Energy
for the
21st Century

Energy for the 21st Century

A Comprehensive Guide to Conventional and Alternative Sources

Roy L. Nersesian

M.E.Sharpe
Armonk, New York
London, England

Copyright © 2007 by M.E. Sharpe, Inc.

All rights reserved. No part of this book may be reproduced in any form
without written permission from the publisher, M.E. Sharpe, Inc.,
80 Business Park Drive, Armonk, New York 10504.

Library of Congress Cataloging-in-Publication Data

Nersesian, Roy L.
Energy for the 21st century : a comprehensive guide to conventional and alternative sources / Roy L. Nersesian.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-0-7656-1323-3 (cloth : alk. paper)
ISBN-10: 0-7656-1323-9 (cloth : alk. paper)
1. Power resources. 2. Energy policy. I. Title. II. Title: Energy for the twenty-first century.

HD9502.A2N465 2007
333.79—dc22 2006011880

Printed in the United States of America

The paper used in this publication meets the minimum requirements of
American National Standard for Information Sciences
Permanence of Paper for Printed Library Materials,
ANSI Z 39.48-1984.

∞

IBT (c) 10 9 8 7 6 5 4 3 2 1

To Friends of the Family:
Daisy, Rue, Heather, Ginger, and Quincy

Contents

Chapter 7. Natural Gas 225

List of Tables and Figures

Tables

Figures

Preface

Many of life's greater expectations remain unfulfilled, but on a more mundane level our lives are ones of fulfilled expectations. We expect that the radio alarm will awaken us in the morning, that there will be light when we turn on the light switch, that water will flow when we turn on a faucet, that the food in the refrigerator will still be cold, that the burner on the stove will fry an egg. We expect a newspaper at our front doorstep and something to look at if we turn on the television. If we commute, we expect that the bus or train will be running, or that the car will run and that the gas station will be open for refueling. In our office, we expect that the mail will be delivered, that the computer boots up when we turn it on, that our e-mail has received all messages that had been sent to us, and that there is a dial tone when we pick up the telephone. There is an endless list of fulfilled expectations that are necessary for our modern way of life to continue for another day.

All this depends on energy: electricity for the lights, refrigerator, computer and communications; natural gas or electricity for the stove; gasoline or diesel fuel for the car, bus and train; jet fuel for the airplane; heating oil or natural gas to heat a home or a building. Electricity itself is derived for the most part from burning coal, natural gas, and oil, and to a lesser extent, nuclear and hydropower. A rather miniscule, but growing, contribution is made from alternative sources such as wind, solar, geothermal, biomass. Of these, wind has made the most progress in becoming a meaningful alternative supply of energy, but most of its progress has been limited to Europe, and the hope is that at least 10 percent of European electricity will be from wind. This still leaves the bulk of energy demand for electricity generation to be fulfilled by conventional means. Despite all the hoopla, the hydrogen economy, the Green Answer to the world's burgeoning energy needs, has far to go technology-wise before it is commercially feasible.

This book examines the role of the principal sources of energy both in the aggregate and by specific types (biomass, coal, oil, natural gas, hydro and nuclear power, and sustainable sources) for a balanced view on energy. Nations exhibit enormous variance in energy consumption both in amount and the degree of reliance on different types of energy. Moreover their energy plans to satisfy future needs vary markedly. Energy diversity on a national level is too great for the global community of nations to adopt a common policy approach to energy. The only international convention that bears on energy is the Kyoto Protocol. However, one can argue that the Protocol is more environmental in orientation in that its primary concern is reducing greenhouse gas emissions. However, in complying with the Protocol, nations will tend to favor natural gas, wind, and solar over coal and oil.

Having divergent and perhaps mutually exclusive energy policies prevents integration into a single, coherent, and consistent policy toward energy; the world will have to live with a portfolio of energy policies that fit each nation, not one that applies globally. While it may be possible to develop regional energy policies, such as the European Union or North America, even here there is a great deal of divergence among the individual nations as to their dependence on various types of energy.

I have tried to present a balanced view on energy without succumbing to the temptation to tell one side of the story. I have probably failed from time to time. In preparing to write the book I discovered to my amazement divergence of opinion rather than consensus on simple matters such as where does oil come from, the relationship between global warming and the rising concentration of carbon dioxide in the atmosphere, and whether we are running out of oil. My approach has been to try to represent both sides of a point. Although I showed partiality at times such as leaving out those who espouse that energy is infinite, consumption does not matter, and pollution, though a nuisance, is nothing to be concerned with.

I would like to express my gratitude to Richard Howard, who read the manuscript and made a number of suggestions, particularly in giving Gulbenkian, an individual who disdained oil exploration, production, refining, and distribution, his due in developing the oil industry. I would like to mention Neal Dougherty and John Altenau also as lifelong friends. I would like to thank Fred Kelly, Dean of the Business School at Monmouth University, the Monmouth University Business Council, and Hurst Groves, Director of the Center for Energy, Marine Transportation, and Public Policy at the School for International Affairs, Columbia University, for their support.

I accept all errors as my responsibility. However, I would like to warn the reader that there is a wide range of opinion on energy issues. Some are far from settled and others are more like questions begging for answers. Lastly I plan to make good the promise I made to Maria, my wife, that I will spend more time with her and less with the computer.

Energy
for the
21st Century

1

Are We on Easter Island?

Energy is a natural resource. Although exhaustion of fossil fuels (coal, oil, and natural gas) is not imminent, we have a history of responding to natural resources in danger of exhaustion. We exhausted forests in Europe at the start of the Industrial Age in our quest for making glass and metals and we nearly drove whales to the point of extinction during the nineteenth century in our quest for whale oil. Fortunately we found ways to avert what could have been a terminal crisis. The forests in Europe were saved from the axe by the discovery of coal as an alternative to wood in glass and metal making. The whales were saved from extinction by finding an alternative source for lighting in the form of kerosene. In the twentieth century we took effective action to rejuvenate a threatened species of marine animal life, but at the same time we discovered the technology to strip-mine the open oceans of fish life. As we exhaust open-ocean fishing, an alternative has been found in aquaculture or fish farming. Aquaculture is similar to relying on sustainable energy whereas open-ocean fishing, when fish are caught faster than they can reproduce, is similar to exhausting fossil fuels.

In the case of energy, it is true that immense energy reserves have been found that have kept up with our horrific appetite for energy, making mincemeat of Theodore Roosevelt's prediction that we will soon, from the vantage point of the early twentieth century, exhaust our natural resources. A key question facing us is whether the future pace of discovery can keep ahead of our growing appetite for energy; that is, will Roosevelt ultimately be proven right? Just because we run short of a natural resource does not necessarily mean that we can find an alternative. That is the tragedy of Easter Island.

Easter Island

Easter Island is over 2,000 miles from Tahiti and Chile. To the original inhabitants, Easter Island was an isolated island of finite resources surrounded by a seemingly infinite ocean. What happened on Easter Island when it ultimately exhausted its finite resources is pertinent because Earth is an isolated planet of finite resources surrounded by seemingly infinite space. Whether we admit it or not, we are in danger of exhausting our natural resources. Rough, and some deem optimistic, estimates of forty years for oil, eighty years for natural gas, and a few hundred years for coal are not particularly comforting when viewed from the perspective of civilization continuing for thousands of years. Like the Easter Islanders who had nowhere to go, this is our home planet now and for the foreseeable future. Space travel is a long way off and flying away to Mars to escape a manmade calamity on Earth is not a particularly inviting prospect.

Examination of the soil layers on Easter Island, or Rapa Nui to the present inhabitants, revealed an island with abundant plant and animal life for tens of thousands of years. Then around 400 CE the island was discovered and settled by Polynesians, who originally named the island Te Pito O Te Henua

(Navel of the World). The natives survived on the bounty of natural animal and plant life on the island and fish in the surrounding waters. Critical for survival was the Easter Island palm, which grew to eighty feet, provided sap and nuts for human consumption and wood to make canoes for fishing. The palms also provided the means to move the massive stone Moai, the stone figures for which the island is famous, who now stand in mute testimony to an ecological calamity that unfolded around 1500. By then the estimated population had grown to somewhere between 10,000 and 15,000 inhabitants.

This sounds like an awful lot of people descended from a few settlers, but this is the nature of exponential growth. If a party of ten people originally settled on Easter Island and grew at a relatively modest 1 percent per year (less than the current growth in world population), the number of Easter Islanders would double about every seventy years. There were nearly sixteen doublings of the population in the 1,100 years from 400 to 1500 CE. Double ten sixteen times and see what you get. In theory the population would have grown to 567,000, a mathematical consequence of compound exponential growth at 1 percent per year over 1,100 years. It would never have reached this level because, as proven in 1500, a population in excess of 10,000 was sufficient to exhaust the island's natural resources.

A growing population increased the demand for meat, which eventually led to the natives feasting on the last animal. More people and no animals promoted more intensive tilling of the land, which first had to be cleared of the palms. With fewer palms, erosion increased and, coupled with the pressure to grow more crops, soil fertility declined. Of course the palms did not go to waste as they were needed to support the leading industry on Easter Island: the construction and moving of the Moai plus, of course, making canoes. Fish became more important in the diet as the population grew, the animals disappeared, and crop yields fell. Around 1500 the last palm tree was cut down. Bloody intertribal warfare, cannibalism, and starvation marked the demise of a civilization.

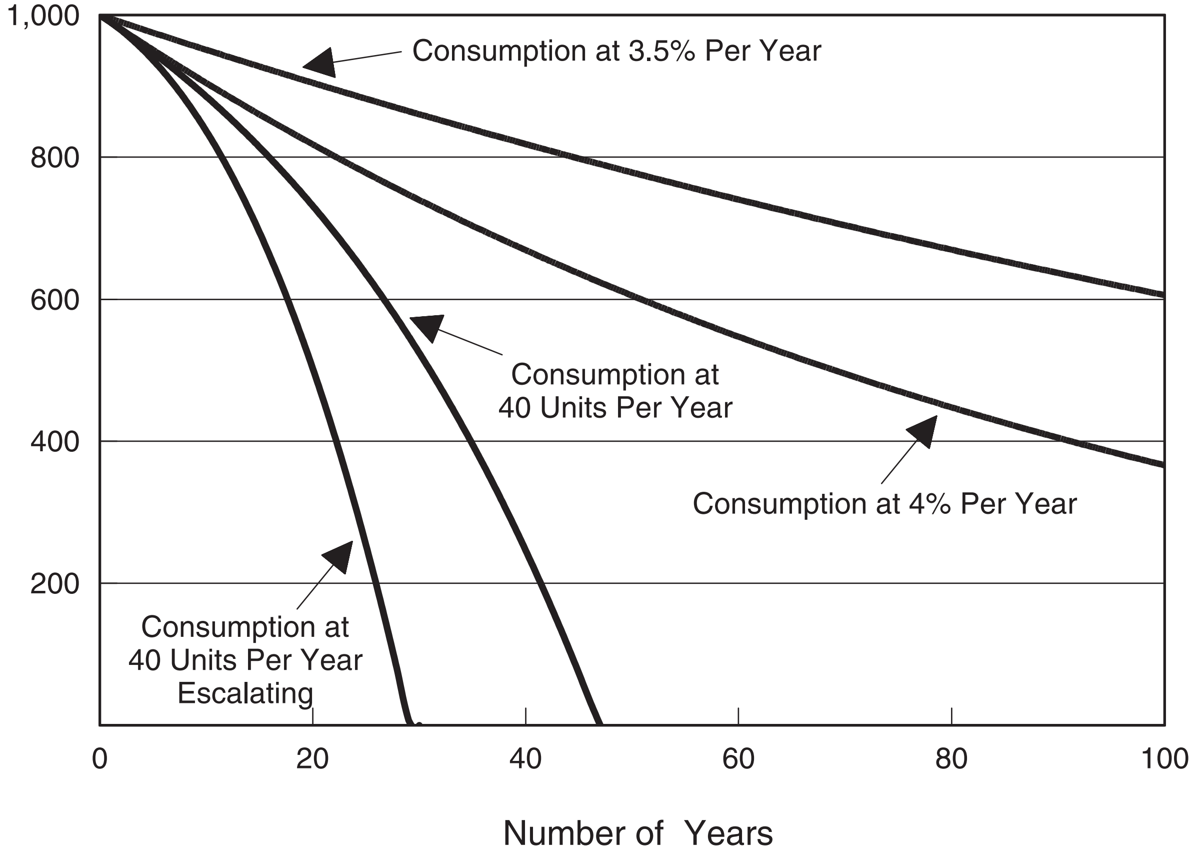
On Easter of 1722, the Dutch explorer Jacob Roggeveen rediscovered the island. The island presented a great mystery as the few surviving and utterly impoverished natives had no memory of the tragedy nor did they understand the meaning of the Moai. The gift of Western civilization—disease—ultimately reduced the native population to a remnant of 111 by 1800. In 1888, when the island was annexed by Chile and renamed Rapa Nui, the population had risen to 2,000 (a growth rate considerably in excess of 1 percent!).

The Mathematics of Extinction

Suppose that we depend on a forest for supplying wood for fuel and building material. If the forest grows at 3 percent per year and we remove 2 percent of the forest per year, the resource will last forever. The forest lasts forever if 3 percent is removed per year, but care has to be exercised to ensure that removal does not exceed 3 percent. If consumption exceeds 3 percent as a consequence of a growing population that needs more wood for fuel and shelter, or if a new technology is introduced that consumes a great deal of wood, such as glass and metal making, then the forest will eventually be consumed.

Suppose that a forest consists of 1,000 units of usable wood that increases naturally by 3 percent per year. Figure 1.1 shows what happens to the forest as a resource in terms of units of usable wood when consumption is 3.5 and 4 percent per year. In a century the forest has been reduced to 600 and 380 units, respectively. One hundred years in the history of humanity is not very long; for

Figure 1.1 **Exhausting a Resource (Supply Increasing 3 Percent Per Year)**



a growing minority, it is a single lifetime. But this does not accurately describe the situation. What is wrong with this projection is that demand declines in absolute terms over time. For example, in the first year the forest gains 30 units and consumption at 4 percent is 40 units, leaving 990 units for the next year. When the forest is down to 800 units, consumption at 4 percent has been reduced from 40 to 32 units.

This is not realistic; there is no reason for demand to decline simply because supply is dwindling. Suppose that consumption remains constant at 40 units with 3 percent growth in forest reserves. Then, as Figure 1.1 shows, the forest is transformed to barren land in forty-seven years. The final curve is the most realistic. It shows what would happen if consumption climbs at 1 unit per year; 40 units in the first year, 41 units the second, and so on. Now the forest is gone in thirty years, a single generation.

Yet, even this projection is not realistic. Consumption, initially growing by 1 unit a year, declines in relative terms over time. For instance, when consumption increases from 40 to 41 units, growth is 2.5 percent; from 50 to 51 units, growth has declined to 2 percent. Another curve could be constructed holding consumption growth at 2.5 percent, based on a starting point of 40 units per year. But the point has already been made: The resource is exhausted within a single generation.

Before the resource is exhausted, other mitigating factors come into play. One is price, a factor not at play on Easter Island. As the forest diminishes in size and consumers and suppliers realize that wood products are becoming increasingly scarce, their price would increase. The more serious the situation becomes, the higher the price. Higher prices dampen demand and act as an incentive to search for other forests or alternative sources for wood such as coal for energy and plastic for wood products (neither option available to the Easter Islanders).

Price would certainly have caused a change of some sort to deal with the oncoming crisis, but would not have affected the eventual outcome. A very high price for the last Easter Island palm would not have saved a civilization from extinction. The individual who became rich selling the last palm would have spent his last dime buying the last fish. Easter Island is not the only civilization that collapsed from a shortage of natural resources. It is believed that the fall in agricultural output from a prolonged drought caused the demise of the Mayan civilization in Central America. Ruins of dead civilizations litter the earth, a humbling reminder of their impermanence.

Progress Is Our Most Important Product

About one-third of the earth's population still depends on wood as a primary energy source. Unfortunately, removing forests to clear land for agriculture is often considered a mark of progress. Where the cleared land stopped marked the boundaries of the Roman Empire. Agriculture transformed war-loving hunter-gatherers into law-biding agrarians. The resuscitation of civilization during the Middle Ages was evidenced by forests and abandoned lands transformed into vineyards and other forms of agricultural enterprise by monks. Removal of forests to support a growing population became too much of a good thing. The first energy crisis occurred early in the Industrial Revolution when wood demand for housing, heating, and the new industries of glass and metal making exhausted a natural resource. A crisis turning into a calamity was averted by the discovery of coal in England and forests in North America.

The growth of the United States as a nation can be traced by the clearing of a large portion of the forest covering the eastern half of the nation for farmland. This was a visual sign of progress for pioneers seeking a new life in the Americas. Despite environment protestations to the contrary, clearing forests is still considered a sign of progress. We are intentionally burning down and clearing huge portions of the rain forests in the Amazon and in Southeast Asia for cattle grazing and other forms of agriculture.

Burning wood or biomass faster than it can be replaced by natural growth adds carbon dioxide to the atmosphere. Burning fossil fuels (coal, oil, and natural gas) releases carbon dioxide previously removed from the atmosphere by plant, animal, and marine life millions of years ago. The increasing concentration of carbon dioxide in the atmosphere is blamed on both the continuing clearing of forests and our growing reliance on fossil fuels. At the same time, clearing forests and consuming energy are signs of economic progress to raise living standards. While there is intense public pressure to reduce carbon dioxide emissions in Europe, Japan, Australia, and New Zealand, over half of the planet's population lives in nations in South America and Asia, where governments are doing everything they can to increase carbon dioxide emissions in pursuit of economic development.

Unlike the Easter Islanders, there are countervailing measures being taken to compensate for burning down vast tracts of the world's tropical forests. Carefully managed forest reserves in various parts of the world replant saplings and seeds after the removal of mature trees by lumber and papermaking companies. Tree farms are commercial agricultural enterprises that supply raw materials for the lumber and papermaking industries. These activities are driven by both environmental and commercial concerns over the long-term consequences of clearing trees without replacement. A number of public service organizations are dedicated to planting trees to combat the rising concentration of carbon dioxide in the atmosphere. An increasing carbon dioxide concentration in the atmosphere itself promotes plant growth that would be a natural countermeasure or negative feedback system. Yet despite human efforts to the contrary, the world's resource of forests continues to dwindle and the carbon dioxide concentration in the atmosphere continues to climb.

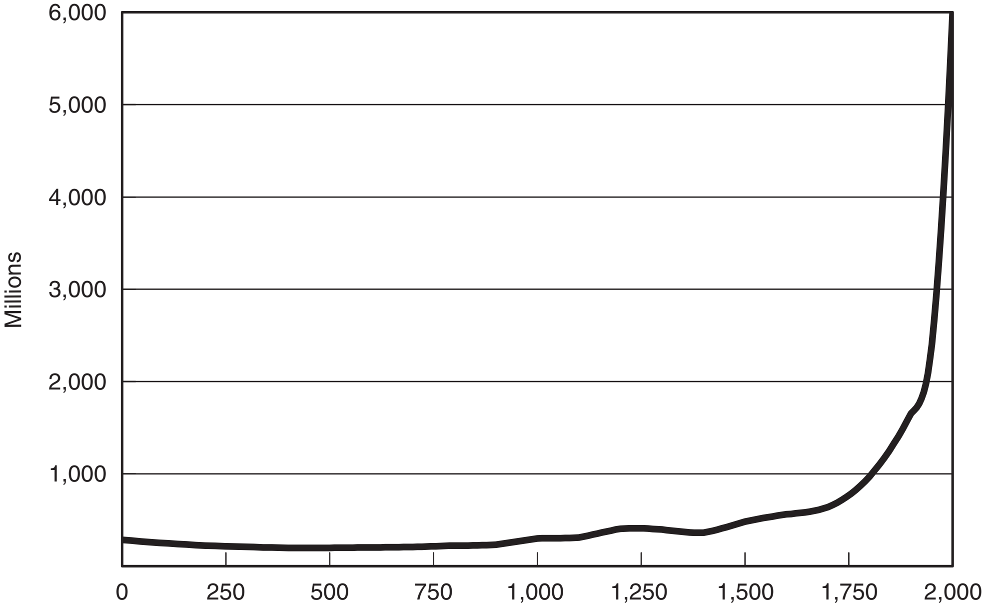
The Unremitting Rise in Population

Both energy usage and pollution can be linked directly to population. Indeed, there are groups who advocate population reduction as the primary countermeasure to cut pollution of the land, air, and water. These groups have identified the true culprit of energy exhaustion and pollution, but their suggested means of correcting the problem does not make for comfortable reading. Figure 1.2 shows the world population since the beginning of the Christian era and its phenomenal growth since the Industrial Revolution.[1]

The world's population was remarkably stable up to 1000 CE. The Dark Age of political disorder and economic collapse following the fall of the Roman Empire around 400 CE did wonders in suppressing population growth. The high death rate for infants and children and the short, dirty, brutish lives of those who survived childhood, coupled with the disintegration of society, prevented runaway population growth. After the Dark Age was over, the population began to grow until the Black Death, starting around 1350, with several subsequent recursions over the next hundred years, suppressed population growth. Somewhere between one-third and one-half of Europe's population was wiped out, and as much as two-thirds in certain areas. It took over a century for the population in Europe to recover to preplague levels.

The totality of human history was required to reach the first billion of population around 1840. The second billion was reached around 1930, only ninety years later, despite the horrendous human losses during World War I, the Russian Revolution, and Civil War, and the Spanish Flu, a pandemic that wiped out 30–40 million lives. This pandemic cost more in human life than the combined efforts of those involved with perpetrating war and revolution and numerically, but not percentage wise, exceeded the Black Death. Despite this, it only took thirty years for the world population to reach its third billion in 1960, despite Stalin's execution of tens of millions of his own people by starvation and firing squad, Hitler's extermination of 12 million Jews and Slavs, plus the deaths of untold military personnel and civilians during World War II. The fourth billion was reached fourteen years later in 1974, the fifth billion thirteen years later in 1987, and the sixth billion twelve years later in 1999.

Figure 1.2 **History of the Global Population**

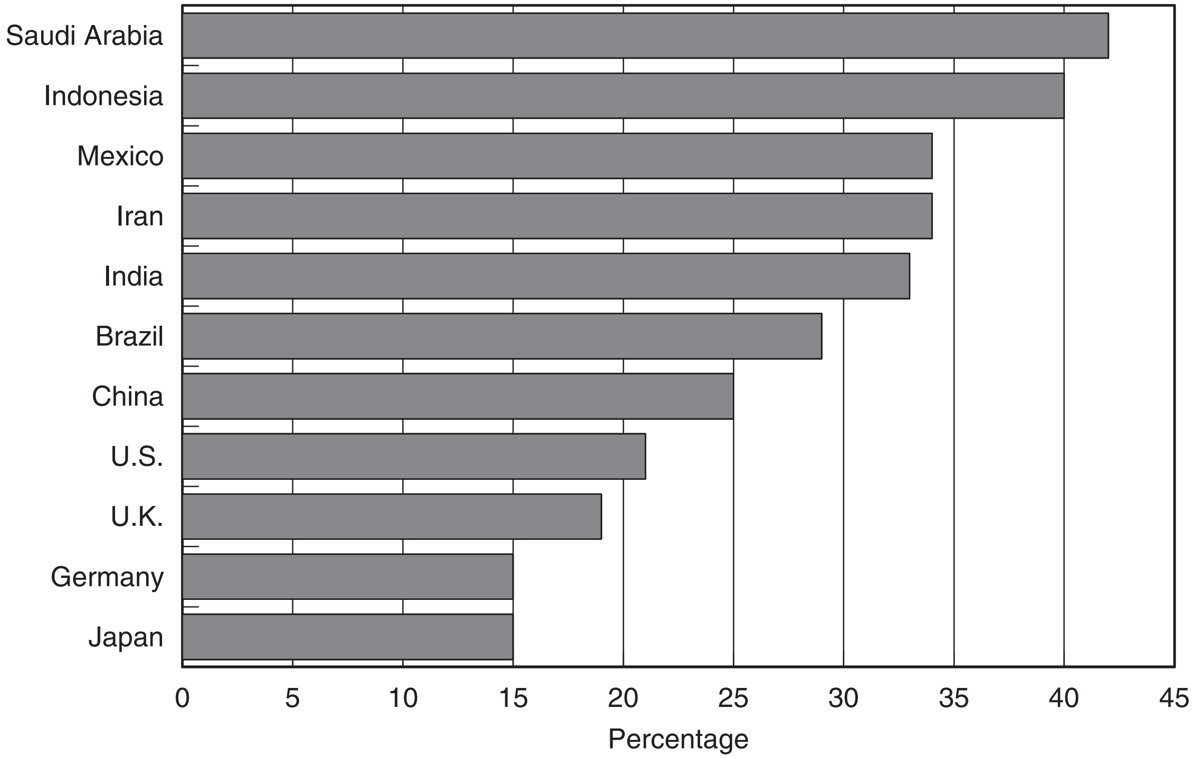


We should be justifiably proud of the medical advances that have drastically reduced the mortality rate of infant and childhood diseases. No one espouses going back to the days of Queen Anne (1665–1714), ruler of England from 1702 until her death. Anne had the best medical care that royalty could buy. Yet she had the misfortune of having around six stillbirths plus another twelve who survived birth, but not her. She died at forty-nine without an heir to the throne. As much as we are grateful for advances in treating disease, there are mathematical consequences.

The quickening pace of adding increments of a billion to the population is not an increase in the growth rate but a property of the mathematics of growth. Going from 1 to 2 billion is a 100 percent gain in population, from 2 to 3 billion is a 50 percent gain, from 3 to 4 billion 33 percent, 4 to 5 billion 25 percent, 5 to 6 billion 20 percent, and 6 to 7 billion is 17 percent growth. Eventually only a 10 percent growth in population would be necessary to go from 10 to 11 billion. Thus, each billion increment of the world's population occurs more quickly for a given population growth rate.

The earth is rapidly getting more crowded, yet there are some who say that we can sustain a much larger population. If every human being were to stand next to one another, how much of the earth would be covered with people? If we place every individual in a $3' \times 3'$ square, enough space sufficient for everyone to stand but not to lie down, we can get slightly over 3 million people into a square mile. The area to accommodate 6.3 billion people is 2,100 square miles, or a square about forty-six miles on a side. Thus, the world's population could fit, standing room only, in Delaware,

Figure 1.3 **Percentage of Population Below Age 15**

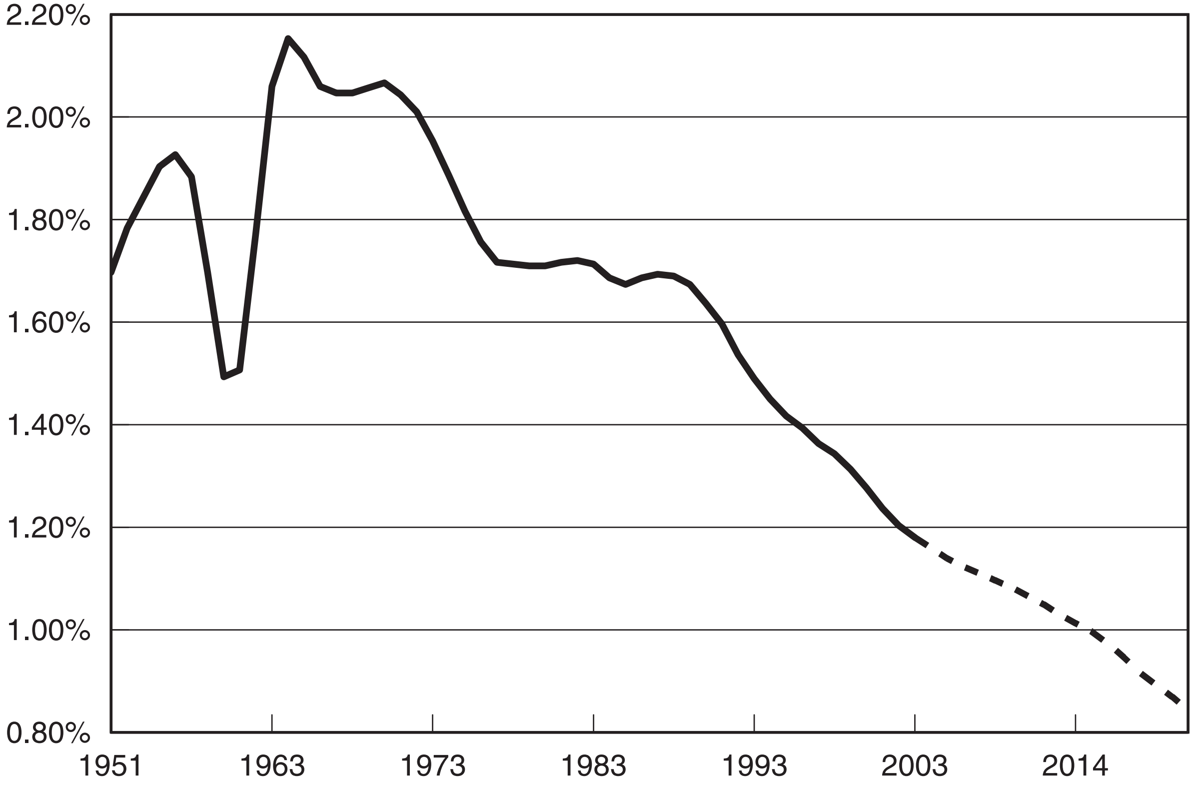


the nation's second smallest state at just under 2,500 square miles, with a little room to spare.

One way to judge the future population is to calculate the portion of a nation below fifteen years of age. A disproportionately high youthful population portends higher than average population growth as this segment reaches the child-bearing years. A somewhat arbitrary selection of nations in Figure 1.3 shows that future population growth will be centered in the Middle East, Asia (excluding Japan), and South America.

Europe and Japan exhibit essentially stagnant population growth. Some years ago China took draconian efforts to contain its population growth at 1 billion people by restricting families to one child through forced abortions and financial, and even physical, forms of punishment for having more than the authorized number of children. Families restricted to one child preferred boys, which resulted in abortions or abandonment of baby girls. China's population is unique in having a higher percentage of males (51.5 percent) to females—111.3 boys are born for every 100 girls. Eventually this in itself may create a social problem as large numbers of males find themselves unable to find mates. Despite Herculean efforts to the contrary, the social experiment to contain the nation's population has obviously failed; China's

Figure 1.4 **Annual Change in World Population**



current population is around 1.3 billion and climbing.

Figure 1.4 shows the annual percentage change in the world population since 1951. Ignoring what may have been a statistical aberration in 1961, the rate of population growth expanded between 1951 and 1964, peaked, and then started a long-term decline, which is currently projected to continue. On a global scale the average number of children per family has to decline to reduce the population growth rate. But other forces are at work that may effectively cut the growth in population growth, if not the size of the population itself. The fall of communism in 1991, and the subsequent economic turmoil, brought about a decade of a declining birthrate and a shortening of the average life span, resulting in a negative population growth rate in Russia. Diseases such as HIV/AIDS are severely reducing the population in sub-Saharan Africa, along with social disintegration, civil upheaval, tribal warfare, and, on occasion, holocausts. Some rapidly growing nations such as Bangladesh must be close to, or have already exceeded, their capacity to feed, clothe, and shelter their populations. Various forms of flu seem to be on the verge of jumping from animals to humans, which could bring on a new Spanish flu-type pandemic. Modern means of travel make it difficult to isolate or quarantine an outbreak of disease. Weapons of mass destruction and terrorism are other threats to human survival. Considering all these factors, the projected population of 9 billion people by 2050, a 50 percent increase from current levels, is not a foregone conclusion.

The Case of Double Exponential Growth

An oil company executive once observed that the oil industry benefits from two exponential curves: population and per capita energy consumption. Both work together to promote a greater volume of consumption of oil products and, presumably, greater corporate revenues and profits.

To illustrate double exponential growth, suppose that the population is growing at 1 percent per year and per capita energy consumption is growing at 2 percent per year. Further, suppose that total annual consumption of energy for 100 people is 500 barrels of oil, or 5 barrels of oil per person per year. At the end of twenty-five years energy consumption would have doubled from 500 barrels to 1,021 barrels for a composite annual growth rate of 2.89 percent. In 100 years energy consumption would be 9,510, nearly twenty times the original amount—the miracle of double exponential growth. This actually happened in the oil industry, if not more so. Does this description of double exponential growth still ring true?

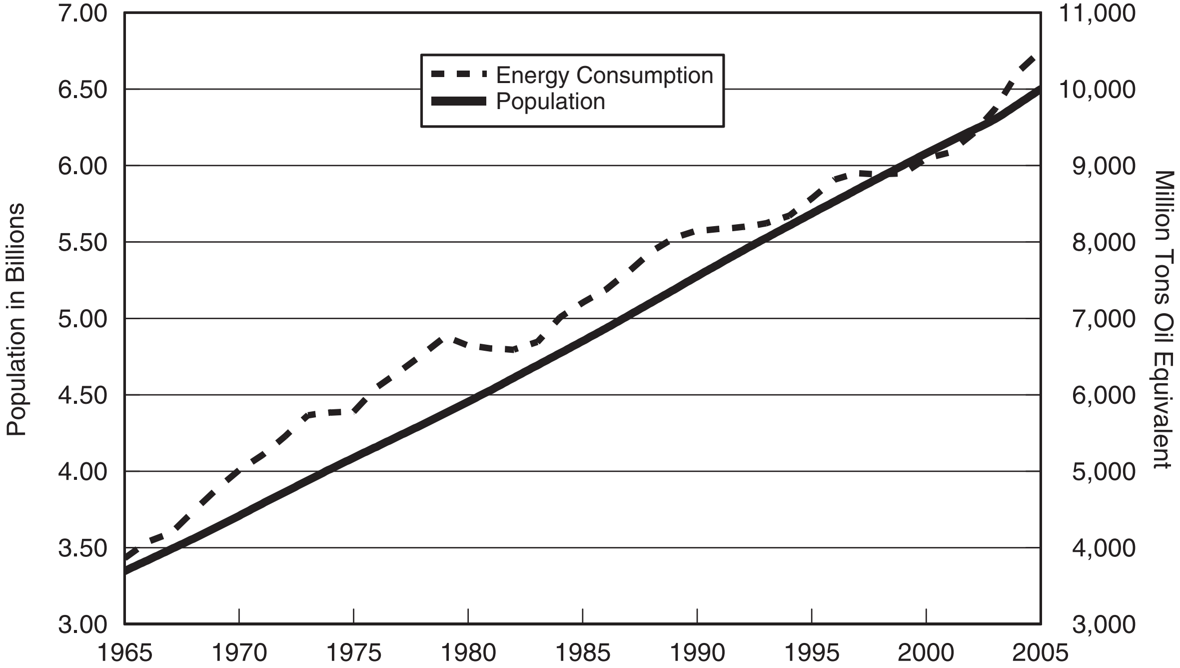
Figure 1.5 shows the growth in the world's population and the consumption of energy. Energy consists of conventional sources of energy (oil, natural gas, coal, nuclear and hydropower), excluding biomass and alternative energy sources. Energy is expressed in terms of barrels of oil equivalent, which is the equivalent amount of oil that would have to be consumed to release the same amount of energy.

While the world population continues its unremitting climb, energy consumption, which was rising faster than the population prior to the oil crisis in 1973, slowed in the wake of the crisis, which had pushed prices for all forms of energy to a high plateau. The 1973 oil crisis proved that energy consumption is price-sensitive, causing a significant change in consumption patterns. However, the recent resurgence in energy growth exceeding population growth reflects the rapid rate of economic development in China and India, home to one-third of humanity. It appears that the world has returned to two exponential curves where growth in energy consumption exceeds that of population.

Figure 1.6 is total energy consumption divided by total population to obtain per capita energy consumption. Tons of oil equivalent is changed to barrels of oil equivalent at seven barrels per ton. A barrel is forty-two gallons, or about three refills of a tank of gasoline.

Figure 1.6 shows that per capita energy consumption was growing at a relatively rapid pace until cut short by the energy crisis in 1973 and the subsequent era of high energy costs until the oil price collapse in 1986. Since then per capita oil consumption has been growing at a much lower pace, even to the point of leveling off. Much of the steadying of per capita consumption has been caused by a major decline in energy usage in the former Soviet Union after the 1991 fall of communism. Russia is now emerging from its era of economic turmoil, which may increase its per capita energy consumption. Another factor at work is the decoupling of economic and energy growth in the developed world (United States, Europe, and Japan). Economic activity is less dependent on

Figure 1.5 **Energy Consumption and World Population Growth**

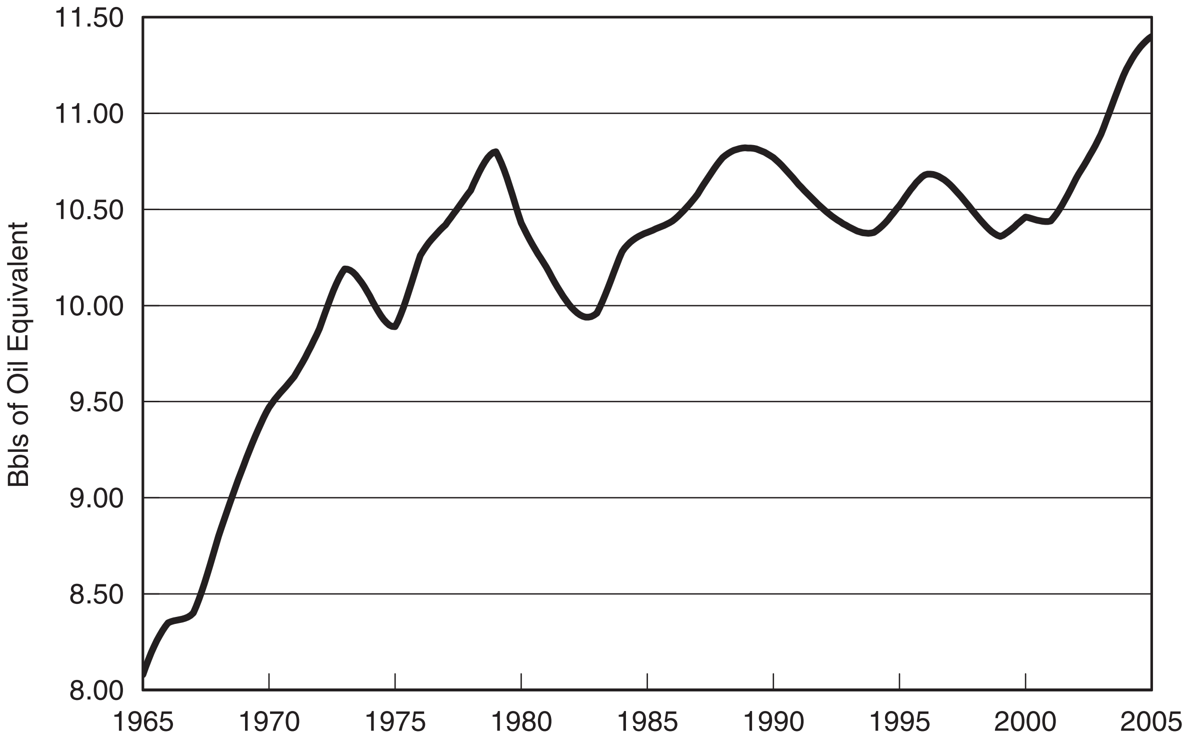


energy than it was in the past. Where once 5 percent growth in economic activity was accompanied by a 5 percent growth in energy consumption, now it is 3 percent growth in energy consumption in the United States and even less in Europe. This decoupling of economic and energy growth can be explained by heavy industries, large consumers of energy, moving from the United States and Europe to developing nations in Asia, the rise of the service industry at the expense of manufacturing, and greater efficiency in industrial processes, motor vehicles, and home appliances. However, these mitigating factors can no longer compensate for the emergence of India and China as major economic powers as seen in Figure 1.6. The return of the double exponential growth curves in Figure 1.5 and the resumption of growth in per capita energy consumption in Figure 1.6 have major implications for energy suppliers, consumers, and those responsible for formulating energy policies.

A Lesson in Fish

Fish is a finite resource that has the property of being sustainable or nonsustainable, depending on the volume of the catch. If it does not seem possible that resources can disappear in a relatively short time, as on Easter Island, or in one generation, as illustrated in Figure 1.1, ponder the world output of fish. It has been transformed from a sustainable to a nonsustainable resource in one generation. Until this generation, the annual fish catch in the world's oceans was less than the reproduction rate, which maintained the fish population.

Figure 1.6 **Global per Capita Energy Consumption**



In one generation—our generation—the population of fish, particularly in open-ocean waters, has been severely diminished.

The Grand Banks off Newfoundland is a series of raised submarine plateaus in relatively shallow waters where the cold southbound Labrador Current interacts with the warm northbound Gulf Stream. It is a living paradise for marine life. When John Cabot discovered the Grand Banks 500 years ago, codfish were so plentiful that they were caught by hanging empty wicker baskets over the ship's side.[2] A century later English fishing skippers reported cod shoals so thick that it was difficult to row a boat through them. Individual fish were six and seven feet long and weighed as much as 200 pounds. Other signs of abundant life were oysters as large as shoes, children collecting ten- to twenty-pound lobsters with hand rakes during low tide, and rivers choked with salmon, herring, squid, and other sea life. Now the cod are gone and the rivers and streams are quiet, essentially void of marine life.

Cod fishing became a victim of modern technology. In 1951 a trawler four times larger than a conventional fishing vessel sailed into the waters of the Grand Banks. Large gantry cranes supported cables, winches, and gear to operate huge nets that were let out and then pulled up a stern ramp to dump the fish straight into an onboard fish processing plant. Automated filleting and fishmeal rendering machines made short work of the catch. The trawler had several crews to allow fishing twenty-four hours a day, seven days a week, for weeks on end. Schools of fish were quickly located with fish-finding sonar, greatly enhancing the vessel's productivity. As time went on trawlers increased in number, size, and technological sophistication until they could tow nets with gaping openings 3,500 feet in circumference that swallowed and hauled in 100–200 tons of fish per hour. The trawlers maintained essentially uninterrupted operations with awaiting tenders to transfer crews and fish. By the 1970s more than 700 high-tech trawlers were in operation around the world, strip-mining the oceans of marine life.

