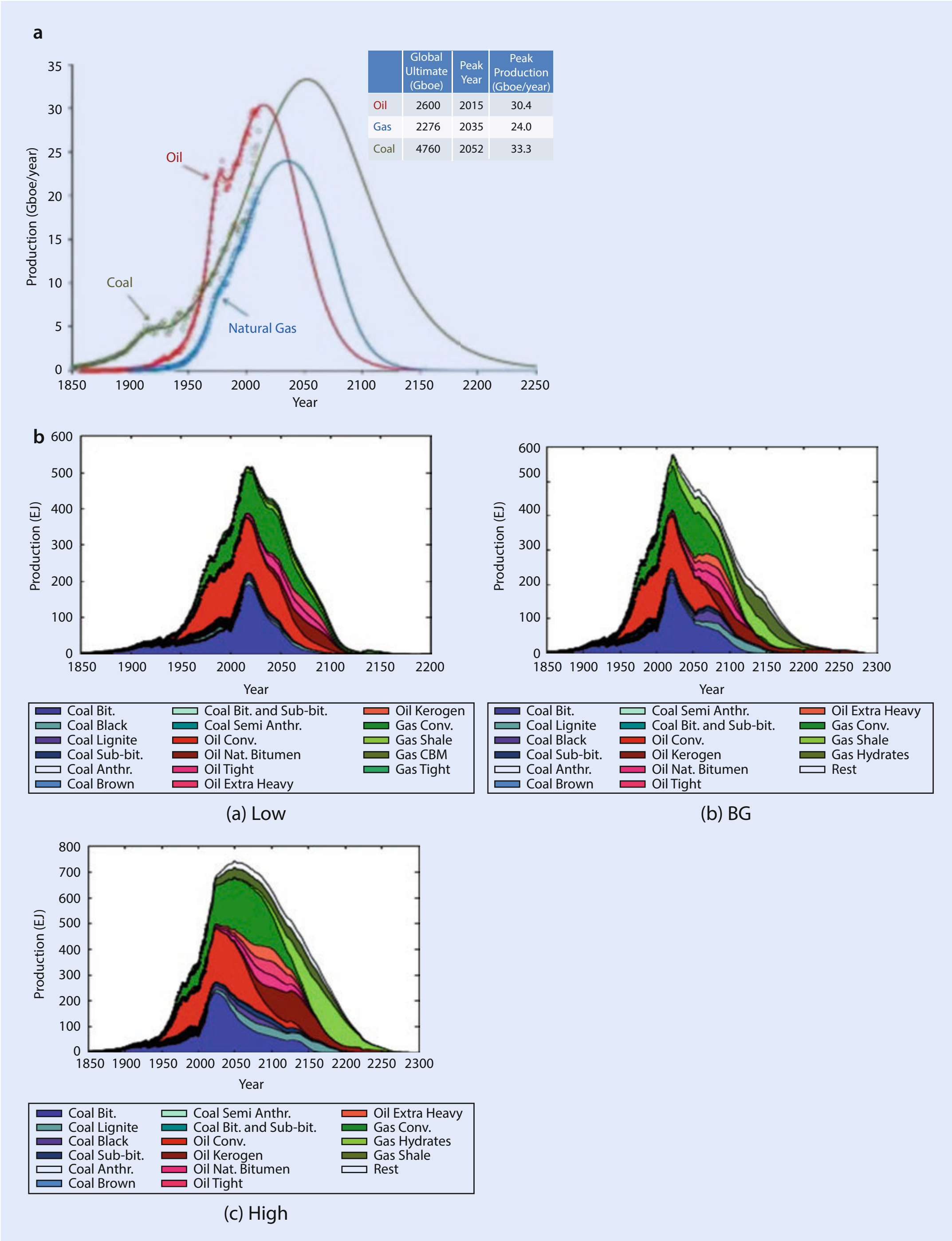


Fig. 24.3 Two new estimates of fossil fuel supplies for the future: **a** from Maggio and Cacciola [16]; **b** low, medium, and high estimate from Mohr et al. [13]

continue to grow as long as we use renewable energy. For the most vociferous adherent of the free-market approach, this will be assured as long as regulations that stifle entrepreneurial innovations are removed. The idea that a mass consumption society represents the zenith of human development is deeply entrenched in the American psyche, perhaps most explicitly enunciated by Walt Whitman Rostow in *The Stages of Economic Growth: A Non-Communist Manifesto* [19; see criticism by Thorstein Veblen]. This idea that humans are capable of dominating nature is itself long lived, perhaps dating to Jehovah's encouragement to the ancient Hebrews to "be fruitful, multiply and subdue the earth." This Promethean view is a fundamental tenant in western philosophy from Francis Bacon to present-day Marxists, who argue that technological change will be sufficient to overcome declining resource quality and degradation of Earth systems [20].

But "subduing the Earth" without violent repercussions is a large and imposing task. The perspective from conventional wisdom is that incremental changes, reliance on abiding faith in technology, and a belief that the necessary changes will somehow be found by compromise. Unfortunately the science, partially summarized in ▶ Chap. 23, leads one to the conclusion that such incremental steps will be insufficient to cope with the recurrent and potentially destabilizing crises that seem to be accelerating in the global system. The task is made all the more difficult by the existence of nonreversible tipping points that we do not comprehend fully at this time. So far, the degradation of natural and social systems has simply overwhelmed our piecemeal attempts at reform within the system [21]. We are especially concerned about the poorly understood connections between biophysical processes and social systems, as brilliantly laid out by Ahmed [22].

Scientists, especially natural scientists, are often uncomfortable about prescribing policy alternatives. But our job is not always over when we publish a book or an article in a respected journal. We need to confront the messy arena of human volition, as well as the more ordered world of the controlled laboratory experiment. In order to attempt to achieve something called sustainability the best we can do is to provide some suggestions. They are derived from our analysis of nature and of the economy, and they reflect our idea of what would make a good society.

Howard Odum was our mentor and guide, and we respected his contributions to systems analysis, ecological modeling, ecological energetics, and an understanding of the relation of humans to nature and to energy enormously. He understood how the world worked in so many fundamental ways. So it is fitting to choose the title of his last book "A prosperous way down" as a guide for where we should be going.

Odum believed that a lower-energy future was inevitable as fossil fuels peaked and declined. He did not write too much about the details, for to him it was just a fact. But he was interested in how humans might respond to this. He believed that a lower-energy future could be a good future, even as its title indicates a prosperous time. The authors of this book agree, for as we said we grew up in the United States during a time when per capita US energy use was only about a quarter or a third of what it is now. Our childhoods were great, our parents drove us almost nowhere (except occasional family vacations), and whatever we wanted we had to get for ourselves. If we wanted to be somewhere else, we got on our bicycles and pedaled there. If we wanted to play sports (which we did nearly every day), we joined our neighborhood friends and played whatever the season dictated on local school fields or sandlots. We had plenty of friends within walking or certainly biking distance because the automobile did not isolate us from our neighbors. There were plenty of places for Charlie to fish in and to explore and for Kent to swim and even surf. Charlie grew up on fresh vegetables his father grew and fish he caught locally. Life was good, even idyllic. Our houses were not opulent to say the least, our parent's cars (one per family) were not bought new and were not driven many miles in a year, and the only place we took vacations was to go to see relatives, who did not live too far away.

So here are some aspects that might actually be better in an energy-constrained world but only one where people had adjusted well to this new reality.