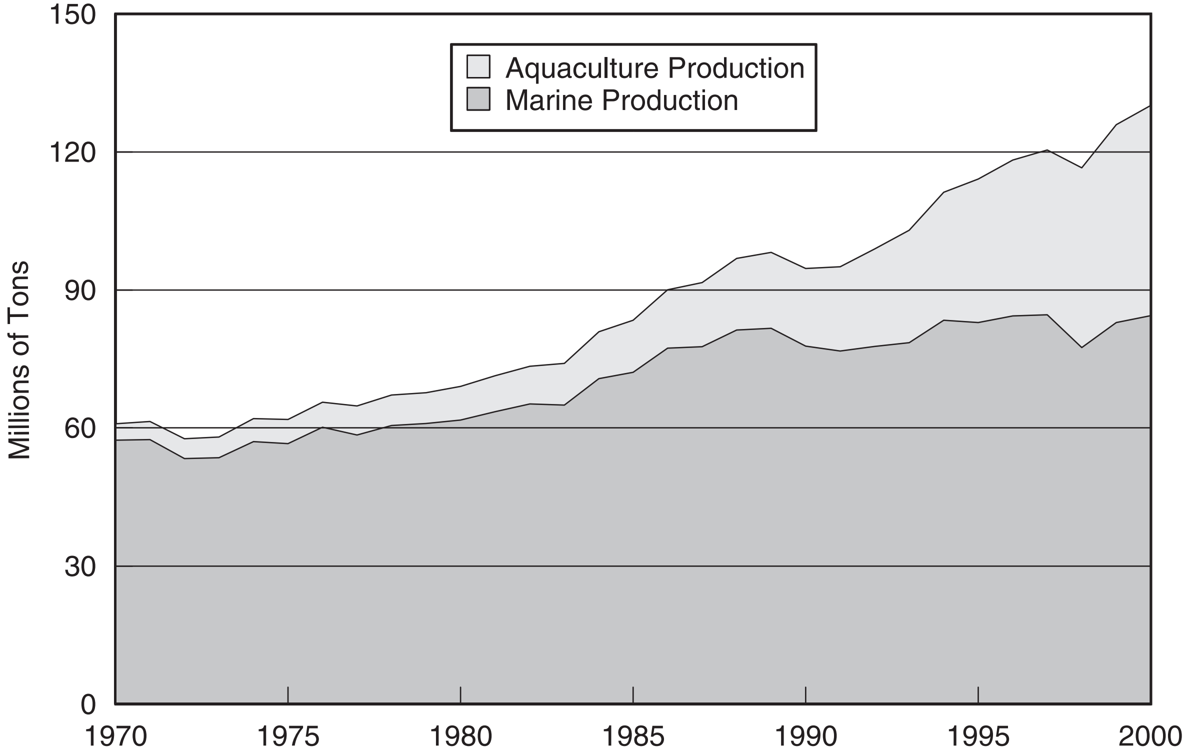
While it would be convenient to blame this situation on the greed of capitalist-owned fishing companies, over half these trawlers were from the Soviet Union. Both the Soviet Union and capitalist nations, through government subsidies to build and finance trawlers, were heavy promoters of developing and building ever-larger trawlers. These vessels, equipped with longer and wider nets, greater fish-processing capacity, and more accurate fish finders, were built to bring home larger quantities of protein to feed a growing population. The outcome was predictable, or at least it should have been, since the trawlers could scoop up fish far faster than they could reproduce. The cod catch peaked in 1968 at 810,000 tons, three times that in 1951 when the first generation of technically advanced fish trawlers made its appearance at the Grand Banks. To combat the precipitous decline that set in after 1968, Canada unilaterally extended its territorial waters from 12 to 200 miles. While other nations were prohibited from entering these waters, Canadian fishermen took advantage of the situation by adding modern fishing vessels to their fleets. The catch fell to 122,000 tons in 1991, forcing the Canadian government to close the Grand Banks to allow fishing stocks to replenish. This devastated the Canadian fishing industry, which in its heyday employed tens of thousands of people.

Much to everyone's surprise, the codfish population never recouped. Dredge-like trawls, designed to harvest marine life from the bottom of the Grand Banks by scouring an area the size of a footfall field, permanently ruined the habitat for juvenile cod. Without a habitat for juveniles, from whence do the adults come? However, one benefit was that the catch of shrimp and crab, the food for juvenile cod, has improved, providing an alternative source of revenue for fishermen.

In the twenty-first century, sonar technology can locate schools of fish and fine-mesh fishnets tens of miles long can strip the ocean of all life above the size of minnows. The population for some species of fish has dropped to a point where males are finding it difficult to locate females, or vice versa, in the vast ocean spaces. The catch of North Sea cod has fallen so precipitously in recent years that scientists feared that there might not be enough mature fish left to maintain the population. But Scottish fishermen met this advice to stop fishing to allow the cod population to recuperate with outrage.[3]

Figure 1.7 shows that world marine production (fish caught on the high seas) peaked in 1989 and then declined for a few years. This peak was thought to be the start of a permanent decline, but world marine production recouped and seemingly stabilized.[4] The leveling off of the fish harvest in open ocean waters was not caused by a rising fish population, but by greater numbers of even more efficient fishing trawlers harvesting a diminishing resource. Long lines with thousands of baited hooks stretching for eighty miles and

Figure 1.7 **Marine Life Harvest**



drift nets up to forty miles in length, responsible for the death of countless birds and sea mammals, have got to be the ultimate in open-ocean strip-mining.

While the number of larger and more efficient trawlers doubled between 1970 and the early 1990s, the total catch remained more or less stable. This infers that the average catch per vessel fell despite the greater capital investment in capacity and technology. Indeed, tons of catch per registered gross ton of fishing vessels fell from 5.4 tons in 1970 to 3.6 tons in 1992, a one-third decline. Lower productivity does not necessarily mean a smaller return on investment if the price of fish escalates enough to compensate for smaller catches. For many nations, fish is an important part of the diet and governments are not willing to cut off this supply of food, regardless of the long-term consequences (shades of Easter Island?). Another reason why so many governments are reluctant to put effective international controls on fishing is to protect their investment in the form of government subsidies for trawler acquisition and financings.[5] This is what is going on in oil. More money is being spent on wells to tap diminishing reserves in more difficult environments, but profitability can and is rising, despite higher costs and lower output, as long as there is a compensating increase in price.

Unlike the Easter Islanders, we have not stood idly by in the face of depleting fish resources in the world's oceans. Many maritime nations have enacted programs to preserve the fish population within their territorial waters by regulating the timing, size, and volume of fish and other marine life that can be caught. As an example, marine biologists can survey the egg population of herring in Alaskan waters from the air because the untold billions of herring eggs make the normally dark waters milky white. From their observations they can estimate the herring population and, through a government regulatory agency, mandate the area and timing of the herring harvest. The harvest of herring is controlled on the basis of preserving a resource to ensure its long-term viability rather than depleting a resource in the quest for short-term profits. The volume of the herring harvest is not regulated, only the allowable area and the permitted time for harvesting. Licenses control the number of fishing boats and the nature of the technology being employed. If it is felt that the volume harvested is too great and endangering the population, more stringent restrictions on area and timing and licensing can be enacted. If the population of herring is rising, then the restrictions can be relaxed.

Regulating fishing in a nation's territorial waters is practiced throughout much of the world. But it is a palliative, not a cure, as fish are free to migrate in and out of territorial waters. Little headway has been made in passing an international convention on open-ocean fishing that would place fishing on the basis of sustaining rather than depleting a natural resource. The politically correct-sounding International Convention on Fishing and Conservation of the Living Resources of the High Seas went into force in 1966, but apparently its provisions are either not effective or not effectively enforced to sustain the long-term viability of commercial sea life in international waters.

Those who maintain that nothing can be done to protect open-ocean fish resources should look at the international regulation of the whaling industry. The International Convention for the Regulation of Whaling was proposed in 1946 and came into force in 1948, after the requisite number of nations had ratified it. The Whaling Commission meets annually and determines protected and unprotected species, open and closed seasons and waters, designation of sanctuary areas, size limits and the maximum catch for each species, and permissible methods, types, and specifications of whaling gear. The Commission also controls the methods of measurement and maintains the requisite statistical and biological records, and places tough restrictions on location and season to ensure that the catch does not exceed certain limits. The whale population for various species is monitored to see if any adjustments have to be made to relax or strengthen restrictions on whaling activities.

The international convention on whaling has been responsible for not only preventing the extinction of various species of whales but also in promoting their recovery. This successful convention not only had a highly desirable environmental impact on sea life, but could also serve as a model for the administration and enforcement of an international convention on open-ocean fishing. However, in all fairness, trying to monitor and control fishing fleets numbering in the millions scattered throughout the world through a system of licenses and quotas would, for all practical purposes, be impossible. The task of the Whaling Commission is made easier by virtue of there being relatively few whaling fleets, restricted to only a handful of nations. However, if the world's maritime nations had the collective will to assume the responsibility for monitoring fishing vessels calling on ports within their jurisdiction, or outlawing those practices that essentially strip-mine the ocean of fish, then it might be possible to reverse the further diminishment of a valuable resource.

From Hunting to Farming

Figure 1.7 also shows a successful countermeasure to overexploiting a natural resource. The rising price of fish has given rise to a new industry: fish farming, agricultural enterprises dedicated to raising fish. Fish pens in protected waters in Norway and Canada supply much of the salmon found in the world's marketplace. Decades ago, farmers in the southern part of the United States could not make a living growing and selling grain until they discovered that they could make a living by throwing grain into a pond and selling the catfish. Trout, tilapia, and shrimp are also farmed. Tilapia originated in Africa and was farmed in ponds and rice paddies in Asia for generations; now tilapia is farmed throughout the world. As seen in Figure 1.7, the tonnage of aquaculture production, or fish farming, has grown thirteenfold in thirty years, providing about one-third of total fish consumption.

To be sure, there is opposition to aquaculture, or fish farming, including the potential environmental impact of thousands of tons of waste collected under fish pens, the biological treatments necessary to prevent the spread of disease in crowded fish pens, plus the consequences of fish escaping from an artificial to a natural environment. On the other hand, the human population needs to be fed, and providing fish from fish farms rather than depleting the world's open-ocean resources seems to have an inherent advantage that should not be ignored.

Energy Depletion

The point of all this is that a trend line indicating that a resource will be exhausted in thirty years need not happen. Unlike fish, which could recuperate in numbers if allowed to, fossil fuels such as coal, oil, and natural gas cannot replenish themselves (ignoring for now speculation about the possible nonorganic origin of natural gas from deep within the earth). With regard to energy, sustainable sources of energy (solar and wind) would be akin to fish farming. The viability of sustainable sources of energy can be assured by a high price on a diminishing source of nonsustainable (fossil) fuel.

We have already set the precedent of having an international convention to preserve a natural resource. International conventions can be successful in preserving natural resources (whales) or unsuccessful (open-ocean fish). As with whales and, perhaps, someday with fish, some form of cooperative action may be necessary to preserve a worldwide resource—energy—if only to prevent the inevitable result of doing nothing: having a planetary-scale Easter Island blowout.

Notes

1. Population statistics are from the U.S. Census Bureau Web site and energy statistics are from the *BP Statistical Review of World Energy* (London: British Petroleum, 2005).
2. Colin Woodard, "A Run on the Banks." *E-Magazine* (March–April 2001).
3. News item from Web site www.independent.co.uk (October 24, 2002).
4. Food and Agricultural Organization of the United Nations (FAO), as listed on the World Resources Institute Web site at www.earthtrends.org.
5. United Nations Food and Agriculture Organization (FAO) statistics as reported at www.home.alltel.net/bsundquist1/fi0.html.

2

Electricity and the Utility Industry

The primary fuel sources are coal, oil, natural gas, nuclear, hydro, biomass, and renewables, but these are not the forms of energy we encounter most often. Other than driving a car and heating a home, the form of energy we are most accustomed to is that which comes from turning on a switch. We use electricity for lighting and running all sorts of electrical appliances and equipment. It is absolutely essential to running a modern economy. Yet, electricity is a secondary form of energy derived from primary fuel sources. This chapter deals with electricity as a source and use of energy, its origin, the organizational structure of utilities, system operation, and models for determining rates.

It Takes Energy to Make Energy

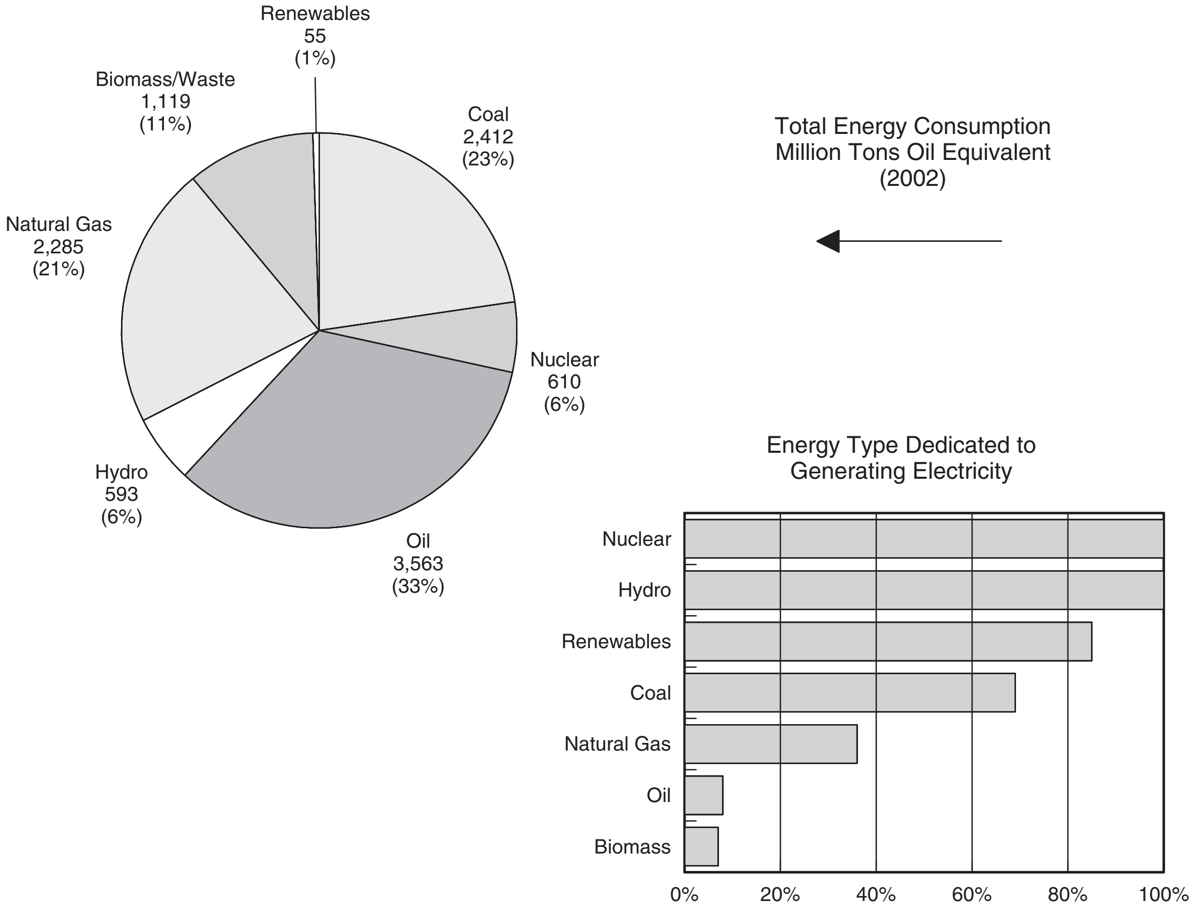
Burning a ton of oil releases a certain amount of energy, but burning a ton of coal does not release quite the same amount of energy and burning 1 million cubic feet of natural gas—well, that is hard to relate to a ton of oil or coal. In order to have a common measure for comparison purposes, it is convenient to talk about 1 ton of oil equivalent. This is the amount of energy released by various sources that is equivalent to the amount of energy released by 1 ton of oil. Energy terminology can be quite confusing, but total world primary energy consumption in 2005 can be expressed simply as 10.5 billion tons of oil equivalent.[1] While this is a large number that probably taxes one's imagination, it can be brought into perspective by remembering that there were about 6.5 billion of us in 2005. Thus, total consumption works out to about 1.62 tons of oil equivalent per person per year, or 11.3 barrels of oil equivalent (7 barrels to 1 ton), or 475 gallons (42 gallons in a barrel) of oil equivalent, or 1.3 gallons of all primary sources of energy expressed in oil equivalents per person per day. Per capita consumption, however, is not equally divided among the world's population—there is an enormous difference between the amount consumed by an American and what everybody else in the world consumes.

As shown in Figure 2.1, the largest single source of energy is oil, which satisfies 33 percent of world demand, followed by coal (23 percent), natural gas (21 percent), biomass (11 percent), nuclear (6 percent), hydro (6 percent), and renewables including geothermal, wind, and solar (1 percent).

In 2004, about 35 percent of energy including biomass and renewables, or 3.7 billion tons of oil equivalent, was dedicated to generating 16,074 terawatt hours of electricity. As seen in Figure 2.1, electricity generation consumes all nuclear and hydro sources, nearly all renewables, about two-thirds of coal, and a little under 40 percent of natural gas. Oil has far more value as a motor vehicle fuel than a boiler fuel. However, a portion of the bottom of the barrel, the residue left after refining crude oil, and certain waxy crude oils too difficult to refine, are burned as fuel for generating electricity.

Using the relationship that 12 terawatt hours of electricity have the same energy content as 1 million tons of oil equivalent, the world generation

Figure 2.1 **Energy Consumption**



of 16,074 terawatt hours of electricity has the same energy content as 1.34 billion tons of oil equivalent. Thus, an input of 3.7 billion tons of oil equivalent of primary energy sources generated an output of electricity with an energy content of 1.34 billion tons of oil. The output is only 36 percent of the input. What happened to the difference?

The answer is that 64 percent of the energy consumed in generating electricity is thrown away. Most of it is passed to the environment by heating the atmosphere or the water in rivers, lakes, bays, and oceans. Electricity produced by steam, be it fossil or nuclear power, is inherently energy-inefficient. Water is heated to produce steam, steam drives a turbine, which in turn drives an electricity generator. The latent heat of vaporization is the heat absorbed to transform water from a liquid to a vapor state, a considerable amount of energy. Steam produced in a boiler is fed to the high-pressure end of a turbine. The efficiency of generating electricity can be enhanced by maximizing the steam pressure differential between the high- and low-pressure ends of a turbine. This is done by increasing the steam pressure entering the high-pressure end of the turbine and reducing the low-pressure end of the turbine to a vacuum by condensing the spent steam to a liquid. Water from a river, lake, bay, or ocean passes through a condenser, absorbing the latent heat of vaporization from the spent steam, which is then passed to the environment. Sometimes the warmed condenser water is cooled for recirculation by transferring heat to air in a cooling tower or to a cooling pond. Either way the latent heat of vaporization present in the spent steam is lost.

The condensed steam, now in the form of hot water, is pumped back to the boiler where it must reabsorb the latent heat of vaporization to become steam again for another cycle through the turbine. Thus, a typical generator's output of electricity is only about 25–35 percent of the energy input to produce the steam. Another 5–10 percent of the energy content of electricity is lost in transmission and distribution. Heating water in a teakettle using electricity takes three to four times the amount of energy required than heating the water directly by flame. Heating is not a good use of electricity. A better use is providing lighting and running appliances. Heating a home can be much more efficiently accomplished with heating oil and natural gas, or even coal and firewood, than by electricity.

Energy for Electricity Generation

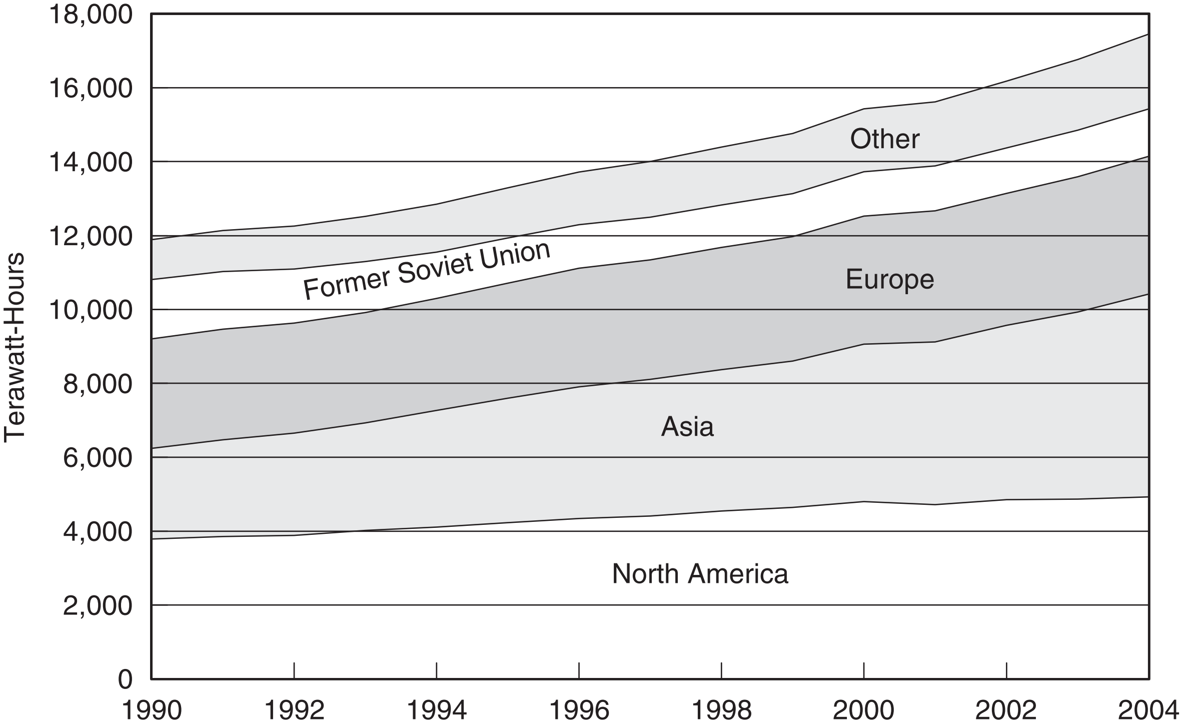
Figure 2.2 shows that the average annual growth in world electricity demand was 2.5 percent per year from 1990 to 2004.

Rates of growth on a regional basis vary considerably. The former Soviet Union had a negative growth rate after the fall of communism in 1989, reflecting the economic floundering that plagued this group of nations. The decline in electricity generation reached its nadir in 1998 and has been slowly recuperating since then. Europe has the second smallest growth in electricity consumption at 1.5 percent per year, followed by North America at 1.7 percent per year. The aggregate growth of other regions (South America, Africa, and the Middle East) is 4.3 percent, with Asia having the highest growth rate, 5.5 percent. These growth rates may not seem high, but they make a difference over time. In 1990 Asia consumed a little over half of what North America consumed, but in 2003 Asia beat North America as the world's largest consumer of electricity. The gap between Asia and North America is going to widen rapidly. World electricity consumption grew by nearly 50 percent in fourteen years, roughly doubling every twenty-eight years. In the not unusually long lifetime of an octogenarian, electricity demand doubles three times for an increase of a factor of 8, or 800 percent.

The relative importance of the world's leading electricity-generating nations is shown in Figure 2.3. In 2002, the United States led, generating 24 percent of the world's electricity, followed by China with 13 percent. Together these two nations accounted for 37 percent of the world's electricity generation; the top five nations accounted for half, and the dozen nations in Figure 2.3, in the aggregate, accounted for 70 percent of world electricity generation.

Energy experts project that the demand for electricity will double again between 2002 and 2030. If this projection holds true, then 4,800 gigawatts (GW) of electricity capacity will have to be built, of which the member nations of the Organization for Economic Cooperation and Development (OECD) will have to build 2,000 GW of capacity, including an allowance for replacing

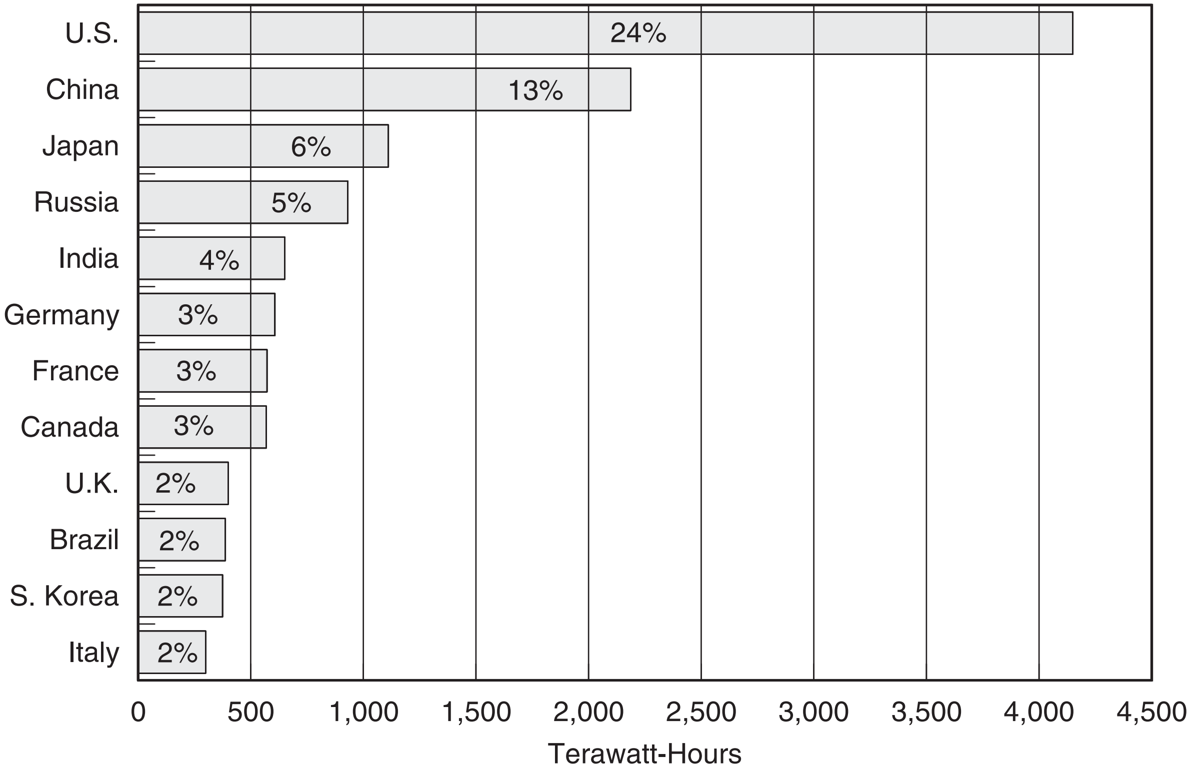
Figure 2.2 **Growth in World Electricity Demand**



one-third of the existing installed capacity slated for retirement. To put a GW, or 1,000 megawatts, or 1 billion watts into perspective, a one-GW plant can supply a city of about 1 million people. This increase in electricity-generating capacity for the OECD nations will require an investment of $2 trillion in power generation and another $1.8 trillion in new and replacement transmission and distribution systems. The developing nations will require 2,800 GW of new capacity, representing $5.2 trillion in new investments in generation, transmission, and distribution systems.

About one-third of humanity is not connected to an electricity grid. Economic development projects are commonly dedicated to bringing electricity to places where much of human effort is spent in the daily drudgery of trekking twenty miles for wood and dung for fuel to heat and cook and manually lifting water from deep wells. While some of the $5.2 trillion in capital spending will expand the electricity grid to those without access to electricity, many will still not be connected. Even so, trillions of dollars in capital investments to upgrade and expand the global electricity generation and distribution system raise the issues of where these funds are going to come from and, for the developing world, the ability of people to pay electricity rates sufficient to service these enormous capital outlays. Figure 2.4 shows the expected role of the various primary energy sources

Figure 2.3 **World's Leading Nations in Electricity Generation**

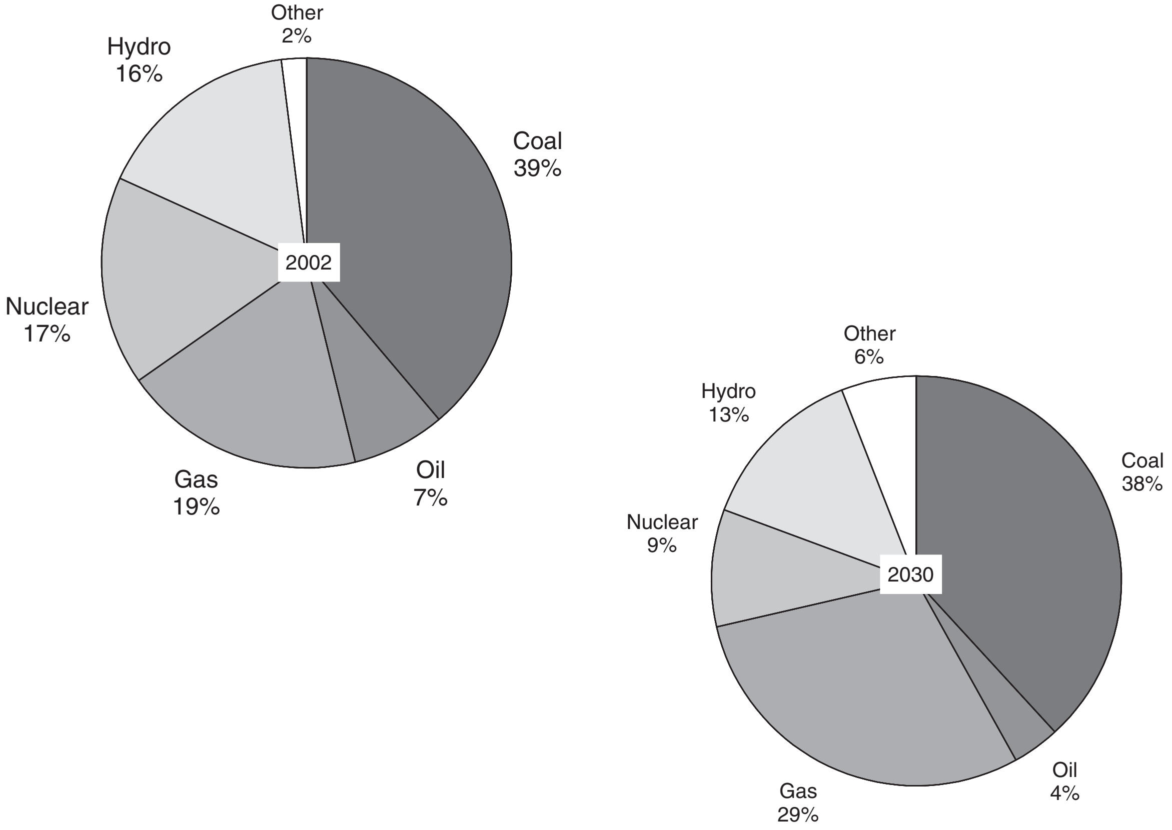


for satisfying a doubling of electricity generation needs by 2030 as compared to 2002.

Of the 16,074 terawatt hours generated in 2002, 39 percent was generated from coal, 19 percent from natural gas, 17 percent from nuclear power, 16 percent from hydropower, 7 percent from oil, and 1 percent each from biomass and renewables. Clean-burning natural gas is growing as the preferred fuel for electricity generation. In the United States, nearly all new generating capacity for two decades since the early 1980s was fueled by natural gas until the run-up in natural gas prices in 2004. At the same time in Europe, natural gas and nuclear power were replacing coal and oil as the favored fuels for electricity generation. "Other" consists of biomass and renewables. Biomass is wood, wood waste, other burnable industrial waste materials; renewables are geothermal, wind, and solar. Each contributed little to electricity generation in 2002.

Despite efforts to reduce the role of coal as a fuel for generating electricity, its expected share in electricity generation is only slightly diminished by 2030. Considering that electricity demand is doubling, coal consumption has to almost double to maintain its share. The roles of oil, nuclear power, and water are cut in relative, but not absolute, terms when taking into account the doubling of electricity generation by 2030. Natural gas takes up the slack

Figure 2.4 **Energy Sources for Electricity Generation**



from stagnant growth of oil, nuclear, and hydro in electricity generation. The tripling of "other" from 2 percent to 6 percent, after taking into consideration the doubling of demand, represents rapid growth of wind, and to a lesser extent solar power, for generating electricity.

Enhancing Efficiency in Electricity Generation

Depending on the age and nature of the electricity-generating plant, typical efficiencies for a utility plant range from 25–35 percent. This can be substantially increased for an industrial plant that consumes both electricity and hot water. The plant can buy electricity from a utility to run the plant and to heat water. Alternatively, a cogeneration plant can be installed on the site that generates electricity required by the plant. The heated water containing the latent heat of vaporization is not thrown away to heat the air or some body of water as with a utility, but consumed internally as a source of hot water. Being able to utilize the latent heat of vaporization and eliminate transmission-line losses by having an on-site electricity-cogeneration plant can significantly increase the efficiency of generating electricity. The combined-cycle cogeneration plant has the highest efficiency of all electricity-generating plants. Natural gas first fuels a gas turbine that drives an electricity generator. The hot exhaust gases from the turbine heat water in a boiler for conventional steam generation of electricity. By capturing the heat of vaporization in the condenser water for internal purposes, the resulting efficiency can be 60 percent or more.

Unfortunately, only about 3 percent of U.S. electricity generation is by cogeneration. Cogeneration units are sized to the needs of an industrial plant and, therefore, lack the inherent economies of scale of the much larger-capacity electricity generators of a utility. The economics may favor buying a lot more electricity from a utility to run a plant and heat the water than producing electricity at the plant with a cogeneration unit and utilizing the waste energy as a source of hot water. The irony is that the price of buying electricity from a utility reflects the cost of passing the latent heat of vaporization to the environment. Despite the efficiency of cogeneration plants, most electricity-generating utilities are not keen about losing a customer. While government policies favor cogeneration plants to enhance energy efficiency, utilities have little incentive to cooperate. Some degree of governmental coercion, or some form of incentive, must be provided to overcome a utility's reluctance to lose a customer to a cogeneration unit.

Some cogeneration plants burn coal to produce steam, which can also be superheated by burning natural gas. The boiler fuel can also be biogas from animal or human wastes, methane from landfills, and agricultural and industrial waste products. Cogeneration plants in the pulp and paper industry consume waste (bark, wood scraps, and the processing residue of papermaking called black liquor) as fuel to generate electricity. Hot water generated from condensing spent steam is consumed internally in the papermaking process. Sugar mills burn waste sugarcane to generate electricity and utilize the hot water from the condenser to produce sugar. The steel industry burns gases given off from open-hearth steel-making facilities to produce steam for powering blast furnace air compressors and for generating electricity. If air rather than water captures the latent heat of vaporization from a steam turbine, the heated air can be used for drying agricultural crops and lumber.

Early History of Electricity

Although one might think that the definition of electricity is obvious—it is what turns the light bulb on—the nature of electricity is not that easy to comprehend. Electrical energy that keeps a light bulb shining is similar to X-rays, light, microwaves, and communication signals except that it has a much lower frequency (60 cycles per second in the United States and 50 cycles per second in Europe). Electrical energy is also known as electromagnetic energy, consisting of magnetic and electrostatic fields that move in the vicinity of wires near the speed of light in response to the movement or vibration of electrons.

Matter is made up of atoms, which consist of protons, neutrons, and electrons. Protons have a positive charge and electrons a negative charge. When present in equal numbers, there is no net charge; when imbalanced, there is a charge called static electricity. The nature of matter was not known in 600 BCE when Greeks generated a spark by rubbing amber with fur. The term *electricity* is derived from the Greek word for amber. The discovery in the thirteenth century that lodestones align themselves with the magnetic north pole when free to turn made long sea voyages beyond the sight of land possible, opening up the Age of Exploration. The connection between electricity and magnetism would not be known for another six centuries.[2]

Among the eighteenth-century scientists interested in electricity was Benjamin Franklin, who established the convention of positive and negative charge and the presumed directional flow of electricity. In this, he made an error, scrupulously preserved to this day. His experiment—flying a kite in a lightening storm with a key attached to the end—could have killed him. Lucky for him, a passing bolt of lightening was far enough away to induce only a small current in the kite string that reached the key, causing a spark to jump to Franklin's hand. He concluded that the spark from lightening was the same as a spark from rubbing amber with fur.

The knowledge of electricity advanced one step at a time during the nineteenth century. Charles Coulomb discovered that a charge of electricity weakened with the same inverse square law that Newton had discovered for gravity. For this, his name was attached to a measure of electrical charge, the coulomb. Luigi Galvani, while dissecting a frog, noted that its legs began to twitch. Galvani thought that there might have been some lightning in the vicinity. After some experiments, Alessandro Volta decided that two dissimilar metals, the knife and the tray holding the frog, was the actual cause of the twitching. From this he developed the voltaic pile, the forerunner of the battery made of disks of two dissimilar metals such as zinc and copper or silver separated by paper soaked in saltwater. Electricity flowed through the pile when Volta completed the circuit with a copper wire. Although Napoleon honored Volta by making him a count, Volta was memorialized by having his name attached to a measure of the electromotive force represented by a difference in a given electrical charge, the volt. Galvani was not forgotten and was memorialized in the phrase *galvanic action*, the corrosive interaction of dissimilar metals, and in the verb *galvanize,* to stimulate action as if by electric shock. Andre-Marie Ampere was one of the first to establish a relationship between magnetism and electricity, inaugurating a new field of study called electromagnetism. As with other discoveries, Hans Oersted simultaneously and independently performed experiments that demonstrated the same relationship. But the honor went to Ampere by having his name attached to a unit of electrical current, the ampere. Oersted was not completely forgotten: The unit of magnetic intensity, the oersted, was named after him. Georg Ohm originated Ohm's Law, which linked electrical potential, a volt, with electrical current, an amp or ampere, passing through a unit of electrical resistance named after him, an ohm.

Michael Faraday, considered one of the greatest experimenters of all time, started out as a bookbinder's assistant who read the books that he was supposed to be binding. He developed an interest in chemistry, but was encouraged to switch to electromagnetism, where his repetition of Oersted's experiments led to his discovery of electromagnetic induction. Faraday could induce an electric current by either moving a magnet through a loop of wire or moving a loop of wire over a stationary magnet. From this he developed an electric dynamo, the progenitor of the modern electricity generator. Faraday's experimental work and "lines of force" inspired James Maxwell to describe the behavior of electricity and magnetism mathematically. Faraday postulated that light was a form of electromagnetic wave; a conclusion also reached by Maxwell when he discovered that the speed of electromagnetic waves was close to that of light. Maxwell's equations became the building blocks for Einstein's theories.

Generating Electricity

An electric current is the flow of electrical charge. When an electric current passes through a conductor such as copper wire, the electrons in the copper are forced to either flow quite slowly in one direction for direct current or vibrate back and forth for alternating current. Electrical energy travels by moving or vibrating the electrons within metallic conductors much as sound waves are propagated by vibrating air molecules. The speed of sound, of the order of 345 meters per second or 770 miles an hour, is not caused by air molecules moving at the speed of sound, but by the speed of sound waves propagated by vibrating air molecules. Similarly, electrical energy is not propagated by electrons moving at the speed of light through wires; but by electromagnetic fields traveling close to the speed of light that are associated with the vibration or movement of electrons.

Electricity is a charge measured in coulombs; electrical energy is measured in joules, and electrical power, the flow of energy over time, is measured in watts (one joule per second), a term that honors James Watt for his work in measuring the equivalent horsepower output of a steam engine. As has become customary, if not traditional, the terms *electricity* and *electrical energy* are used interchangeably even though they are technically quite different.

Generators cannot make electricity (technically electrical energy) because electricity is a property of matter. An electricity generator "pumps" an electrical charge back and forth inside a wire sixty times per second and the electromagnetic fields created around the wire are electrical energy. Electrical energy flowing through a motor, heater, or light bulb over a period of time becomes electrical power that turns the rotor and warms or lights a room.[3] The electricity industry converts various sources of energy into a single indistinguishable product that is easily distributed for lighting, heating, and running machinery, equipment, and appliances. This is quite unlike the oil industry, which converts a single source of energy into a wide variety of products ranging from motor vehicle fuels to plastics.

To produce electricity, a turbine is rotated to drive an electricity generator. Steam is the most common motive force of rotating a turbine and is produced by burning coal, oil, natural gas, biomass, or from a geothermal or nuclear source. Falling water, tidal currents, river flow, wave action, and wind are other motive forces to rotate a turbine. The only manmade source of electricity not created by rotating a turbine is a solar photoelectric cell that converts sunlight directly into electricity. Our capacity to generate electricity pales into insignificance when compared to nature. Enormous circulating electrical currents surrounding the core create the earth's magnetic field. Lightning occurs when the buildup of static electricity at different cloud levels, or between a cloud and the earth, creates a voltage differential large enough to overcome the resistance of air to conduct electricity. If lightning could be harnessed, it would easily fulfill humanity's dream of unlimited and free electrical energy.

Generating Electricity Commercially

Thomas Edison invented the electric light bulb in 1878 by trial and error. It was the end result of innumerable attempts to find a filament material that could conduct electrical current to the point of incandescence without burning up. Edison was also an astute businessman and founded the first investor-owned utility in 1882. The Pearl Street station lit up lower New York with direct current from electricity generators powered by reciprocating coal-fueled steam engines. He was challenged by George Westinghouse, who backed Nikola Tesla's alternating-current electricity generators. In this contest between two industrial giants, Edison publicly backed the idea of an alternating-current electric chair for the state of New York to demonstrate its inherent danger. He won the contest of having an alternating-current electric chair, but lost the larger contest as to whether homes and businesses would be fed by direct or alternating current. The problem with direct current was that it could not be distributed over a wide area without significant line losses. In Edison's world, direct-current electricity would be distributive in nature with many plants, each serving a small area.

Alternating current was superior to direct current in that it could be transmitted over long distances at a high voltage with relatively small line losses. Alternating current allows for a small number of large centralized generating plants with their inherent economies of scale to serve a wide area via long-distance transmission lines. In 1895 Westinghouse built the first commercial alternating-current electricity-generating plant at Niagara Falls, the progenitor for all future electricity-generating plants. Although Niagara Falls used hydropower to turn the generators, Westinghouse also spearheaded the development of the steam turbine, the brainchild of William Rankine. The substitution of the steam turbine for the steam engine increased the thermal efficiency for generating electricity from 5 percent in converting coal to electricity with a reciprocating steam engine ultimately to 35 percent for a modern steam turbine.