First, is wealth as measured by GDP necessarily something that leads to happiness? In fact, this has been studied (which is not easy) considerably. The answer is yes, but that other things are more important. For example, Richard Layard of the London School of Economics found a peak in US happiness in 1956, which is not too different from the results that the NGO Redefining Progress came up with a "genuine progress indicator" that found a peak for the United States in 1977 (◘ Fig. 24.4). Inglehart and Inglehart and Klingemann (and others) [e.g., 17] have measured subjective estimates of happiness in the world and found that after a given minimum level of income, there was no correlation between either income or long-term growth in income and personal happiness (Fig. 24.5). The countries with the largest number of happy people—Ireland, Nigeria, Mexico, and Venezuela—were certainly not the wealthiest, and the countries with the least number of self-described happy people, Russia, Armenia, and Romania, were not among the poorest. Instead, happiness seemed to depend a great deal on a sense of personal freedom and control over one's life. The "Eurobarometer" ranking of the happiness index, that is, how much people enjoy their life as a whole on scale 0 to 10, again found little correlation with GDP. Here are the rankings from this study: Colombia 8.1, Denmark 8, Malta 8, Switzerland 8, Iceland 7.8, Ireland 7.8, Ghana 7.7, Canada 7.6, Guatemala 7.6, Luxembourg 7.6, the United States 7, France 6.6, Nigeria 6.5, Bulgaria 4.5, Russia 4.4, Belarus 4.3, Georgia 4.1, Georgia 4.1, Armenia 3.7, Ukraine 3.6, Moldova 3.5, Zimbabwe 3.3, and Tanzania 3.2. So, overall, the answer to this question appears to be that some level of wealth, as measured by GDP, is a necessary component of personal happiness if you are poor but has little importance above some minimum level (Fig. 24.5). We can start educating our young people to this perspective now.

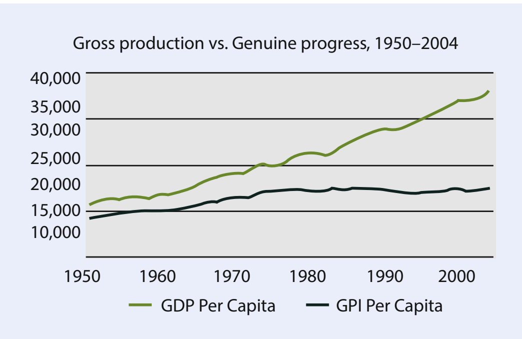
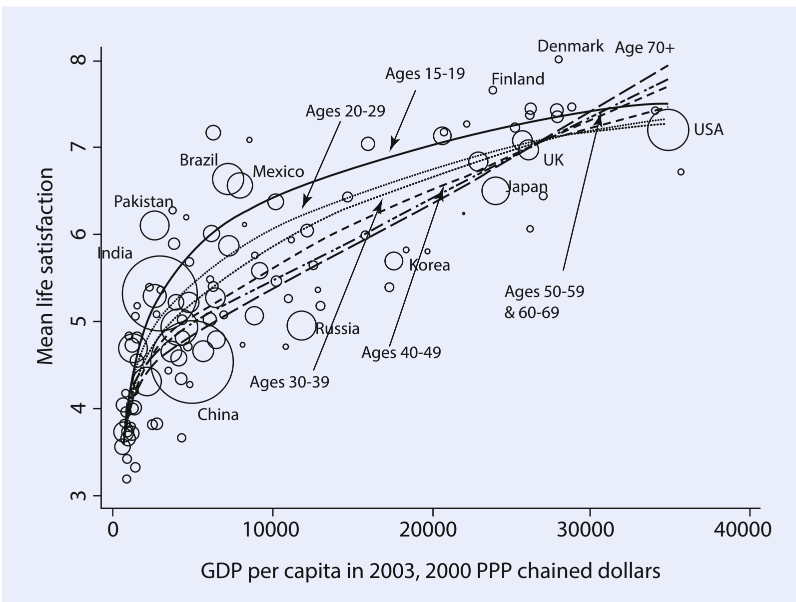


Fig. 24.4 The "genuine progress indicator" has remained constant, while the official estimates of GDP have increased substantially

Second, there are many indications that a less energy-intensive lifestyle can be one of much greater community. This is the explicit objective of various grassroots groups such as "The New Road Foundation" and "the evolution of transition" groups [18, 19], where transition means a transition to a post-peak oil world. Surely our present success-driven, affluence-seeking, status-driven world is not one that generates the greatest happiness and respect for others.