Generation of electricity (the capital investment plus operating and fuel costs) makes up about one-half of the delivered cost of electricity; the remainder is transmission and distribution costs. Transmission lines, bundles of copper or aluminum wires usually above, but sometimes below ground, carry electrical current at several hundred thousand volts from generators to local distribution companies. Transmission makes up 5–15 percent of the delivered cost of electricity to cover capital, operating, and maintenance costs. Electrical energy heats up transmission lines, which expand (lengthen), causing transmission lines to droop noticeably when under heavy load. The dissipation of this heat to the environment is known as line losses. In the United States, the average line loss is 7 percent of generated electricity, but actual losses between two points vary with the amount of electrical energy passing through the transmission lines, their design characteristics, the surrounding environmental conditions, and the distance traversed. The remainder of the delivered cost of electricity is local distribution that steps down the voltage through transformers, routes the electricity to individual homes, businesses, and industries, and bills customers, whose payments support the entire financial edifice of the industry.

System Operation

Unlike fossil fuels, there is no way to store electricity (batteries are incapable of storing the amount of electricity required to support the operations of a utility). The electricity business is the only one that operates without an inventory. Some maintain that water in back of a dam can be considered stored electricity. Using this logic, a pile of coal sitting outside a generating facility is also stored electricity. But water and coal are not electricity until they provide the motive force to rotate a turbine to drive a generator. Once generated, electrical energy flows at the speed of light between the generator and the consumer when the switch is turned on and stops just as abruptly when the switch is turned off. Unlike oil or natural gas in a pipeline, throttling a valve does not direct the flow of electricity. Electricity follows the path of least resistance. If that path leads to overloading a transmission line and melting the wires, so be it. Although breakers protect transmission lines from overloading, the usual way to decrease the flow of electrical energy through transmission lines between points A and B is to raise the generating output of electricity at B and cut the output at A.

During times of low demand, when transmission systems are not limited by capacity constraints, the price of electricity is fairly uniform throughout a region. During times of heavy demand, transmission capacity constraints create local price disparities. For instance, reducing the output at A to prevent overloading the transmission system may involve shutting down a low-cost electricity generator and increasing the output at B, as a substitute source may involve starting up a high-cost electricity generator. This creates a price disparity between points A and B. Another cause of a price disparity between two points is line losses; that is, what goes into a transmission system is not what comes out.

With no way to store electricity, the system of generating and transmitting electricity must adjust to variation in demand on an immediate basis. There is significant variation in demand over the course of a day, when peak daytime demand may be double that of nighttime, over the course of a week, when more electricity is consumed on weekdays than weekends, and over the course of a year, when electricity demand peaks from air conditioning during summer hot spells. In cold climates in areas where electricity heats homes (e.g., eastern Canada), peak demand occurs during the winter. With some exceptions, such as where there is sufficient hydro plant capacity to satisfy peak loads, peaking generators have to be purchased. These are usually combustion turbines (modified jet engines) fueled by natural gas that may run for only a few days or a week or so during an entire year. Amortizing the cost of peaking generators over such a short period of operating time makes for extremely expensive electricity. Yet, if peaking generators are not available, blackouts ensue unless other arrangements have been made to curb demand.

In one such arrangement, operators of office buildings and factories are paid by the utility to disconnect during times of peak seasonal demand and supply themselves with power by operating their emergency backup generators. Another arrangement is for heavy users of electricity to slow operations during times of peak demand and shift some of the load to times of reduced demand. Some plants (e.g., aluminum smelting) have their own electricity-generating capacity. During times of peak demand, it may be more profitable for these plants to curtail production and sell excess electricity to utilities. Other companies pay a lower rate for an interruptible supply of electricity and are willing to be disconnected for a few hours a day during times of peak demand to benefit from lower-cost electricity for the rest of the year. Companies desiring uninterrupted service pay a higher rate to ensure that there is always enough generating capacity available to sustain their operations.

Another way to handle peak demand is to institute demand management. One form of demand management is installing timers on appliances, such as hot-water heaters, so that they operate only during nighttime lulls in electricity demand. The more common form of demand management is time-of-day metering, with electricity rates varying by the hour in response to demand. More costly electricity during times of peak demand creates an incentive for individuals and businesses to reduce their electricity load by turning off their air conditioning for an hour or two when rates (and temperatures) peak. With flat rates that now exist for most consumers, the cost of running an air-conditioning unit is the same regardless of the time of day or night. Consumers are not sensitive to fluctuations in market-priced electricity. Demand management makes them sensitive to the time of day and the day of the week. Running washers and dryers at night and on weekends can significantly reduce electricity bills. The last resort for accommodating peak demand without sufficient generating capacity is controlled rolling blackouts that cover different areas for relatively short periods of time, announced or otherwise, to reduce demand below generating capacity. Failure to take effective action when demand exceeds supply will result in a loss of system control and an unplanned blackout.

System operation is a critical function that controls the output of generators to satisfy demand in real time, where electrical energy flows at the speed of light over the path of least resistance, without the benefit of being able to draw down on inventory. The system operator must deal with imbalances between supply and demand, congestion (overloading transmission lines), and ancillary services. The latter includes power needed to run the system, reserve capacity to meet unexpected demand, backup power plants in case an operating generator fails or a loss of motive force, for example, hydro dams without an adequate water supply, wind turbines when the wind calms, or solar arrays on cloudy days. The system operator is responsible for scheduling (planning the future starting and stopping of generators) and dispatching (real-time starting and stopping). Scheduling and dispatching have to be carefully coordinated to prevent overloading transmission lines while maintaining system stability with continually fluctuating demand. Overloading transmission lines and/or losing system stability are the root causes for blackouts that can spread over large areas of a nation through utility interconnections and last for extended periods of time.

Methods of Rate Regulation

The century-old approach to electricity was to regulate the industry as a natural monopoly. Multiple transmission and distribution lines from a number of generators, each connected to individual households and businesses to give consumers a choice of provider, would be inordinately expensive. The investment would be much more than having a single wire entering a household or business from a single generator. This would result in high electricity rates to amortize a huge investment in grossly underutilized assets. A natural monopoly comes into being once a decision is made to have only one wire from a single generator connected to each consumer. Once a monopoly is established, a company might be tempted to take advantage of the situation and raise the price of electricity to the point where it would become cheaper to have competitive suppliers with multiple generators and transmission and distribution lines.

There is no inherent impediment to keep monopolists from charging high rates other than their conscience, usually cast aside in the process of becoming monopolists, and the threat of consumers throwing the switch and doing without. To prevent a natural monopoly from behaving like an actual monopoly, government bodies granting franchises to create natural monopolies also established regulatory agencies to govern rates and oversee business and operating practices. Rates set by regulators cover operating costs and provide a fair rate of return on the monopolists' investments. A fair rate of return takes into consideration the return that can be earned by investing in other businesses of similar risk. A regulated utility serving a franchise area, with rates set to cover costs and provide a return on investment, has little risk compared to a manufacturing company. A manufacturing company must compete against others for the consumer's dollar with little in the way of consumer allegiance if a competitor brings out a better product with a lower price. An adequate return reflecting the inherent risks ensures that a regulated company can attract sufficient capital to build its asset base to satisfy customer demand.

Integrated utility companies provide the complete package of generation, transmission, and distribution for a designated area where a single rate covers all costs. Integrated utility companies can obtain high credit ratings if the regulators ensure that rates provide ample cash coverage of interest expenses. A high credit rating results in lower interest rates on debt issued for capital expenditures, which in turn reduces interest expense and, thus, electricity rates. Too high a rate would be reflected in a higher return on investment than warranted for the business risks faced by a regulated utility. If, on the other hand, regulators are too eager to squeeze electricity rates for the benefit of consumers, they are also reducing cash flow coverage of interest expense that can lead to a cut in a utility's credit rating. This results in higher interest rates as investors compensate for the greater perceived risk of default by demanding a higher return. Rates then have to be increased to compensate for the increased interest expense. If the regulators squeeze rates too far for the benefit of consumers, then a utility may not be able to raise the necessary capital to sustain its asset base, making it unable to meet its obligation to provide an ample supply of electricity to a growing population with higher expectations.

Too little pressure results in high electricity rates and a return on the utility's investments above that of a fair return, reflecting its inherent risks. Too much pressure on rates by regulators can threaten a utility's operational as well as its financial viability. Regulators must walk a fine line in approving rates, balancing the opposing needs of providing low-cost electricity to the public and ensuring that a utility has the financial wherewithal to carry out its obligation to the public.

On the surface, regulation of rates based on cost plus a reasonable return appears to be a sound approach for ensuring that a natural monopoly is properly funded, enabling it to provide its intended service at a reasonable cost. However, there are two problems associated with regulation of cost-based rates. The first is the absence of any incentive to be efficient because all operating costs are rolled into the rate base. In fact, there is an incentive to be a little inefficient when rates are being negotiated to obtain a higher rate, then improving efficiency after the rate has been set to enhance profitability. The second is the incentive to overinvest in plant capacity as the return is not only competitive but also more or less guaranteed by the rate-setting mechanism. To combat these drawbacks of cost-based rates, regulators review a utility's operations. Regulators have the power to replace management if operations become too inefficient. With regard to overinvestment, regulators normally insist that a utility clearly demonstrate that new electricity-generating capacity, or any significant capital investment, is needed before approving the expenditure of funds. Despite the best attempts by regulators, who are themselves subject to influence by those being regulated, a lingering suspicion existed that cost-based rates were higher than necessary.

This turned out to be the case when prices fell after the privatization of the British electricity industry. In 1988, concerned over what was perceived to be overpriced electricity from cost-based rates, the British government under Margaret Thatcher announced its intention to privatize the government-owned and -operated electric utility industry. The transformation of a socialized industry to several competing commercial enterprises as part of a national energy policy began in 1990 and was essentially completed by 1999. During this period, consumers experienced a 20 percent decrease in retail prices, 34 percent for small industrial customers and 7–8 percent for medium and large industrial consumers. The overall decline in wholesale electricity prices averaged 2.1 percent per year, demonstrating the ability of market pricing to lower electricity costs to consumers over regulatory cost-based pricing.[4]

In the United States, the roots of deregulation—some prefer to call it liberalization because the electricity utility industry is still highly regulated under deregulation—go back to the 1973 oil crisis. President Nixon's Project Independence was aimed at reducing the nation's dependence on oil and natural gas by switching to other fuels and encouraging energy efficiency and conservation to cut overall energy demand. At that time, oil and natural gas each contributed 20 percent of the fuel consumed in electricity generation. Project Independence sought to cut oil consumption in electricity generation to reduce imports. While natural gas was indigenous, there was a belief that a natural gas shortage might develop if there were a significant switch from oil to natural gas for generating electricity. Project Independence focused on the development of nuclear power, coal, and renewables for generating electricity.

The electricity-generating industry operated under the Public Utility Holding Act of 1935, a law that had dismantled the pyramid utility holding companies of the 1920s that went bankrupt during the Great Depression of the 1930s. The Act restored the utility business to its original state in which a single corporate entity provided electricity (and natural gas) to a franchise or specified area protected from competition. Within their franchises, utilities were lords and masters of generation, transmission, and distribution, subject, of course, to regulatory authorities. This cozy arrangement ended after the oil crisis. Congress, fearing that utilities would resist adopting new technologies, passed the Public Utility Regulatory Policies Act (PURPA) of 1978. PURPA required state regulatory commissions to establish procedures for qualifying facilities (QFs) that were not utilities to sell electricity made from renewable energy sources, waste, and cogeneration plants run on natural gas to utilities. Cogeneration plants, as noted, have a high thermal efficiency, double that of a conventional plant, because they can utilize waste heat as a source of hot water for industrial or processing purposes. PURPA could be viewed as a form of government coercion in support of cogeneration plants and renewable energy sources.

Utilities were obliged to buy electricity from QFs paying the "avoided" cost, the amount that a utility would have to pay for replacement electricity if it did not buy electricity from the QF. If the avoided cost made it profitable for independent power producers (IPPs) to invest in qualifying electricity-generating facilities whose output had to be purchased by utilities, then so be it. Some states, most notably California, required utilities to buy electricity at a price above avoided cost in order to jump-start new electricity-generating technologies involving solar, wind, and biomass.

The overall effect of PURPA was to raise the price of electricity and, by this narrow definition, could be considered a failure. But PURPA was the first intrusion of independent third parties into the monopoly of electricity generation by inadvertently taking the first step toward liberalization. PURPA also unintentionally challenged the concept of having a few large nuclear and coal-fired plants supplying a wide area through long-distance transmission lines. These centralized plants were burdened with billions of dollars of cost overruns, resulting in a cost of electricity far higher than originally envisioned. As these plants established the avoided cost, PURPA opened the door to having a more distributive system in which smaller capacity generating plants fueled by renewables and cogeneration plants run on natural gas served a more limited area.

The fear that utilities would exercise their monopoly control over transmission lines to make it difficult for QFs to develop a competitive market for electricity was dealt with in the Energy Policy Act of 1992 and the Federal Energy Regulatory Commission (FERC) Order 888 of 1996, which began the transformation of electricity transmission into a common carrier. These three legislative acts (PURPA, the Energy Policy Act of 1992, and FERC Order 888) established the opportunity for the emergence of wholesale competition in electricity within the regulatory framework governing natural monopolies.

Deregulation/liberalization entails the unbundling of generation, transmission, distribution, and system operation. As an integrated utility, one rate covered all operating and capital costs associated with generating and delivering electricity. Since competition was concentrated in third-party access to electricity generation, and did not cover transmission and distribution, it became necessary to break a single cost into three separate cost components for generation, transmission, and distribution, an accountant's delight to say the least. But shifting the price of generating electricity from cost-based to market price created an immediate problem for integrated utilities. Under the old regulatory regime, cost overruns such as those associated in building nuclear-powered plants were simply rolled into the rate base. Discounting the future stream of profits of rates based on cost resulted in a generating asset being carried on the balance sheet at a book value that reflected cost overruns.

With third-party access to generation permissible and with IPPs relying on more energy-efficient, lower capital cost generators run on natural gas (whose price had fallen when the natural gas "bubble" appeared, and hung around for two decades after the first energy crisis), market rates for electricity fell below cost-based rates. The market rate of electricity for nuclear power and other large plants plagued with huge cost overruns, when discounted into the future, did not create a book value for these generating assets that was even close to covering their capital costs. This would have necessitated writing down the book value of the assets to their market value, resulting in a diminution and, in some cases, an elimination of shareholders' equity. This difference in asset value between cost-based and market-priced electricity rates was given a name: stranded costs. To save utilities from having their creditworthiness impaired—resulting in lower bond ratings and higher interest rates—an incremental charge was added to electricity rates to cover stranded costs. This increment, charged to all sources of electricity including IPPs, was paid to the affected utilities until their stranded costs were liquidated. The existence of stranded costs was proof positive that rates based on costs were not the most economical way to produce electricity.

Operating Models in an Era of Deregulation/Liberalization

Where once there had been one model for the electricity business, now there are four. The first model is the traditional vertically integrated monopoly still operating in many parts of the world where rates are regulated to cover costs and provide an acceptable rate of return on capital assets. The second model resulted from the PURPA legislation in 1978 that gave IPPs third-party access to utilities. This initial step in liberalizing the industry took the form of a utility entering into a long-term contract to buy the entire output of the IPP generating plant. An IPP was forced to enter into a life-of-asset contract with the utility; otherwise, its investment was at risk. The utility was the IPP's only customer because the IPP did not have open access to transmission lines. Without such access, the IPP could not compete with the utility by entering into a contract with individual customers to supply their electricity needs.

In addition to the United States, this model has been adopted fairly widely in Asia and South America as a means to attract private capital for increasing the generating capacity of state-owned utilities. The creditworthiness for financing the building of the generating plant relies primarily on the nature of the contract between the IPP and the state-owned enterprise, not on the IPP's creditworthiness. Of course potential investors scrutinize the IPP to ensure that it can carry out its operational responsibilities, but security for repayment of debt lies almost exclusively on the sales contract. By issuing what essentially is a self-funding life-of-asset contract to buy the entire output of the IPP's generating plant, a state-owned enterprise does not have to tap external sources of funds or borrow from the government to increase its generating capacity.

The third model gives IPPs access to the transmission system and the ability to enter into contracts with large consumers such as distribution companies and large industrial enterprises and, thereby, compete with the utility. This model was first put into effect in Chile, followed by Argentina, the United Kingdom, and New Zealand. The United States does not have a national policy on how the utility industry is to operate, but the third model was instituted in parts of the United States through utility pools.

In the third model, each generator, whether owned by a utility or an IPP, pays a fee for the use of the transmission system. The fee covers the operating and capital costs of the transmission system, in effect, converting the transmission system into a common carrier in operation, if not in actuality. The rate can be a "postage stamp" rate, which is the same regardless of the distance of transmission, or be based on distance. The latter is preferable because it provides a better means of funding the installation of new, or replacement of old transmission capacity. Each generator becomes an independent supplier regardless of its ownership, selling electricity under a variety of contractual arrangements with buyers. Term contracts run for a period of time, fixing the cost of electricity for consumers and the revenue for providers. Term contracts account for the bulk of generated electricity and are arranged directly between the generator owner, either a utility or an IPP, and the consumer, a utility's distribution company or a large industrial consumer, or indirectly through intermediate market makers. The remaining electricity is bought and sold on the spot market. Consumers submit bids on what they are willing to pay for specified amounts of electricity that cover their needs in a specified time frame and providers submit bids for what they are willing to sell the output of particular generating units during the same time frame. These bids can apply to hourly intervals in the day-ahead market and the current spot market, but shorter time frames can be used. A computer program determines the clearing price at which supply meets demand for each hour or specified time frame of the day-ahead market and for the current spot market.

The day-ahead spot market is a contractual arrangement that meets anticipated needs. The spot market handles differences between the planned and actual use of electricity. These differences arise from a buyer not needing all, or needing more, electricity than anticipated, unexpected generator problems that reduce availability of contractual electricity, transmission congestion, and actions necessary to keep the system stable. Day-ahead and spot prices are not determined for an entire region, but at node points (sites of generating capacity) or defined zones. Price disparities between nodes or zones are primarily determined by system transmission-capacity constraints, line losses, and differences in generating costs. Price disparities provide useful economic signals for determining the size and location of additional transmission and generating capacity; something entirely missing under cost-based rates.

Obviously, generators have different fixed and variable costs depending on their capital investment, depreciated value, efficiency for translating energy to electricity, type and price of fuel, operating and maintenance costs, and the nature of ownership. Generating plants owned by the U.S. government and by state and municipal authorities operate in a tax-free environment and have access to more favorable financing alternatives than investor-owned utilities and privately owned generators. Some of these factors affect marginal costs, which play an important role in the rate-setting mechanism. Marginal costs normally reflect the fixed costs of operation including fuel with the investment considered a sunk cost. Rates fixed at the marginal cost generate the minimum revenue necessary to meet the cash operating needs of a utility, but not its capital costs in terms of servicing debt or paying dividends. Continual operation at marginal costs will eventually drive a utility out of business. Marginal cost is like a taxi fare that only covers the costs for the driver, fuel, insurance, and maintenance, with no funds available to pay the cab's financing charges or being accumulated to buy a replacement cab.

Coal, nuclear, and water-powered plants have the lowest marginal costs because of their relatively low fuel costs (hydro plants have no fuel costs). Nuclear power and coal-fired plants do not respond well to fluctuating demand because they require considerable time to ramp output up and down. These plants normally supply base-load electricity and enter into term contracts for most of their output. Generators whose output is more easily adjusted (natural gas and hydro) tend to be more exposed to fluctuating demand. Coal and nuclear plants bid at their marginal costs in order to secure employment at night, locking out higher marginal-cost generators fed by natural gas. As demand increases, natural gas generators submit bids for different quantities of electricity at their higher marginal costs that are tallied until demand is satisfied. The wholesale electricity rate is determined by the rate necessary to clear the market; that is, the rate submitted by the last generating plant whose output, when added to all the others, is sufficient to clear the market (satisfy demand). This rate becomes the single rate paid to all providers, regardless of their bids. Now coal and nuclear generators no longer receive their low bid based on marginal costs that kept them busy all night, but a rate based on the marginal cost of the highest cost-generating facility that cleared the market. This higher wholesale rate provides the additional revenue to repay sunk costs (service debt to the bondholders and dividend payments to the equity investors).

There is a real risk in this business. If too many low-cost base-load plants are built, electricity rates will reflect their marginal costs for longer periods of time during a twenty-four-hour day. As a result, electricity rates may not remain above marginal costs long enough for these plants to recoup their sunk costs (return of and on investment). Moreover, higher-cost units built for transient demand will see their hours of employment and profitability curtailed when there are too many low-cost base-load plants. Independent power producers risk their capital when operating in the third model, which has led to bankruptcies that would have rarely occurred in a fully regulated environment.

The fourth model gives IPP access not just to principal utility customers, but also direct access to individual households and small businesses, which are handled by distribution companies under the third model. In the third model, the transmission company becomes a *de facto* common carrier to serve a buyer, the distribution company, which, in turn, supplies thousands or millions of individual consumers. In the fourth model, the distribution company also becomes a *de facto* common carrier and is paid a tariff that covers its operating and capital costs. Individual consumers select a provider (utility or IPP) to supply their needs. An electricity bill then has three components: the contractual arrangement with the provider, a common-carrier charge from the transmission company, and another from the distribution company.

The great advantage of the fourth model is the introduction of demand management. Demand management can only occur if time of day automated reading meters are installed in order for rates to reflect what is being paid to providers. This gives individuals and businesses an incentive to reduce electricity usage during hours of peak demand, when the price of electricity is high, by shifting a portion of electricity demand to periods of base demand, when the price is low. In this way part of the load associated with hot water heaters, washers, and clothes dryers can be switched from times of high electricity rates to times of low electricity rates, and consumers receive lower electricity bills by managing their load.

The third and fourth models of direct access to large and small consumers require separate control over transmission independent of generation. In England, it was relatively easy to separate transmission from generation, and transmission from distribution, during privatization of a government-owned industry. The UK government simply organized new corporate entities to serve these three functions as they wished, without much ado other than from those directly involved in managing and operating the proposed companies. Restructuring the electricity utility business is much more complicated in the United States, where generating plants are owned by private and public institutions. Municipal utilities own generating plants along with investor-owned utilities and also the U.S. government, which owns hydro- and nuclear-powered plants under the Tennessee Valley Authority in the east and hydro plants (Hoover, Glen Canyon, Grand Coulee, and others) in the west. The dichotomy of ownership also exists for transmission lines in the United States. Transmission systems within an integrated utility's franchise area are owned by the utility. Interconnecting transmission systems may be owned piecemeal for those portions of the system passing through a utility's franchise area or by a separate corporate entity or the U.S. government. The U.S. government owns transmission lines associated with its generating plants and has also played an active role in providing loans and grants to utilities to build transmission lines to rural America under the Rural Utilities Service (formerly the New Deal Rural Electrification Administration).

In a deregulated system, utilities continue to own, operate, and maintain transmission lines, but they cannot have any real or perceived influence or control over their usage. If utilities could influence or exercise control over the transmission system, then the transmission system could be employed to their advantage and to the detriment of IPPs. This would hinder the formation of a wholesale market for electricity in which rates are determined by supply and demand, not by those who have control over access to the transmission system.

In addition to open access to the transmission system, the rate-setting mechanism for wholesale providers of electricity utilities and IPPs cannot allow any single provider to dominate the market. Studies have indicated that market domination might occur if any single participant has more than a 20 percent share of the business. This implies that a market free of manipulation that responds only to underlying shifts in supply and demand must consist of at least a half dozen independent and somewhat equally sized participants. Of course, the more participants there are, the better the market in terms of depth (volume of transactions and the number of parties buying and selling) and freedom from potential manipulation. The mechanism for determining price should be efficient (similar to a stock exchange), liquid (easily transferable obligations to buy or sell electricity), and transparent (transactions displayed and known to all participants). Besides equal access to transmission and the right to compete in order to get the business of distribution companies and major consumers, no cross-subsidies (regulated activities underwriting unregulated activities) can be allowed, and a mechanism for dealing with environmental issues must be instituted that does not interfere with the workings of the marketplace.

Deregulation requires a restructuring of integrated utility companies, separating generation from transmission; if not in ownership, certainly in operation. Historically, the system operator was responsible for the operations of a single integrated utility that owned the generating units, transmission, and distribution systems within its franchise area. The allegiance of a system operator cannot be dedicated to a utility when IPPs are trying to cut deals with the utility's customers. An Independent Systems Operator (ISO) must be established that acts impartially and is not beholden to any provider. The ISO is responsible for scheduling and dispatching (turning generators on and off), for accommodating demand—taking into consideration bilateral sales agreements between buyers (consumers) and sellers (owners of generating units), transmission constraints, and system stability. ISOs in the United States and Canada are responsible for the operation of groups or pools of utilities that cover a number of states and provinces. ISOs also control the system operation of large areas of Australia and China and entire nations such as Argentina, England, France, Mexico, and New Zealand.

The transmission system acts as a separate company under the operational control of the ISO, in effect a common carrier giving no preference to users and charging a regulated tariff that covers operating and capital costs. It would be preferable if the transmission company were truly independent with the general public, financial institutions, utilities and IPPs owning shares. The present ownership of transmission companies in the United States is split among utilities, corporations, and the U.S. government, with each owning various sections of the national transmission grid system. This arrangement makes decision making on expanding capacity cumbersome. Decision gridlock has not been the cause of limited building of new transmission lines in the United States; rather, the cause is local opposition or BANANAS (building absolutely nothing anywhere near anybody syndrome), which affects many industries. Without building new transmission lines, the United States is consuming the spare capacity of an aging system that can only result in future trouble.

The distribution function of an integrated utility becomes a separate operation in the third and fourth models. While presently owned by utilities, a better alternative for the third and fourth models would be for distribution companies to become independent entities with their own shareholders. In the third model, the utility cannot influence the distribution company's electricity purchases. But there is nothing that precludes a distribution company from entering into a term contract with its owning utility as long as other IPPs have been given equal access to bid for the business. The distribution company charges a regulated tariff that covers its operating and capital costs plus its electricity purchases. At the present time, the regulated rates for distribution companies cover all costs, including the cost of electricity. There really is no incentive for distribution companies to buy from the lowest cost source because the cost of purchasing electricity, no matter what it is, becomes part of the rate base. There has been some movement by regulators to set up an incentive system that rewards distribution companies if they can demonstrate that they have been more successful in seeking the best deal for their customers than other distribution companies. This reward could be in the form of incremental profits based on a portion of the difference between actual purchases and the average purchases by other distribution companies.

Competition in a deregulated, or liberalized, environment is primarily focused on electricity generation—transmission and distribution are still regulated activities essentially free of competition. Ideally, transmission and distribution companies would become separate corporate entities that own the assets rather than the assets being owned by investor- and government-owned utilities. They would charge a regulated rate to cover their operating and financial costs. Under the third model, the rate would also include the cost of electricity because the distribution company is responsible for selecting the electricity provider for its customers. Under the fourth model, the customers must select a provider, thereby reducing the role of the distribution company to that of a conduit or common carrier for direct sales between generator owners (utilities and IPPs) and individuals and businesses. This places the responsibility for purchasing electricity squarely on the shoulders of consumers, not the distribution companies.

For this model to work, consumers need time-of-day meters that automatically communicate electricity use to the utility. This way the utility can charge for electricity by time increments that reflect rates charged by suppliers, which, in turn are influenced by supply and demand. Europe is paving the way in the development of demand management, which has been set up, at least partially, in England and Wales, Scandinavia, and Spain. The Nordic electricity grid (Norway, Sweden, Finland, and Denmark) installed 1 million AMR (automated meter reading) meters in 2005, which are expected to grow to 5–8 million meters by 2010. These meters read electricity consumption in five-minute increments and send the information via wireless satellite to providers.

AMRs save money by not having to employ an army of meter readers. Signals from the meter readers keep providers informed of customer usage and whether or not they are receiving service. This allows providers to quickly identify power interruptions and initiate action to restore service. AMRs also benefit providers beyond billing, collections, and customer service. The wealth of information gathered by AMRs can be integrated into asset management, energy procurement, operational control, risk management, and field operations.

Demand management benefits consumers by shifting electricity loads from high- to low-cost periods. Only a portion of the peak day load can be shifted to night, but whatever that portion is, it represents significant savings to the consumer. Demand management also benefits providers. With the shifting of a portion of demand from peak- to low-demand periods, the base load of the utility is increased, with a commensurate reduction in peak demand and the need to invest in peaking generators. Demand management also encourages aggregators to represent groups of consumers. An example of the power of aggregators to lower costs can be seen in some office buildings that aggregate telephone service for all their tenants into one account. The office building enters into a single contract with a communications company. The communications company bills only the office building, which then breaks down the billing to the individual tenants within the building, and receives a fee for this service that represent a portion of the savings. This gives office buildings, as aggregators of phone service, a powerful negotiating presence when dealing with competing communications companies. In the same way, aggregators of electricity representing a group of industrial and commercial users can increase the group's bargaining power with providers in contracting for electricity services. Aggregators could someday represent hundreds or thousands of individual households and small businesses as a single bargaining group.

Even though customers have a choice of providers, the distribution company itself must be a provider because some customers might refuse to make a decision, leaving the distribution company as provider by default. Moreover, external providers are not obligated to serve a customer with a checkered credit history. This leaves the responsibility of supplying electricity to less than desirable customers, creditwise, to the distribution company. This has been dealt with in England by installing pay-as-you-go meters where customers pay in advance for their electricity through insertion of a prepaid card into a meter, just as phone cards enable customers to pay in advance for calls.

While the fourth model envisions distribution companies as regulated common carriers, the distribution company is still a provider for customers who refuse to choose a third-party provider and/or have been rejected by third-party providers. This makes a distribution company a buyer of electricity. Distribution companies normally do not have generating facilities because they are on the wrong side of the step-down transformers fed by transmission lines from large-scale electricity generators. But renewable sources of electricity, namely wind, solar, and micro and mini hydro plants generate relatively small quantities of electricity at the stepped-down voltage. Should distribution companies be allowed to install generating capacity to serve its customers in competition with third-party providers—on condition that a customer can always switch to the third-party provider? If so, suppose that a distribution company, unburdened of transmission costs, could provide electricity more cheaply than other providers. How much generating capacity should a distribution company be allowed to install? Suppose that a distribution company could build enough generating capacity to cover the needs of its customers at a lower rate than third-party providers, such as by wind or solar or small hydro or nuclear plants. As long as its customers are free to switch to other providers, should a distribution company be allowed to transform itself into a resurrected integrated utility serving the needs of a local community? Would it be permissible for the distribution company to add even more generating capacity and transform itself into an IPP selling power into the grid as a provider rather than buying power from the grid as a consumer?

As these hypothetical questions suggest, deregulation (liberalization) is an evolving concept with a moving target. In the United States, the pace and shape of deregulation vary from state to state. Regulators, depending on the state, favor one of the first three models, although the third model is growing at the expense of the other two. The fourth model, direct access between generating facilities and consumers, appears to be the regulators' ultimate goal for the utility industry.

Utility Pools

Utility pools predate deregulation. Pools were not set up to challenge the concept of an integrated utility, but as a means of increasing system reliability among independent integrated utilities. In this way surplus capacity for one utility would be available to meet demand for another utility short on capacity. A *tight pool* is a pooling of utilities with membership restricted to a specific region. Two tight pools, one for New England and one for New York, were formed to share generating capacity among pool members for greater system reliability. New transmission lines were built to interconnect the pool members to allow one utility with surplus electricity generating capacity to support another facing a shortage. This ability to share capacity enhanced reliability, reducing the risk of blackouts. It also increased the productivity of generating capacity and reduced the need for peaking generators. But the exchange of electricity between utilities required an agreement on a rate for settling accounts, thereby creating a wholesale market between utilities in addition to the retail market between utilities and their customers.

In addition to the New England and the New York pools, three other pools were organized. The PJM pool is an open pool originally organized with utilities in Pennsylvania, New Jersey, and Maryland. The Texas pool is a closed pool limited to utilities in Texas. The California pool is an open pool including utilities in the western part of the United States and Canada. Having pooled their generation and transmission resources and created a wholesale market, it was relatively easy for the pools to admit IPPs when required by PURPA legislation.

Pools had a major impact on the involvement of FERC in electricity markets. As independent utilities serving their franchise areas, utilities are exclusively under state or municipal regulation. FERC only has jurisdiction over wholesale buying and selling of electricity between utilities and interstate transmission. While the Texas pool operating within the state of Texas escaped FERC oversight for interstate transmission, FERC had jurisdiction over wholesale transactions between pool members. For this reason, the spread of pools increased FERC involvement in electricity markets via wholesale deals and interstate transmission to a degree that was not envisioned when FERC was first established.

FERC's limited jurisdictional authority to act was corrected by the Energy Policy Act of 2005, which also abolished the Public Utility Holding Act of 1935. This legislative change better reflected the reality that electricity generation and transmission were no longer a local matter best handled by local regulatory authorities. Electricity generation and transmission had become a regional matter and a growing national matter as a result of the increased tying together of transmission grids and generating stations through pools and utility-to-utility marketing arrangements. Large portions of the nation's electricity grid on either side of the Rockies are integrated, but a large-capacity cross-Rocky Mountain transmission system would have to be built to fully integrate the electricity grid of the nation. This would allow electricity to flow from where it is least needed to where it is most needed across the country. Moreover, there is an increasing flow of electricity both ways across the borders with Mexico and Canada that could result in a continental or international electricity grid.

Two pools deserve mention. One is the highly successful PJM pool, an excellent example of how to organize and run a pool, and the California pool, an excellent example of what not to do. The PJM pool was the world's first electricity power pool, formed by three utilities in 1927 to share their resources. Other utilities joined in 1956, 1965, and 1981, which led to the PJM pool covering most of Pennsylvania, New Jersey, and Maryland. Throughout this period, system operation was handled by one of the member utilities. In 1962 PJM installed an online computer to control generation in real time, and in 1968 set up the Energy Management System to monitor transmission grid operations in real time. The transition to an independent neutral organization began in 1993 and was completed in 1997 with the formation of the PJM Interconnection Association, the nation's first fully functioning ISO approved by FERC.[5]

PJM also became the nation's first fully functioning Regional Transmission Organization (RTO) in 2001 in response to FERC Order 2000. RTOs operate transmission systems on a multistate or regional basis to encourage development of competitive wholesale power markets. PJM Interconnection coordinates the continual buying, selling, and delivery of wholesale electricity throughout its region, balancing the needs of providers and wholesale consumers as well as monitoring market activities to ensure open, fair, and equitable access to all participants. The PJM Energy Market operates much like a stock exchange with market participants establishing a price for electricity through a bidding process that matches supply with demand.

The Energy Market uses location marginal pricing that reflects the value of the electricity at specific locations and time. During times of low demand and no transmission congestion, prices are about the same across the entire grid because providers with the lowest-priced electricity can serve the entire region. During times of transmission congestion that inhibit the free flow of electricity, location marginal price (LMP) differences arise that can be used for planning expansion of transmission and generation capacity. The Energy Market consists of day-ahead and real-time markets. The day-ahead market is a forward market for hourly LMPs based on generation offers, demand bids, and scheduled bilateral transactions. The real-time market is a spot market where real time LMPs are calculated at five-minute intervals, based on grid operating conditions. The spot market complements that portion of the total market not covered by term contracts between buyers and sellers and unforeseen adjustments that have to be made for buying and selling transactions originally made on the day-ahead market.

PJM has expanded from its original base in Pennsylvania, New Jersey, and Maryland to include Virginia, West Virginia, the District of Columbia, a large portion of Ohio, and smaller portions of Indiana, Illinois, Michigan, Kentucky, and Tennessee. It is the world's largest competitive wholesale market, serving a population of 51 million with over 1,000 generating plants with total capacity of 164 GWs, 56,000 miles of transmission lines, and 350 members, and still growing. In addition to creating and serving this market, PJM is also in charge of system reliability including planning for the expansion of transmission and generator capacity. PJM has become a model emulated elsewhere in the world such as in Colombia, where the national pool has been expanded to include utilities in Ecuador and Peru, with plans to bring in utilities in Bolivia, Venezuela, and Brazil.

When Demand Exceeds Supply

Deregulation assumes that competitive interactions between independent suppliers of electricity (utilities, IPPs) would act in the consumers' best interests by lowering prices through greater efficiency of operations, productivity gains, and investments in capital assets that can generate electricity cheaply. Competition lowers the overall return on investment to a level that sustains the investment process without overly impoverishing or greatly enriching the investor. This assertion only holds true, however, when supply exceeds demand. The devil in free enterprise rears its ugly head whenever demand exceeds supply. When supply exceeds demand, the price of electricity falls to the marginal costs of the last provider needed to clear the market. When demand exceeds supply, there is no real impediment to how high prices rise other than individuals and companies pulling the plug.

The California electricity crisis of 2000 illustrates what can happen when demand exceeds supply. While demand exceeding supply affected not just California, but the entire western part of the United States, the peculiar regulatory framework set up in California provided a launch pad for a rocket ride that sent one major public utility company into financial oblivion and reduced others to a precarious state of illiquidity, squandered a state surplus, caused the issuance of bonds to ensure that tomorrow's taxpayers will pay for yesterday's utility costs, created unpaid receivables that will forever remain unpaid, caused financial distress among energy traders and merchants, locked the ratepayers into long-term, high-cost electricity contracts, and set off an avalanche of allegations, investigations, lawsuits, and countersuits to provide guaranteed lifetime employment to a gaggle of lawyers. In short, the California electricity crisis is a classic case study of how not to deregulate the electricity industry.[6]

As background, investor-owned utilities (mainly Pacific Gas and Electric, Southern California Edison, and San Diego Gas and Electric) supplied 72 percent of electricity to California customers; 24 percent was supplied by municipal utilities and the remainder by federal agencies. The investor-owned and municipal utilities had historically operated as vertically integrated monopolies generating, transmitting, and distributing electricity in their franchise areas. The California Public Utilities Commission (CPUC) set the rates for the investor-owned utilities on the basis of covering costs and providing a fair rate of return on vested capital while local authorities regulated the rates for municipal utilities. The CPUC was particularly aggressive in implementing PURPA regulations, opening up third-party access to electricity generation. CPUC forced the investor-owned utilities to enter into contracts at higher rates than what would have applied for conventional sources to justify third-party QF investments in wind farms, biomass- and waste-fueled generators, and cogeneration plants run on natural gas. By 1994, 20 percent of electricity-generation capacity in California was from cogeneration (12 percent) and renewables (8 percent), the highest proportions in the nation. Electricity rates to jump-start renewables, coupled with cost overruns on nuclear power plants, resulted in an average cost of nine cents per kilowatt-hour for California residents in 1998 versus a nationwide average of nearly seven cents per kilowatt-hour. Electricity rates in Hawaii, Alaska, New Jersey, New York, and New England were higher than in California.

In the belief that deregulation (liberalization) would lower retail prices, the CPUC aggressively set out to deregulate the electricity industry to give major customers a choice among competing providers of electricity. The CPUC's *Order Instituting Rulemaking on the Commission's Proposed Policies Governing Restructuring California's Electric Service Industry and Reforming Regulation* (R.94-04-031), commonly referred to as the *Blue Book*, in 1994 started the process of liberalization by first recognizing that existing utilities had stranded costs, such as nuclear power cost overruns, that had to be taken care of before the electricity market could be deregulated. New IPPs with no history of cost overruns could build a plant and offer electricity at rates that would bring financial ruin to existing utilities stuck with stranded costs. The *Blue Book* dealt with stranded costs by creating a rate increment that would be paid by all electricity buyers no matter what the source. The revenue would be directed to the appropriate utility to pay for stranded costs until they were liquidated. There was nothing wrong with this approach other than the CPUC capped retail rates until stranded costs were liquidated. The rationale for capping retail rates was that the CPUC believed that wholesale prices under deregulation would fall. As they fell, a larger portion of the difference between the capped retail price and the wholesale price would be dedicated to repaying stranded costs, hastening the time when stranded costs would be liquidated and the retail price cap removed. However, if wholesale prices rose, a smaller portion would be available for stranded costs, delaying the lifting of the retail cap. The financial strength of the utility would not be affected with capped retail prices and the repayment of stranded costs would offset changes in wholesale prices—as long as wholesale prices did not rise above the retail price cap. Since the unanimous belief was that deregulation would result in an overall lowering of wholesale prices, no one envisioned a situation in which wholesale prices would rise above the retail price cap.

The *Blue Book* was followed by a Memorandum of Understanding that created an independent system operator, the California Independent System Operator (CAISO) with the sole responsibility of managing the electricity grid, and an independent power exchange (PX) with the sole responsibility for managing the spot market in electricity. The only allowable markets were an hourly day-ahead and an hourly spot market with transparent prices and transactions. Whereas deregulation elsewhere called for a tightly integrated structure of managing the grid and overseeing the wholesale trading of electricity, the CPUC made these separate and independent functions with no coordination and limited information flow. This administratively imposed barrier on the interchange of information between CAISO as operator and PX as market maker created inefficiencies that became made-for-order profit opportunities for energy traders and independent merchants. The *Blue Book* and the Memorandum of Understanding set the stage for the passage of Assembly Bill 1890 in 1996, which became effective in 1998. Although the investor-owned utilities still owned generating units, transmission lines, and distribution systems, they could not translate ownership to operational control. Control of transmission would be handled by CAISO and all generated electricity would be sold to the PX. An investor-owned utility supplying its customers would first have to sell its electricity to the PX and then purchase electricity from the PX with CAISO handling the transmission details.

Within the western region, California accounted for 25 percent of electricity consumption. The state was a net importer of electricity during the summer from the increased air-conditioning load and a net exporter to the Pacific Northwest during the winter. Thus, the California utilities were net consumers of electricity when the crisis occurred in the late spring and summer of 2000, purchasing more electricity from the PX than they supplied. In the dubious belief that the only way to create a market with substantial depth to reflect the true value of electricity was to channel all sales through the spot market, the CPUC prohibited investor-owned utilities from entering into term contracts to fix the cost of their purchased electricity. This prohibition was put into effect by the PX mandated as the only conduit for sales and purchases of electricity by the investor-owned utilities. But the PX was limited to buying and selling electricity on the day-ahead and current spot markets. This made it impossible for the investor-owned utilities to enter into term contracts, but municipal utilities could act independently and enter into term contracts with providers because they were not regulated by the CPUC.

The PX operated on a day-ahead basis, accepting bids from each generator to sell its output at some offering price and each investor-owned utility's distribution company or major customer indicating the amount of electricity to be purchased on an hourly basis. Offering-price bids were ranked from the lowest to the highest and their volumes accumulated until they met demand. The price at which the amount of electricity from the accumulated bids by suppliers equaled the amount of electricity required by purchasers became the hourly market-clearing price for all bids. All sellers received the same market-clearing price even if they had bid less than the clearing price. Sellers who had bid more than the market-clearing price would have no market outlet for their generating units. The underlying rationale for this pricing mechanism was that the risk of bidding too aggressively would result in idle generating units. This fear of idle capacity would encourage bidders of generating capacity to price electricity close to the marginal cost of each generating unit. This rationale held true as long as supply exceeded demand.

A computer system was set up to handle twenty-four separate markets for each hour of the current day and the day-ahead market. Providers basically had to guess at what would be the appropriate bid for each hour and those with multiple generating units would be playing an hourly price-bidding game for each of their units to try to maximize company revenues. Owners of various type plants would bid low on those units that best served base needs to ensure their employment, and higher on those units whose output could be more easily changed to try to capture incremental revenues.

What was not envisioned was how the system would behave if nearly all the generating capacity was needed to satisfy demand. Under these circumstances, owners of generating capacity became emboldened to bid more aggressively for their units that were dedicated to satisfying variable demand. There was less risk of being left with idle capacity because most units had to operate to meet demand. Moreover, the financial loss of being left with an idle unit was less because of the higher clearing price for the operating units. In a tight market with little leeway between system demand and system capacity, meaning few idle generators, a new pricing pattern emerged that was never seen before. It was dubbed the hockey stick pattern. When surplus capacity was plentiful, the price for electricity rose slowly in response to large increments in demand. When surplus capacity became scarce, the price rose sharply in response to small increments in demand. The combination of these two price patterns as demand approached the limits of supply looked like a hockey stick.

Jumps in the spot price were particularly harmful to investor-owned utilities in California who, as net consumers of electricity, were forced to buy and sell exclusively in the spot market. Municipal utilities in California and utilities in other states and provinces of the western region outside the jurisdiction of the CPUC had entered into fixed-rate term contracts for the bulk of their electricity purchases, thereby escaping the financial carnage faced by the California investor-owned utilities. While spikes in spot prices starting in California spread throughout the western region, they had limited impact on the aggregate cost of electricity throughout the system because most electricity needs were filled by fixed-price term contracts. "Throughout the system" was, of course, true everywhere and for everyone except the investor-owned utilities in California whose net electricity purchases were funneled entirely through the spot market.

Another adverse consequence of prohibiting investor-owned utilities from entering into term contracts affected the construction of new electricity generating capacity in California. Investors could not reduce their financial risk by entering into term contracts with the investor-owned utilities that made up 72 percent of the market. They could, of course, enter into term contracts with municipally owned utilities to assure at least partial employment, but that excluded much of the market. Without assurance of employment, investors were generally reluctant to bear the financial risk of building new capacity. On top of this, plants under construction in California faced public hearings, permitting inspection hurdles that delayed the start of construction by as much as two or more years compared with other western states.

The separation of responsibilities between the system operator, CAISO, and the PX, as market maker, and the prohibition for these two organizations to coordinate their activities and interchange information, forced CAISO to become a buyer on an immediate spot basis. This was the only way for CAISO to handle mismatches between supply and demand that CAISO was not allowed to communicate to PX. Thus, there came into being two spot markets: one run by PX and the other by CAISO. With limited information transfer between the two, energy traders and merchants had a field day taking advantage of price disparities between these two separate markets. To make gamesmanship a little easier for energy traders to play one market (PX) off the other (CAISO), the computer coding for the CAISO model for determining the price of electricity was in the public domain.

Shortly before the emergence of the crisis in 2000, the three investor-owned utilities were 57 percent reliant on natural gas to run local generating units, 12 percent on nuclear power produced locally, 13 percent on hydropower imported from the Pacific Northwest, 5 percent on imported electricity from coal-burning plants in the Southwest, and the remaining 13 percent renewables (wind, geothermal, biomass, and solar). Growth in natural gas consumption for electricity production was beginning to strain pipeline delivery capacity and the surplus of generating power throughout the western region had been eroded by demand growing faster than supply in the preceding years. A drought in the Northwest forced a reduction in hydro output. This became the precipitating event that led to an overall shortage of capacity to satisfy California electricity demand just as it began to climb toward the seasonal peak in the late spring of 2000.

Before the crisis erupted, the wholesale price in the western region varied between $25–$40 per megawatt-hour, equivalent to $0.025–$0.04 per kilowatt-hour. The average retail price of $0.09 per kilowatt-hour also included distribution and stranded costs pertinent to California. Remembering that retail prices in California were capped and all purchases and sales by the investor-owned utilities had to be transacted through the spot market on the PX, and that they were net buyers of electricity during this time, an intolerable cash-flow squeeze occurred when wholesale prices jumped to $75 per megawatt-hour in early May. This was followed by a decline, then a surge to $175 in mid-May, followed by a decline, then a surge to $300 in early June, a decline, then an all-time record spike of $450 in mid-June, again a decline, then another surge to $350 in late July. At these prices, aluminum smelters and other industrial concerns in the Northwest laid off their workforce in order to sell electricity that was either produced at the facility or had been purchased cheaply on long-term contracts in the spot market. This shutdown of industrial output actually increased the supply of electricity in the western region and contributed to limiting the crisis (the laid-off workers had another view of the situation).

With a tight market in which nearly every generator had to be employed to meet demand, providers, knowing that few of their operating units would remain idle, became extremely aggressive in their bidding. A provider with multiple units could afford to bid high on a couple of units as the probability of ending up with an idle unit was pretty low. Moreover, since the highest bid that cleared the system would apply for all bids, the financial loss of having an idle unit with the rest employed at high rates would be an acceptable outcome. This change of attitude—from fear of idle capacity giving way to unrestrained greed—was reflected in the hockey stick price pattern. To add misery to woe, environmental rights to emit pollution had been issued in California, based on actual nitrous and sulfur oxide emissions in 1993. The intention was for the issuing authority to slowly decrease the availability of such rights. The staged retirement of these rights to emit pollution resulted in a higher price, providing an economic incentive for utilities to build new and cleaner-burning plants or add equipment to existing plants to reduce pollution emissions. This program was successful in gradually reducing pollution emissions by utilities.

But in 2000, with California experiencing rolling blackouts (although these blackouts gained national notoriety, only six occurred, each affecting only a small portion of the population for a relatively short period of time), every plant in California had to be put into operation to generate electricity. This involved reactivating previously mothballed plants with high pollution emissions. These could not be operated without purchasing emission rights. The shortage in emission rights sent their price through the ceiling and added to the cost of generated electricity that could not be recouped from customers. In the midst of the electricity crisis, legal actions were being taken against utilities for not having the necessary pollution rights to cover their emissions. The utilities were faced with an impossible choice: to fulfill their public obligation to supply electricity by breaking the law (not buying the requisite rights to cover total emissions) or to obey the law (buying the requisite rights at extremely high prices and thereby aggravating their cash drain). While it was possible for them to cut back on their electricity generation to reduce their need for emission rights, this would have caused more chaos in the market, more extended blackouts, higher rates, and charges of manipulating the market. Not only did politicians stand fast on doing nothing to increase the volume of pollution emission rights under these dire circumstances, but they also stood fast in making sure that retail price caps remained intact.

Thus, the investor-owned utilities were drained of all their liquidity by buying high and selling low, leading to the bankruptcy of one and the insolvency of the others. CUPC's insistence on not allowing retail price relief was challenged as a violation of the due-process clause of the Constitution to no avail (the state presumably is not allowed to rob shareholders of their wealth without giving them due process for redress). Electricity providers became increasingly unwilling to accept payment other than cash in advance from utilities rapidly becoming insolvent. Refusing to sell electricity through the PX to investor-owned utilities, the state of California was forced to step in and buy electricity for the investor-owned utilities. Now it was California's turn to buy high and sell low, which quickly squandered its entire surplus. Although California had prohibited utilities from entering into term contracts when wholesale spot prices were low, now California itself entered into term contracts with sellers for large quantities of electricity when wholesale spot prices were at record-breaking highs.

During this entire crisis, retail customers, other than being inconvenienced by an occasional rolling blackout, had no economic incentive to reduce consumption. The only action the state took to reduce demand was to order state office buildings to cut electricity usage and initiate a program to subsidize the introduction of energy-efficient fluorescent light bulbs—hardly a palliative for the ongoing crisis. There was a concerted effort on California's part to ensure that state inspectors did all they could not to unnecessarily delay the completion of electricity-generating plants already under construction. The fact that they did hasten the completion of construction is a bitter commentary on their performance prior to the crisis. The crisis began to cool, along with the weather, in the fall of 2000, which reduced the air-conditioning load and the need to import electricity. Eventually, the completion of additional electricity-generating plants in California, and elsewhere in the western region, added enough capacity to restore a surplus and a semblance of order to what really should be a very orderly business.

Having bankrupted one utility and left others stripped of cash, California had to issue bonds to restore the surplus squandered by buying high and selling low. This enabled California taxpayers to foot the bill plus interest over the long term for what they did not have to pay in the short term. And, to complete the picture, the term contracts entered into by California, while attractive when they were inked with record-high wholesale spot prices, became decidedly unattractive when wholesale spot prices fell to precrisis levels. The people of California are not only saddled with repaying billions of dollars of bonds to restore the state's liquidity, but also spending more needless billions of dollars for high-priced electricity fixed on term contracts. Not all the purchases of high-priced electricity were paid. The bankruptcy of the PX in January 2001 left those holding PX receivables with something good only for papering their bathroom walls.

Eventually retail electricity rates were raised substantially, but in a manner that had limited-impact households consuming less than a baseline amount of electricity. Those who consume above the baseline amount face significant step-ups in rates. FERC eventually banned utilities from having to buy and sell all their power through the PX or CAISO, restoring the old world in which utilities could make deals in the forward markets, enter into term deals, and dedicate their generated electricity to supplying their customers. FERC attempts to rectify matters in other areas were resisted by state authorities, who are ultimately responsible for the regulation of utilities under their jurisdiction.

At this point, the California electricity industry is basically under state control. The crisis is past, but its legacy will go on for a long time in terms of repaying bonds, honoring high-priced contracts to buy electricity, rejuvenating financially crippled utilities, dealing with unpaid receivables, plus the accusations and investigations, suits, and countersuits. In 2006 the estimated cost of the California debacle to the state was $70 billion, of which $6.3 billion in settlements had already been made. Sixty different investigations of market manipulation and a host of criminal and civil trials were still in the works. For example, evidence was found of a generating plant that had been shut down to fix a boiler that did not need to be fixed; the only remaining reason was to further restrict supply in order to increase price. E-mails and tapes have been discovered that point to rather unsavory behavior on the part of some energy traders and merchants. But this is only the result of a fatally flawed market design set up by regulators giving suppliers the opportunity to take advantage of a resulting shortage. From the start of the energy debacle, and at all times during the debacle, everything that could have made a bad situation worse was done and everything that could have alleviated a bad situation was not done; truly the worst of all possible worlds.

The Real Lesson of California

The real lesson to be learned from the California electricity debacle is that rates become unstable when demand gets too close to supply. When supply is ahead of demand, rates are reduced to marginal costs, which is beneficial to consumers. Deregulation means lower prices only as long as supply exceeds demand. When demand gets too close to supply, rates for electricity—and prices for anything, oil, copper, gold, grain, you name it—do not escalate by a little but by a lot. All commodity traders know about the hockey stick pattern. There is little to moderate prices as buyers attempt to outbid one another for what is perceived to be a commodity in short supply. Escalating panic among buyers is matched by growing greed among sellers. This, of course, is the classic economic signal to increase capacity. The problem is that capacity cannot be added in a fortnight.

The original regulation of electricity—determining rates by covering operating costs and guaranteeing a reasonable rate of return on investment—also guaranteed surplus capacity. Indeed, this has been frequently cited as one of the drawbacks of regulation: With a guaranteed return, the temptation to build excess capacity is overwhelming. This was not limited to the number and size of generating plants, but anything that could be thrown into the rate base. The drawback of letting the market decide electricity rates is that the market does not reward spare capacity, but punishes the company that builds too much capacity by making it difficult or impossible to earn enough revenue to recoup its investment. As a consequence, companies tend to use modest growth rates for projecting demand when deciding on investing in additional capacity as a means to avoid the mistake of building too much capacity. While the market system reduces rates by minimizing the amount of capital invested in excess generating facilities, it also forces supply to be close to demand. This leaves little room for accommodating shocks to the system.

There is also a lesson to be learned from the experience of Colombia. Colombia has hydropower plants that supply a large portion of its needs. It is by far the lowest-cost source of electricity. But droughts can affect hydropower output. To accommodate this potential shock, the electricity-generating authorities of that country have entered into contracts for backup fossil-fueled electricity-generating capacity to be built, but not operated as long as hydropower is available. The operators of these plants are paid regardless of the output for these plants. Electricity rates reflect money spent for idle capacity built just in case it is needed. Some careful attention to this means of establishing spare capacity should be given in market-driven systems to reduce system vulnerability to shocks and avoid the pandemonium that breaks out when demand gets too close to supply.

Notes

1. The statistics on energy consumption for the primary sources of energy are from *BP Energy Statistics* (London: British Petroleum, 2005) and the statistics on biomass and renewables and the shares of various energy sources consumed for electricity consumption, electricity consumption, and capital costs are from *World Energy Outlook* (Paris: International Energy Agency, 2004).
2. See www.electricityforum.com and www.wikipedia.org for information on the history of electricity.
3. See www.amasci.com for articles on the common confusion over terminology concerning electricity.
4. Sally Hunt's *Making Competition Work in Electricity* (New York: John Wiley & Sons, 2002) is well worth reading for a more comprehensive view of deregulation.
5. See the PJM pool Web site at www.pjm.com.
6. James L. Sweeney's *The California Electricity Crisis* (Stanford, CA: Hoover Press, 2002), which describes what happens when a system breaks down, is also worth reading.

3

Biomass

In the twenty-first century, energy is not as it always was. Yesterday's world was entirely dependent on biomass, particularly wood for heating and cooking. A century ago biomass was eclipsed by fossil fuels. Biomass is generally viewed with disfavor as something associated with abject poverty. Yet there is another side to biomass; there is now something of a resurgence going on. As fossil fuel prices increase, biomass promises to play a more active role as a utility fuel, a motor vehicle fuel, and a supplement to natural gas. Biomass will never replace fossil fuels, other than on the margin, nor is there any hope that we can return to a world where biomass plays a significant role in satisfying society's energy needs. This chapter examines the past and present roles and then the potential for biomass as tomorrow's energy fuel.

Yesterday's Fuel

Until about 300 or 400 years ago, the world depended nearly exclusively on biomass as a source of energy. The population was low in relation to the number of trees. Nature simply replaced those that were chopped down for heating and cooking. The environmental impact was minimal because carbon dioxide released by burning wood was absorbed by the plant growth that replaced the burnt wood. With no net loss of tree resources, carbon dioxide was recycled, described by contemporary proponents of biomass as a closed carbon cycle or a sustainable system. Fossil fuels, on the other hand, release carbon dioxide that was locked away eons ago as partially decayed plant life and marine organisms.

Despite the environmental benefits of recycling carbon dioxide and emitting less nitrous and sulfur oxides than coal and oil, pollution—in the form of smoke from burning wood—would have filled the cave, tent, hut, or dwelling before someone devised the chimney. Smoke is a health hazard for the respiratory system and an irritant to the eyes. Early explorers observed that smoke from American Indian fires filled the Los Angeles basin with smog long before the automobile age. Now smoke from burning biomass contributes to the brown cloud overhanging much of southern Asia and to serious health problems in India and elsewhere in Asia where emissions from burning biomass are largely confined within living quarters.

Biomass maintained its dominance as a fuel source up to the beginning of the nineteenth century. With the advent of the Industrial Revolution, coal entered the picture first in Britain, followed by the United States and Germany, and then Japan. Even as late as 1850, coal only made up 10 percent of the energy mix and biomass provided 90 percent. By the mid-1870s biomass still contributed twice as much to fulfilling energy needs as coal. With industrialization proceeding at a rapid pace, biomass and coal were about evenly split by the end of the nineteenth century. Coal replaced charcoal for producing steel and split wood for fueling railroad locomotives and heating homes. Most "natural" gas piped into homes and businesses was actually manufactured gas from coal.

What little energy demand remained after biomass and coal was filled by hydropower (water mills turning shafts that, via belts, powered machinery) and oil. The automobile age had not yet begun, and oil was used mainly as kerosene for illumination and lubrication of machinery.

Today's Fuel

Biomass is still a major source of energy, though often excluded from energy statistics because of the inherent difficulty of gathering reliable data from remote areas where biomass is the principal source of energy. For many, biomass is a noncommercial source of energy freely gathered from the local environment. In recent years, biomass has been gaining ground as a commercial fuel purchased as charcoal for cooking, firewood for heating, and crops grown specifically for their energy content.

The estimate for biomass and waste for 2002 was 1.1 billion tons of oil equivalent in addition to the 9.5 billion tons of oil equivalent for all commercial sources of energy consumed.[1] Thus, the total commercial sources of energy—plus biomass and waste—was 10.6 billion tons of oil, of which the role of biomass and waste was about 10 percent. This is a rather impressive amount of energy.

Biomass takes many forms. It is carried on the heads of native women in semi-arid regions of Africa and Asia. Many of these women must trudge ten or twenty miles each day to find limbs of dead trees and dung of camels and other animals. Animal dung must be dried in the open sun before being burned and is a preferred energy source for mud ovens because it burns slowly and evenly and releases a great deal of heat. But the demand for dung from a growing human population is beginning to exceed the supply of droppings from camels and other animals that wander the countryside. Dung burned for fuel also robs the ground of a valuable fertilizer. Introducing an energy-efficient oven would reduce the demand for biomass fuels, but an individual who depends on dung or wood for cooking would most likely not have the financial wherewithal to acquire the latest model. While treks into the hinterland for wood and dung make for interesting TV documentaries and fascinating photographs in *National Geographic*, the reality is not so attractive. How many of these women would give up the romance of gathering wood and animal droppings for a small kerosene stove that could heat their hut and cook their food?

Biomass is organic matter primarily in the form of wood, crop residues, and animal waste, in that order of importance. Biomass as wood is readily available in temperate and tropical regions or, as mentioned, is collected with great personal effort in semi-arid areas. The great advantage of biomass is that it is free, and in temperate and tropical regions, freely available. Wood can be burned directly or be first transformed into charcoal through pyrolysis: the heating of wood in the absence of sufficient oxygen to prevent full combustion. Organic gases and water are evaporated, and leave charcoal, which is nearly pure carbon. Burning the released gases provides the fuel for pyrolysis and can be used to dry fresh wood. Any backyard barbecue hamburger-flipping aficionado can recite the virtues of charcoal over wood: higher heat content, cleaner burning, and conveniently transportable.

Generally speaking, since biomass is "free," it is inefficiently utilized as a residential, or commercial, fuel. For instance, about two-thirds of the energy content of wood is lost when it is transformed into charcoal in developing nations. What does inefficiency mean other than greater personal effort when the wood is freely gathered from the local environment? Most proposals for utilizing biomass in developing nations emphasize energy efficiency to reduce the input of biomass to produce the same output.

While biomass is estimated to make up 10 percent of all energy consumed, its pattern of consumption varies enormously from nation to nation. The industrialized nations rely on biomass for only about 3 percent of their energy needs. Biomass is burned for heating homes during the winter in New England and other parts of North America and northern Europe. Biomass can be firewood split from logs or bark and edgings residue from a lumber mill. Fireplaces burning split logs provide an attractive background setting in the living rooms of millions of homes. Unfortunately, conventional fireplaces allow most of the heat to escape up the chimney. Some fireplaces may actually increase heating needs by acting as a pump transferring warm air from inside to outside the house. When people depend on biomass to heat their homes, the wood is burned in specially designed space heaters where relatively little heat escapes, along with the products of combustion, to the outside.

Wood residue is an important source of biomass. As much as 75 percent of a tree becomes residue, beginning with the leaves, tree top, branches, and stump left in the forest, to the bark, edgings, and sawdust produced when a log is transformed into lumber, and to the shavings, edgings, and sawdust of making lumber into a finished product. Bark and other wood residue can be used for residential heating, as an industrial fuel by supplying power for lumber mills and other manufacturing activities in the developing world, and for producing electricity in developed nations such as Finland and Germany.

Some sub-Saharan African nations such as Burundi and Rwanda are over 90 percent reliant on energy from biomass while others are 70–80 percent reliant on it for their total energy needs, which includes commercial and industrial as well as residential demand. In terms of residential demand, nearly all rural households in Kenya, Tanzania, Mozambique, and Zambia rely on wood, and 90 percent of urban households rely on charcoal for cooking. Heavy biomass users in Asia are Indonesia, the Philippines, Thailand, Myanmar (formerly Burma), Vietnam, Bhutan, Laos, and Cambodia, and in the western hemisphere Guatemala, Honduras, Nicaragua, and Haiti.

Of the world population of 6.2 billion people (in 2002), an estimated one-third does not have access to electricity. Almost by definition, those without access to electricity depend on biomass. Even with access to electricity, many cannot afford to buy electricity and therefore remain dependent on biomass. With or without access to electricity, it is estimated that 2.4 billion people, or 38 percent of the world population, depend primarily on biomass in the form of wood, agricultural residues, and dung for cooking and heating. Half of these live in India and China, but sub-Saharan Africa has the world's highest per capita dependency on biomass. Not only does heavy reliance on biomass pose health problems, but it also contributes to ecological problems such as deforestation, which is occurring in parts of Africa, India, and elsewhere, in addition to the loss of dung as a fertilizer. As one may surmise, there is a direct link between poverty and dependence on biomass.

China

About 56 percent of the population relies on biomass in the form of wood and agricultural residues for cooking and heating. Most biomass is consumed in rural areas. The estimated 2002 consumption of biomass was about 213 million tons of oil equivalent, not far behind the 260 million tons of oil consumed in China. If we add 213 million tons of oil equivalent in the form of biomass to the total of commercial forms of energy of 1,035 million tons of oil equivalent, then biomass made up 17 percent of China's total energy consumption. Biomass consumption is expected to remain flat for the foreseeable future, balanced between a rising population and continued migration to urban areas, where commercial fuels are more widely and more efficiently used. Shifting from biomass to commercial fuels is considered beneficial because it reduces local pollution as well as aggregate energy demand.

India

As is the case in China, 58 percent of the people in India depend on biomass for heating and cooking. With similar populations, India consumed a little less biomass than China, less than 200 million tons of oil equivalent, which greatly exceeded the 111 million tons of oil consumed in India in 2002. Future biomass consumption is also expected to remain relatively flat. When 200 tons of oil equivalent of biomass is added to 340 million tons of oil equivalent for commercial forms of energy, biomass made up an impressive 37 percent of India's total energy consumption. Rural areas of India are almost entirely dependent on biomass, which is leading to widespread deforestation. Of course, where biomass consumption results in deforestation, biomass is no longer a closed carbon cycle or sustainable source of energy. By definition, deforestation means that more carbon dioxide is being released into the atmosphere than is being absorbed by replacement growth.

India has initiated an afforestation program in an area stripped of its indigenous evergreen forests. The aim of the program is to transform what has become wasteland back into forestland. If successful, the forest will reduce soil erosion and increase groundwater. The improved fertility and productivity of the soil will benefit agriculture in the surrounding area while the forest itself will provide employment opportunities and fuel. If this program is successful, similar afforestation projects will be undertaken elsewhere where deforestation has resulted in land degradation. The goal of the National Forestry Action Program is afforestation of a significant portion of the nation, with the local population supplying the labor and the government supplying the material.

Indonesia

Many remote and isolated islands of Indonesia and other island nations of Southeast Asia are not well served by commercial forms of energy. About 74 percent of the population of Indonesia depends on biomass for heating and cooking. In 2002, the total consumption in Indonesia of biomass, mainly wood, was 47 million tons of oil equivalent, slightly behind 53 million tons of oil. Comparing the consumption of biomass with the 104 million tons of oil equivalent of commercial sources of energy, Indonesia is 31 percent dependent on biomass. Biomass is used as an industrial fuel to provide steam for running lumber mills. Biomass would be an ideal fuel for micro-electricity-generating plants that could bring the advantages of electricity to isolated islands of Southeast Asia. While the most likely fuel is wood, it could also be bagasse, a residue from processing sugarcane, and rice husks.

Brazil

Brazil's estimated consumption of biomass in 2002 was about 44 million tons of oil equivalent compared to the total of 178 million tons of oil equivalent for commercial forms of energy. Thus, Brazil was 20 percent dependent on biomass. Consumption of biomass in Brazil is radically different from that of China, India, Indonesia, and other nations. As in these nations, biomass as wood and charcoal is consumed for cooking and heating in rural areas as a residential fuel; but in Brazil over half of biomass is consumed as a commercial or an industrial fuel. Companies in mining, cement, paper and ceramic making, and food processing rely on biomass as a fuel. Another unusual feature of biomass consumption is that most nations use coal to make steel, but Brazil has little in the way of coal reserves suitable for steel production. While Brazil imports some metallurgical coal, 80 percent of its charcoal output is dedicated to replacing coal in steel production.

Brazil is not alone in using biomass as a commercial or industrial fuel. Biomass is used in developing nations for smoking fish, curing tobacco, processing food, and drying bricks, lumber, furniture, and ceramics. However, Brazil stands out because of its greater reliance on biomass for commercial and industrial fuels and is unique among nations for its reliance on biomass as a motor vehicle fuel.

Biomass as a Motor Vehicle Fuel

Sugarcane grown in Brazil is consumed as a raw material for making ethanol, which, in Brazil, is a motor vehicle fuel. Ethanol, or ethyl alcohol, is what is found in alcoholic beverages. Ethanol, sometimes referred to as bioethanol (to reflect its origin), is made by fermenting sugar from sugarcane and sugar beets or by converting starch in grains such as barley, wheat, and corn (maize) first to sugar, then to alcohol. Ethanol from cellulose plant life such as wood chips and grass is currently a complex and costly process and research is being conducted to improve the process. Iogen Corporation, a Canadian biotechnology company supported by Shell Oil, Petro-Canada, and the Canadian government, operates the world's only facility that can convert cellulose biomass to ethanol using an enzyme technology. Making ethanol from agricultural wastes, wood chips, and fast-growing grasses would expand its availability without affecting food supplies.

The easiest way to make ethanol is to start with sugarcane or sugar beets. Grains such as barley, wheat, and corn must first have their starch content converted to sugar before converting the sugar to alcohol. This is accomplished by first milling grain with hammers into a fine powder called meal. Then the meal is mixed with water and an enzyme (alpha-amylase) and then cooked to liquefy the starch and eliminate naturally occurring bacteria. The resultant mash is ready for a process called saccharification whereby the mash is cooled and a secondary enzyme (gluco-amylase) is added to convert the liquefied starch to a fermentable sugar (dextrose). Then special yeast is added to ferment the sugar, producing ethanol and carbon dioxide. Fermentation continues until the bacteria "drown" in their alcoholic waste.

The fermented mash, called beer, contains about 10 percent alcohol as well as all the nonfermentable solids from the sugar or grain and yeast. The beer then passes through a distillation unit that separates the alcohol from the solids and water. Alcohol leaving the distillation unit is 96 percent ethanol and 4 percent water, the same as 192-proof country brew. To be fit for use as a motor vehicle fuel, the alcohol must pass through a dehydration system to remove any remaining water. The 200-proof pure anhydrous (without water) ethanol is made unfit for human consumption by adding a small amount (2–5 percent) of gasoline or another adulterant. Pure ethanol has a strong tendency to absorb water and care must be exercised in storage and transportation to keep water away to preserve its integrity as a motor vehicle fuel.

The two main by-products of ethanol production are carbon dioxide and the residue of fermentation. Making ethanol produces prodigious amounts of carbon dioxide that is cleaned of any residual alcohol, compressed, and sold to producers of carbonated beverages and meatpackers for flash freezing. This carbon dioxide, along with the carbon dioxide released as exhaust from automobiles burning ethanol, eventually enters the atmosphere. But the system is sustainable because an equivalent amount of carbon dioxide is absorbed from the atmosphere by the next crop of ethanol-producing plants. If grain is the raw material, rather than sugarcane or sugar beets, distillers grain, either wet or dried, with essentially all the protein and other nutrients of the original grain, plus the yeast residue, is highly valued as livestock feed. Ethanol production is one of those rare processes in which there is little or no waste, similar to baking bread on a cold winter's day, when waste heat escaping from the oven warms the kitchen.

Ethanol was the basis for a truly successful social and economic development program in Brazil. In the 1980s the country was plagued with both high unemployment in the northeastern region and costly oil imports. Brazil aggressively embarked on a program to utilize ethanol as a motor vehicle fuel. A conventional automobile can burn pure ethanol with minor adjustments to the carburetor. The only precaution that has to be taken is to ensure that the seals in the fuel system are made of a material that can withstand the greater corrosive properties of ethanol. No engine modifications are necessary for gasohol, a mixture of gasoline and ethanol. Ethanol has the advantage of having a high octane rating that enhances engine performance, but the disadvantage of a lower energy content than gasoline or diesel fuel. About 1.5 gallons or liters of ethanol have to be burned to obtain the same energy output as 1 gallon or liter of gasoline. As a result, automobile mileage drops to the extent that ethanol is added to gasoline despite improved engine performance. Ethanol produces less pollution in the form of nitrous oxides, but more in the form of acetaldehydes.

Brazil created an economic incentive to induce automobile owners to switch to ethanol in the form of a motor vehicle tax differential between gasoline and ethanol, which made ethanol more attractive to buy than gasoline, even after taking into account its lower mileage. With a price differential that favored ethanol, the price of gasoline was increased until the price for ethanol reached a level that made it profitable for individuals and corporations to invest in sugarcane plantations and sugar and ethanol production facilities. Consumers, not the government, provided the bulk of the funds for advancing biomass as a motor vehicle fuel.

Growing and harvesting sugarcane is labor-intensive, providing job opportunities for large numbers of unemployed workers. Businesspeople converted marginal and idle land to sugarcane plantations and built sugar and ethanol production facilities and operated them profitably. Thus, both workers and businesspeople benefited from this program. Furthermore, there was no displacement of agricultural output in terms of food production because the land dedicated to growing sugarcane was lying idle. Ethanol-fueled tractors and trucks planted, harvested, and shipped sugarcane to the sugar mill. The residues of sugarcane (bagasse) and the fermentation process, other than that portion fit for animal feed, were burned to produce steam and electricity for operating sugar mills and ethanol plants. Thus ethanol production in Brazil is entirely divorced from petroleum and is sustainable with little or no net gain of carbon dioxide emissions to the atmosphere. The only exception to this may be in the petroleum or natural gas component (if any) in fertilizers used to grow sugarcane. This biomass program has the environmental advantage of being essentially a closed carbon cycle for the production of motor vehicle fuels, but also contributed to the social and economic development of a poverty-stricken region of Brazil and reduced the trade deficit associated with importing oil.

In recent years Brazil has successfully discovered and developed oil fields and substantially reduced the need to import oil; in fact, Brazil emerged as a net oil exporter in 2005. This brought to the forefront the reality that has to be faced by ethanol advocates: Gasoline is simply less costly to produce than ethanol. With growing domestic supplies of petroleum, price supports for ethanol were reduced, making ethanol less popular as an automobile fuel. Currently, pure ethanol-burning automobiles are being phased out, but motor vehicle fuel in Brazil is gasohol, a mixture of 24 percent ethanol and 76 percent gasoline. With Brazil's goal of becoming the Saudi Arabia of ethanol, ethanol will always have a place as a motor vehicle fuel even though its percentage share in gasohol may change, depending on its availability.

The availability of ethanol is also a function of the world price of sugar, which, as with any commodity, fluctuates over time. As prices change, sugarcane processors in Brazil must decide whether to sell their output to international sugar buyers or to local ethanol producers. If the portion of ethanol in gasohol increases when the price of sugar is low, then ethanol becomes a price stabilizer for sugar growers. If other major sugar-growing nations of the world had gasohol programs similar to Brazil's, the interplay between ethanol production and the availability of sugar would help stabilize the world price of sugar. In the same way, ethanol production can act as a price stabilizer for corn and grain growers as well. Agricultural interests in the United States are strong proponents of increasing the ethanol content in gasoline for this reason, particularly growers of corn (maize) for animal feedstock. Automobile owners who buy gasohol (10 percent ethanol and 90 percent gasoline) in the corn-growing regions of the United States do so on the premise that their purchase of the higher-priced ethanol component of gasohol benefits the local economy.

Table 3.1

Top World Ethanol Producers

	Billion Liters	Percent of Total
Brazil	11.9	38
United States	7.6	24
China	3.1	10
India	1.8	6
Russia	1.2	4

In 2001, world production of ethanol was 31.4 billion liters, distributed as shown in Table 3.1.[2]

Most ethanol is produced from sugarcane, except in the United States, where ethanol is produced from corn (maize), and Russia, where ethanol is produced from sugar beets. Ethanol production is quite small compared to the overall consumption of gasoline in the United States. Compared to gasoline consumption of nearly 9 million barrels per day (bpd) in 2001, 7.6 billion liters of ethanol are equivalent to about 130,000 bpd. Thus, if ethanol were totally dedicated to being an automobile fuel, it would only account for 1.5 percent of the gasoline stream at that time. Other than the gasohol available in the corn-growing belt, the role of ethanol in gasoline is a consequence of federal regulations to improve air quality.

The U.S. Clean Air Act of 1990 contained two programs to deal with air pollution from motor vehicles. One required an oxygenate to be added so that the gasoline used in urban areas in winter would burn cleaner and, thereby, reduce carbon monoxide emissions. The other program was a year-round requirement for reformulated gasoline in urban areas suffering from the worst smog pollution. The type of oxygenate was not specified in the Act. Originally oil refiners could choose either ethanol or methyl tertiary butyl ether (MTBE) as oxygenates. MTBE is produced as a by-product of oil refinery operations or directly from methanol and isobutylene. It was first used in the 1980s to replace tetraethyl lead. Tetraethyl lead was an antiknock additive that improved engine performance, but had an undesirable environmental impact in terms of lead emissions in automobile exhaust. Ironically, decades earlier tetraethyl lead had replaced ethanol as an antiknock additive.

MTBE was favored over ethanol as an oxygenate for its lower cost, superior blending characteristics, and that it could be shipped and stored in existing oil product pipelines and storage tanks. However, in recent years various state and federal environmental protection agencies began to oppose MTBE because MTBE, a carcinogen, was finding its way into water supplies. The principal pathways were leaking underground gasoline storage tanks that were contaminating groundwater feeding water reservoirs and wells, and as unburned fuel in the exhaust of two-cycle outboard motors that powered boats on reservoir waters. The only alternative to MTBE, a petro-based oxygenate, was ethanol, a bio-based oxygenate.

Switching to ethanol, however, was not without its problems. The new market for ethanol called for more ethanol production capacity, but this problem never materialized because agricultural companies rose to the occasion. Ethanol is highly absorbent of water and, for that reason, cannot use existing oil product pipeline, storage, and distribution systems. This problem was solved by using railcars to ship ethanol from production facilities to refinery terminals where it was “splashed” into gasoline just before delivery to gas stations. As ethanol displaced high-octane MTBE, other gasoline components had to be adjusted to deal with ethanol’s higher vapor pressure. Moreover, ethanol was much more expensive to produce than MTBE and this incremental cost, along with ethanol’s more expensive handling requirements, had to be covered somehow. The Environmental Protection Agency (EPA) estimated that gasoline prices rose between four and eight cents per gallon to cover ethanol as an oxygenate. However, most of the incremental cost was made up by the ethanol component in gasoline not being subject to federal highway taxes. A 1 percent ethanol content in a gallon of gasoline reduces highway taxes by 10 percent; a 3 percent ethanol content reduces federal taxes by 30 percent, and gasohol that is 10 percent ethanol—sold in the corn-growing region of the United States—incurs no federal highway taxes. In addition, corporate tax benefits accrue to those building ethanol production facilities. Consequently, the public saw relatively little of the extra cost of ethanol production, estimated to be around fifty cents per gallon. Costs have to show up somewhere, and here most of the cost appeared as smaller balances in the highway trust funds and less corporate tax revenue.

Starting around 2004, several states (California, Connecticut, Kentucky, Missouri, and New York) initiated a ban on MTBE representing about half its demand of 250,000 bpd. In response, the U.S. Energy Policy Act of 2005 called for a phaseout of MTBE in favor of ethanol to be accomplished in 2006. This would have doubled demand for ethanol, increasing its content to about 3 percent of the gasoline stream as an oxygenate. However, bowing to state opposition to oxygenates, particularly from California, and a number of studies showing that oxygenates are unnecessary because gasoline can be formulated to burn just as cleanly without them, the EPA no longer requires their addition to gasoline.

The new role for biofuels incorporated in the Energy Policy Act of 2005 calls for an increase in plant-based fuels, such as ethanol and biodiesel, from 4 billion gallons in 2006 to 7.5 billion gallons per year by 2012. The Act no longer stipulates that ethanol must come from corn and leaves the door open for other sources of ethanol such as cellulose (wood chips, grasses) and for imports (the fifty-cent per gallon tariff on imported ethanol was quietly dropped in 2006 to increase ethanol availability as MTBE was phased out). Rather than stipulating required percentages of ethanol in EPA regulations, the government is leaving it up to the oil industry to decide how infusing almost 500,000 bpd (7.5 billion gallons per year) of biomass motor vehicle fuel into the gasoline stream is to be accomplished (one oil company in the Northeast was already selling gasohol of 10 percent ethanol in 2006). As a point of comparison, 2005 consumption of gasoline and diesel fuel was about 15 million bpd. Thus, in terms of current consumption, a half million bpd of biofuels would be about 3 percent of the motor vehicle fuel pool, approximately the same as under EPA regulations.[3]

There is an ongoing debate in the United States concerning whether ethanol made from corn (maize) is, or is not, consuming more fossil fuel for its production than the fossil fuel being replaced. Diesel fuel and gasoline are consumed by tractors in preparing the ground for growing corn, planting, and harvesting and by trucks transporting corn to an ethanol facility, where more energy is consumed to convert corn to ethanol. Rail transport is employed to transport ethanol to refinery terminals for mixing into gasoline. Moreover, oil and natural gas are used to make fertilizers and pesticides for growing corn. The contention of some studies is that the net energy value of ethanol made from corn is positive, meaning that the energy extracted from ethanol exceeds the energy content of oil consumed to produce it. Other studies indicate that the net energy value is negative, meaning that more oil is being consumed to produce ethanol from corn than is being replaced.[4]

The disparity in the extreme results of these studies stem from differences in assumptions on corn yields, the effectiveness of ethanol conversion, the oil content of fertilizer and pesticides, the value of co-products from ethanol production, and the energy required for ethanol's production. Both sides accuse each other of "bad science": the selection of data and an analytical methodology that supports a preconceived conclusion. One would think that the extent to which ethanol replaces fossil fuels, whether positive or negative, would be a settled matter, particularly when it is incorporated into the nation's energy policy. Apparently it is not. There is no similar dispute in Brazil because all vehicles associated with producing ethanol run on ethanol. The electricity consumed in converting sugar to ethanol is obtained by burning bagasse, the residue of sugarcane processing. With virtually no petroleum products consumed to produce ethanol—other than, perhaps, fertilizer—there is no question that ethanol replaces petroleum.

The United States is the world's largest grain exporter, and diverting large areas of agricultural land to ethanol production would adversely affect the availability of grain and other food exports for food-importing nations. Four billion gallons per year of ethanol requires 1,460 million bushels of corn equivalent to 13.5 percent of annual corn production. Hence 7.5 billion gallons per year of biofuels in 2012, assuming all of it being ethanol, would consume a quarter of the nation's current corn production. For this reason, there is a moral as well as an economic consideration associated with the large-scale substitution of ethanol for gasoline. This assertion is not true in Brazil, where idle land was converted to growing sugar for ethanol production. Other nations planning to increase the role of ethanol as a motor vehicle fuel are China and Thailand. These nations want to transform surplus grain into ethanol in order to reduce their oil imports. India plans to do the same, but, unlike China and Thailand, it is a food-importing nation. Thus, dedicating more agricultural land to growing sugarcane would increase food imports unless marginal or idle land is used. Moreover, India has a large molasses industry, which is not at all interested in seeing its raw material, sugar, diverted to making ethanol.

As the world population continues its inexorable climb, more food is required. Land dedicated to growing food need not rise proportionately to population growth if crop yields continue to be enhanced by better agricultural practices, more effective fertilizers, and higher-yielding plants created by genetic selection or engineering. Thus, some agricultural lands can be dedicated to biomass as fuel for motor vehicles or for generating electricity without affecting the availability of food. But there is a limit. The first priority of agricultural lands is to feed a growing world population, followed by other uses such as growing cotton for clothing and, in last place, biomass for electricity generation and motor vehicle fuels. Ideally, increased use of biomass as an energy source should have no negative impact on food production other than absorbing surplus production to support prices. While one may argue that establishing a floor price on commodities in effect increases food costs, the counterargument is that the financial liquidation of agriculture as a viable business from low commodity prices is not really in the best interests of the world population.