Third, our economy is so wasteful that it should be easy to use only half as much energy and maintain something very much like the same lifestyle. For example, our railroads could be electrified, generating less energy-intensive freight transfer [17]. Sedans that deliver essentially the same services on half the gasoline can

Fig. 24.5 Asymptotic relation of happiness and wealth



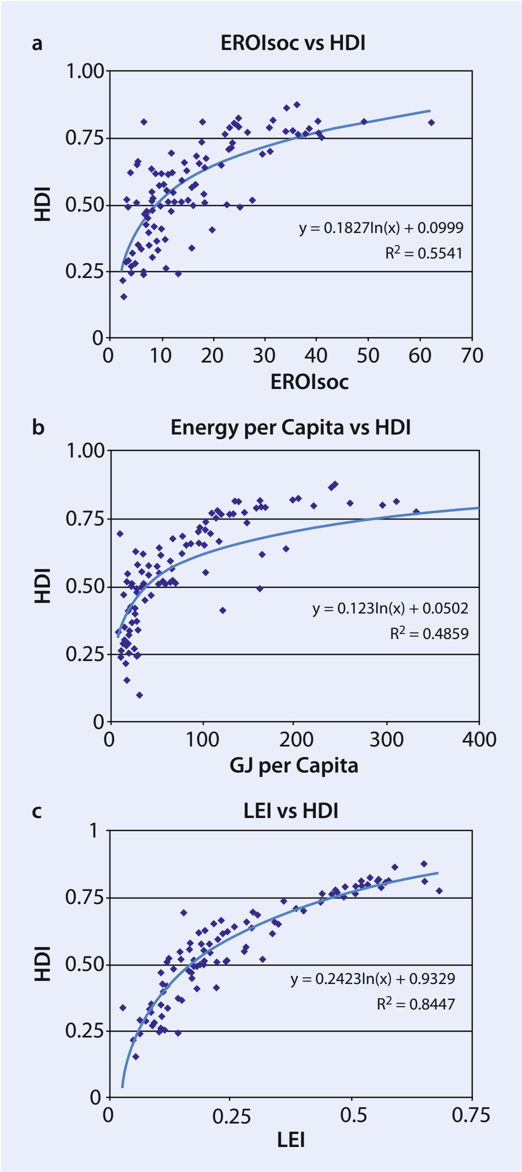


Fig. 24.6 Asymptotic relation of three indices of energy use and human welfare (From Lambert et al. [25])

easily be used, and older buildings can be retrofit with insulation (Fig. 24.5).

Another analysis that shows that human welfare becomes asymptotic with increasing wealth, or in this case energy use, is provided by examining the HDI, the human development index (devised as an alternative to GDP as an index of human well-being) vs. an index of energy use (Fig. 24.6).

24.3.1 What We Must Do if We Are to be Truly Sustainable

We are besieged nearly daily by many different "green" plans that promise, usually through some kind of technology or improvement in efficiency, sustainability, or at least progress in that direction. There is a certain logic and appeal of such plans because they offer indefinite "sustainability" with less impact on the Earth or the supplies of its critical resources.

Unfortunately, we believe that most such technologies are in fact counterproductive because of some manifestation of Jevons' paradox. Stanley Jevons originally believed that given the ultimate depletion of England's coal it was necessary to make the machines that used it more efficient [24]. But, in fact, he found that in the past such efficiency improvements made the use of steam engines cheaper so that more uses were found for them and technical changes designed to save coal actually ended up causing more coal to be used. More recent examples are that more efficient automobiles have led to more miles driven, more efficient refrigerators to larger refrigerators, more insulation to larger houses, and so on. Even cheap solar energy, should that be obtainable, allow the continued exacerbation of all the global problems given in Figs. 23.4 and 23.5. While we do think that efficiency improvements of many sorts certainly do have their place, *they must be implemented within the context of constraints of total use, or they are likely to be counterproductive*.

Thus to continue as a species with reasonable prospects for a decent life for the next hundred years, humanity must do two very difficult things. We must learn to live in harmony with nature and with one another. Neither of these can be obtained in a world where growth of human populations or human economies is the goal or indeed is even allowed. In order to move towards these goals, the changes given in Table 24.1 need to be implemented. Suggestions 1–3 involve ending growth and fundamentally altering present social relations, and are likely to be seen as our most controversial suggestions. Nonetheless, if we do not make these changes *first*, then other changes will probably be ineffective. Suggestions 4–5 focus on technological change, which can be effective if changes 1–3 are operational. Suggestions 6–9 involve transforming ourselves.