A biomass substitute for diesel fuel for motor vehicles is biodiesel, made from vegetable oils, a renewable resource. Biodiesel is produced from the chemical reaction of vegetable oils with methanol, in the presence of a catalyst, to form esters and glycerols. The esters become biodiesel and glycerols are used in pharmaceuticals and cosmetics. Biodiesel can be consumed by conventional diesel engines without any modifications and can be entirely substituted for oil-derived diesel or mixed with it in any proportion without adverse effects. Pollution in the form of nitrous oxides is about the same as oil-based diesel fuel, but there are no sulfur emissions. Biodiesel's higher oxygen content allows for more complete combustion, resulting in a substantial reduction in emissions of unburned hydrocarbons, carbon monoxide, and particulates (soot). Soot may become a major environmental issue in coming years. There is currently scientific speculation that the increased concentration of soot in glacier ice in the northern hemisphere leads to more rapid melting. Soot decreases the reflection of sunlight, thus increasing energy absorption, which then accelerates glacier melting in addition to rising temperatures. The primary sources of soot in the atmosphere are smokestack emissions from burning coal for electricity generation and as an industrial fuel and the exhaust of diesel-fueled motor vehicles.

Going full circle, it is ironic that the first diesel engines ran on vegetable oils. Rudolf Diesel's first diesel engine prototype in 1898 was fueled by peanut oil. He envisioned that a variety of vegetable oils would be suitable as fuels; but in the 1920s petroleum-based diesel fuel proved to be less costly, more efficient, and more readily available.[5] Soybean producers have taken the lead in supporting biodiesel (SoyDiesel) to counteract surplus production and falling prices. Soybeans are grown in great abundance, and the United States accounts for 36 percent, Brazil 27 percent, Argentina 18 percent, China 8 percent, India 4 percent, and others 7 percent of world production in 2003.[6] In recent years soybean output in Brazil and Argentina has been greatly increased with further significant growth slated for the United States.

Biodiesel has made minor inroads in the United States as a motor vehicle fuel. There are 200 truck and bus fleets running on biodiesel including those operated by the U.S. Postal Service, the U.S. military, metropolitan transit systems, agricultural concerns, and school districts. Some gas stations in Europe offer biodiesel as an alternative fuel. As with ethanol, the case against biodiesel is one of cost. About 7.3 pounds of soybean oil are required to produce one gallon of soy diesel. The feedstock cost, at about $0.20 per pound, is around $1.50 per gallon before processing, marketing, distribution, and overhead costs, not to mention an allowance for profit and highway taxes. An alternative to soybeans (20 percent oil) is rapeseed, which has double the oil content. However, the economics of utilizing plants with higher oil content to produce biodiesel must also consider their cost.

Just as it is ironic that Rudolf Diesel espoused vegetable oil as a fuel for diesel engines, Henry Ford espoused ethanol as a fuel when he started manufacturing Model Ts. Early Model Ts were built with an adjustable spark advance on the carburetor that allowed fueling by ethanol, gasoline, or kerosene. Ford saw ethanol as a means to support farm product prices during lean times when farmers were plagued with low prices from excess production, a nearly endemic condition during the early part of the twentieth century. As Ford saw it, ethanol would not only boost farm income but would also provide more employment opportunities for rural America.

Three factors came to bear against ethyl alcohol as a motor vehicle fuel: one was that it was drinkable. The government depended on alcohol taxes as an important source of revenue, but had no mechanism to separate alcohol produced for personal consumption from that used to fuel motor vehicles. The alcohol tax made ethanol too expensive as a motor vehicle fuel, but did not prevent ethanol, with its antiknock properties, from being used as a gasoline additive to enhance engine performance. The second factor that weighed against the use of ethanol was the development of lower-cost tetraethyl lead as an antiknock additive. The third was the success of the growing temperance movement in getting state legislatures to prohibit the production and sale of alcohol for any purpose, including its use as an automotive fuel.

The historical facts are that Standard Oil was involved with a blended fuel that contained alcohol sold on a limited basis as Alcogas in the 1920s and that the oil industry lobbied against the blending of alcohol and gasoline in the 1930s.[7] But one can imagine that John D. Rockefeller, a Baptist teetotaler, would support the prohibition movement publicly for moral reasons but privately, for business reasons, as a way to rid himself of a potentially competitive fuel, just as he had previously rid himself of competitive refinery operators.

For ethanol and biodiesel to be acceptable as motor vehicle fuels, some sort of mechanism has to be set up to deal with biofuels being more expensive than gasoline and diesel fuel. One way is to recognize oil as a wasting resource and incorporate a tax that reduces reliance on something that will someday disappear. Ethanol and biodiesel would not pay this tax because they are renewable resources. The wasting-resource tax would be at a level that would make ethanol and biodiesel commercially viable, mixed in some specified proportion with gasoline and diesel fuel. This is similar to what was done in Brazil to make ethanol economically attractive to consumers. In this way, ethanol and biodiesel could enter the motor vehicle fuel stream free of direct government subsidies because the extra cost would be passed on to consumers. The Energy Policy Act of 2005, in mandating a minimum volume of biofuels in gasoline and diesel fuel, could fund the extra costs as the government does now by exempting biofuels from paying highway trust-fund taxes plus other tax-depreciation gimmickry associated with building ethanol plants. Alternatively the government could phase out these tax breaks and make the oil companies pay the full cost of biofuels and pass their higher cost straight through to consumers in the form of higher gasoline prices. With high priced crude oil, the difference may not be that noticeable.

Biomass for Electricity Generation

The chief argument against biodiesel and ethanol is its impact on the cost and supply of food for human consumption. Large-scale use of agricultural land for motor vehicle fuels is not possible without adversely affecting food supplies. This places an upper limit on the use of grain and corn as motor vehicle fuels, although lands currently in nonagricultural use in the tropics can be converted to growing sugar as was done in Brazil. Land dedicated for growing biomass for electricity generation should preferably be unused or marginal land unfit for agricultural use. Such land should be suitable for fast-growing trees (poplars, willows) and grasses (switchgrass). An example of innovation for biofuels is a newly developed grass (*Miscanthus* x *giganteus*), a hybrid (indicated by the "x" in the species name) of an Asian variety related to sugarcane. The plant sprouts each year, requires little water and fertilizer, thrives in untilled fields and cool weather, and grows rapidly to thirteen feet tall. After its leaves drop in the fall, a tall bamboolike stem can be harvested and burned to generate electricity. It is estimated that if marginal land in Illinois, accounting for 10 percent of the state's area, were dedicated to growing this grass, it could provide half of the state's electricity needs without affecting its food output.[8]

Without a technological breakthrough to improve the transformation of cellulose to ethanol, trees, grasses, waste from growing grain (stalks, leaves, husks), and other cellulose plant matter are better consumed as a biofuel for electricity generation. The benefits of biofuel for electricity generation are that biomass:

- is plentiful, with large regions of the earth covered by forests and jungles;
- can be increased by planting marginal lands with fast-growing trees and grasses;
- stabilizes the soil and reduces erosion;
- is a renewable and recyclable energy source that does not add to carbon dioxide emissions;
- stores solar energy until needed, then is converted to electricity, whereas solar panels and wind turbines generate electricity, whether needed or not, and then only when the sun is shining and the wind is blowing;
- does not create an ash waste-disposal problem since the ash can be spread in the forests or fields to recycle nutrients and not be directed to landfills as is ash from burning coal;
- creates jobs in rural areas.

Some environmentalists are critical of biomass plantations because they deplete nutrients from the soil, promote aesthetic degradation, and increase the loss of biological diversity. Growing biomass as a fuel depletes the soil of nutrients, but spreading ash from combustion replenishes the soil with what was removed with the exception of nitrogen. Interspersing nitrogen-fixing plants among the biomass plants can replenish nitrogen rather than using nitrogen-based fertilizers made from fossil fuels. On the plus side, biomass plantations can reduce soil erosion and be managed in a way that minimizes their impact on the landscape and on biological life. In fact, there is no reason why biomass plantations cannot make a barren landscape more attractive and encourage biological life. Another argument against biomass as a fuel for making electricity is smoke emissions during combustion. This can be sidestepped by gasifying biomass to feed gas turbines. For this to occur, gasification technology has to be perfected and made cost-effective before it can be adopted for large-scale electricity generation.

Biomass for electricity generation can be forest residues including imperfect commercial trees, noncommercial trees thinned from crowded, unhealthy, or fire-prone forests, dead wood, and branches and other debris from logging operations. While "free," there is the cost of collecting and shipping a thinly dispersed energy source from remote locations to an electricity-generating plant. More promising from a logistical point of view is collecting the bark, edging, and sawdust residues at lumber mills. Lumber mills are generally located closer to population centers and collect logs from a wide area. They concentrate wood residues at a few sites making transport to an electricity-generating plant easier and less costly. Furniture manufacturing facilities are also concentrated sources of wood waste. However, some of this waste from lumber mills is already burned to supply power to the lumber mill, and in some cases, as in northern Europe, electricity to the general population. Waste from paper pulp manufacturing is also being burned to power paper pulp plants.

A second source of biomass is the residue of harvesting agricultural crops. These include wheat straw, corn stover (leaves, stalks, and cobs), orchard trimmings, rice straw and husks, and bagasse (sugarcane residue). Sugarcane is harvested and shipped to a sugar-processing plant that concentrates bagasse at a single location, and it is then often burned to supply power to the sugar-processing plant. Other agricultural wastes are generally left in the field and decay to become part of the soil. The high cost for collecting and shipping would make agricultural wastes as commercially unattractive as forest residue. Furthermore, agricultural wastes are seasonal, although they could be combined with wood residues to feed a biomass electricity-generating plant. Total removal of agricultural residues, however, would also have adverse consequences on soil nutrition.

A third source is so-called energy crops grown specifically for fuel. These crops are preferably fast-growing, drought- and pest-resistant, and readily harvestable by mechanical means. Depending on growing conditions, hybrid poplars and willows can be harvested every six to ten years. Trees can be cut and shipped to the utility plant as wood chips, shipped whole and converted to chips at the plant prior to burning, or be burned whole in specially designed boilers. Switchgrass does not require replanting for up to ten years. It is cut, baled, and shipped to a utility plant and ground up prior to burning. However, none of these sources is strictly sustainable in that the fuel burned in tractors and trucks in growing, harvesting, and shipping these biomass fuels adds to carbon emissions. Overall carbon emissions would still be lower than burning fossil fuels because of the carbon dioxide absorbed by replacement plant growth.

Another option is to use biomass as a co-fuel in existing coal-burning plants instead of burning it in specialized electricity-generating plants. These facilities would have dedicated storage and material-handling systems for biomass in addition to their existing facilities for handling coal. Biomass would be mixed with coal in proportion to the plant's load factor. The higher the load factor, that is, the closer the plant is operating to its rated capacity, the greater will be the portion of coal because of its higher energy content. Technical problems begin to emerge when too much biomass is mixed with coal in a conventional coal-burning plant. For this reason, the typical mix for plants combining biomass and coal is generally less than 10 percent biomass. These technical problems have to be dealt with before higher portions of biomass can be mixed with coal. There are a few specially designed facilities that can burn either 100 percent coal or 100 percent biomass. Again, at high load factors, coal is favored for its higher energy content.

Vegetable oils, used motor vehicle lubricating oils, and paper trash have also been suggested as fuels for electricity-generating plants. Used vegetable oils are not available in the quantities necessary to run an electricity-generating plant and the cost and effort of collecting used vegetable oils from a million and one hamburger franchises would be an inhibiting cost factor. However, there are some imaginative owners of diesel trucks who have discovered that they can stop at friendly fast-food restaurants for a bite of food and then do a favor for the proprietors by disposing of their used vegetable oil free of charge. The concoction of vegetable oil and diesel fuel supposedly burns just as efficiently as pure diesel fuel plus odorizes the exhaust with French fries, fried chicken or fish. Maybe an entrepreneur will some day take a more organized approach to collecting used vegetable oils for recycling as diesel fuel, a higher order of use than burning it as a utility fuel.

Because many states prohibit the dumping of used motor vehicle lubricating oils into the environment, used motor vehicle oils are collected for a fee and recycled as re-refined lubricating oil, again a higher order of usage than burning it for its heat content. The same is true for paper trash that has been processed and sold as recycled paper and cardboard products.

Biomass energy accounts for less than 1 percent of U.S. electricity generation and 2 percent in Europe, where much of the available biomass is waste from lumbering operations in Finland and Germany. Most existing biomass electricity-generating facilities are small, dedicated to meeting the needs of a local industry or community. Their most important contribution is that they demonstrate the potential for biomass to generate electricity and serve as platforms for improving technology. One such plant in Vermont burns waste wood from nearby logging operations, lumber mill waste, and discarded wood pallets. In addition, there is a low-pressure wood gasifier capable of converting 200 tons per day of wood chips to fuel gas, which is fed directly into the boiler that burns the waste wood. Hot water waste from generating electricity is pipelined to nearby buildings for internal use. Net carbon dioxide emissions, taking into consideration the sustainable growth of biomass, have been cut by over 90 percent compared to burning fossil fuels. Other research activities center on developing more effective technologies to gasify municipal and animal wastes, wood and agricultural wastes, and black liquor waste from papermaking to fuel electricity-generating plants. Gasification eliminates smoke emissions and is an efficient means of delivering biomass to distant electricity-generating plants.[9]

For special circumstances, biomass can economically produce electricity, but the economic viability of large-scale use of biomass to generate electricity remains questionable. Growing, harvesting or collecting, and shipping biomass are costly compared to the alternative of mining and shipping coal, which can be looked upon as concentrated biomass. Biomass electricity-generating facilities built in the United States in response to the oil crisis that were economically sound when oil was $35 per barrel in the early 1980s became financial albatrosses when oil prices fell after 1985. At this point, biomass for large-scale electricity-generating facilities does not appear to be in the cards because biomass is generally more expensive than fossil fuels. This could change if current research and development efforts result in a technological breakthrough that radically changes biomass energy economics or the price of fossil fuels rise to levels that make biomass for electricity generation financially attractive. All the same, for biomass to make any difference in reducing reliance on fossil fuels for electricity generation, its contribution has to double, and then double again, just to show up on the radar screen. At this point, there appears to be relatively subdued interest for significantly expanding the capacity of biomass-generated electricity.

If we take the position that carbon emissions contribute to climatic change, which of itself represents a cost, then one can justify a tax on carbon emissions. A carbon emissions tax placed on burning fossil fuel to generate electricity would make sustainable biomass energy consumption economically attractive because of its 90 percent or so reduction in carbon emissions. If something on its own merits is not economical, then it can be made so by discriminatory taxation. Nevertheless, it would be preferable if technology could make something that is environmentally desirable also economically attractive.

Even without technological breakthroughs, biomass energy is ideal for electricity generation in isolated areas in the temperate and tropical regions, such as the island nations of Southeast Asia and in South America and Africa that are not connected to electric power grids. Micro-electricity-generating plants could serve the local needs of such communities. Unfortunately, areas already facing deforestation would be worse off if biomass were to become a source of energy for generating electricity unless it were based on sustainable tree plantations. Micro-electricity-generating plants that depended on sustainable sources of biomass fuel would provide basic services, such as lighting, to a village and encourage cottage enterprises to provide basic amenities. This, of course, presumes that the people consider this a desirable outcome. Some indigenous people would rather continue living the way they have for countless generations than adopt the ways of modern society. And who is to say that they are wrong?

Biogas

In the presence of dissolved oxygen, aerobic microorganisms decompose biodegradable organic matter releasing carbon dioxide, water, and heat. However, in the absence of dissolved oxygen, an anaerobic digestion process takes place that releases carbon dioxide and methane, which can be collected as a fuel. Aerobic digestion normally occurs in compost heaps. Anaerobic digestion occurs wherever concentrations of organic matter accumulate in the absence of dissolved oxygen such as the bottom sediments of lakes and ponds, swamps, peat bogs, and in deep layers of landfill sites.

As with making ethanol, a number of steps involving different microorganisms are necessary to produce biogas. It starts with a hydrolytic process that breaks down complex organic wastes into simpler components. Then fermentation transforms these organic components into short chains of fatty acids plus carbon dioxide and hydrogen.

Next the syntrophic process converts the short chains of fatty acids to acetic acid, thereby releasing heat and more carbon dioxide and hydrogen. One type of bacterium converts the acetic acid to methane and carbon dioxide, while another combines hydrogen with carbon dioxide to produce more methane. Still another bacterium reduces any sulfur compounds to hydrogen sulfide, which in turn reacts with any heavy metals that may be present to form insoluble salts. The simple process of decay turns out to be biologically complex.

The resulting biogas from anaerobic decay is approximately two-thirds methane and one-third carbon dioxide and can be made from sewage, animal manure, and other organic matter such as wood chips, household refuse, and industrial organic waste. Biogas production is very slow at ambient temperatures, but it can be sped up by raising the temperature of the organic matter to a specified range. The energy for heating is generated from organic decomposition, and, if necessary, a portion of the biogas production can be siphoned off and burned to further increase the temperature.

Gasification of raw sewage involves an initial screening to remove inorganic objects before being pumped into sedimentation tanks, where the solid organic matter settles as sludge. The sludge is pumped into large anaerobic digester tanks where decomposition takes place at a heightened temperature that hastens the process. In about two months, half the sludge has been converted into gas.[10] The remaining sludge can be dried and used as a fertilizer, burned as a fuel, or dumped into a landfill. The public does not accept sludge from human waste as a desirable fertilizer for the backyard tomato patch and relatively little is burned as a fuel. Most sludge from human sewage is buried in a landfill or dumped at sea.

Biogas is not a high-quality fuel and is usually burned locally in a turbine or fed as a gaseous fuel into a specially adapted internal combustion engine to drive a generator that produces electricity for local consumption. Biogas generating systems are being set up where animal and chicken manure present a disposal problem. In the past, chicken, beef, and pig farms had sufficient land to grow crops for farm animals that were fertilized with their waste. Now runoff from these fields is considered a contaminant of local streams and rivers. Even more importantly, modern chicken, beef, and pig farms are more like factories and buy most, if not all, of their feed. This modern industrial approach to agriculture minimizes the amount of manure that can be spread on fields. Biogas generators are not only a source of energy for running the farm, but also reduce the volume of organic waste by half. What is left can be spread on fields, if permissible, dried and used as fertilizer, burned as a fuel, or disposed of in a landfill.

The carbon dioxide emissions from generating and burning biogas are considered a closed carbon cycle. Human sewage comes from eating plants either as grain, vegetables or fruits, or meat from plant-eating animals. The source of animal waste is plant food fed to the animals. Biogas is not a completely closed carbon cycle because growing and harvesting crops, processing and distributing food, and manufacturing fertilizer require a great deal of energy in the form of gasoline, diesel fuel, and electricity, much of which is generated by burning fossil fuels. Nevertheless, biogas reduces carbon dioxide emissions by substituting for fossil fuels. Europe has taken the lead in producing biogas from organic matter, but biogas contributes less than 0.5 percent to the generation of electricity.

As organic matter decays in a landfill, biogas normally finds its way to the surface and disperses to the atmosphere. If the landfill is covered with a layer of clay to prevent escape to the atmosphere, biogas can be extracted by sinking tubes into the landfill. The biogas can fuel an internal combustion engine or turbine to generate electricity locally. The problem here is that a landfill covered by a layer of clay is probably full and the investment must be justified by the amount of biogas generated from a finite and nonreplenishable source.

In addition to disposing of sludge, disposing of garbage is a major problem for the principal population centers of the world. Ocean dumping and landfills are not desirable ways to dispose of garbage. Ocean dumping off New York City has created a marine dead zone and fish that live nearby it have a high incidence of cancer and/or suffer from various grotesque mutations. Landfills near metropolitan areas are usually undesirable although they have a role to play in urban development. While LaGuardia Airport in New York City is built on top of a landfill, a residential development built on top of a landfill might be a hard sell. Marshes buried under enormous mounds of garbage capped with a layer of soil are becoming less available near populated areas and are negatively perceived by the public. Now landfill sites may be hundreds of miles away from metropolitan areas.

There is an alternative to ocean dumping and transforming picturesque countryside into landfills. Modern garbage disposal starts with people separating recyclables such as paper, cardboard, and items made from plastic, glass, aluminum, tin, and other metals. Recycling reduces the energy intensity of a society because glass and aluminum require 90 percent less energy when made from recycled glass and aluminum than from sand and bauxite. Paper and cardboard made from paper trash and steel made from scrap also require a lot less energy than making paper from trees and steel from iron ore and coal. After removing recyclable waste, what remains can be collected and burned at an electricity-generating plant to produce steam, which can be superheated by burning natural gas to enhance turbine efficiency. The garbage is ultimately reduced to ash, a small fraction of its former volume, which can be buried in a landfill.

While this may be considered an attractive means of disposing of garbage, it is also costly to build an electricity-generating plant that disposes of garbage. The fuel is not only free, but a charge for garbage collection becomes another source of revenue in addition to generating electricity. Even so the revenue from selling electricity and collecting garbage may not be sufficient to justify the investment. Burning garbage does not generate nearly the same amount of electricity as burning coal or natural gas. Communities may still find it cheaper to dump the garbage in the ocean or ship it to a distant landfill rather than pay for it to be burned under controlled conditions for generating electricity.

The problem is that there is no cost associated with environmental damage wrought by ocean dumping or landfills. As long as ocean dumping and landfills are the cheapest alternatives, waste disposal will continue to degrade the environment. What has to be done is to place a tax for environment degradation on ocean dumping and landfills, a solution that would internalize the external cost of degrading the environment. Once this externality has been internalized, then the economics of generating electricity from garbage will have a more favorable hue.

Some time ago a garbage-burning electricity-generating plant, to be used by the surrounding communities, was built in a U.S. metropolitan area. The charge for disposing of garbage reflected the capital cost of building the plant and its operating costs, net of the revenue of selling the electricity, divided by the tons of garbage processed. While the ton charge was reasonable at full capacity, there was no obligation on the part of the surrounding communities to use the plant. One community discovered that there was a nickel's worth of savings in trucking the garbage several hundred miles to a landfill site rather than using the facility and opted out. The loss of volume from this community caused the unit-processing fee to increase to cover the fixed components of capital and operating costs. This hike in fees induced another community to opt out, which in turn raised the processing fee again, inducing still another community to opt out until, at the end of the day, there were virtually no communities using the facility. What should have been an environmentally desirable and economical way of disposing of garbage turned out to be a financial fiasco.

Having no cost associated with dumping garbage into the ocean or in transforming the countryside into landfills, other than shipping and dumping fees, it is economically attractive for municipalities to continue doing business as usual. An environmental degradation tax would internalize the cost of dumping garbage in the ocean or in landfills and make these options more costly. If this were done, then sharp-eyed accountants determining whether to pay the shipping and dumping or landfill fees, with an associated environmental degradation tax, versus using a garbage-burning electricity-generation plant without an environmental degradation tax might have a change of heart. As long as accountants are weighing the relative merits of alternatives strictly in terms of dollars and cents, then internalizing an externality (putting a cost on environment degradation) is a way to sway these individuals to select an environmentally sound way to dispose of garbage. Persuasive arguments and appeals to their better nature mean little when there is a cheaper, though less desirable, alternative. It is unfortunate that accountants make such decisions; but in some ways this makes it relatively easy to shape their decisions. All that has to be done to make a desirable outcome financially attractive is to ensure that it is the low-cost alternative, which an environmental degradation tax would accomplish. Moreover, the proceeds of the tax can be dedicated to funding the building of environmentally sound garbage-disposal plants whose output of electricity would reduce the need to burn fossil fuels. Such a simple solution to a complex problem seems to escape human attention. Unfortunately, the NIMBY (not in my backyard) syndrome has made it difficult to site plants that produce electric power by burning garbage. Another drawback has been the discovery that these plants emit mercury and other noxious metal fumes from burning discarded batteries that have found their way into household trash.

Tomorrow's Fuel

The projection of biomass consumption is summed up in Table 3.2.[11] As indicated, biomass will make only a marginally greater contribution in satisfying future energy needs in absolute terms.

Table 3.2

Contribution of Biomass

Biomass Consumption in Million Tons of Oil Equivalent

	1995	2020	Percent Annual Growth
China	206	224	0.3
East Asia	106	118	0.4
South Asia	235	276	0.6
Latin America	73	81	0.4
Africa	205	371	2.4
Other nonOECD	24	26	0.3
OECD	81	96	0.7

Biomass satisfied about one-third of the energy needs of developing nations and 3 percent of developed (OECD) nations in 1995, and these portions will decline to less than one-quarter of the energy needs of developing nations and 2 percent for developed nations. The relative decline in the contribution of biomass to satisfying energy needs is that, while biomass increases in absolute terms, its rate of growth lags behind that of commercial sources of energy.

Although the projected contribution of biomass is less than robust, government and private funding for research and development aim to make biomass energy technically feasible and economically justifiable. Environmental groups provide a great deal of moral support because substituting biomass from sustainable sources reduces the need for burning fossil fuels along with their carbon dioxide emissions. Yet, all this hoopla has produced relatively little in results. For instance, while hundreds of individual biogas projects in Europe supply energy to industry, agricultural enterprises, and towns, their aggregate contribution to satisfying overall energy consumption is, as stated previously, less than 0.5 percent. In the United States, the contribution is even less.

Ethanol cannot be looked upon as a substitute for gasoline because massive ethanol production in food-growing regions would have dire effects on food production and prices. Yet ethanol can play a role in those nations that imitate the Brazilian model. Having said that, one must still ask why the Brazilian model, for all its benefits, has not been adopted by Cuba and the Philippines. Both are sugar-producing, oil-importing nations with large numbers of unemployed workers and plenty of nonagricultural land suitable for growing sugarcane. Moreover, these nations would benefit financially if sugar were diverted to the production of ethanol as a price support mechanism.

With regard to biomass energy for producing electricity, yes, there are examples of relatively small dual-fired utility plants that consume biomass along with coal that have successfully demonstrated the feasibility of the concept. While Europe plans to double its biomass-generated electricity capacity, the United States, despite having a number of ongoing biomass projects, has taken a less aggressive stance. Even if power plants that burned biomass to produce electricity were built on a large scale, the relatively high cost of using biomass to generate electricity remains a problem. Unless there is a massive hike in the price of fossil fuels in the future, biomass plants cannot compete economically without support from government subsidies or a carbon tax on fossil fuels.

Brazil has been a model nation in having a national energy policy aimed at reducing the consumption of fossil fuels, and leads the world in utilizing biomass as a motor vehicle fuel. Brazil nearly eliminated fossil-fueled electricity generation by developing its hydropower resources. Its original goal was to have hydropower supply all its electricity needs until a drought caused severe power outages throughout the nation. Faced with the need to find alternative ways to fuel electricity plants, Brazil turned to natural gas. The irony is that Brazil has little in natural gas reserves unless new discoveries are made. Pipelines have been built, or are under consideration, to tap natural gas fields in Bolivia, Argentina, and in remote areas of the Amazon, and to import natural gas in a liquefied state. Rather than pursue electricity-generation plants fueled by natural gas, why not plant sustainable tree farms to supply electricity-generating plants? Biomass seems to be a neglected fuel for large-scale generation of electricity. Perhaps Bolivia's nationalization of its natural gas reserves in 2006, where Petrobras, the national oil company of Brazil, stands to lose the most money, will induce Brazil to view biomass in terms of energy security much as we view coal.

The United Nations Development Program (UNDP) is responsible for the implementation of the UN conventions on biological diversity and climate change. The Global Environment Facility (GEF), the financing arm of the UNDP, is funding, along with private corporate support, the development of a biomass integrated gasification/gas turbine (BIG/GT) in Brazil fueled by wood chips from tree plantations. Brazil already leads the world in having huge pine and eucalyptus tree plantations, but these are dedicated to making paper pulp, not generating electricity. BIG/GT transforms wood chips into a clean-burning gas and steam, both of which could be used to generate electricity. At this point, BIG/GT can produce electricity at about the same cost of building a hydropower plant, but would create many times the number of jobs in planting and harvesting trees. If proven commercially and technologically feasible, BIG/GT installations can be sized to serve local communities and built wherever there is land fit for growing trees on a sustainable basis to avoid deforestation.

A centralized electricity-generating system requires high-density population centers to financially support the construction of large conventional plants with their long-distance transmission lines that serve the surrounding area. Such systems cannot economically serve remote areas of the Philippines, Indonesia, Malaysia, and Africa, but a distributive electricity-generating system, such as BIG/GT, can be fueled by sustainable tree farms to neutralize carbon dioxide emissions. Yet, micro-electricity biomass-fueled plants capable of serving the needs of about 2 billion people living outside the main power grids have made little progress. Even those villages with biomass-fueled plants for local industrial activities such as lumber mills and food-processing plants are, for the most part, without electricity for light and comfort. The absence of electricity prevents the development of cottage industries that could provide basic amenities. People without electricity are hopelessly locked in poverty because, without electricity, there can be no factories and without factories, there can be no jobs.

One would think that building a micro-electricity-generating plant fueled by freely available biomass in a remote village would be a high-priority item for governments in pursuit of social and economic development, but this is apparently not the case. All one sees is a fairly uniform lack of progress. However, if the BIG/GT technology proves commercially and technologically feasible, distributive BIG/GT installations serving local needs (in conjunction with solar and wind) could contribute to the economic development of large areas of the world that cannot be served by conventional electricity-generating systems.

In the future, it is possible that biomass will be used to produce biopetroleum by utilizing a thermal conversion process that mimics the geological and geothermal processes of nature to produce gas and oils. The difference is that the process does in minutes what it takes nature thousands or millions of years to accomplish. Organic feedstock such as sewage sludge, tires, plastics, and animal and agricultural refuse are pulped and mixed with water, heated under pressure, flashed to a lower pressure to separate its components, reheated to drive off light hydrocarbons and water, followed by separation of the end products. This system breaks down long chains of organic polymers and reforms them into a new combination of solid carbon, a liquid oil similar to diesel fuel, and gases. The end products are determined by the raw materials being processed. A pilot plant being built to consume 210 tons per day of residue from a turkey-processing plant is expected to produce 70 tons per day of diesel-type oil plus 7 tons per day each of carbon and a low-energy flue gas that can be burned for fuel. This pilot project, if successful, will change the perception of organic waste and sludge. Towns and cities can build "refineries" at their sewage and organic waste-collection facilities and sell motor vehicle fuels to the public in competition with the oil companies![12]

Notes

1. The statistics for commercial sources of energy are from British Petroleum, *BP Energy Statistics* (London: Author, 2005); the source for the statistics related to biomass consumption in the aggregate and per country are from the International Energy Agency (IEA), *World Energy Outlook* (Paris: Author, 2004). All figures are for the year 2002.

2. The statistics provided in Table 3.1 are from Christoph Berg, *World Ethanol Production, 2001,* available online at www.distill.com/world_ethanol_production.html.

3. These statistics are available online from the U.S. Department of Energy at www.eia.doe.gov and the U.S. Environmental Protection Agency at www.epa.gov.

4. Hosein Shapouri, et al., *The Energy Balance of Corn Ethanol: An Update,* Agricultural Economic Report No. 814 (Washington, DC: U.S. Department of Agriculture, Office of the Chief Economist, Office of Energy Policy and New Uses, 2002).

5. A History of Biofuels at www.ybiofuels.org/bio_ fuels/history_diesel.html

6. U.S. Department of Agriculture Annual World Production Summary at www.fas.usda.gov/wap/circular/2003/03-02/Grains.pdf.

7. William Kovarik, "Henry Ford, Charles Kettering, and the Fuel of the Future," Automotive History Review 32 (Spring 1998), pp. 7–27. Reproduced on the Web at www.radford.edu/~wkovarik/papers/fuel.html.

8. For further information about Miscanthus x giganteus, contact Professor Stephen Long, University of Illinois at Urbana-Champaign, 1401 West Green Street, Urbana, Illinois 61801.

9. J. Goldemberg and T.B. Johansson, eds., Energy as an Instrument for Socio-Economic Development (New York: United Nations Development Programme, 1995).

10. The information on converting raw sewage into usable fuel comes from the U.S. Environmental Protection Agency's Web site at www.epa.gov/epaoswer/ non-hw/compost/biosolid.pdf) and Zia Haq, "Biomass for Electricity Generation," available online at the U.S. Department of Energy's Web site at www.eia.doe.gov/oiaf/analysispaper/biomass/.

11. The summary of biomass consumption presented in Table 3.2 is based on International Energy Agency (IEA), *World Energy Outlook* (Paris: Author, 2004). East Asia includes Thailand, Indonesia, the Philippines, and Malaysia; South Asia consists primarily of India, Pakistan, and Bangladesh and projections are based on Sylvie Lambert D'Apote, *Biomass Energy Analysis and Projections* (Paris: International Energy Agency, 1998).

12. Changing World Technologies, Inc., see the Web site at www.changingworldtech.com.

4

Coal

Coal suffers from an incredibly bad image. It has few advocates other than the hundreds of thousands whose livelihoods depend on mining and burning coal by the trainload for generating electricity. No one strikes it rich in coal; that metaphor is reserved for oil. For some, coal brings back an image of coal miners who go in hock to buy a set of tools when they are young and quit decades later with black lung, still in hock to the company store. That might be one of the better images. Another would be the mangled bodies of miners caught in mine mishaps or those trapped by cave-ins awaiting their fate in pitch blackness. Still another would be youngsters harnessed to sleds dragging coal up narrow underground passageways on their hands and knees like pack animals or straddling precariously above fast-moving conveyor belts of coal picking out the rocks. For still others the image of coal is as a pollutant of the first order that has to be eliminated under any or all circumstances. Nothing short of unconditional surrender can appease these environmental militants.

Yet, at the same time, this biomass fuel from ages past is irreplaceable and absolutely essential to ensure that the lights go on when we flick the switch. World coal consumption, essentially stagnant during the 1990s, surged by 30 percent between 2000 and 2004. Not only is the world consuming more coal, but its share of the energy pie increased from 23.4 percent in 2000, its historical low point, to 27.8 percent in 2005, reversing a fifteen-year trend of losing relative standing. Coal is becoming more important as a primary source of energy, not less as many people desire. Wishful thinking will not make coal go away, but there are ways to alleviate the worst of its adverse environmental consequences. This chapter reviews the history of coal, its importance in today's economy, and what is being done to overcome its principal drawbacks.

The First Energy Crisis

The first energy crisis was associated with living biomass (wood). It was an on-and-off-again crisis that extended over centuries. One of several reasons why the natural growth of forests could not keep up with the ax was glassmaking. Glassmaking has a long history, going back to about 3000–3500 BCE as a glaze on ceramic objects and nontransparent glass beads. The first true glass vases were made about 1500 BCE in Egypt and Mesopotamia, where the art flourished and spread along the eastern Mediterranean. Glassmaking was a slow, costly process and glass objects were considered as valuable as jewels; Manhattan Island was purchased from the Indians for $24 worth of glass beads and Cortez was able to exchange glass trinkets for gold!

The blowpipe was invented in Syria around 30 BCE. Using a long thin metal tube to blow hollow glass shapes within a mold greatly increased the variety of glass items and considerably lowered their cost. This technique, still in practice today, spread throughout the Roman Empire and made glass available to the common people. Transparent glass was first made around 100 CE in Alexandria, which became a center of glassmaking expertise, along with the German Rhineland city of Köln (or Cologne). During the first golden age of glass, glassmaking became quite sophisticated. For example, glassmakers learned to layer transparent glass of different colors and then cut designs in high relief. All these achievements in glassmaking were lost in the 400s with the fall of the Western Roman Empire.

The so-called Dark Ages take on new meaning with the disappearance of glassmaking, but vestiges of glassmaking remained in Germany, where craftsmen invented the technique for making glass panes around 1000 CE. These were pieced together and joined by lead strips to create transparent or stained glass windows for palaces and churches. The second golden age of glass started in the 1200s when the Crusaders reimported glassmaking technology from the eastern Mediterranean. Centered in the Venetian island of Murano, glassblowers created Cristallo glass, which was nearly colorless, transparent, and blown to extreme thinness in nearly any shape. In the 1400s and 1500s, glassmaking spread to Germany and Bohemia (Czech Republic) and then to England, with each country producing variations in type and design of glass objects. The ubiquitous glass mirror was invented comparatively late, in 1688 in France.[1]

Glass is made from melting a mixture of mostly sand (silicon dioxide) plus limestone (calcium carbonate) and soda ash (sodium carbonate) in a furnace, along with glass waste, at a temperature of around 2,600°F–2,900°F. Considering what has to be heated to such high temperatures, clearly glassmaking was an energy-intensive process that consumed a lot of wood. As forests were cleared, glassmaking furnaces were moved to keep close to the source of energy rather than moving the source of energy to the furnaces. The first energy crisis began when English manors for the rich and famous were built with wide expanses of glass panes that opened up their interiors to sunlight. Not only did this put a strain on wood resources for making the glass, but also for heating since interior heat passes more easily through a glass pane than a stone wall covered by a heavy wool tapestry.

The growing popularity of glass was not the only villain responsible for deforestation. Part of the blame lies with the increased demand for charcoal used in smelting iron, lead, tin, and copper. Consumption of these metals increased from a growing population, greater economic activity, and an improving standard of living as humanity emerged from the deep sleep of the Dark Ages. Deforestation started around London in 1200 and spread throughout the kingdom. By the 1500s metal ores had to be shipped to Ireland, Scotland, and Wales for smelting, deforesting these regions in turn. One of the economic drivers for the founding of the Jamestown colony in Virginia in 1607 was to take advantage of the New World's ready supply of trees to make glass for export to England. The rapidly escalating price of firewood, the economic consequence of deforestation, provided the necessary incentive to search for an alternative source of energy. The final answer to the energy crisis was not deforesting the living biomass of the New World, but burning the long-dead biomass of the Old World.

The Origin and History of Coal

Switching from wood to coal had an environmental consequence. Living plants absorb carbon dioxide from the air, which is released when they decay. For sustainable biomass energy, carbon dioxide is simply recycled between living and dead plant matter and its content in the atmosphere remains unchanged. One way to decrease the amount of carbon dioxide in the atmosphere is to increase the biomass such as planting trees on treeless land (afforestation), but this is neutralized when living and dead plant matter are once again in balance. The other way is to interrupt the decay process. And this is what happened eons ago when huge quantities of dead plants were quickly submerged in oxygen-starved waters. This delayed onslaught of decay interrupted the natural carbon dioxide cycle.

The partially decayed plants submerged in swamps first became peat. Peat has a high moisture content that is squeezed out if buried by silt of sand, clay, and other minerals from flowing water. Continued burying, either by the land submerging or the ocean rising, added sufficient weight to transform the original deposits of sand and clay to sedimentary rocks and the peat to coal. Three to seven feet of compacted plant matter is required to form one foot of coal. Some coal veins are 100 feet thick, which gives one pause to consider how much plant life is incorporated in coal. Most coal was formed 300–400 million years ago during the Devonian and Carboniferous geologic epochs when swamps covered much of the earth and plant life thrived in a higher atmospheric concentration of carbon dioxide. The interruption of plant decay by the formation of massive peat bogs removed huge amounts of carbon dioxide from the atmosphere, clearing the way for a more hospitable environment for animal life. However, some coal is of more recent vintage, laid down 15–100 million years ago, and the newest coal has an estimated age of only 1 million years. When coal is burned we are completing a recycling process interrupted eons ago, or much more recently for those who believe that coal stems from Noah's Flood.

Peat bogs are found in Ireland, England, the Netherlands, Germany, Sweden, Finland, Poland, Russia, Indonesia, and in the United States (the Great Dismal Swamp in North Carolina and Virginia, the Okefenokee Swamp in Georgia, and the Florida Everglades). The high water content has to be removed before peat can be burned as a biomass fuel whose heat content is much lower than coal. Peat is burned in Ireland for heating homes and in Finland for heating homes and generating electricity as a substitute, along with wood waste, for imported fossil fuels. Peat is also mixed with soil to improve its water-holding properties and is a filter material for sewage plants. Once removed, fish can be raised in the resulting pond or, if the peat bog is drained, agricultural crops can be grown, or the peat bog can simply remain fallow. There is always the possibility that these peat bogs may one day become coal beds if buried by hundreds of feet of silt and water.

As in many other areas, the Chinese beat out the Europeans in burning coal. Coal from the Fu-shun mine in northeastern China was consumed for smelting copper and casting coins around 1000 BCE. In 300 BCE the Greek philosopher Theophrastus described how blacksmiths burned a black substance that was quite different from charcoal. From evidence in the form of coal cinders found in archeological excavations, it is known that Roman forces in England burned coal as a fuel before 400 CE. Although the Romans did not record burning coal, they did record a "pitch-black mineral" that could be carved into trinkets for adorning the human body. That pitch-black mineral was an especially dense type of coal. Like glassmaking, burning coal for heat and blacksmithing and offerings to the gods, plus carving into trinkets for the fashionable of Rome, disappeared along with the Roman Empire. We presume that ever-expanding human knowledge being passed on to following generations has always been ongoing, is ongoing, and will always be ongoing. This, as history clearly shows, is an unwarranted presumption.

The English rediscovered coal in the 1200s during an early episode of deforestation around London, about the same time that the Hopi and Pueblo Indians began burning coal to glaze their ceramic ware in what is now the U.S. Southwest. After the coal gatherers picked up the coal lying on the ground on the banks of the River Tyne near Newcastle, they began chipping away at the exposed seams of coal in the nearby hillsides. Coal mining started when holes became tunnels that bored deep into the thick underground seams of coal. A new profession and a new class of people emerged, ostracized by the rest of society by their origin (displaced peasants) and the widely perceived degrading nature of their work. Coal miners as individuals were at the mercy of the mine owners until they learned to band together for their mutual benefit and protection, giving birth to the modern labor movement.

And there was plenty of incentive for miners to band together as the coal miners bored deeper into the earth. Mining is a very dangerous occupation. Cave-ins can trap the miners. If not immediately snuffed out by the falling rock, they remained trapped, awaiting rescue or dying from asphyxiation or starvation. To combat the peril of cave-ins, miners bonded with huge rats that lived in coal mines by sharing their meals with them. Miners remained alert to the comings and goings of the rats on the theory that rats could sense a cave-in before it occurred, not unlike rats deserting a sinking ship. Perhaps miners' casualty lists best document the perspicacity of rats to sense impending disaster.