Table 24.1 Actions required for true sustainability

Fundamental social changes
1. Stabilize population growth
2. Stabilize economic growth
3. Create a more just distribution of income
Some technological changes
4. Improve energy efficiency
5. Move towards renewable power
Changing ourselves
6. Raise consciousness about embodied energy
7. Truth in labeling to include energy and resource requirements
8. Restore the dignity of meaningful work
9. Adjust our expectations

The fundamental social changes that we see as necessary are the direct opposite of the conventional wisdom regarding economic and political objectives and goals. However, we see most conventional sustainable objectives by others as also eventually doomed to failure without also achieving these goals. We are not naïve enough to believe that the world is about to abandon its growth mania, but believe without that, any economic policy will be insufficient to produce as society in which we can live well within nature's limits. Stated frankly, we have exceeded nature's limits as of today. We must shrink to live within them. At the same time, a market economy must grow to sustain capital accumulation. It is difficult to attain both goals for we cannot grow our economy and shrink our impact at the same time. While it is highly unlikely that the conventional political process itself will institute these changes, it is very likely that nature will make them for us, as may be occurring already (see Figs. 24.3 and 4.6).

Stabilize and reduce population. As can be seen from Figs. 23.1–23.5, humanity has already exceeded crucial planetary boundaries and is rapidly approaching even more. The stability of the Holocene climate most probably cannot withstand the continued growth of socioeconomic systems and the exponential degradation of the Earth systems. Moreover, the ability to feed more than seven billion people depends largely on fossil fuels. The Green Revolution commenced when there were only about 3 billion people on the planet. If our ability to use 10 calories of fossil energy per calorie of food disappears, then we can feed only the number living before the worldwide commitment to fossil agriculture. We cannot get to that number, in the absence of mass starvation or genocide, unless we voluntarily control fertility. We believe that voluntary control of fertility is a vastly superior alternative to the more Malthusian options of mass starvation and genocide. At the same time, we do not expect this process to be smooth and stable. The individual right to conceive and raise children is among the most dearly held of human rights – but it is enormously detrimental to Earth and the human population. Developed nations that have reduced their population growth rates below replacement rate have witnessed a declining and aging population. This leads to its own problems. In the long term, this means a smaller, and most often a less affluent, population of the young must try to support a growing population of the elderly that can no longer work as they once did. If, as the cheese slicer model implies, more of our national income must be spent to acquire energy, where will we get the money to support our old in the absence of economic growth?

Stabilize and reduce economic activity. Even if we reduced the world's population by eliminating the poorest half, the impact upon climate and other planetary boundaries would be minimal. Nearly all the impact comes from already existing rich nations, and from the rapidly industrializing nations, such as India and China. Put simply: a system in overshoot cannot grow its way into sustainability. The signs of the human economy are everywhere. Wealthy nations use up 3–5 planets' worth of resources to maintain their lifestyles. Every location where hydrocarbon development occurs is an environmental sacrifice area. As we have seen in ► Chap. 11, every increment of economic activity requires a more or less proportional increase in energy use, most often with an additional release of climate-modifying gases. The weight of plastics will exceed the weight of fish in the world's oceans by 2050.We can only live within nature's limits by shrinking the economy, then maintaining the smaller economy indefinitely. It is important to note that the Limits to Growth studies also could not generate a stable future without eliminating investments: i.e. growth.

Create a more just distribution of income. One way to meet the requirements of the poor of the world and the retirees mentioned above is through more equitable distribution of such wealth as is produced. More generally, a sustainable society must be a just society. A world in which a small wealthy elite live so opulently as islands of prosperity in a sea of misery cannot persist indefinitely. Attempts to create a fortress world will cause social breakdown and barbarism. We are already seeing this everyday. While it is called "terrorism" and blamed on "the other" with different customs and religion, the basis of social breakdown in unequal access to energy needs to be explored to a much greater degree [22].

Improve energy efficiency, but within constraints. Our present system of burning fossil fuels in concentrated locations and transmitted over long distances is pressuring our remaining fuel sources, is environmentally destructive, and creates unjust and inequitable access to energy. At the same time, social mechanisms must be put in place to avoid Jevons' Paradox, whereby increases in efficiency lead to greater resource use. We are unaware of any technological change in the twentieth century that did not improve "efficiency" without also increasing resource and energy use.