In addition to cave-ins, coal miners had to contend with poisonous gases. Mining could release pockets of carbon dioxide or carbon monoxide, odorless and colorless gases of plant decay trapped within the coal seam that quickly killed their victims by asphyxiation. Canaries were the best defense since their chirping meant that they were alive. When they stopped chirping, they were already dead, a dubious warning system at best. A third colorless and odorless gas was methane, also released by mining operations when they exposed pockets of natural gas embedded in the coal seam. Unlike carbon dioxide and monoxide, methane is lighter than air and combustible. As methane accumulates along the ceiling of a mine, it eventually comes in contact with a lighted candle where it either burns or sets off a horrific explosion, depending on its concentration. A new professional, called, euphemistically, a fireman, would wrap his wretched body with wet rags and crawl along the bottom of the mine holding up a stick with a candle at the end, hoping he would discover methane before it was sufficiently concentrated to set off an explosion. Now all he had to do was hug the mine floor while the methane blazed above him.

Coal found in the hills around the River Tyne was moved down to the river and loaded on vessels for shipment to other parts of coastline England, notably London. Access to water provided cheap transportation on ships whereas the overland movement of coal on packhorses was prohibitively expensive. Roads hardly existed and, where they did, deep ruts made them impassable for heavily laden horse-drawn wagons. By 1325 coal became the first internationally traded energy commodity when exported from Newcastle to France and then elsewhere in northern Europe. Thus, coal saved not only the English but also the European forests from devastation. The saying "carrying coals *to* Newcastle" originally referred to something only a simpleton would do since Newcastle was the world's first and largest and most famous coal-exporting port. Six and a half centuries later coal was carried to Newcastle when Britain began importing coal.

Burning coal made an immediate impression on the people. In 1306, the nobles of England left their country estates to travel to London to serve in Parliament, as was their custom. This time there was something else in the air besides the stench of animal dung, raw sewage, and rotting garbage. The nobles did not like the new pungent aroma spiced with brimstone (sulfur) and succeeded in inducing King Edward I to issue a ban on burning coal. It is one thing for a king to issue a ban, and quite another to enforce it, the classic limit of power faced by parents of teenagers. Regardless of the king's edict, the merchant class of newly emerging metallurgical enterprises had to burn coal because wood was not available in sufficient quantities around London, and what was available was too expensive. Simple economics overruled the king's ban. The fouling of the air of London and other English cities remained for centuries to come. It is hard to imagine that the charming English countryside we know, speckled with quaint towns, cottages, and farms was once, like the eastern United States, nearly one continuous forest.

From the beginning, coal was a matter of dispute between the church, which happened to own the land where the coal was found, the crown, which coveted this natural resource, and the merchant class that transformed coal into a considerable amount of personal wealth. As church, crown, and capital struggled over who would reap the financial benefits, merchant vessels were built to ship coal on the high seas. This, in turn, necessitated building naval vessels to protect the merchant fleet from marauders and pirates. The English also imposed a tax, which greatly favored the building and manning of English ships, on non-English vessels carrying coal exports. In this way, coal contributed to making England a sea power and is, therefore, partly responsible for the emergence of England as the world's greatest colonial power. Growth of sea power put more pressure on forests for lumber to build ships and, in particular, trees fit for masts, which eventually would be harvested in English colonies in the New World.

The Black Death did not enhance coal's reputation as its victims turned black smelling brimstone in the air from burning coal, widely interpreted as to where they might be heading. The Black Death wiped out between one-third and a one-half of Europeans. The depopulation of London meant less coal had to be burned, improving the quality of its air, and forests regained a toehold in the countryside. The reign of Elizabeth I was marked by population and economic recovery after the Black Death, increasing the demand for firewood. She greatly expanded the English Navy to defend the kingdom against the Spanish Armada, increasing the demand for lumber and masts to build warships and charcoal for smelting iron for ship armament. This again put pressure on the kingdom's forests, resulting in widespread deforestation throughout England and another steep rise in the price of firewood.

The adoption of the chimney in London homes in the 1500s allowed for the conversion from wood to coal for heating in the early 1600s, a conversion already completed by industry. While the ability of chimneys to keep the heat inside and channel smoke outside was an advantage for those who dwelt inside, the same could not be said for those who ventured outside. Appalling amounts of acrid smoke eroded and blackened stone in statues and buildings, stunted plant life, affected the health of the population, and made black and dark brown the colors of choice for furnishings and fashion.

London was not the only city that suffered from severe air pollution. During the rapid advance of the Industrial Revolution in the nineteenth century, Manchester became the center of British textile manufacturing and Pittsburgh the center of American steelmaking. The former suffered mightily from coal burned in steam engines to run the textile machines and the latter from coal consumed in making steel. Not all cities suffered equally. Philadelphia and New York were spared at first because of rich anthracite coalfields in eastern Pennsylvania. Anthracite, a hard coal of nearly pure carbon, burns with little smoke. Unfortunately, anthracite reserves were in short supply when coal-burning electricity-generating plants were built at the end of the nineteenth and early twentieth centuries. These plants burned cheaper and more available bituminous coal. New Yorkers staged an early environmental protest against the fouled air that the utility managers could not ignore, so they switched to anthracite coal to appease people while they were awake, but switched back to bituminous while they slept.

We tend to think of air pollution caused by burning coal as a nineteenth-century phenomenon affecting London, Manchester, and Pittsburgh. Yet, only a little over a half-century ago, for four days in early December 1952, a temperature inversion settled over London, trapping a natural white fog so dense that traffic slowed to a crawl and the opera had to be cancelled when the performers could no longer see the conductor. Then coal smoke, also trapped in the temperature inversion, mixed with the fog to produce an unnatural black fog that hugged the ground and cut visibility to less than a foot. Perhaps unbelievably from our vantage point, 4,000 Londoners died from traffic accidents and inhaling sulfur dioxide fumes. Parliament subsequently banned the burning of soft coal in central London, bringing to an end a quaint 700-year-long tradition. In the twenty-first century, Beijing, Shanghai, and other cities in Asia have picked up where London left off. While the results of living in a cloud of polluted air is not as calamitous as in London, nevertheless dwellers in Asian cities suffer from various health impairments.[2]

Coal and the Industrial Revolution

Coal played an important role in England's emergence as the world's greatest seafaring nation and, subsequently, as the world's leading trading nation and colonial power. It also played an important, if not a pivotal, role in bringing about the Industrial Revolution and England's subsequent emergence as the world's greatest industrial power.

At first coal mines were above the River Tyne and narrow downward shafts dug from the mines to the outside world took care of removing water seepage from rain. As the coal seams bent downward, it was only a matter of time before mining took place under the River Tyne and the North Sea. This opened up a whole new peril for the miners: death by drowning. Even if mining did not breach the river or the sea, water was continually seeping in through the ground, threatening to flood the mines, though not necessarily the miners. For many years the chief way to prevent flooding was to have men haul up buckets of water to the mine surface. As mines went deeper into the earth, a vertical shaft was dug where a continuous chain loop with attached buckets brought water from the bottom of the mine to the surface. Water wheels and windmills powered a few of these continuous chain operations, but most were powered by horses. The capital cost in chain loops, along with their attached buckets and the operating cost of feeding and tending to the horses, encouraged the development of bigger mines employing larger numbers of miners in order to produce the greater quantities of coal needed to cover the higher capital and operating costs. Concentrating coal mining in a smaller number of larger operations meant even deeper mines, perversely exacerbating the problem of water removal.

By the 1690s, Britain's principal industry of providing 80 percent of the world's coal was threatened with a watery extinction. The nation's intellectual resources were focused on solving what seemed to be an overwhelming challenge: how to prevent water from flooding the ever-deeper mines. Denis Papin proposed the idea of having a piston inside a cylinder where water at the bottom of the cylinder would be heated to generate steam under the piston that would drive the piston up. Then the heat would be removed, and a pressure differential would be created between the top and bottom of the piston as the steam condensed to form a vacuum. Atmospheric pressure on top of the piston would drive it down and then the water in the bottom of the cylinder would be reheated to generate steam to drive the piston back up. The up-and-down motion of the piston could power a water pump. Thomas Newcomen, who may or may not have heard of Papin's idea, worked ten years to develop a working engine that did just that.

The Newcomen engine was a piston within a cylinder. Steam from burning coal was fed into the cylinder space below the piston, forcing it up. Then a cold-water spray entered the cylinder space and condensed the steam to create a vacuum and a pressure differential between the top and the bottom of the cylinder. Atmospheric pressure on top of the piston would drive the piston down. Simultaneously, an exhaust gate would open, allowing the water from the spray and condensed steam to drain from the cylinder space. Then the exhaust gate would close and steam would reenter the cylinder space. This continual cycle of feeding steam followed by a spray of water into the bottom of the cylinder kept the piston moving up and down. A crossbeam connected the moving piston to a water pump. Mines could now be emptied of water without horses and chain loops with attached buckets, which by this time had reached their limits of effectiveness. By 1725 Newcomen engines were everywhere and had grown to prodigious size, but the alternate heating and cooling of the lower cylinder walls during each cycle of the piston movement made them extremely energy-inefficient. With coal cheap and plentiful, the Newcomen engine had no technological rival for sixty years. As energy-inefficient as Newcomen engines were, they nevertheless saved the English coal-mining industry from a watery grave and enabled England to maintain its preeminence in coal mining for another century.

Thus, coal or, to be more exact, the threat of coal mines filling with water, brought into existence the first industrial fossil-fueled machine that delivered much more power with far greater dependability than wind or water. The fickleness of the wind makes wind power vulnerable and waterpower is constrained by the capacity of a water wheel to translate falling or moving water into useful power and by the occurrence of droughts. The Newcomen engine had no such limitations.

The building of Newcomen engines required iron and smelting iron consumed charcoal, another contributor to the deforestation of England. The pressure on forests was lifted in 1709 when Abraham Darby, who also advanced the technology of casting pistons and cylinders for Newcomen engines, discovered that coke from coal could substitute for charcoal from wood in smelting iron. It is a bit ironic that coke itself had been discovered some sixty years earlier, in 1642, for brewing beer. London brewers needed a great deal of wood to dry malt. As wood supplies dwindled, they first experimented with coal, but quickly found out that sulfur in coal tainted the malt and, thus, the flavor of the beer. The brewers discovered coke by copying the process of making charcoal from wood, which is essentially baking coal in the absence of oxygen to drive out volatile elements and impurities. Coke is harder than coal, almost pure carbon, and burns at a high temperature without smoke. Malt dried with coke produced a pure, sweet beer.

In 1757 James Watt, an instrument maker for the University of Glasgow, was given an assignment to repair the University's model of the Newcomen engine, which spurred his lifelong interest in steam engines. Watt soon realized that the shortcoming of the Newcomen engine was the energy consumed in reheating the cylinder wall after each injection of cold-water spray. His idea was not to cool the steam in the hot cylinder, but to redirect the steam to another cylinder, or condenser, surrounded by water, where the steam could be condensed without cooling the cylinder wall. Rather than a valve opening to allow a cold spray to condense the steam, a valve opened to allow the expended steam to escape from the cylinder to the condenser. The condensed steam created a vacuum in the bottom of the cylinder, which allowed atmospheric pressure on top of the cylinder to push the piston down. In this way the power cylinder wall would remain hot throughout the operation of the engine, improving its thermal efficiency.

James Watt was assisted by the moral and financial support of Matthew Boulton, a well-known Birmingham manufacturer. After obtaining a patent, the first two steam engines were built in 1776. One pumped water from a coal mine and the other drove air bellows at an iron foundry. The foundry owner, John Wilkenson, invented a new type of lathe to bore cylinders with greater precision, a device that would prove useful for manufacturing steam engines. The final version of the Watt engine came in 1782, when Watt developed the double-acting engine. In this model steam entered either end of the piston. Steam entering one end of the cylinder drove the piston in one direction, while a valve opening on the other end of the cylinder allowed the spent steam from the previous stroke to exhaust into a condenser. This operation was reversed to drive the piston in the opposite direction. Valves for allowing live steam to enter the cylinder space or spent steam to enter the condenser were opened and shut by the movement of the piston. To further enhance energy efficiency, steam was admitted inside the cylinder only during the first part of the piston stroke, allowing the expansion of the steam to complete the stroke. To further cut heat losses a warm steam jacket surrounded the cylinder and a governor controlled the engine speed. With these enhancements, the Watt steam engine could operate with one-quarter to one-third the energy necessary to operate an equivalent Newcomen engine.[3] Both the Newcomen and Watt engines spurred technological advances in metallurgy to improve metal performance and in manufacturing technology to make cylinders and pistons, lessons not lost on the military for building bigger and better cannons.

Watt's intention was to improve the energy efficiency of the Newcomen engine for pumping water out of mines. Boulton saw Watt's invention as something more than a more efficient Newcomen engine or a more reliable means of powering his factories than water wheels. Boulton was a visionary who saw the steam engine as a means to harness power for the good of humanity. In Boulton's vision, steam engines would not only drain mines of water, but power factories that could be built at any location where coal was nearby.

Goods made by machines powered by steam engines would free humans from the curse of drudgery and poverty that had plagued them throughout history.

The world's first industrialized urban center was Manchester, England. Manchester became the textile center of the world, processing cotton from slave plantations in the United States. Coal was consumed in making iron that went into constructing factory buildings, steam engines, and textile-making machines. Coal also fueled the steam engines that powered the machines and gas given off by heating coal was piped into the factory buildings and burned in lamps to allow round-the-clock operations. All this coal burning smothered Manchester in a thick black blanket of smoke that rivaled pollution in London and, later, Pittsburgh.

The demand for coal from mines near Manchester was so great that narrow shaft seams, which only children could fit into, were brought into use. They had to crawl on their feet and hands dragging heavy sleds of coal behind them like pack animals. Many of these children lived like animals in abandoned portions of mine shafts, separated from their families and daylight. For workers in the Manchester factories, the long hours, the harsh working conditions, the poor pay, the putrid stench of the atmosphere, their appallingly poor health and high death rates, and the breakdown of the family had to be an Orwellian nightmare at its worst, not Boulton's vision at its best. What Freidrich Engels saw in Manchester was recorded in his work *The Condition of the Working Class in England* (1844), which in turn helped Karl Marx shape *The Communist Manifesto* (1848).

Coal and Railroads

The amount of coal a horse can carry on its back is limited, but its carrying capacity can be improved by having it pull a wagon. The dirt roads of the day, with their deep muddy ruts, were impassable for horses hauling heavy wagonloads of coal. A horse's capacity to move cargo jumps by several orders of magnitude when, instead, it pulls a barge on still water. Canals, not roads, could move large volumes of coal to inland destinations. One of the first canals in Britain moved coal to Manchester from nearby coalfields where horses pulled barges from towpaths alongside the canal. This began the canal-building boom in England where, by the early 1800s, canals were used not only to move coal, but all sorts of raw materials and finished goods to and from cities. Since the nature of the terrain and the availability of water restricted canal construction, wagon ways, where horses were harnessed to cargo-laden carriages riding on wooden rails, complemented canals. Rails made using horses to move coal more efficient than pulling loaded wagons on muddy, rutted, dirt roads.

Rails also improved coal-mine productivity. It turned out that getting coal out of the mine was as labor-intensive as mining coal. Often human pack animals were responsible for hauling coal on its journey to the mine surface. One human pack animal would pick up a small wagonload from another human pack animal, tow it a bit, and pass it on to still another human pack animal, then walk back to get the next. Lifetimes were spent hauling coal out of mines and, sometimes, living in mines. Mine operators did what they could to make hauling coal easier, but not strictly for altruistic reasons. Installing rails reduced operating costs by having the same work done by fewer human pack animals, thus improving productivity and, incidentally, profitability. Most rails were made of wood, but a few were made of iron.

Because the use of rails had solved the problem of how to move heavy loads, the concept of the railroad was in place when George Stephenson, the father of railways, put together the elements of iron track with a high-pressure Watt's steam engine on a locomotive platform with flanged iron wheels that pulled flanged iron wheeled carriages. Fittingly, the world's first railroad connected a coal town with a river town twenty-six miles away. The Age of the Railroad began in earnest a few years later, in 1830, when a train on its inaugural run between Liverpool and Manchester hit a top speed of an unbelievable thirty-five miles per hour. By 1845 Britain had 2,200 miles of track, a figure that tripled over the next seven years. While the building of railroads meant relatively cheap and fast transportation between any two points in England, the iron for the rails was not cheap.

Coal and Steel

The Iron Age began sometime around 2000 BCE, perhaps in the Caucasus region, where iron first replaced bronze. Iron is harder, more durable, and holds a sharper edge longer than bronze. Iron is also the fourth most abundant element, making up 5 percent of the earth's crust. Iron ore is made up of iron oxides plus varying amounts of silicon, sulfur, manganese, and phosphorus. From its start, smelting iron consisted of heating iron ore mixed with charcoal until the iron oxides began reacting with the carbon in the charcoal to release its oxygen content as carbon monoxide or dioxide. Adding crushed seashells or limestone, called flux, removed impurities in the form of slag, which could be separated from the heavier molten iron. This left relatively pure iron, intermixed with bits of charcoal and slag that could then be hammered on an anvil by a blacksmith to remove the remaining cinders, slag, and other impurities. The result of the hammering produced wrought (or "worked") iron with a carbon content between 0.02–0.08 percent. This small amount of carbon, absorbed from the charcoal, made the metal both tough and malleable. Wrought iron was the most commonly produced metal throughout the Iron Age.

By the late Middle Ages, European iron makers had developed the blast furnace, a tall chimneylike structure in which combustion was intensified by a blast of air pumped through alternating layers of charcoal, flux, and iron ore. The medieval ironworkers harnessed water wheels to power bellows to force air through the blast furnaces. Centuries later, this would be one of the first tasks for James Watt's steam engines, in addition to pumping water out of coal mines. The blast of air increased the temperature, which allowed the iron to begin absorbing carbon, thereby lowering its melting point. The product of this high-temperature process was cast iron, with between 3–4.5 percent carbon. Cast iron is hard and brittle, liable to shatter under a heavy blow, and cannot be forged (that is, heated and shaped by hammer blows). The molten cast iron was fed through a system of sand troughs, formed into ingots, which reminded people of a sow suckling a litter of piglets, and became known as pig iron. Pig iron was either cast immediately or allowed to cool and shipped to a foundry as ingots, where it was remelted and poured directly into molds to cast stoves, pots, pans, cannons, cannonballs, and church bells.

These early blast furnaces produced cast iron with great efficiency and less cost than wrought iron. However, the process of transforming cast iron to more useful wrought iron by oxidizing excess carbon out of the pig iron was inefficient and costly. More importantly, what was desired was not wrought iron from cast iron, but steel. Steel is iron with a carbon content between 0.2–1.5 percent, higher than wrought iron but lower than cast iron. Crucible steel, named after its manufacturing process, was not only very expensive but the extent of the oxidation of carbon, and therefore the carbon content, could not be controlled. Regardless of its cost, steel was preferred over wrought iron because it was harder and kept a sharp edge longer (the best swords were made of steel) and was preferred over cast iron because it was more malleable and resistant to shock.

Early rails made from wrought iron were soft and had to be replaced every six to eight weeks along busy stretches of track. Steel, in contrast, is perfect for rails because it is harder than wrought iron and more malleable than cast iron. Steel rails, however, were prohibitively expensive. The man of the hour was Henry Bessemer, who was not responding to the needs of the railroad industry, but the military. Bessemer had invented a new artillery shell that had been used in the Crimean War (1853–1856). The army generals complained that the cast iron cannons of the day could not handle Bessemer's more powerful artillery shell. In response Bessemer developed an improved iron-smelting process that involved blasting compressed air through molten pig iron to allow the oxygen in the air to unite with the excess carbon and form carbon dioxide. Ironically, Bessemer's invention, patented in 1855, was similar to the method of refining steel used by the Chinese in the second century BCE.

In 1856 the first Bessemer converter, large and pear-shaped with holes at the bottom for injecting compressed air, was completed. Other individuals contributed to improving the Bessemer converter by adding manganese to the converter's ingredients of coal (coke), iron ore to get rid of excess oxygen left in the metal by the compressed air, and adding limestone to get rid of any phosphorus in the iron ore and whose presence made steel excessively brittle. Limestone becomes slag after absorbing phosphorus and other impurities and floats at the top of the converter where it is skimmed off before the steel is poured out. Bessemer converters were batch operations to which iron ore, coke, and limestone were added; within a short period of time, molten steel was on the bottom and slag was floating on the top. After removing the slag, the converter was then emptied of its molten steel and then reloaded to make another batch.

The economies of large-scale production utilizing the Bessemer converter transformed undesired wrought-iron rail at $83 per ton in 1867 to desired steel rail at $32 per ton by 1884. It was not long before the Bessemer process had a technological rival: the open-hearth furnace. The open-hearth furnace, while it took longer, could make larger quantities of steel because raw materials were continuously added and slag and steel continually removed. Moreover, steel could be made with more precise technical specifications and scrap steel could be consumed as feedstock along with iron ore, coal, and limestone. Improvements in the chemical composition of steel had increased the life of steel rails and their weight-carrying capacity several fold by 1900, when the open-hearth furnace had largely replaced the Bessemer converter. Another man of the hour, Andrew Carnegie, organizationally shaped the steel industry and, in so doing, reduced the price of steel rail to $14 per ton by the end on the nineteenth century. Carnegie also introduced the I-shaped steel girder for building skyscrapers, a major addition to steel demand once the Otis elevator was perfected.

By 1960, the basic oxygen furnace had, in its turn, replaced the open-hearth furnace. The basic oxygen furnace is essentially a modification of the original Bessemer converter. The first step is feeding iron ore, coke, and limestone into a furnace with air blasted through the mixture to produce molten iron, which is periodically tapped from the bottom of the furnace while the molten slag is periodically removed from the top. The molten iron then goes into the basic oxygen furnace where steel scrap and more limestone are added, along with a blast of oxygen to produce almost pure liquid steel.

In making steel, coking coal supplies carbon to remove the oxygen in the iron ore and heat to melt the iron. Coking, or metallurgical coal, must support the weight of the heavy contents in a furnace yet be sufficiently permeable for gases to rise to the top and molten steel to sink to the bottom of the furnace. Thus, coals are divided into two types: thermal coal fit only for burning and coking coal fit for steelmaking. The liquid and gaseous by-products in producing coke from metallurgical or coking coal find their way into a host of products such as synthetic rubber, ink, perfume, food and wood preservatives, plastics, varnish, stains, paints, and tars.[4]

The basic oxygen furnace produces 63 percent of the world's crude steel production—about 1,129 million tons in 2005—incidentally consuming 592 million tons of coal. The world's largest steel producers are China (349 million tons, up an amazing 25 percent from 2004), Japan (113 million tons), the United States (94 million tons), Russia (66 million tons), South Korea (48 million tons), and Germany (45 million tons). Most of the remaining steel production is made from a more recent innovation, the electric arc furnace.[5] The raw material for electric arc furnaces is scrap. Incidentally, steel is the most recycled commodity on Earth: Fourteen million cars in the United States alone are recycled annually. Whereas 1 ton of steel made from raw materials requires, in round terms, 2 tons of iron ore, 1 ton of coal, and a half ton of limestone, 1 ton of recycled steel needs a bit more than 1 ton of scrap. While coal is absent as a raw material in making steel with the electric arc furnace, an electric arc furnace uses a lot of electricity, as one can imagine, which is mainly generated by burning coal augmented by capturing the waste heat of steelmaking. Thus, coal is consumed directly in making steel with the basic oxygen furnace and indirectly in making steel with the electric arc furnace.

Coal played a vital role in shaping the world as we know it today. Coal was needed as a substitute for wood for producing glass and smelting metals after the forests were cut down. Coal became a major export item for England, spurring the development of the English navy. The challenge posed by flooding coal mines frantically called for a solution—the Newcomen engine—the first industrial power-generating machine not dependent on wind or water. The Newcomen engine spurred further advances in metal and toolmaking and led directly to Watt's steam engine. Watt's steam engine powered the Industrial Revolution with coal, steel, and railroads. Coal, then, is at least partly responsible for England becoming a world sea power, a colonial power, and, after the birth of the Industrial Revolution, the world's first and mightiest industrial power. This lasted for over half a century before being challenged by the emergence of rival centers of industrial power in the United States, Germany, and Japan.

The Rise and Fall of King Coal

While early steam locomotives were fueled by wood, it was not long before they switched to coal. One reason was deforestation; the other was the availability of coal as the most commonly carried commodity. Coal became the sole source of energy for fueling locomotives, which for decades before the automobile age was the sole source of transportation on land other than horses. Robert Fulton invented the first steam-driven riverboat, the *Clermont*, which propelled itself from

New York to Albany in 1807. While wood could be burned on riverboats, oceangoing vessels burned coal, a more concentrated form of energy that took up a lot less volume. The famed clipper ships of the waning decades of the nineteenth century marked the final transition from a source of power that was undependable, renewable, and pollution- and cost-free to one that was dependable, nonrenewable, polluting, and not cost-free. Now coal had it all on land and sea. Thomas Edison's first electricity-generating plants were fueled by coal, although hydropower was soon harnessed at Niagara Falls. Coal and hydropower were the principal sources of energy for generating electricity during the first half of the twentieth century.

Coal's share of the energy pie peaked at 60 percent in 1910. Oil, natural gas, and hydropower contributed another 10 percent, and biomass 30 percent. After 1910, things began to change for King Coal. Coal maintained its preeminence in passenger transportation until Henry Ford put America, and the world, on gasoline-driven wheels. In 1912, the *Titanic* had 162 coal-fired furnaces fed continuously by 160 stokers working shifts and shoveling as much as 600 tons of coal per day. This might work well for passenger vessels, but coal-burning warships were constrained in fulfilling their primary mission by the large portion of the crew dedicated to shoveling coal, rather than manning guns, and the amount of space dedicated to holding coal rather than carrying ammunition. Moreover, warships with a heavy cargo of coal moved slowly and their pillars of smoke signaled the enemy as to their whereabouts. Admiral Sir John Fisher, head of the British Navy, spearheaded the transformation from marine boilers powered by coal to oil in the years prior to World War I. Naysayers scoffed at the idea, but as soon as the obvious advantages of oil over coal were demonstrated in higher speed, greater firepower, and less emissions to betray a vessel's presence, it became a race to dump coal in favor of oil. As ships made the transition from coal to oil, the worldwide network of coal-bunkering stations supplied by coal colliers was converted in tandem to handle oil supplied by tankers.

Coal and wood remained the chief sources of energy for cooking until the advent of the electric stove in the 1920s, along with stoves that burned natural gas and liquid propane. About this time homes began a slow conversion from coal to heating oil and natural gas. Automobiles were taking passengers away from electric trolleys, whose electricity was generated from coal, for inner-city transportation. Intercity railroad passenger train traffic, powered by coal-fueled locomotives, declined as a network of roads sprang into existence. When the fall of King Coal from preeminence sped up during and after World War II, one individual stood out: John L. Lewis, a former coal miner and president of the United Mine Workers. A contentious personality who had the audacity to defy President Franklin Delano Roosevelt by leading a coal miners' strike during World War II, Lewis was instrumental in raising the pay and improving the health and retirement benefits and working conditions for coal miners. As laudable as these well-deserved benefits were, they also increased the price of coal and, in so doing, hastened its demise. Perhaps no better proof of this was Perez Alfonso, a Venezuelan oil minister, who wanted to erect a statue to honor Lewis for boosting the market for Venezuelan oil exports.

The rise in the price of coal from John L. Lewis's success was an added inducement for homeowners to switch from coal, which had to be shoveled into a furnace (from which ashes had to be removed and disposed of) to the much greater convenience of heating oil, propane, and natural gas, which did not require the hard labor associated with coal. In cooking, the switch was already far advanced from coal to electricity and natural gas and propane.[6] While oil-driven automobiles, buses, and airplanes were diverting people from coal-burning passenger trains, and trucks had taken over local distribution of freight, railroad freight trains still carried the bulk of the nation's intercity freight. Trucks were unable to cut deeply into intercity freight traffic because the road network was relatively undeveloped and better fit for automobiles than trucks. All this changed with the launching of the interstate highway system by President Dwight D. Eisenhower.

A large steam locomotive pulling a loaded freight train burned 1 ton of coal per mile, which required a fulltime fireman to continually shovel coal. Railroads were enormous consumers of coal and railroad executives displayed equally enormous reluctance to abandon steam locomotives when the diesel engine first appeared in the late 1930s. Steam locomotives had become an intimate part of railroading folklore. Distinct in design and operating nuances, they had to be maintained by a dedicated crew that became inseparable from the locomotive, which required a lot of downtime for maintenance and repair.

Railroaders were unwilling to switch from steam to diesel, even though diesel locomotives had inherent advantages. Diesel engines were fuel-efficient because they burned gallons of diesel fuel per mile, not a ton of coal per mile. The diesel engine avoided the inherent energy inefficiency of a steam engine from which the latent heat of vaporization was passed to the atmosphere. In a diesel engine, fuel sprayed into the cylinder space above a piston is ignited by heated compressed air. The expansion of the gases of combustion powers the first downward stroke. After the power stroke, the piston is forced up to expel the exhaust gases, then down to draw in fresh air, then up to compress the air. The heated compressed air ignites another spray of fuel whose expanding gases of combustion powers another downward stroke. Thus, every other downward stroke is a power stroke that, through a crankshaft connected to the other pistons, drives an electricity generator that powers electric motors attached to the engine wheels.

Diesel engines have other advantages as well. They are more reliable because they require less maintenance and repair, both in downtime and cost; less manpower, because no coal has to be shoveled; and less frequent refueling. Steam locomotives of various horsepower have to be built to handle freight trains of different sizes, whereas a different number of standard sized diesel engines can be hooked together to obtain the requisite horsepower. In short, the only reason to keep steam locomotives once diesel engines made their appearance was management's reluctance to change.

The advantages of the diesel engine could no longer be ignored when John L. Lewis's success in improving the lot of coal miners increased the price of coal. The first diesel engines were restricted to moving freight cars around freight yards and were excluded from long intercity runs, the exclusive domain of the steam locomotive. Steam locomotives could persevere as long as all railroad managers agreed to use steam locomotives on intercity freight trains, ensuring equal inefficiency in operations for all. But this holding action could not ignore the competitive threat of a growing volume of trucks gaining access to intercity traffic made possible by the interstate highway system. If any railroad bolted to diesel for hauling intercity freight, then the inherent efficiencies and advantages of diesel locomotion would give that railroad a competitive edge over the others. And that is what happened: One railroad bolted. As soon as one made the switch to diesel for intercity freight trains, it was a race to convert locomotives from coal to oil similar to the race to convert ships from coal to oil. Despite efforts by steam locomotive aficionados and railroad executives to hold the fort, the steam whistle and the chugging locomotive spewing steam, smoke, and at times blazing ashes disappeared within a decade.

Adding to King Coal's woes, electricity-generating plants built after World War II were designed to run on oil, natural gas, and nuclear power in addition to coal and hydro. King Coal was no longer king in transportation, electricity generation, heating houses and commercial buildings, and home cooking. The king, however, was not dead and even enjoyed a reprieve during the second energy crisis: the oil crisis of 1973.

Types of Coal

There are four types of coal aside from peat, a precursor to coal. The lowest quality of coal and the largest portion of the world's coal reserves is lignite, a geologically young, soft, brownish-black coal, some of which retains the texture of the original wood. Of all coals, it has the lowest carbon content, 25–35 percent, and the lowest heat content, 4,000–8,300 British thermal units (Btus) per pound. The next step up is sub-bituminous coal, a dull black coal with a carbon content of 35–45 percent and heat content between 8,300–13,000 Btus per pound. Both lignite and sub-bituminous coals, known as soft coals, are primarily thermal coals for generating electricity. Some sub-bituminous coals have lower sulfur content than bituminous coal, an environmental advantage.

Next are the hard coals, bituminous and anthracite. Bituminous is superior to soft coal in terms of its carbon content, 45–86 percent, and energy content 10,500–15,500 Btus per pound. Bituminous coal is the most plentiful form of coal in the United States and is used both to generate electricity (thermal coal) and, if it has the right properties, as coking or metallurgical coke for steel production. Anthracite coal has the highest carbon content, 86–98 percent, and a heat content of nearly 15,000 Btus per pound. Anthracite coal was closely associated with home heating because it burned nearly smokeless. As desirable as anthracite is, it is also scarce. In the United States, anthracite is found in only eleven counties in northeastern Pennsylvania.[7]

Coal Mining

Coal mines have historically been subterranean regions where accidents and black lung have taken their toll. Mining coal in the twenty-first century is an activity carried out differently than it was in the past. In developed nations, no gangs of men swing pickaxes to remove the over- and underburden of rock to gain access to the coal, then again to chip out the coal. No gangs of men shovel the rock or coal into small wagons or carts for the trip to the surface. Now the most popular way of removing coal is continuous mining machines with large, rotating, drum-shaped cutting heads studded with carbide-tipped teeth that rip into a seam of coal. Large gathering arms scoop the coal directly into a built-in conveyor for loading into shuttle cars or a conveyor for the trip to the surface. Continuous cutters ripping and grinding their way through coal seams can do in minutes what gangs of miners with pickaxes and shovels took days to accomplish.

The next most popular method for removing is a machine resembling an oversized chain saw that cuts out a section of coal in preparation for blasting to allow for its expansion. Holes are then drilled for explosives that blast large chunks of coal loose from the seam. Loaders scoop up the coal into conveyors that fill shuttle cars to haul the coal out through the shaft. For both methods of mining, long rods or roof bolts are driven into the roof of the mine to bind layers of weak strata into a single layer strong enough to support its own weight. If necessary, braces are used for additional support. Wood is favored for this because it makes a sharp cracking sound if the roof begins to weaken.

An increasingly popular and efficient means of mining introduced into the United States from Europe in the 1950s is longwall mining: a rotating shear moves back and forth in a continuous, smooth motion for several hundred feet across the face or wall of a block of coal. The cut coal drops into a conveyor and is removed from the mine. Some of the rock on top of the coal also collapses, which is then removed to the surface or piled in areas where the coal has been removed. The main supports for the rooms created by longwall mining are pillars of solid coal, which are the last to be mined before a mine is abandoned.

Regardless of the type of mining technology employed, mine shafts for transporting miners and coal either slope down to coal beds that are not too deeply located in the earth or are vertical to reach beds of coal more than 2,000 feet beneath the surface. Huge ventilation fans on the surface pump air through the mineshafts to reduce the amount of coal dust in the air, prevent the accumulation of dangerous gases, and ensure a supply of fresh air for the miners.

In recent decades, surface mining has gained prominence over subterranean mining. In the western part of the United States, 75 percent of the coal produced is obtained from surface mines with coal deposits up to 100 hundred feet thick. Surface mining also occurs in Appalachia. Surface mines produce 60 percent of the coal mined in the United States, while the remaining 40 percent comes from underground coal mines located primarily in Appalachia. While there are large open-pit mines in other parts of the world, such as Australia and Indonesia, globally speaking about two-thirds of coal comes from underground mines.

A few utility plants are located at the mouths of mines, but most coal is loaded on barges and trains for transport to electricity-generating plants or export ports. In the United States, about 60 percent of the coal mined is moved by railroad to the consumer, often in unit trains of a hundred automatically unloading coal cars, each holding 100 tons of coal, or 10,000 tons of coal in a single trainload. Coal is unloaded by hoppers in the bottom of coal cars that open to drop the coal onto a conveyor belt located below the rails or by a rotating mechanism that empties 100 tons of coal by turning the coal cars upside down as though they were toys. Coal is still a major revenue generator for railroads around the world. Coal in the United States that is not moved by rail is primarily moved by barge on 25,000 miles of inland waterways. One unconventional way to move coal is to pipeline pulverized coal mixed with water from a coal mine to a power station, where the water is decanted and the pulverized coal is fed directly into a boiler.

After mining, coal is processed to ensure a uniform size and washed to reduce its ash and sulfur content. Washing consists of floating the coal across a tank of water containing magnetite for the correct specific gravity. Heavier rock and other impurities sink to the bottom and are removed as waste. Washing reduces the ash and pyretic sulfur-iron compounds clinging to the surface of the coal, but not the sulfur chemically bonded within the coal. Washing can also reduce carbon dioxide emissions by 5 percent. Magnetite clinging to the coal after washing is separated with a spray of water and recycled. Coal is then shipped by rail or barge to power plants. Some power plants run off a single source of coal while others buy various grades of coal that are mixed together before burning in order to obtain optimal results in heat generation, pollution emissions, and cost.

Coal-mining operations are highly regulated in the developed world. In the United States, a company must comply with hundreds of laws and thousands of regulations, many of which have to do with the safety and health of the miners and the impact of coal mining on the environment. Legal hurdles may require ten years before a new mine can be developed. A mining company must provide detailed information about how the coal will be mined, the precautions taken to protect the health and safety of the miners, and the mine's impact on the environment. For surface mining, the existing condition of the land must be carefully documented to make sure that reclamation requirements have been successfully fulfilled. Other legal requirements cover archaeological and historical preservation, protection and conservation of endangered species, special provisions to protect fish and wildlife, forest and rangeland, wild and scenic river views, water purity, and noise abatement.

In surface or strip mining, specially designed draglines, wheel excavators, and large shovels strip the overburden to expose the coal seam, which can cover the entire top of an Appalachian mountain. Coal is loaded into huge specially designed trucks by large mechanical shovels for shipment to a coal-burning utility or to awaiting railroad cars or barges. Surface mining has lower operating and capital costs and provides a safer and healthier environment for the workers than underground mining. After the coal is removed, the overburden is replaced and replanted with plant life to restore the land as closely as possible to its original state. Reclaimed land can also be transformed into farmland, recreational areas, or residential or commercial development, as permitted by the regulators.

Critics of surface mining point out the damage done to the landscape when the overburden removed from the top of a mountain or hill is dumped into nearby valleys, called “valley fill.” In addition to the destruction of the landscape and vegetation, valley fills become dams creating contaminated ponds of acid runoff from sulfur-bearing rocks and heavy metals such as copper, lead, mercury, and arsenic exposed by coal mining. They also object to the dust and noise of strip-mining operations and “fly-rocks” raining down on those unfortunately residing nearby. Another problem is abandoned underground mines, which eventually fill with water. The water can range from being nearly fit for drinking to containing dangerously high concentrations of acids and metallic compounds that may end up contaminating ground and drinking water.

Of course, the record also shows that there are large established companies mindful of their legal obligations to restore the landscape and protect the environment. There are instances of reclamation carried out so effectively that, with the passage of time, there is no apparent evidence that strip mining had ever taken place. Aside from corporate ethics, there are sound business reasons for being a responsible corporate citizen such as the desire to remain in business for decades to come. For these companies, the extra costs in protecting the health and safety of the miners and safeguarding the environment generate huge payoffs by allowing them to remain in business over the long haul. Private ownership is a right granted by governments on the basis that the conduct of business is better handled by businesspeople than government bureaucrats. If in reality, or if in the perception of the electorate, the supposed benefits of private ownership are not being achieved, then private ownership itself is threatened.

Table 4.1

Employment, Productivity, and Safety

	Employment (2000)	Miners per Million Tons Output	Deaths	Deaths per Million Tons Output
Australia	18	76	4	0.02
United States	77	96	38	0.05
United Kingdom	8	241	4	0.05
South Africa	54	298	30	0.17
Poland	158	1,561	28	0.28
India	456	2,171	100	0.48
Russia	197	1,195	137	0.83
China	5,000	5,501	5,786	6.36

There has been environmental degradation, but much of this lies with fly-by-night companies that fold without meeting their light-of-day responsibilities. While critics of coal extraction in developed nations abound, the developing nations, most notably China and India, seem to exist on another planet. Coal mining, particularly in the tens of thousands of small mines, violates elemental concerns over health and safety of the workers and the environment. No one in those countries seems to care about spontaneous combustion of coal-mining residues that burn on forever or drinking water and agricultural lands permanently contaminated with poisonous metal compounds.

Employment of coal miners has changed drastically in recent decades as machines have replaced labor. While there are 7 million coal miners in the world, 5 million are in China and another half million are in India, where the use of picks and shovels is the dominant coal-mining technique. Table 4.1 shows employment, productivity, and safety in terms of the number of miners per million tons of output, the number of miners' deaths, and deaths in terms of a million tons of output for 2000.[8] The table shows the enormous disparity in worker productivity and mortality rates between the developed and developing worlds. Coal mining in the United Kingdom, where it all began, is now a faint vestige of its former glory.

Needless to say, the lowest fatality rates occur in nations where there is the strongest commitment to health and safety standards for miners and for workers in general. China has the most abysmal safety record, and that may be a gross understatement. Most casualties are associated with small mines employing women and children, not the large state-owned mines. Methane explosions from lack of proper ventilation and gas monitoring are responsible for half of the deaths. These figures reflect mine mishaps, not deaths from health impairment from mining. A nonfatal occupational risk for miners and for many other industrial workers is loss of hearing. For coal miners, loss of hearing, caused by explosives used to dislodge coal and machinery noise in close quarters, occurs slowly and often without the miner's awareness. With regard to fatal occupational risks, the most common disease is pneumoconiosis, commonly known as black lung disease. Black lung disease has dropped precipitously for mines with ample ventilation to reduce coal dust, but still remains a problem in China and India and other nations where relatively little is invested in protecting the workers' health. China's terrible record in protecting miners extends to the end users. Drying chilies with coal contaminated with arsenic was responsible for thousands of cases of arsenic poisoning. Drying corn with coal contaminated with fluorine caused millions to suffer from dental and skeletal fluorosis.

Coal in the Twenty-first Century

Coal's fifteen-year retreat in relative standing among other energy sources has ended. Coal is here to stay and is gaining ground in absolute and relative terms. Despite criticisms leveled against coal, it does have virtues that cannot be ignored such as being:

- abundant, frequently reserves are measured in hundreds of years;
- secure, in that coal is available in sufficient quantities without the need for large-scale imports for most coal-consuming nations;
- safe (does not explode like natural gas);
- nonpolluting of water resources as oil spills are (although there are other adverse environmental consequences of mining and burning coal);
- clean burning, but at a cost (carbon dioxide emissions are greater than for oil and natural gas);
- cost-effective, by far the cheapest source of energy.