Along with increased efficiency, we must move towards renewable power. We will have neither the availability of high-quality fossil hydrocarbons at reasonable cost, nor the assimilative capacity of the atmosphere to accommodate the fossil economy for more than the next half century, if that. Moreover, since the construction of the solar economy depends upon fossil fuels to produce and move the wind turbines, concrete pads upon which to locate them, and the photovoltaic panels to generate electricity, we need to start now, and not wait until fossil fuels are in desperately short supply.

Raise consciousness about embodied energy. Few people living in wealthy, energy-intensive societies think about the energy embodied in their day-to-day actions. How many extra tons of carbon are emitted when an able-bodied person uses the electric door opener mandated for the handicapped, or when one does not turn off their computer at night? How many people have actually calculated the volume of water used in a shower, or the amount of electricity needed to run the pumps, or the fuel used to heat the water?

Truth in labeling. Along with calories ingested by consumption on food labels, we should include calories used to produce the foodstuff. Energy returns on investment should be displayed explicitly on all consumer products.

Restore the dignity of meaningful work, which allows each and every worker to combine the brain work with the manual work to produce something of value that improves society. Although this will raise the price of consumer goods, it will also go a long way in reducing inequality and waste. Few psychologists believe that more consumption leads to more happiness, beyond a minimum of survival. The community of meaningful work among associated producers could easily produce a happier society, even if this means longer hours of physical labor. The human body was not designed to sit behind a screen for long hours. Get moving!

Adjust our expectations. We cannot conspicuously consume our way into happiness. In the United States, only about 1–2% of our energy is produced by renewables. Would the elements on the Periodic Table exist in sufficient quantities to produce the same level of output for all people of the Earth that citizens of the wealthy nations now consume? We doubt it. Perhaps we need to realize that our comfort, convenience, profits, and income are not as important as the proper functioning of the Earth's biophysical systems.

► Chapter 23 showed that many of our socio-economic and earth systems are already in overshoot, and a system in overshoot simply cannot grow its way into sustainability. Yet our present economic system requires continual economic growth in order to maintain employment and provide income. We are convinced that we will not achieve sustainability simply by recycling more. We must transform the economic system from one that is growth dependent to one that can provide a decent standard of living without growth. John Bellamy Foster enunciated our challenge well when he said:

> » To achieve these things we will need to break with "business as usual," that is, with the current logic of capital, and introduce an entirely different logic, aimed at the creation of a fundamentally different social metabolic system of reproduction [20].

24.3.2 Why We Are Not Entirely Optimistic

While we believe that a relatively smooth transition to a lower-energy future with a good lifestyle is quite possible, we are not necessarily optimistic that it will occur. The first reason is that the American public is almost completely ignorant about peak oil—which indeed is the simplest part of the dilemma, "the energy mess," that we have inherited [21]. Quite curiously, neither the press nor the national funders of science (NSF, DOE, etc.) have any particular interest in this issue and, if anything, have attempted to suppress any research or discussion on the subject [22]. This is quite surprising considering the enormous amount of attention given to possible climate change. While we wish in no way to belittle the importance of the attention paid by both the press and the science community to climate change, we find it curious that peak oil, a situation that seems to be more immediate, more certain, and perhaps more devastating, receives essentially zero press or funding, at least as of 2017. As part of this problem, the public is fed a constant stream of advertisements and programs promising green clean energy when the quantitative nature of the contributions—all trivial—is never mentioned. Likewise, the energy cost of so many "green" things, from trips to ecotourism sites to LEED buildings, is rarely mentioned.

A second reason we are not optimistic is that Americans (and most others in the world) have been conditioned by a lifetime of television and other advertisements all indicating that happiness and sexual fulfillment, you name it, are possible only through a never-ending stream of purchases. This seems to be so ingrained in our culture and our economy that it is hard to imagine it otherwise.