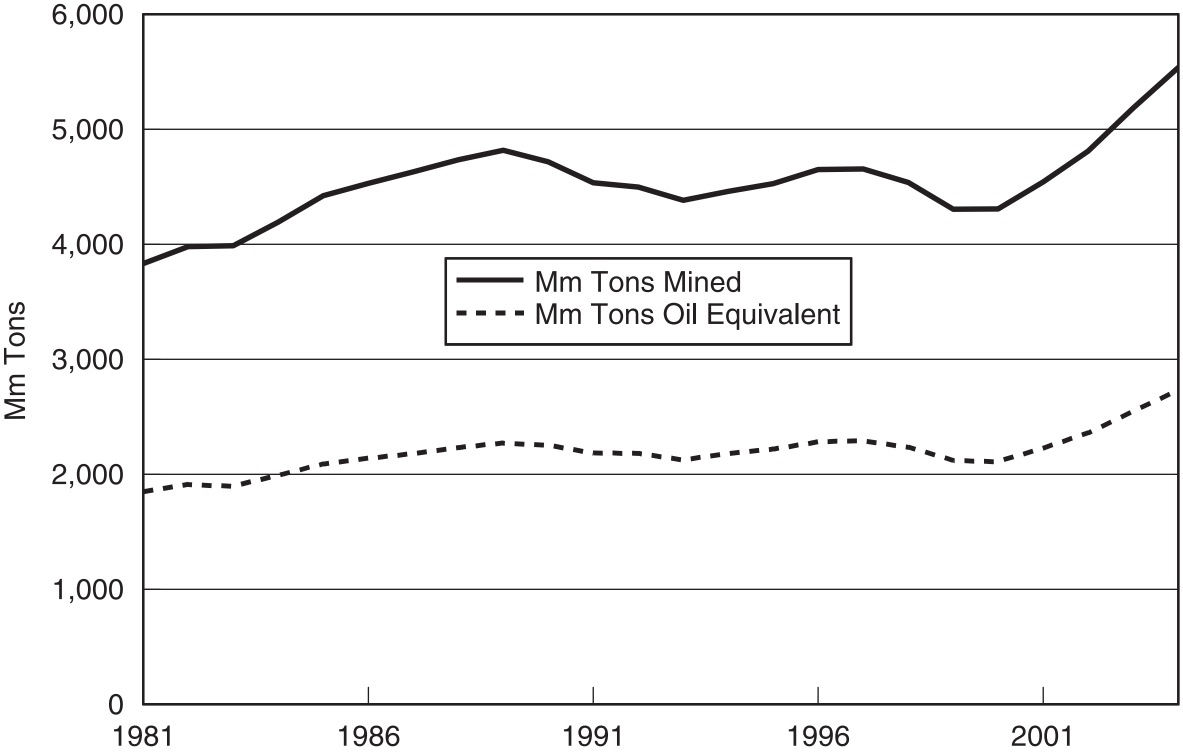
As seen in Figure 4.1, the volume of coal production leveled out in the 1990s, but is heading upward again. The top line is coal mined in physical tons and the bottom line is coal production expressed in terms of the equivalent amount of oil that would have to be burned to match the energy released by burning coal. As the figure shows, close to 2 short tons of coal have to be burned to obtain the same energy release as burning 1 metric ton of oil.[9]

Figure 4.2 shows the consumption of coal in terms of tons of oil equivalents and the relative contribution that coal makes in satisfying world energy demand for commercial sources, excluding biomass. Other than a rebound between 1973 and 1985, coal has been in a general retreat relative to other primary sources of energy, declining from 38 percent in 1965 to a low of under 24 percent in 2000, before the recent upturn. The latest shift in direction of relative standing was largely caused by China's fueling its nearly explosive growth in economic activity, followed by India and Japan. India is also undergoing rapid internal economic development, although at a slower pace than China. Japan's resurgence after nearly two decades of economic stagnation was caused by a significant rise in capital goods exports to China. This makes China the principal driver in the world coal business.

King Coal was in freefall in terms of its relative standing among other energy sources up to the second energy crisis, the oil crisis in 1973. For coal to be losing its share of the energy pie, other energy sources such as oil and natural gas must be gaining. The hike in the price of oil in 1973 that accompanied Saudi Arabia's oil embargoes against the Netherlands and the United States suddenly made the public's perception of coal more favorable as a secure source of energy not subject to the vagaries of foreign potentates. In the aftermath of the oil crisis, coal consumption increased at a faster rate than other primary energy sources, enhancing its share of the energy pie until 1985. With the shock of the oil crisis pretty much over, coal's share of the energy pie began to dwindle from increasing reliance on nuclear and hydropower and natural gas as preferred sources of energy for generating electricity.

Figure 4.3 shows the world's largest consumers and producers of coal in 2004 in terms of millions

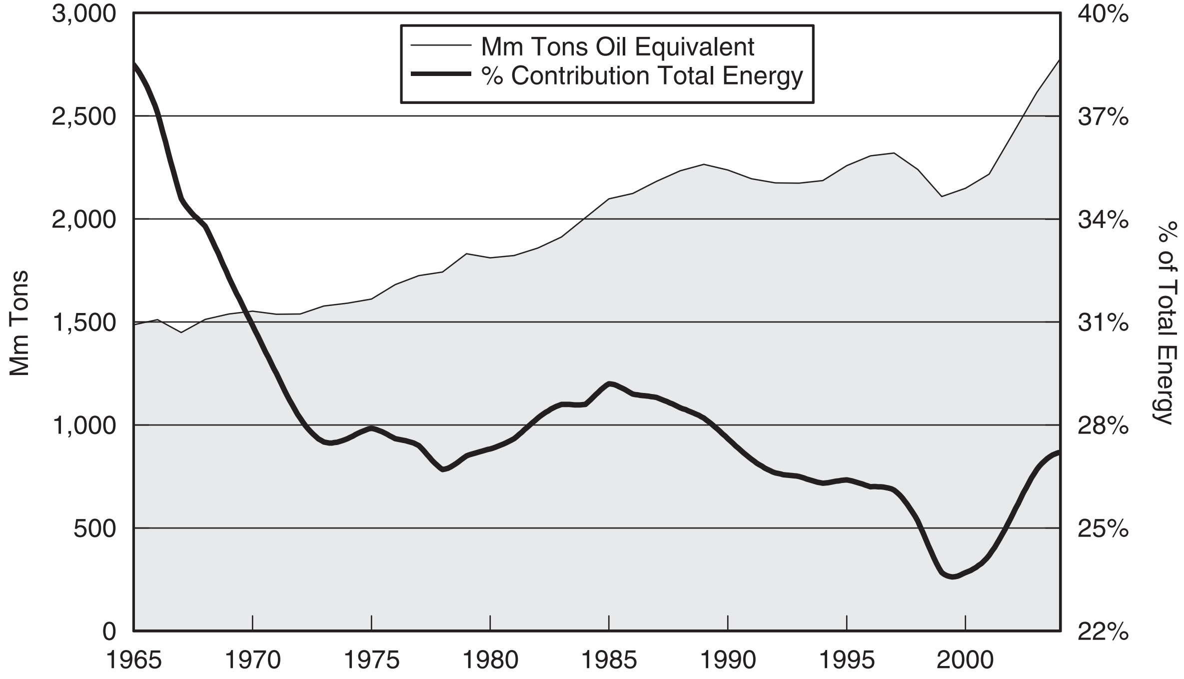
Figure 4.1 **Global Coal Production**



of tons oil equivalent. The approximate physical tons of coal can be obtained by doubling the indicated figures, but this figure has to be qualified because nations that rely on lower-grade coals would require more than a doubling, whereas nations that rely on higher-grade coals would be less than a doubling to translate tons of equivalent energy to physical tons.

China is the world's largest consumer and producer of coal, with production exceeding consumption. Even though China is a net exporter, the nation both exports and imports coal. China suffers from a poorly developed internal logistics system. Movement from inland distributions to coastline population centers relies heavily on China's river systems. Movement of goods and commodities along China's long coastline, where a number of its principal population centers are located, is by water rather than by land. As a substitute for moving commodities along its coastline, China selectively exports and imports. With regard to coal, China imports coal to utilities located on its coast from Australia and Indonesia and exports coal to neighboring countries such as North and South Korea and Japan. The steam locomotive has not entirely gone the way of dinosaurs. China, India, and South Africa still utilize steam locomotives because of their enormous

Figure 4.2 **Global Coal Consumption**

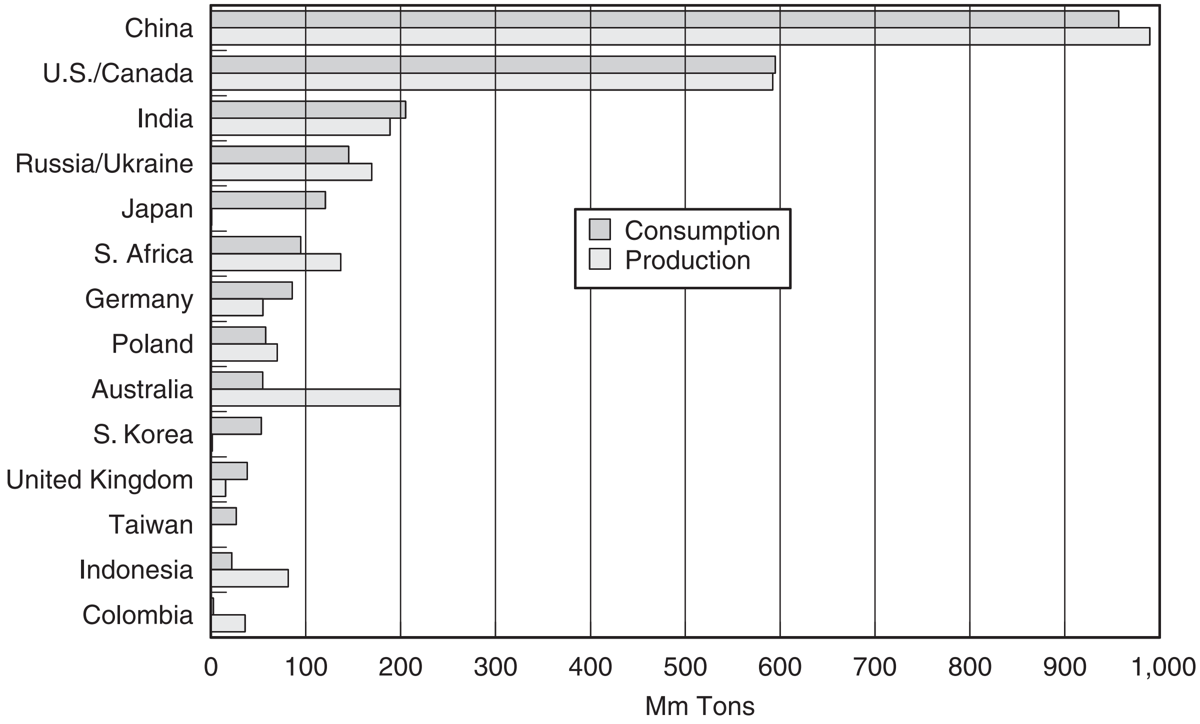


domestic supplies of coal, coupled with limited domestic supplies of oil.

The relative importance of the United States, along with Canada and China, as consumers and producers of coal can be seen by the huge step down to the third largest consumer and producer, India. All three nations are roughly in balance between consumption and production, but China and the United States are net exporters of coal and India a net importer. The largest steam and coking coal exporters in 2004 were Australia (219 million tons), Indonesia (107 million tons), China (86 million tons), the United States and Canada (70 million tons), South Africa (67 million tons), Russia (65 million tons), and Colombia (52 million tons). The largest importers were Japan (183 million tons), South Korea (79 million tons), Taiwan (60 million tons), Germany (39 million tons), and the United Kingdom (36 million tons). Japan, South Korea, and Taiwan favor coal as a means of reducing their reliance on Middle East oil.

South Africa has abundant coal resources and limited oil resources, and oil-exporting nations were reluctant to trade because of its past apartheid policies. As a consequence, South Africa became a world leader in producing petroleum products (synthetic fuels) and chemicals from coal, using the Fischer-Tropsch technology that originated in the 1920s.[10] The Germans relied on this technology to make gasoline from its plenteous supplies of coal during World War II to compensate for not having indigenous oil resources to run its war

Figure 4.3 **World's Leading Consumers and Producers of Coal**

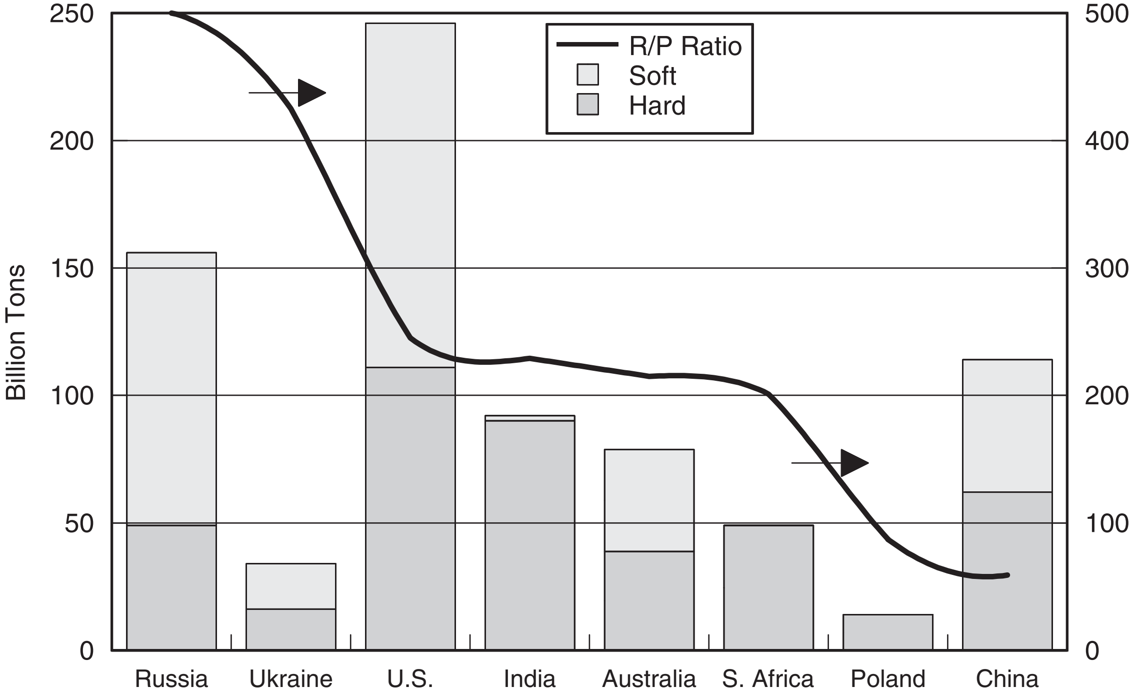


machine. This technology is very much alive. China is building a coal-liquefaction plant in Inner Mongolia that will consume 5 million tons of coal annually to produce 50,000 barrels per day of motor vehicle fuel plus other oil products.

The Fischer-Tropsch process transforms low-quality coal to high-quality liquid fuels. The coal is first gasified to yield a mixture of hydrogen and carbon monoxide, which, after passing through iron or cobalt catalysts, is transformed into methane, synthetic gasoline or diesel fuel, waxes, and alcohols, with water and carbon dioxide as by-products. Synthetic fuels from coal are higher in quality than those made from oil. For instance, diesel fuel made by the Fischer-Tropsch process has reduced nitrous oxides, hydrocarbons, and carbon monoxide emissions with little or no particulate emissions compared to oil-based diesel fuels.[11]

Unlike oil, where the world's total proven reserves divided by current consumption equal only forty years, over a century would be required for current consumption to eat away at proven coal reserves. The reserves to production (R/P) ratio has to be handled gingerly as we have a knack for discovering new reserves. (Theodore Roosevelt estimated that oil reserves would be exhausted in twenty years, given consumption and known reserves in the 1910s.) Moreover, reserves are made up of known reserves plus estimates of probable reserves, and as such are subject to error.

Figure 4.4 **Known Coal Reserves and R/P Ratios**

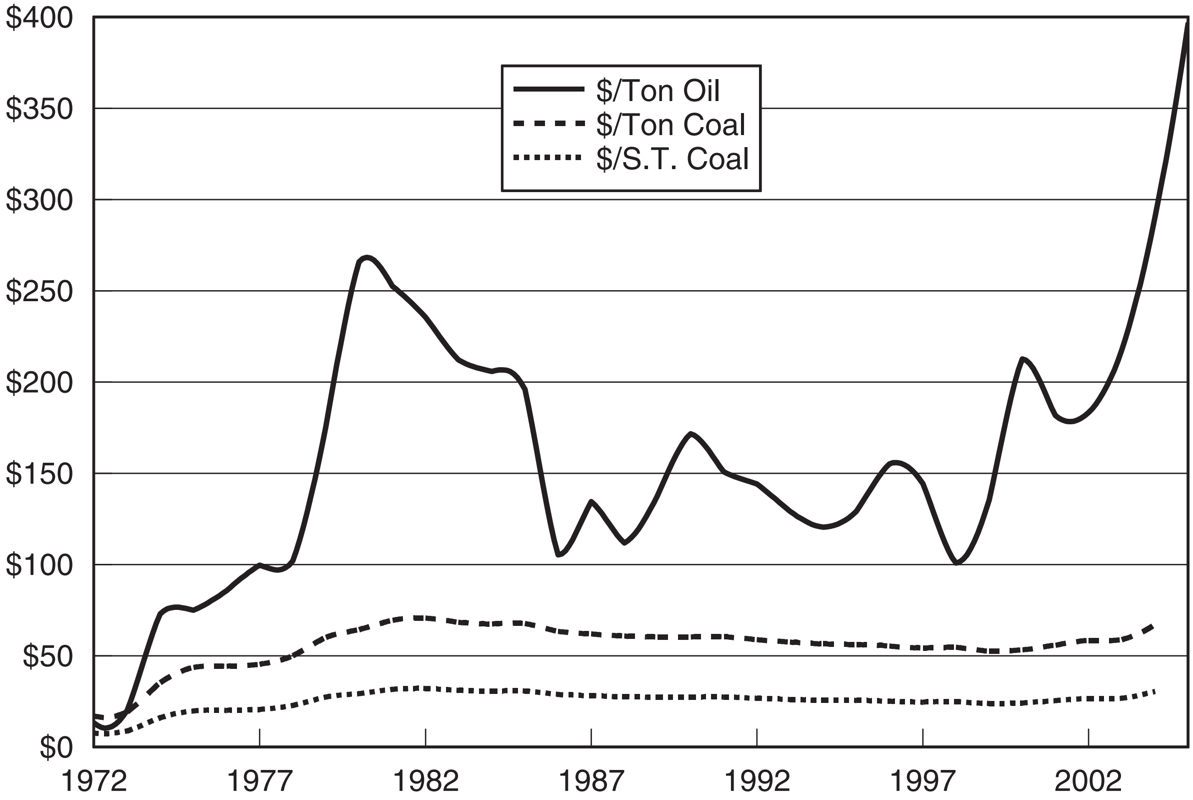


Some criticize R/P ratios because they are based on current, not future, consumption and to that extent overestimate the life of existing reserves. On the other hand, they do not take into account future discoveries and so underestimate the life of existing reserves. Figure 4.4 shows the world's largest known coal reserves in terms of size, ranked by how long they will last at the present rate of consumption.

The United States has the world's largest reserves of coal, followed by Russia and China. Of course, the nature of the reserves does not reflect the type of coal actually being mined. As already mentioned, soft coals are lignite and subbituminous and hard coals are bituminous and anthracite. Premium bituminous coal for making coking coal for steel production is found in Australia, the United States, Canada, and South Africa. Significant portions of reserves in Russia, Ukraine, and China are soft coals, generally perceived to be greater pollutants than hard coals. India has only hard coal, but of poor quality in terms of heat, ash, and sulfur content. Both China and India burn coal with virtually no environmental safeguards. Ash, the residue of burning, is released to the atmosphere in the form of airborne particulates (soot) and sulfur is released as sulfur dioxide gas.

The United States's enormous reserves of coal enhance the nation's energy self-sufficiency. Its reserves can last nearly 250 years at the present rate of production. Coal is quite unlike oil, of

Figure 4.5 **U.S. Oil and Coal Prices**

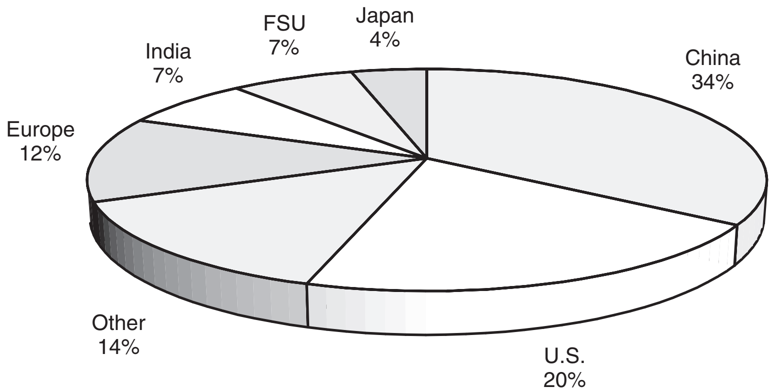


which nearly two-thirds is imported and the R/P ratio is only about eleven years. Some of the imported oil is from volatile and unstable and, at times, distinctly unfriendly nations. Coal does not demand an enormous overseas military presence to ensure security of supply. Moreover, coal has other virtues: It is cheap and its price is much more stable compared to oil as shown in Figure 4.5.[12]

A picture is worth a thousand words: Since the oil crisis of 1973, coal prices have been much lower than oil and much more stable. But a picture does not include everything. What cannot be seen is that coal is a reliable domestic source of energy not subject to the whims of potentates.

The picture for Europe would reflect higher mining costs for coal than in the United States. The picture for Japan would reflect higher shipping costs since all coal must be imported. The picture for China and India would reflect lower mining costs in terms of lack of investment in mechanization, near-slave wages for miners, with little spent for personal safeguards for their health and safety and for environmental safeguards to protect the population from pollution. This heavy

Figure 4.6 **Coal Consumption by Nation in 2002**



reliance on low-cost coal affects the competitive position of China and India in world trade since the cost of energy is an element in the price of exported goods.

The Role of Coal Among the Major Consumers

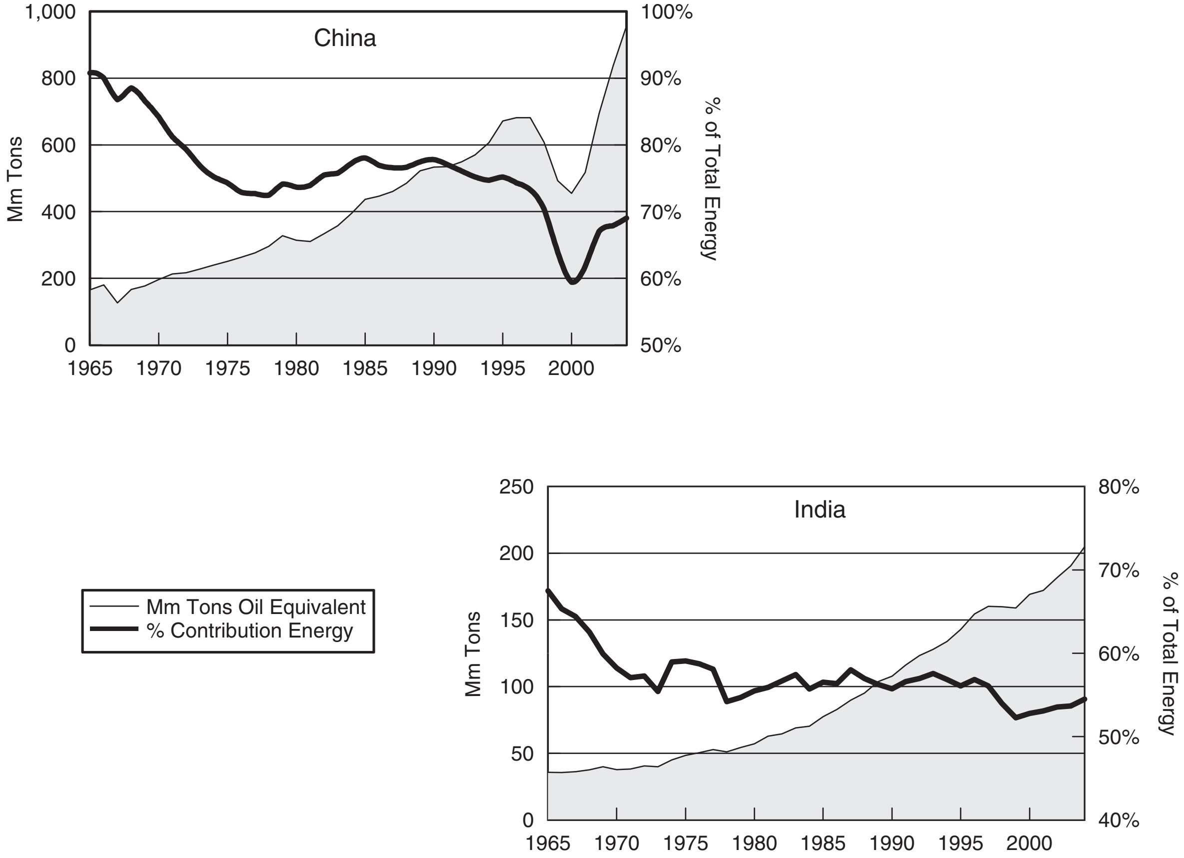
The primary use of coal is in electricity generation. Electricity and cleaner-burning heating oil and natural gas heat homes and cook food in developed nations, but coal (and biomass) are still burned for heating homes and cooking food in China and India. The five leading consumers of coal in 2002 were China, the United States, Europe, India, Russia, and the other former member nations of the Soviet Union, and Japan, as shown in Figure 4.6.

As seen in Figure 4.7, China and India rely heaviest on coal as a source of energy. The dip in Chinese coal consumption in 2000 was the result of an order emanating from Beijing to close 50,000 small and inefficient mines for safety and economic reasons. The official data released by China on coal consumption presumed that these mines were closed and no longer producing coal. However, just as King Edward I's ban on burning coal in London was not heeded on the streets of London, it turned out that orders emanating from Beijing were not carried out in the provinces. China, without much in reserves of oil and natural gas, depends on coal as an industrial and residential fuel. Without replacement energy, thousands of small inefficient mines could not be closed, although the official statistics presumed that they were, resulting in reported coal consumption taking a sharp dip.

The failure of thousands of mines to close when ordered to do so also underscores a critical problem in China; its relentless pursuit of economic development is driving energy consumption through the roof. As much as China desires to diversify its energy sources to reduce the nation's reliance on coal, it cannot cut coal consumption without suffering severe economic dislocation. As long as China's economic locomotive speeds faster and faster, coal will play an increasingly important role. China's building of hydropower dams and nuclear power plants will cut into coal consumption, but it will be years before their construction is completed.

On the surface, India is in a better position than China because it is less dependent on coal,

Figure 4.7 **Role of Coal in China and India**



although its dependence has been slowly climbing from its low point in 1999. From another perspective, India is in a worse position than China. China has an enormous trade surplus that is being used to develop alternative sources of energy to coal (natural gas, hydro and nuclear power). India suffers from a negative trade balance and is less able to finance development of alternative energy sources or the import of energy such as natural gas. Thus, greater coal consumption, and possibly greater biomass consumption, may be the primary solution to India's growing energy needs rather than importing clean energy such as natural gas; unless energy providers are willing to accept rupees rather than dollars (one liquefied natural gas import scheme calls for rupee payments). Until there is a slowing of their economic locomotives, coal consumption in India and China will continue to expand both in volume and relative share of the energy pie.

The role of coal in the United States and Europe is different than China and India as

Figure 4.8 **Role of Coal in the United States and Europe**

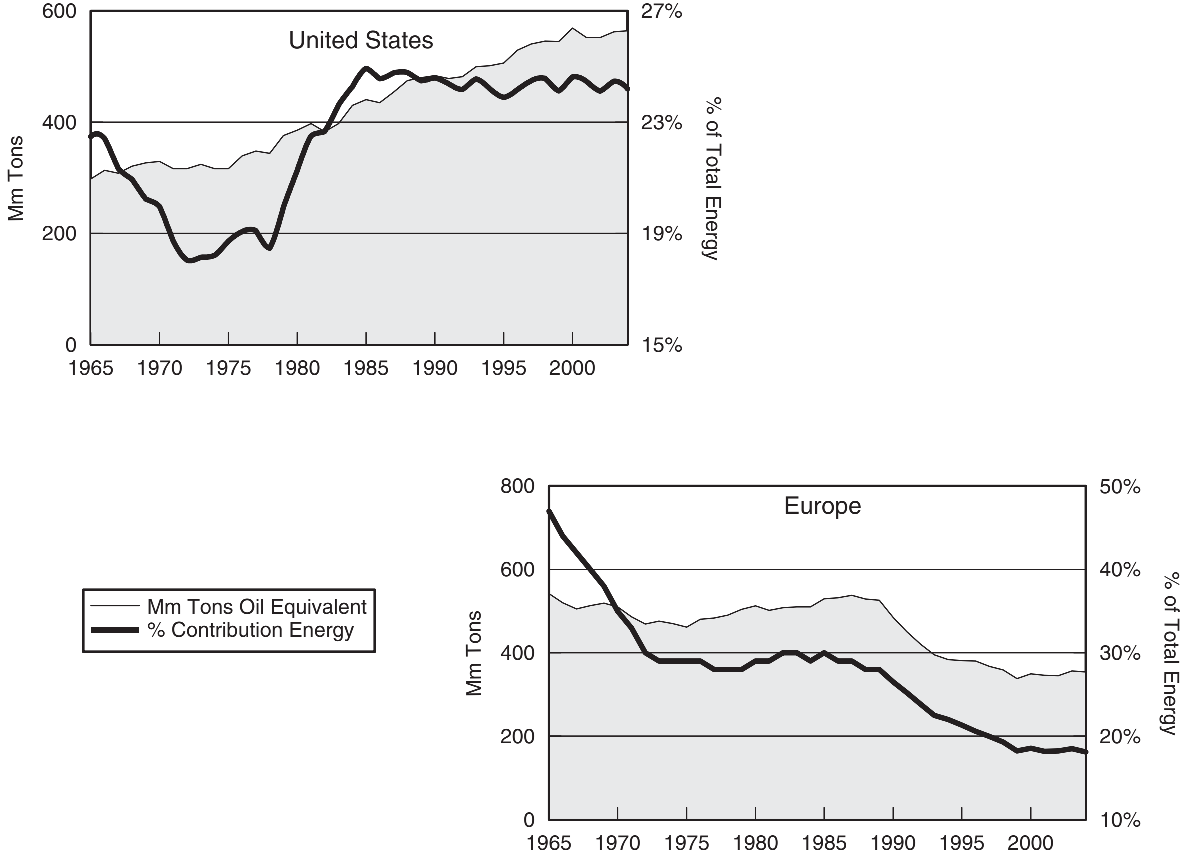


Figure 4.8 shows. Europe includes all nations west of Russia including Poland, a major coal producer.

King Coal continued to be unceremoniously dumped in the United States until the oil crisis in 1973, as seen in the fall in coal's share of the energy pie to a low point of 18 percent. Even so, coal consumption in absolute terms was still rising slowly. Consumption accelerated after the oil crisis as coal found a ready market to replace oil as a fuel for electricity generation, rising to 24 percent of the energy pie in 1985, where it has remained ever since. A slowing in the growth of electricity consumption and the collapse in oil prices in 1985 removed the financial incentive for building large coal plants. For environmental and economic reasons, there was a major shift in favor of building lower capital cost and smaller natural gas-burning electricity-generating plants, which better fit growth patterns.

Since 1985 coal's share of the energy pie has been relatively constant, yet coal consumption increased in line with total energy consumption.

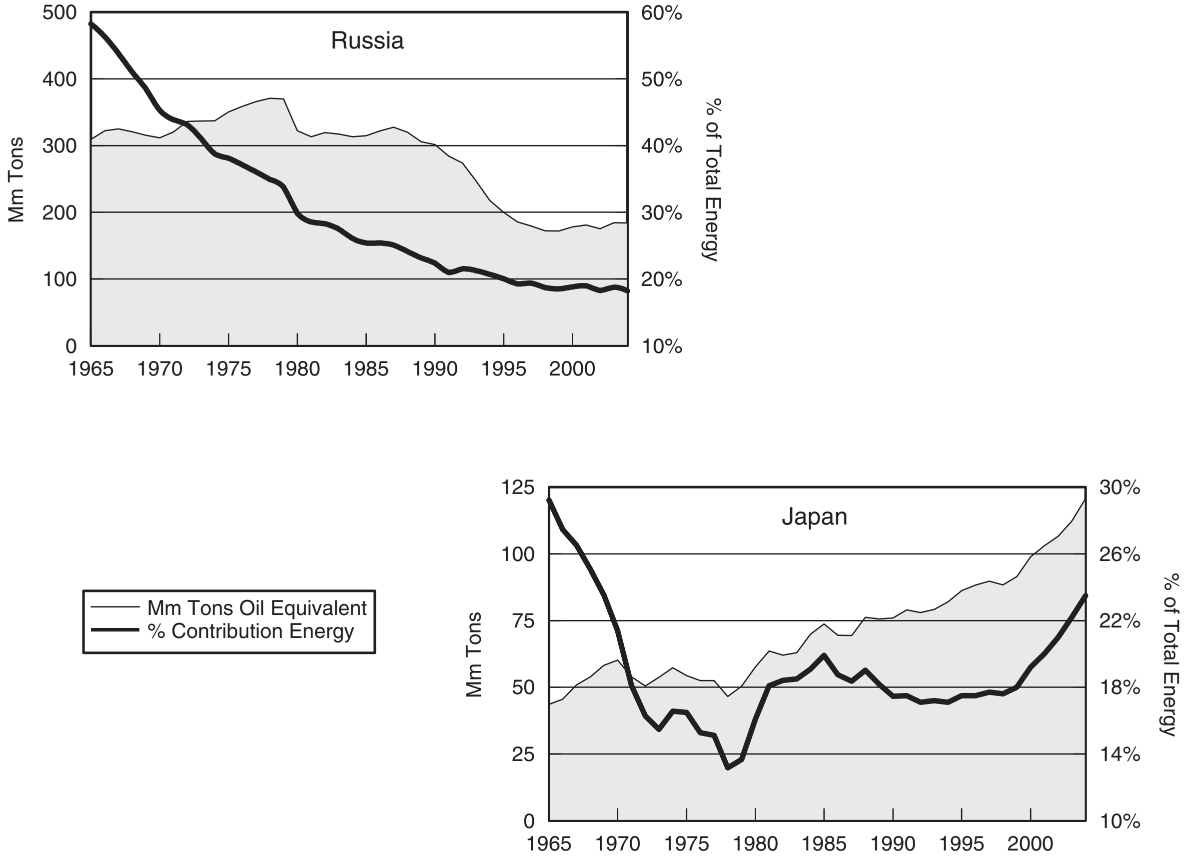
This is quite remarkable considering that nearly all new electricity-generating plants in the United States in the 1990s and early 2000s were fueled by natural gas. With virtually no coal-burning plants built during these years, the only conclusion one can reach in examining the upward trend in consumption is that existing coal plants must be operating at higher average utilization rates. This near-total reliance on natural gas during the almost twenty-year natural gas "bubble" of low natural gas prices burst in 2003 when demand finally outstripped supply. As natural gas prices rose to record levels, utilities took a second look at the idea of constructing coal-fired plants, and a number have been ordered. With the building of these plants, coal's share of the energy pie may slowly begin to increase. Despite the bad publicity coal receives in the United States, it is still viewed as a national asset, plentiful, cheap, and secure, providing half the nation's electricity. Existing coal-fired electricity-generation plants and new ones being built will keep the coal industry a viable business and ensure the employment of tens of thousands of coal miners for a long time to come.

Europe is one place in the world where coal is in retreat, both in relative and absolute terms. Coal consumption was slowly declining and its share of the energy pie was dropping fast until the oil crisis in 1973. Then coal's share leveled off as coal consumption increased to displace oil in electricity generation. Since 1985, it has been downhill for coal as it has been replaced by nuclear power and natural gas. Nuclear power has been aggressively pursued in Europe, particularly in France. A natural gas pipeline grid has been built connecting the gas fields of Russia, the Netherlands, North Sea, and Algeria (two underwater trans-Mediterranean pipelines connect Algeria with Spain and Italy) with customers throughout Europe. Nuclear power and natural gas have largely displaced coal and oil for electricity generation and as an industrial fuel. Moreover, the Europeans are intent on ensuring that the role for coal is not resurrected by relying on wind and natural gas to meet incremental electricity needs. (The interruption of Russian gas supplies in 2006 as a result of its pricing dispute with the Ukraine may temper European reliance on natural gas for electricity generation.)

Coal mining is a heavily subsidized industry in parts of Europe. Given an average import price of $40 per ton range, the average subsidy per ton of coal produced in Germany is estimated to be $144 per ton and $75 in Spain. France has an even higher subsidy rate, but its coal production is small. Subsidizing industry has been losing its allure for the last few decades. The United Kingdom has done away with coal subsidies by closing its most inefficient and heavily subsidized mines and significantly increasing the productivity of those remaining. Moreover, UK coal must compete with other forms of energy after the UK privatized its electricity-generating industry, including imported coal. European coal production is still alive for steel-making and for still-existent coal-burning electricity-generation plants. The role of coal is not dead, it simply lacks promise of growth.[13]

Before the fall of communism in 1989, coal consumption in Russia was fairly steady, although nuclear power and natural gas were eroding coal's share of the energy pie. After 1989, the reduction in coal consumption was primarily caused by the fall in electricity demand that accompanied the collapse of the Russian economy. During the 1990s, oil consumption for electricity generation was sharply curtailed to make room for exports, slowing the decline in the role of coal. Looking into the future, as the Russian economy strengthens, coal consumption may rise marginally, but the primary beneficiary for satisfying incremental demand for electricity will be natural gas. The organizational

Figure 4.9 **Role of Coal in Russia and Japan**



and financial restructuring of coal mines in Russia, the Ukraine, Kazakhstan, and Poland have resulted in the closing of the most inefficient and heavily subsidized mines and enhanced productivity of those remaining. The restructuring has basically stabilized aggregate coal production for these nations.

Japan does not look at coal as a pollutant as much as a means to diversify energy sources to reduce its reliance on oil, most of which comes from the Middle East. Consumption of coal is increasing, partly as a result of nuclear power plant shutdowns after cracks were discovered in reactor piping in 2002. In addition to thermal coal for generating electricity, Japan, as a major world steel producer, also imports coking or metallurgical coal. As in North America and Europe, coal is burned in an environmentally sound manner in Japan. The role of coal in Japan was stable at around a 17–18 percent share of the energy pie. As is clear in Figure 4.9, growth in coal consumption is accelerating and its share is now close to 24 percent. Coal is having a bit of a revival in Asia, besides China and India, as a means of energy diversification. South Korea has built coal-fired electricity-generating plants and coal-fired plants are under consideration in Thailand, Malaysia, and the Philippines.

The Case Against Coal

The case against coal can be put simply, in a word, pollution. Pollution from lower-grade coals, whether soft or hard, is greater than higher-grade coals in terms of the quantities of ash and nitrous and sulfur oxides released during combustion. Also, a greater quantity of lower-grade coals has to be burned for the same release of energy. While airborne, nitrous oxides contribute to smog and sulfur oxide droplets collect on the upper surfaces of clouds, enhancing their reflectivity. This reduces the amount of sunshine reaching the earth and, paradoxically, is a counter-pollution measure to carbon dioxide that reduces the amount of heat that can escape from the atmosphere. Eventually, sulfur and nitrous oxides return to Earth in the form of acid rain, which harms plant and marine life and erodes stone buildings and statues. Mercury, arsenic, selenium, and other heavy metals are also released when coal is burned. Surface mining destroys the landscape and, along with residues from underground mining, affects water supplies.

Abandoned coal mines can catch fire and burn underground. Once on fire, there is little that can be done to stop these fires other than entering the mine with earth-moving equipment and taking away the source of the fire: the remaining coal in the mine. In 1962, burning trash near the mouth of a mine near Centralia, Pennsylvania, started an underground inferno that has been spreading ever since despite attempts to extinguish it. The fire is burning at a depth of 300 feet beneath the surface and giving off enough heat to bake the surface, threatening to cremate bodies buried in the local cemetery. There is also venting of poisonous gases and opening up of holes large enough to swallow automobiles. It is thought that the fire will continue for another 250 years in an eight-mile area encompassing 3,700 acres before the fire runs out of fuel. Centralia has been largely abandoned except for a few diehards.[14]

Coal fires are not all the fault of men. When lightning ignites brush fires, spontaneous combustion of coal exposed to the atmosphere can start a mine fire. Burning Mountain in Australia has been burning for an estimated 6,000 years. Most of the thousands of coal mine fires that threaten towns and roads, poison the air and soil, and worsen global warming are, however, inadvertently started by man. The estimate of the amount of coal burned each year in mine fires in China varies between 20–200 million tons per year; the high-end estimate is an appreciable fraction of China's total consumption. As bad as China is, India is even worse: Rising surface temperatures and toxic by-products in the groundwater and soil turn formerly populated areas into uninhabitable wastelands.

Clean-Coal Technologies

Coal is indispensable in the generation of electricity. Nations with the greatest reliance on coal for generating electricity are Poland (95 percent), South Africa (93 percent), China (79 percent), Australia and Israel (77 percent), Kazakhstan and Morocco (70 percent), India (68 percent), Czech Republic (62 percent), Greece (61 percent), and Germany and the United States (51 percent). A great deal of corporate- and government-sponsored research is dedicated to producing a clean coal, termed an oxymoron by critics of coal. Modern coal-burning utility plants remove 99 percent of the ash produced as residue falling to the bottom of the combustion chamber or boiler and by electrostatic precipitators that remove ash from the flue gas. A flue gas desulfurization unit sprays a mixture of limestone and water into flue gas to reduce sulfur oxide emissions by 90–97 percent. Sulfur oxides chemically combine with the limestone to form calcium sulfate, or gypsum.[15] Sulfur emissions have fallen 2–3 percent per year in the United States, despite rising coal consumption, through greater use of scrubbers to remove sulfur and greater reliance on low-sulfur coal.

After mining and washing, coal is transported by train, barge, or truck and piled outside the electricity-generating plant until needed. A conveyor then moves the coal into the plant where it is first crushed and pulverized into a fine powder before being blown by powerful fans into the combustion chamber of a boiler in a conventional plant to be burned at 1300°C–1400°C, which transforms water in tubes lining the boiler to high-pressure steam that is fed to a turbine.