A third important reason we cannot be too optimistic that we will make the needed transition will be the political response to this situation. This of course requires that people understand what is happening and that the political situation can adjust to this new reality in a reasonable way. There are many thoughtful papers that have attempted to examine the potential transition in various and often quite sophisticated ways [21, 23]. All agree that a critical first step is to question a belief in, and policies attempting to promulgate, growth. How this can be undertaken in the current political climate where even far less controversial legislation is stalled is beyond our comprehension. Possibly, peak oil will put some sense into the electorate's head, but more likely there will simply be a blame game for the fact that no political parties can bring back the good old days where the American dream was realized for generation after generation. If there is to be a new American dream, it has to be based on something besides ever more affluence, and that will be tough. But there are simple things we can start doing. Two simple things are to live near where you work and contribute to making sure your neighborhood, and neighborhoods in general, provides the necessities of life to decrease your and our dependence upon automobiles. We like the ideas of Will Allen (Growing Power, Inc.) and others to bring agriculture into the central cities.

We do not believe that simply by "doing simple things" we can save the Earth, nor do we believe that technology alone will save us. We have to do big and complicated things if we want the planet of the future to be similar to the one on which our species evolved. If we do not achieve a stabilization of growth and justice, then the rest of our suggestions will not matter very much. These are difficult and complex changes which will require a fundamental reordering of economy and society.

Another reason that simple changes or some magic technology will not, by themselves, produce sustainability is the need for perpetual economic growth in a capitalist economy to produce profits, avoid poverty, and reduce unemployment. If individuals live within nature's limits the planet, their lives, and especially those of their progeny, will be better off. Yet the economy may collapse from the reduced consumption. While legions of economics teachers implore their students to believe that capitalism is about efficiency, without copious amounts of waste enough economic surplus could not be absorbed to maintain prosperity.

Thus, a good future and even a prosperous way down are, we believe, quite possible for economic and political reasons but very unlikely due to psychological and conditioning issues relating to the attitude of the American people relating to advertisement, growth, and wealth as status. We conclude that what we need most is to create a biophysically based approach and model for economics, one that would serve on at least equal footing with the present firm-household-market-based model. This is our next project and the annual meetings of the International Society for BioPhysical Economics is one important place to start.

24.3.3 Why We Have Reason to Be Optimistic

To begin with, there are a lot of very smart people who are working on these problems. They range from academics to those in nongovernmental organization to political activists. The Occupy movement did more to raise questions about income distribution than did the sum of peer-reviewed academic journals. The Women's March for Science was just one example of how the recent election shocked many people out of their complacency. Student organizations such as Power Shift are raising the issues with today's students about things their lives will depend upon and urging them into action. Organizations such as *Via Campesina* and the Unity Council of the Cayuga Nation are showing those of us in the global North that our ways are not necessarily the ways. Paul Raskin of the Tellus Institute put it well. "The future will depend upon decisions that have not yet been made." The most recent recipient of the Global Development and Environment Institute's Leontief Prize, Joan Martinez-Alier, said that the alternatives will emerge in the struggle. We do not know what sustainability will look like, but we know what it will not look like. A sustainable society will not be the business-as-usual strategy of globalized monopoly finance capitalism guided by neoclassical economics: perpetual growth and resource depletion, no matter the consequences. Neither will it be the top-down Stalinist repression of Soviet-era heavy industry. But there is a lot of room in between or perpendicular to these two poles with many options. The time is now to start exercising them. It might be helpful to remember the words of Margaret Mead. "Never doubt that a small group of thoughtful citizens can change the world; indeed, it's the only thing that ever has." We hope the analyses in our book can help guide you down the path of living the best life possible in a resource-constrained world.

We can envision a future of a stable economy using half the resources of today, but sufficient to provide basic dignity for all while maintaining incentives. However, this cannot be done within the confines of conventional economics and our present social order. What we have presented in this book, BioPhysical Economics, provides the logic and tools to begin the transition to a just and truly sustainable world.

? Questions

1. Are you an optimist or a pessimist about the future? Why? About what?
2. What is Maslow's hierarchy of human needs? Can you list them in order?
3. What are some ways that we can make more jobs available for labor? What would be some good and some bad sides to that?
4. Name five ways that food production depends upon oil.
5. What are your views about the future of coal in the world economy? What factors might be especially important in influencing this?
6. Do you think that GDP is an adequate measure of our wealth? Why or why not?
7. What are some of the advantages that might come from a less energy-intensive lifestyle?
8. What ideas do you have to provide for a better future for all Americans and all people of the world?

References

1. Kasser, T. 2003. *The high price of materialism*. Cambridge, MA: The MIT Press.
2. Galbraith, J.K. 1996. *The good society*. Boston: Houghton Mifflin Company.
3. Raskin, P., T. Banuri, G. Gallopín, P. Gutman, A. Hammond, R. Kates, and R. Swart. 2002. *Great transition*. Boston, MA: Stockholm Environment Institute.
4. Jacobson, M.Z., M.A. Delucchi, M.A. Cameron, and B.A. Frew. 2015. *Proceedings of the National Academy of Sciences of the United States of America* 112: 15060–15065. See also Clack, C., et al. (2017). Proceedings of the national academy of sciences USA 114, 6722–6727; Sole, J., García-Olivares, A., Turiel, A., Ballabrera-Poy, J. (2018). Renewable transitions and the net energy from oil liquids: A scenarios Study. Renewable Energy 116 (2018) 258–271.
5. Maslow, A. 1943. A theory of human motivation. *Psychological Review* 50: 370–396.

6. Hall, C.A.S., D. Lindenberger, R. Kummel, T. Kroeger, and W. Eichhorn. 2001. The need to reintegrate the natural sciences with economics. *Bioscience* 51: 663–673.
7. http://www.regjeringen.no/en/dep/fin/Selected-topics/the-government-pension-fund.html
8. Hall, C. 2000. *Quantifying sustainable development: The future of tropical economies*. San Diego: Academic Press. LeClerc, G., & Hall, C, (2008). Making development work: A new role for science. Albuquerque: University of New Mexico Press.
9. Vaccari, D. 2009. Phosphorus famine: The threat to our food supply. *Scientific American* 36: 54–59.
10. Feiffer, D.A. 2003. *Eating fossil fuels*. Sherman Oaks: Wilderness Publications.
11. Bartlett, Albert. 1997-98. Reflections on sustainability, population growth and the environment - revisited. *Renewable Resources Journal* 15 (4): 6–23.
12. Patzek, T., and G. Croft. 2010. A global coal production forecast with multi-Hubbert cycle analysis. *Coal Energy* 35: 3109–3122.
13. Mohr, S.H., J.L. Wang, G. Ellem, J. Ward, and D. Giurco. 2015. Projection of world fossil fuels by country. *Fuel* 141: 120–135.
14. Rutledge, D. 2010. Estimating long-term world coal production with logit and probit transforms. *International Journal of Coal Geology* 85: 23–33.
15. http://eclipsenow.wordpress.com/2010/05/06/peak-coal-hits-china-richard-heinbergs-article/l
16. Maggio, G., and G. Cacciola. 2012. When will oil, natural gas, and coal peak? *Fuel* 98: 111–123.
17. Freidemann, A. 2016. *When the trucks stop running: Energy and the future of transportation*. New York: Springer.
18. Jacobsen, M., and M.A. Delucchi. 2011. Providing all global energy with wind, water, and solar power, part I: Technologies, energy resources, quantities and areas of infrastructure, and materials. *Energy Policy* 39 (2011): 1154–1169.
19. Rostow, W.W. 1963. *The stages of economic growth: A non-communist manifesto*. Cambridge: Cambridge University Press.
20. Foster, J.B. 2017. The long ecological revolution. *Monthly Review* 69 (6): 1–16.
21. Raskin, P. 2014. *A great transition? Where we stand*. Boston: Great Transitions Initiative.
22. Ahmed, N. 2017. *Failing states, collapsing systems. Biophysical triggers of political violence*. Springer.
23. Klitgaard, K. 2017. *The struggle for meaningful work*. Boston: Great Transitions Initiative.
24. Jevons, S. 1865. *The coal question. An inquiry concerning the progress of the nation and the probable exhaustion of our coal mines*. London: MacMillan and Company.
25. Lambert, Jessica, Charles A.S. Hall, Stephen Balogh, Ajay Gupta, and Michelle Arnold. 2014. Energy, EROI and quality of life. *Energy Policy* 64: 153–167.

Supplementary Information

C.A.S. Hall, K. Klitgaard, *Energy and the Wealth of Nations*,
https://doi.org/10.1007/978-3-319-66219-0

Index

F

G

H

I

U

V

W

Y