In addition to a conventional boiler, a fluidized bed combustion chamber can burn pulverized coal of any quality including coal with a high ash and sulfur content. The pulverized coal is burned suspended in a gas flow with heated particles of limestone at half the temperature (1500°F) of a conventional coal-fired boiler. At this lower temperature, about 90 percent of the sulfur dioxide can be removed by the limestone absorbing the sulfur dioxide to form calcium sulfate or gypsum without the use of an expensive scrubber. In a conventional plant, water tubes in the combustion chamber generate steam to drive a steam turbine. In a fluidized bed combustion plant, both steam and hot combustion gases drive two types of turbines. Steam from the boiler tubes is fed into a conventional steam turbine. Hot combustion gas, after ash and gypsum have been removed, is fed into a gas turbine. Both the steam and gas turbines power electricity generators. The spent combustion gases from the gas turbine pass through a heat exchanger to further warm condensed water from the steam condenser returning to the combustion chamber. The two advantages to a fluidized bed combustion plant are an enhanced energy efficiency of 45 percent and a reduction of about 40–75 percent in nitrous oxide emissions from the lower temperature of combustion. Fluidized bed combustion chambers normally operate at atmospheric pressure, but one currently being developed would operate at a considerably higher pressure.

The first thermal plants built around 1900 were only 5 percent energy-efficient. The current rate of U.S. efficiency averages around 35 percent, with new plants achieving up to 45 percent, depending on the type of design. The average OECD efficiency is 38 percent, but efficiency in China is only 28 percent. Increasing energy efficiency is a major action item for reducing carbon dioxide emissions because the greater the efficiency, the less coal that has to be burned to generate the same amount of electricity.

Coal gasification is a thermochemical reaction of coal, steam, and oxygen to produce a fuel gas largely made up of carbon monoxide and hydrogen. The integrated coal gasification combined cycle (IGCC) is more complicated than fluidized bed combustion, and in some ways is a step back into history. Manufactured gas, the predecessor of natural gas, was the reduction of coal to a mixture of hydrogen, carbon dioxide, carbon monoxide, and methane that was distributed by pipeline to consumers. Similarly, coal is not burned in coal gasification, but processed to produce combustible products.

The process begins with an air-separation plant that separates oxygen from nitrogen. Coal is milled and dried in preparation for being mixed with oxygen and hot water for gasification. Synthetic gases (syngas), mainly carbon monoxide and hydrogen, are then treated to remove solids (ash) and sulfur, producing a pure, salable form. Some of the nitrogen separated out by the air-separation plant is added to the clean syngas prior to burning to control nitrous oxide generation. The syngas is then burned in a combustion chamber to drive a gas turbine and, in turn, an electricity generator. In addition to burning syngas to drive a gas turbine, a steam turbine also runs off steam produced in the gasifier and in cooling the synthetic gas from the gasifier. The spent steam is partly reheated by the exhaust from the gas turbine and fed back into the steam turbine and partly condensed to water to feed the gasifier (the combined cycle part of the IGCC).

The by-products of an IGCC plant can be hydrogen for the hydrogen economy or a range of motor vehicle fuels. The advantages of IGCC are increased energy efficiency of above 50 percent, less generation of solid waste, lower emissions of sulfur, nitrous oxides, and carbon dioxide, and recovery of chemically pure sulfur. In a conventional coal plant, carbon dioxide emissions are mixed with the intake air, which is 80 percent nitrogen. Carbon dioxide emissions from an IGCC plant are pure carbon dioxide that can be sold or captured. The government-subsidized Wabash River coal gasification plant, in operation since 1971, removes 97 percent of the sulfur, 82 percent of the nitrous oxides, and 50 percent of the mercury from plant emissions. The higher thermal efficiency of an IGCC plant reduces carbon dioxide emissions for the same amount of power output produced by conventional coal-fired plants that operate at a lesser degree of thermal efficiency. These plants cost considerably more than conventional plants and represent a higher level of technological sophistication, along with a greater technical challenge in operation.

Advanced hybrid systems that combine the best of both gasification and combustion technologies are under development. Here the coal is not fully gasified, but partially gasified to run a gas turbine with the residue of gasification also burned to run a steam turbine. Again, higher energy efficiencies with even lower emissions are possible. Ultra-low emissions technology is being funded by the ten-year, $1 billion Futurgen project to build the world's first integrated sequestration and hydrogen-production research power plant. Futurgen employs coal gasification technology integrated with combined cycle electricity generation. Futurgen will be the world's first zero-emissions fossil fuel plant capable of transforming coal to electricity, hydrogen, and carbon dioxide. Hydrogen can fuel pollution-free vehicles using low-cost and abundant coal as the raw material. Electricity can be sold as well as the carbon dioxide by-product.[16]

The U.S. government is not the only party funding advanced research in clean-coal technology. ZECA Corporation, formerly the Zero Emission Coal Alliance of coal and utility companies, proposes using coal as a raw material to produce hydrogen without combustion, then converting the hydrogen to electricity using a fuel cell, and permanently disposing of carbon dioxide through mineral carbonation.[17] Magnesium carbonate is a stable depository for carbon dioxide produced by combining carbon dioxide with two plentiful iron-magnesium oxide-bearing minerals, serpentine and olivine. The challenge is getting carbon dioxide to react quickly with the magnesium oxide within the minerals. Some thought has also been given to utilizing the world's oceans, natural absorbers of carbon dioxide, to dispose of the carbon dioxide created by man. Utilizing the ocean as a waste-disposal dump in lieu of the atmosphere is sure to arouse public opposition.

As with everything else that has to do with this planet, nothing is constant. The concentration of carbon dioxide cycles over the ages peaking at 280 parts per million (ppm). Unfortunately, the start of the Industrial Revolution coincided with a cyclical peak. Since then humanity has added over 100 ppm from burning fossil fuels. The current carbon dioxide concentration of about 400 ppm has never occurred before in the known climatic record of the world, which goes back about 400,000 years; thus, there is no precedent for judging its impact.

No practical way exists to capture the 3 tons of carbon dioxide emitted by driving a thirty-mile-per-gallon automobile 10,000 miles.[18] However a stationary coal-fired power plant does lend itself to capturing and storing its carbon dioxide emissions. A typical large coal-burning power plant of 1,000 megawatts produces about 6 million tons per year of carbon dioxide, equivalent to the emissions of 2 million automobiles. There are about 1,000 of these plants in the world. Flue gas is roughly 15 percent carbon dioxide and the remainder mainly nitrogen and water vapor. Rather than passing the carbon dioxide through a smokestack for disposal in the atmosphere, flue gas passes through an absorption tower containing amines that absorb the carbon dioxide. An associated stripper tower heats the amines, releasing the carbon dioxide and regenerating the amines for another cycle through the absorption tower. The question is: What to do with the carbon dioxide from the stripper tower?

If the power plant sits on top of impermeable caprock below which is a horizontal porous sand formation filled with brine, carbon dioxide can be pumped down a vertical pipeline that reaches the porous formation and is then dispersed via horizontal pipelines running through the formation. The brine formation should be more than 800 meters beneath the surface, where the pressure is sufficient for the injected carbon dioxide to enter into a "super-critical" phase where its density is near that of the brine that it displaces. In addition to the carbon dioxide displacing brine, brine also absorbs some of the carbon dioxide. When carbon dioxide saturates an area of the formation, more horizontal pipelines are necessary to open up new areas. Huge volumes of carbon dioxide can be safely stored in this manner, but the geologic formation has to be about six times larger than a giant oil field to contain the sixty-year lifetime plant output of about 100,000 barrels per day of carbon dioxide condensed to a super-critical phase.

Carbon sequestering means that more coal has to be burned for a given level of power generation to dispose of the carbon dioxide, but it may be possible to also get rid of sulfur dioxide along with the carbon dioxide as a side benefit. The cost of carbon dioxide sequestering is equivalent to a $60 per ton surcharge on coal, roughly double the cost of coal delivered to the power plant. This will work its way through the rate structure to the electricity consumer in the form of a rate hike of about two cents per kilowatt hour, or a 20 percent surcharge for consumers paying ten cents per kilowatt hour and more for those paying less.

Carbon sequestration is not without its risks. Lake Nyos in Cameroon sits in a volcanic crater where carbon dioxide seeps into the bottom of the lake where it is held in place by the weight of the overlying water. One night in 1986 the lake overturned and released between 100,000–300,000 tons of carbon dioxide. Carbon dioxide, heavier than air, poured down two valleys, asphyxiating 1,700 individuals and thousands of cattle. Any geologic formation holding carbon dioxide must be an effective lock against escape. Carbon dioxide can also be pumped into depleted oil and natural gas fields. The carbon dioxide associated with natural gas production in certain fields in the North Sea and Algeria is separated and sequestered in nearby porous geological formations.

A payback can be generated if carbon dioxide sequestering increases fossil fuel production. Carbon dioxide pumped into methane-rich fractured coal beds displaces the methane, which can then be gathered and sold. Carbon dioxide can also be pumped into older oil reservoirs, where its interaction with residual crude oil eases its migration through the porous reservoir rock to the production wells. One coal-burning plant already pipelines its flue gas emissions over 200 miles for tertiary oil recovery.

Not all research is space age. One project is exploring the possibility of adding 10 percent biomass to existing coal-burning plants, which may reduce greenhouse-gas emissions by up to 10 percent. Ash represents a disposal problem; most ends up in landfills. Alternatively, ash from burning coal, gypsum from flue gas desulfurization units, and boiler slag can be made into "cinder" construction blocks, which consume less energy and release less pollution than making cement construction blocks. Research is also being conducted on using wastes from burning coal in road construction and in methods to reduce metals emissions, particularly mercury.

Eliminating Coal Not So Easy

Carbon dioxide is the result of a chemical reaction that occurs during burning. Switching from coal to oil or natural gas only reduces, not eliminates, carbon dioxide emissions. For the United States, further reliance on natural gas is now very costly with demand exceeding supply. Switching from coal to oil increases oil imports and U.S. dependence on Middle East oil exporters. Switching to nuclear and hydropower and renewables (wind, solar), and the hydrogen fuel economy would eliminate carbon dioxide emissions entirely, but major impediments have to be overcome. Switching from coal to nuclear power cannot occur unless public opposition to nuclear power is somehow reduced. Switching from coal to hydro is hampered by a lack of suitable sites for damming. Switching from coal to wind and solar, while possible as incremental sources of power, cannot replace coal because generation is dependent on the wind blowing and the sun shining. Switching from coal to hydrogen, while environmentally the best choice—along with solar and wind—is stymied by a less than fully developed and commercially feasible technology.

Much can be done to reduce coal-burning emissions without resorting to clean-coal technologies. Physical washing removes sulfur-iron compounds (pyretic sulfur) on the surface of raw coal, but not sulfur embedded in coal's molecular structure. While coal washing is prevalent in the United States, Europe, Japan, and other developed nations, it is not in China and India, whose high ash and sulfur content coal would benefit most from washing. While China and India are making headway in washing coal, there are capital constraints in establishing washing facilities, and possibly a shortage of available water in certain areas. A shortage of capital might apply for India, but China, with a large balance-of-payment surplus, does not lack capital. In the past, China lacked the national will to deal with pollution because capital invested in pollution controls could not be dedicated to its economic development. Having said that, there is mounting evidence that China is becoming more concerned over the environmental consequences of its economic policies and is starting to take remedial steps.

Closing small and inefficient mines, as espoused in China, can improve the environment. Fewer and larger mines ease inspection efforts by government authorities and larger coal volumes more easily justify investments to protect the health and safety of workers and minimize harm to the environment. Using coal and biomass in home cooking and heating is a major source of uncontrolled pollution in Asia. On the surface, greater amounts of coal would have to be burned to switch home cooking from coal to electricity, but burning coal in a few locations provides the means of monitoring and controlling pollution emissions.

The future of coal is certain: It already plays too significant a role in generating electricity to be dismissed out of hand. What is uncertain is what is going to be done to reduce its adverse environmental impact.

Notes

1. The history of glass information can be found online at www.texasglass.com and www.glassonline.com.

2. Barbara Freese, *Coal A Human History* (Cambridge, MA: Perseus Publishing, 2003); this well-written and highly educating book about the history and environmental impact of burning coal is worth reading.

3. The Web site www.geocities.com/Athens/Acropolis/6914/index.htm covers the development of Newcomen's and Watt's steam engines.

4. Joseph S. Spoerl's "A Brief History of Iron and Steel Production" is available online at www.anselm.edu/homepage/dbanach/h-carnegie-steel.htm.

5. These statistics, as well as statistics on imports and exports and the role of coal in electricity generation, are available from the World Coal Institute at www.worldcoal.org.

6. I remember my father shoveling coal and having to remove and dispose of the ashes from a coal furnace before converting to a heating oil furnace in a residential home on Long Island. I also remember my mother cooking on a combination wood- and coal-burning stove before switching to a propane-fueled stove in an upstate farmhouse. I am not that old!

7. This information is available from the American Coal Foundation's Web site at www.acf-coal.org.

8. The statistics in Table 4.1 are taken from *Sustainable Entrepreneurship* (December 2001), prepared by the World Coal Institute for the UN Environment Program; the figures are mostly for 2000, but a few are for 1999.

9. *BP Energy Statistics* (London: British Petroleum, 2005).

10. This information is available online at the Sasol Corporation Web site at www.sasol.com.

11. This information can be found in the *Clean Alternative Fuels: Fischer-Tropsch Fact Sheet* published by the U.S. Environmental Protection Agency.

12. Sources in Figure 4.5 for the price of coal at FOB (free on board, used to specify that product is delivered and placed on board a carrier at a specified point free of charge) mine mouth is the U.S. Department of Energy's Web site at www.eia.doe.gov/neic/historic/hcoal.htm, for bituminous coal in terms of a short ton. My source for the price of oil is *BP Energy Statistics* for $/bbl (dollar cost per barrel) FOB West Texas Intermediate; $/bbl price was multiplied by 7 in order to obtain $/metric ton (cost in dollars per metric ton). The price of coal was multiplied by 1.1 to convert from short tons to metric tons and then by 2 to convert physical tons to tons of oil equivalent to approximate the relationship between oil and coal in terms of equivalent energy released; these figures do not include shipping costs. Shipping costs for oil by pipeline, barge, and tanker are lower than rail, while barge shipping of coal is comparable to oil; rail shipments of coal are generally shorter in distance than oil.

13. See *International Energy Outlook,* published by the Energy Information Administration of the U.S. Department of Energy (2003).

14. Kevin Krajick, "Fire in the Hole," *Smithsonian Magazine* (May 2005), p. 54ff.

15. See the Web site of the World Coal Institute, London, www.worldcoal.org, which is a principal source of clean-coal technology.

16. For more information on FuturGen, see the U.S. Department of Fossil Energy Web site at www.fe.doe.gov.

17. Information on ZECA is available at their Web site at www.zeca.org.

18. This point is made by Robert H. Socolow in his article, "Can We Bury Global Warming?" *Scientific American* (July 2005), pp. 33–40.

5

The Story of Big Oil

When we think of oil, we think of gasoline and diesel fuel for motor vehicles, but the beginning of the oil industry was kerosene for illumination. Kerosene was the foundation of the Rockefeller fortune and marked the birth of Big Oil. Oil provided an alternative fuel for lighting; if oil disappeared, it would be back to whale oil, tallow, and vegetable oils. Oil was not indispensable or vital to the running of the economy then; now, however, no oil, no economy. The transition from a preferred fuel for lighting to something without which modern society cannot survive started with Henry Ford putting America on wheels in the early 1900s. The transition was complete by World War I when military vehicles, tanks, and fighter aircraft fueled by oil played a pivotal role in securing victory for the Allies. Oil had become as important as armaments and ammunition in the conduct of war. During World War II, one of the principal targets of the Allies' bombing were the coal plants that produced gasoline to fuel the Wehrmacht. As a depleting resource, oil has moved beyond supporting war efforts to being a cause of war. This chapter looks at the historical development of two of the world's largest oil companies and the role that Big Oil may play in supplying the world with energy products as we proceed beyond petroleum.

History of Lighting

Prior to 1800, only torches, lamps, and candles lit the darkness of night. Torches were oil-, pitch-, or resin-impregnated sticks. Lamps—shallow rocks, seashells, or man-made pottery containing a natural fiber wick that burned grease or oil, animal fat, or rendered fat, called tallow—first appeared during the Stone Age. Candles go back to 3000 BCE and were made of tallow until the seventeenth and eighteenth centuries when the favored material became whale sperm oil. Paraffin wax, in use today, made its debut in the nineteenth century. To varying degrees, these modes of illumination produced more smoke than light.

In the early 1800s, the best lamp fuel was whale oil, which became increasingly expensive with the decimation of the whale population. There were plenty of alternatives to whale oil such as vegetable oils (castor, rapeseed, peanut), tallow, turpentine (from pine trees), and a variety of wood and grain alcohols. The most popular lamp fuel was a blend of alcohol and turpentine called camphene. Alcohol was obtained by distillation: Alcohol vapors from a heated fermented mix of grain, vegetables, or fruits were separated, cooled, and condensed into a liquid. The distilling process for making alcohol for lamp fuel or whiskey was adopted in its entirety by the early oil refiners to separate the constituent parts of crude oil.

Another source of lighting in the 1700s and 1800s was coal gas. Gas emissions (hydrogen, carbon monoxide, carbon dioxide, and methane) produced by baking coal in a closed environment were piped to street lamps in cities in Europe and America. Lamplighters lit the street lamps in the evening and extinguished them in the morning. Coal gas was also piped into factories, buildings, and residences for illumination, but the benefits of coal gas were restricted to urban centers. This experience in piping coal gas to streetlights and buildings would be put to good use when natural gas was discovered, along with oil, during the latter part of the 1800s. Nevertheless, despite these advances in lighting, most human activities stopped at sunset.

History of Oil

Asphalt, or tar, was found on the surface in the Caspian Sea, the Middle East, Indonesia, Burma, California (the La Brea tar pit in Los Angeles is a tourist attraction), western Pennsylvania, and elsewhere. Oil was a medicine for various ailments for much of human history. Tar or pitch was mixed with clay as masonry cement in ancient Babylonia, and is still visible today. The Egyptians used tar as an adhesive in mummification. Romans burned oil as a fumigant to get rid of caterpillar infestations. Cracks between a sailing vessel's wooden planks were sealed with tar to prevent water from seeping through and sinking the vessel. Tar is said to have caulked Noah's ark and the bulrush cradle bearing Moses. Oil-soaked soil was burned as a fuel in the tenth century in the Baku region around the Caspian Sea, where Marco Polo also recorded oil seeping from the ground in the fourteenth century. Travelers in Baku in the seventeenth century recorded holes dug into the ground where oil was collected and then transported in wineskins on camels.[1]

Incendiary weapons made of naphtha also have a long history that goes back to the fourth century BCE. Of these the most famous was Greek Fire, which was mechanically projected from flamethrowers installed in the prows of Byzantine ships. Greek Fire was instrumental in turning back two invading fleets against Constantinople in 678 and 718 CE. Similar to modern napalm, it adhered to whatever it struck and could not be extinguished with water. The secret of Greek Fire, thought to be a mixture of naphtha, resins, and sulfur, was passed down from one eastern Roman emperor to the next until it was lost in only about a half century of time, perhaps as a consequence of a less than orderly transfer of power. The Arabs developed a form of Greek Fire to fight Crusader ships and the Chinese developed a similar weapon in the tenth century that was ignited with gunpowder.

In more recent times, Seneca Indians collected oil that seeped from the earth in western Pennsylvania for war paint and caulking canoes. Some of this natural seepage found its way into Oil Creek, giving it an oily sheen long before the discovery of oil. Immigrant settlers in the area dug holes that slowly filled with oil. These small quantities of seep oil, also called rock oil or its Latinized version, petroleum, were sold as medicinal cures for just about everything, first as Seneca Oil, which, when properly pronounced with the accent on the second syllable, became Snake Oil.

As in all human activities, not one, but many individuals made contributions whose aggregate impact was to launch a major industry. In the 1840s, Abraham Gesner, a medical doctor turned geologist, obtained a distilled liquid from coal that he named kerosene, from the Greek words for wax and oil. In 1850 he formed the Kerosene Gaslight Company, which lit houses and streets in Halifax, Nova Scotia. Gesner was convinced that kerosene would one day overtake whale oil if it could be cheaply made.[2] James Young, a Scotsman, patented a process in 1850 for distilling paraffin wax and oil from oil shale and bituminous coal. Paraffin wax was made into candles for the first time and paraffin oil was burned for lighting and heating. By 1862 production of paraffin wax and oil consumed about 3 million tons of oil shale and bituminous coal annually and continued for over half a century before being replaced by distilling crude oil. Distilling oil from oil shale was revived in the United Kingdom during World War II to produce petroleum products. It may resume again if crude prices are driven to a point that can economically justify processing vast deposits of oil shale found in parts of the world.

Western Pennsylvania and the Baku region were not the only areas where seep oil was "mined." During the 1850s seep oil from holes dug in the ground in Galicia and Romania was refined for its kerosene content to light lamps. Refining in the United States was more influenced by the activities of Samuel Kier, a whiskey distiller in Pittsburgh, than by Young or Gesner or the refining activities in Europe. In the 1850s Samuel Kier modified a one-barrel still for distilling seep oil in Pittsburgh, Pennsylvania. He later built a five-barrel distilling unit and bought seep oil by the gallon. The experience gained in developing these first tiny commercial refineries was crucial in the development of the American oil-refining industry.

About this time a group of promoters, headed by George Bissell, in search of something to promote, commissioned Benjamin Silliman, a chemistry professor at Yale College, to examine the commercial potential of oil. Silliman's report noted the superior properties of distilled oil to burn brighter and cleaner compared to other illuminating fuels. Bissell also had the intuitive insight to come up with the idea to drill rather than dig for oil. As with so much else, the Chinese had beaten the West by 2,500 years when they succeeded in drilling for oil using a drill bit attached to bamboo poles.[3] Bissell did not know about drilling for oil in China, nor did he know that a well was drilled, not dug, in 1846 in Baku, thirteen years earlier, nor about an oil well drilled in Canada about the time when he thought of the idea. Reinventing the wheel, so common in the past, is less likely in today's world of global communication and information systems. Based on Silliman's report and his insight to drill rather than dig, Bissell put together a group of investors who bought newly minted shares in the Pennsylvania Rock Oil Company.

The oil industry did not spring from nothing—it was an event waiting to happen. It was an accepted fact of the time that anyone who discovered an abundant and cheap source of oil would "strike it rich." In 1859, the event happened, but not before the Pennsylvania Rock Oil investors backing "Colonel" Edwin Drake, a retired railroad conductor (and never a colonel), had given up hope, one by one, on Bissell's idea to drill for oil. The last remaining investor sent a letter notifying Drake that no further funds were forthcoming and to cease operations.

Drake put Bissell's insight into action and modified a derrick device that drilled either for freshwater or for salt brine for salt manufacture to drill for oil. Drake was first to place a pipe within the drill hole to prevent the ground from closing in and plugging the hole, the forerunner of casing a well, still in practice today. He rigged up a hand-operated water pump to extract oil from the casing within the well if any should appear. As strange as this may sound, his entire approach was ridiculed as Drake's Folly. Anybody with half a brain knew that the only true and tried way of obtaining oil was by digging a hole and extracting the tiny quantities that seeped into the bottom. It seems so strange from our perspective that people who dug for or, in essence, mined liquid oil and drilled for water and salt brine could not make the mental leap to drill for oil.

Besides these technological innovations, Colonel Drake made three strategically important decisions. First, Drake employed William A. Smith (Uncle Billy), who rigged up Drake's contraption so that it would actually work; second, Drake chose to drill in soil saturated with oil; and third, Drake ignored the letter from the last financial backer to cease and desist. He borrowed money to keep the operation going. Despite the fact that he was at the point of financial exhaustion, he doggedly kept going, a story that would be repeated many times in the development of the oil industry, creating fortunes for some and financial ruin for others. For Drake, it would be a bittersweet combination of both when the "crazy Yankee struck oil." Everyone agrees that the well was sixty-nine feet deep, but there is disagreement on the output of the well, ranging from ten to twenty-five barrels per day and on the price fetched in the market, ranging from $20–$40 per barrel. Regardless of the actual flow rate and market price, overnight a new industry was born—Drake's claim to fame. For this singular achievement, Drake was to die a pauper, the first of a small, select group of individuals who would not profit from their success.

Immediately the area around Titusville became a gold rush town typical of the Wild West. A dollar invested in a producing well could yield thousands of dollars in profits. The most despicable and disreputable jostled with the honest and upright to build oil derricks almost on top of each other. The winner of this bonanza would be the individual who had the most wells pumping oil as fast and as furiously as possible. Revenue is price multiplied by volume. Since there is nothing one can do about price, the secret of producing untold wealth was to maximize production before the price fell or the oil field went dry.

Pandemonium reigned: The landscape was disfigured with fallen trees and uprooted vegetation, littered with derricks drilling for or pumping oil, construction gear and equipment tossed hither and yon, with trees, plants, soil, derricks, equipment, and drillers covered in oil. Oil was first stored in pits dug into the ground, soon replaced by wooden, and later, by metal tanks. Barrels originally intended for storing and transporting whiskey were expropriated to get the oil from the pits or tanks to a refinery. The early whiskey turned oil barrels ranged in capacity between thirty and fifty gallons and were standardized at forty-two gallons in the early 1870s. As one might expect, there were insufficient numbers of barrels to carry the oil to market. A barrel boom ensued as cooperage firms employing joiners tried to keep up with the demand. Teamsters moved the barrels of oil on horse-drawn wagons from the oil fields around Titusville to the Allegheny River for loading onto barges that were floated downstream to Pittsburgh, the world's first refining center, thanks to Samuel Kier.[4]

While joiners and teamsters prospered, drillers either made their fortunes or went broke trying. Wells drilled with wild abandon pumping full out soon flooded the market with unwanted oil. Maximizing revenue by maximizing volume works well when supply is less than demand. It is a different story when supply exceeds demand. Oil prices plunged from ten dollars to ten cents per barrel in less than a year, making a barrel more valuable than the oil within. Pumping oil continued unabatedly as prices spiraled downward because individual drillers could still maximize revenue by maximizing production as long as the price of crude oil exceeded the cost of extraction. One driller showing restraint and slowing his rate of production only meant lower revenue for him as others pumped with all their might. Drillers collectively seemed unable to sense the repercussions of what maximizing production today would do to price tomorrow; if they did, there was nothing they could do about it. The oil industry would have to wait for Rockefeller to teach the valuable lesson that practicing restraint today could maximize profits tomorrow.

As boom went bust, overnight fortunes evaporated into a spate of bankruptcies. Collapsing oil prices were not all that brought on the bad times; too many wells operating full out were sucking oil fields dry in no time. Consider the town with the quaint name of Pit Hole, about fifteen miles away from Drake's well in western Pennsylvania. Oil was discovered in what was a sleepy farming "community" of two buildings in January 1865. By September, nine months later, the population had exploded to 12,000–16,000, with fifty-seven hotels to house the flood tide of those looking for honest work along with another of rank speculators, unscrupulous stock-jobbers, reckless adventurers, and dishonest tricksters.[5] Near-valueless land a few months earlier was selling for over $1 million and interests in producing wells for hundreds of thousands of dollars. Considering the value of a dollar in 1865, these were considerable sums. The post office in Pit Hole became the third busiest in the state after Philadelphia and Pittsburgh. With so many oil wells built with such wild abandon pumping with all their might, the oil field soon went dry. Some oil drillers were bankrupt before their equipment arrived. Aided by two major fires, the city was mostly abandoned in a little over a year. Lumber from the remaining buildings was scavenged for construction elsewhere. Today Pit Hole is a ghost town.

Gesner was proven right. While oil is now generally under attack by environmentalists, it was oil or, more exactly, kerosene that saved the whales from extinction. In 1846 the whaling fleet numbered 735 vessels and was making a healthy rate of return, especially when the price for whale oil peaked in 1856 at $1.77 per gallon. By 1865 plentiful supplies of kerosene selling for fifty-nine cents a gallon sharply undercut the price of whale oil. The whaling fleet shrunk to thirty-nine vessels by 1876. The price of kerosene kept declining to a little over seven cents per gallon by 1895, when whale oil was selling for forty cents per gallon. With this price differential, there was no incentive to buy whale oil. Kerosene wiped out the whaling industry.[6]

Enter John D. Rockefeller

Building railroads made a major change in the Oil Regions, which started in Pennsylvania and later spread to Ohio, West Virginia, and Indiana. Oil could now be more cheaply transported to Cleveland by loading barrels on railcars, and later pumping oil into railroad tank cars, than loading barrels onto barges bound for Pittsburgh. The railroads made Cleveland the new world refining center, where John D. Rockefeller, a bookkeeper, happened to reside.

Rockefeller shaped the oil industry more than anybody else. As with many movers and shakers, he started life a nobody. His father, William "Big Bill" Rockefeller, was an itinerant trader taking advantage of price disparities that arose in a world of stationary buyers and sellers who did not know the price of goods over the next hill. Big Bill was a conniver and would play deaf and dumb if it suited his purpose. Considering his business practices, which were at times questionable, and his general behavior toward women, it is strange that he hated tobacco and liquor. It is stranger yet that he ended up marrying a strict Baptist, Eliza Davison, who shared his disdain for liquor but not tobacco (she smoked a corncob pipe). John Davison Rockefeller, their first child, born in 1839, would pick up their mutual aversion to liquor.

The newly wedded couple moved into Bill's cottage where his long-term housekeeper Nancy Brown also lived. Both women gave birth to their first babies about the same time. The Davison family eventually prevailed on Big Bill to send

Nancy away. As an itinerant trader, Big Bill was away from home for months at a time. His trading activities had to be fairly successful because he could support one family with his Rockefeller name at one end of his travels and another family under a pseudonym at the other, which he managed to keep secret for many years. He could also finance John D. Rockefeller's first commercial ventures. Big Bill also taught John D. valuable lessons in business such as picking him up as a toddler and then dropping him to the floor with the stern admonition never to trust anyone, not even his own father![7]

From his earliest days, buying and selling flowed through John D.'s veins. After Big Bill moved his family to Cleveland, Rockefeller enrolled in a commercial school without completing high school. Like his mother, Rockefeller was a strict Baptist and, as a fifteen-year-old, taught Bible class and sang in the choir. He would be an active churchgoer for the rest of his life. At sixteen Rockefeller was beating the pavement looking for his first job. He eventually found one at a wholesale firm dealing in everything from grain to marble. He was a meticulous bookkeeper and a persistent collector of unpaid invoices. After three months of working from six in the morning to ten at night, the firm thought well enough of Rockefeller to put him on salary.

Even at what was low pay for what he did, Rockefeller showed two seemingly contradictory character traits that would be with him throughout his life: frugality and philanthropy. He was frugal with what he spent on himself and he was frugal in the conduct of business; absolutely nothing went to waste. Yet he was generous with those in need. Rockefeller believed that his ability to make money was a gift from God that was not to be neglected without suffering God's damnation. He must have emblazoned in his mind the parable of the talents in which God severely punished the one who did not put his talent to use. Rockefeller also believed that money received was a gift from God and would eventually have to be given back to Him.

Rockefeller seemingly never had an inner personal conflict between being a model family man at all times and a model churchgoer on Sundays, including teaching Sunday School and singing in the choir, with his role as an utterly ruthless businessman for the rest of the week. His approach to business was to unmercifully crush his competition, bringing un-Christian suffering, misery, and distress to many. In his mind, he viewed his business practices as ultimately beneficial to humanity by bringing order out of disorder and eliminating the waste inherent in untrammeled free competition.

Rockefeller carried out his pledge to God that money made as a gift of God would have to be returned. By the time of his death, he had given away nearly all he had earned except for a not-so-small kitty to provide for his old age. Of course, much of what remained was given to his only son, John D., Jr., but major beneficiaries were the first college for African-American women, Spellman College in Atlanta, Georgia, (which says a lot about the man), the Rockefeller Foundation, Rockefeller University, the founding of the University of Chicago, and the building of Rockefeller Center in New York City during the Great Depression. John D., Jr., would continue returning his father's gift to God by funding the restoration of Versailles and the Rheims Cathedral, creating the Acadia and Grand Teton National Parks, donating land for the construction of the United Nations headquarters in New York, and restoring Colonial Williamsburg.[8]

After Rockefeller served his apprenticeship in a trading firm, he formed the firm Clark and Rockefeller with his friend Maurice Clark with a loan from his father. The firm successfully traded in grain and other commodities. In the early 1860s, with Cleveland hosting twenty refineries, oil began to draw Rockefeller's attention and he visited the Oil Regions. There is a photograph of the early movers and shakers of the oil industry. They stand in a group. Off to the side, distinctly separate from the others, stands a solitary figure in the middle of an empty field. It is not known whether this is Rockefeller, but it is thought that it was Rockefeller because Rockefeller stands alone.

Rockefeller was first to recognize the four principal facets of the oil business. One was production, the world of speculative drillers who, collectively, were unable to exercise self-control—a world of boom and bust, depending how supply and demand lined up. The second part of the oil industry was transportation, moving crude from the oil fields to refineries and oil products from the refineries to market. While oil transport first depended on canals and rivers, railroads had taken over much of the transport business by the time Rockefeller arrived on the scene, and railroads did not interest Rockefeller. The third part of the business was oil refining and the fourth, marketing. Refineries were relatively few compared to the number of drillers, and combining refineries under one corporate umbrella was possible, whereas combining drillers under one corporate umbrella was not. By creating a horizontal monopoly, a monopoly that controlled only refining, Rockefeller realized that he could control the entire oil industry. As the sole refiner, he would become the sole buyer of the nation's supply of crude oil and the sole seller to satisfy the nation's thirst for kerosene and lubricating oils.

With financial assistance from Big Bill, Rockefeller formed the firm Andrews, Clark, and Company in 1863 for an investment of $8,000 to get into the refining business. In 1864 he married Laura Spellman, a woman as strong in character and as firm in her religious beliefs as his mother. In 1865 he bought out Clark by carrying out what Clark thought was a bluff and renamed the firm Rockefeller and Andrews. Exercising his God-given penchant for making money, Rockefeller bought and sold oil and the profits rolled in. He brought his brother Will into the firm and opened up the firm's second refinery, the Standard Works. The word *standard* was purposely selected to evoke in the minds of customers the image of a steady and reliable source of oil products made to a consistent standard.

The refineries of the day produced only three products: lubes, or lubricating oils, for machinery; kerosene for lighting; and naphtha. Naphtha is lighter and more volatile than kerosene and could not be used in kerosene lamps without the risk of its exploding. While most refiners dumped naphtha into the nearest stream and burned the heavy end of the barrel for fuel, Rockefeller developed products for the heavy end of the barrel and burned naphtha to fuel his refineries, a sign of his frugality and aversion to waste. It is ironic that what is now considered the most valuable part of a barrel of crude oil, gasoline, which is primarily naphtha, was for four decades a waste product of the refining process. Naphtha would have its day with the coming of the automobile age.

People were awed by Rockefeller's rapid ascent to business prominence. His overwhelming impression was one of power. His blank eyes revealed nothing, yet his eyes seemed to penetrate and read the minds of others. He knew everything going on in the oil business as if God had given him special powers to see "around the corner." Seeing around the corner was a special knack that Rockefeller had for posting paid observers who reported to him all that was happening in the oil patch. Rockefeller was secretive in nature and devised a code for internal communications within Standard Oil. His contracts contained secrecy clauses that voided the contracts if their contents were revealed. With or without God's help, Rockefeller knew everything happening around him and those around him knew nothing about what Rockefeller was up to and, more importantly, his intentions.

Other than providing his family with the accoutrements of success, and giving to charities and to deserving individuals in need, Rockefeller plowed every penny the company earned back into oil. He believed and practiced frugality to an extreme. He knew that a penny saved a million times over was a lot of money that could also be plowed back into the firm. In addition to generating cash, he also knew how to tap bankers' money and borrowed heavily to finance the expansion of his business. In 1867 Rockefeller and Andrews became Rockefeller, Andrews, and Flagler, and by 1869 the three partners employed 900 workers producing 1,500 barrels per day of oil products. With 10 percent of the global refining capacity, they were the world's largest refinery operators. This implies a total global refinery capacity of 15,000 barrels per day, which is about one-tenth to one-twentieth the capacity of a typical modern refinery.

Rockefeller was a trust maker compared to Theodore Roosevelt, a trust buster. In Rockefeller's mind, a trust had certain benefits. It deals directly with the one principal fault of the free market, a lack of stability marked by boom and bust. When supply is short of demand, prices shoot up, bringing on a boom, encouraging overenthusiasm for increasing productive capacity. This lasts until there is an excess of productive capacity, which transforms a shortage into a glut, causing prices to collapse. The ensuing bust lasts until demand catches up with supply, fueling the next boom.

A trust brings stability to an industry in chaos. A trust would never overindulge in building excess productive capacity to bring on a bust because the decision to expand productive capacity is in the hands of a single individual, or a small group of individuals acting as a cartel. A trust, as the sole buyer, would be able to purchase supplies and raw materials at the lowest cost, which means lower prices for consumers. Focusing on oil, a trust would have large-capacity refineries whose inherent economies of scale would further lower costs, which could never be achieved with many independent producers, each operating a small refinery. An oil trust would set prices for its products at levels that ensured the industry's profitability. Steady profits would be able to pay for an adequate supply of oil at a fair price, which, in turn, would provide job security for workers, ensure sound bank loans, and a flow of dividends for shareholders. In essence, Rockefeller wanted to set up a system that outlawed the business cycle along with its layoffs, bankruptcies, stock-market plunges, and banking crises.

After taking over a market by wiping out the competition, Rockefeller did not take advantage of being the sole supplier and set an exorbitant price, as one might expect. Rather, he set a price where he could make a profit, but not a profit high enough to tempt new entrants into building a refinery. Rockefeller could maintain a monopoly by not being too greedy. Too high a price would only invite a new competitor to build a refinery, which Rockefeller would then have to crush by lowering prices below the competitor's costs, forcing the sale of the refinery to Rockefeller. Even so, some individuals were not above building a refinery just to force Rockefeller to buy it.

In 1870 the company was renamed the Standard Oil Company, with Rockefeller, now thirty-one years old, having the largest share (29 percent) of the company's stock. By 1879, in less than a decade, the Standard Oil Company owned

90 percent of the nation's refining capacity, having removed most, though not all, of its 250 original competitors by one of the following methods:

1. Rockefeller's God-given talent for making money when others failed.
2. Rockefeller's penchant for secrecy, preventing others from knowing what he was up to, but through "his men" knowing everything going on in the industry. For instance, railroads had to tell Rockefeller the details of shipments by his competitors including the volume, destination, and shipping rate. Corporate intelligence was a major weapon in Rockefeller's business arsenal for vanquishing his foes.
3. Rockefeller's realization of the inherent economies of scale of large refineries before anyone else. Rockefeller had the best-operated, most efficient, and the largest refineries, making him the low-cost producer. He concentrated his refining at three plants, which at one time represented 75 percent of global refinery capacity. His refining costs were half those of his competitors. Being the low-cost producer was a major card to hold in the corporate game of King of the Hill since Rockefeller could lower his price to a point where his competitors were losing their shirts while he was not. Rockefeller was not above buying a refinery from an independent and closing it, then adding capacity to one of his refineries to replace the scrapped capacity, benefiting from further gains in economies of scale.
4. Rockefeller, as the largest refinery operator, was able, through the efforts of Henry Flagler, to get the railroads to offer a secret rebate for Rockefeller's business.[9] Railroads, as common carriers, were at least morally, though not legally, bound to charge the same rates for everyone. With the cost of shipping crude oil about the same as its value, shipping was an important component in determining profitability. As the industry's largest shipper and also the owner of a fleet of railroad tank cars, Rockefeller took advantage of the intense competition among three railroads to negotiate a secret rebate. This rebate cut Rockefeller's shipping cost by nearly half. Then, on top of this, he negotiated a drawback, a kickback, of the shipping rates charged to his competitors. Rockefeller's competitors had to pay not only twice the shipping rate he did, but, to add insult to injury, part of what they paid to the railroads also flowed to Rockefeller through the innocuous-sounding South Improvement Company.
5. Rockefeller was the first to sell his products in Europe and Asia. From the beginning, a large portion of U.S. kerosene production was exported and Rockefeller made Standard Oil the first multinational oil company and the United States the world's largest oil exporter. With widely dispersed markets throughout the United States and the world, Standard Oil was the only company so positioned that profits in one area subsidized losses in another. This was a great advantage for Rockefeller when it came time to give a competitor a "good sweating." Rockefeller could bring any competitor to heel, domestically or internationally, through discriminatory price-cutting without suffering an overall loss.
6. Rockefeller was not above sabotaging a competitor's refinery if that would bring the competitor under his control quicker. He practiced corporate espionage by paying employees of competitors to spy for him.

He was also a master at corporate deceit. One time Rockefeller purchased a refinery on the condition that the seller would not reveal the purchase. The ex-owner continued to operate the refinery as an "independent" and combined with other independent refinery operators in order to better combat Rockefeller's ruthless takeover of the refining business. The sellers learned too late that they were now within the firm grip of Rockefeller's octopus.

7. Rockefeller knew how to handle bankers and was always able to cajole, when he could not convince, bankers to finance his acquisitions. The bankers were willing lenders because Rockefeller never defaulted on one penny of his borrowings, valuable business advice from his father.

Rockefeller achieved his high-water mark of over 90 percent control of the refining industry in the late 1870s. The lines on his face began to reveal the never-ending stress of working strenuously by day and worrying mightily by night. Even though Rockefeller seemingly held all the cards, it was not a simple matter for him to achieve his objective of total control over the refining industry. The American independents were just as determined to escape from Rockefeller's grasp, survive, and come back to fight again as Rockefeller was to subdue them. The American independents were absolutely determined and dedicated to not ending up as Rockefeller's property just as, a few years later, the Russian independents would be equally determined and dedicated to not ending up as Nobel and Rothschild property.

To combat Rockefeller's control over the railroads and his favorable shipping rates, the independents started building a pipeline to connect the Oil Regions with the east coast market. Rockefeller put every legal impediment in their way that his lawyers could devise. He bought land through which the pipeline would pass with the intent to deny permission for its construction. The independents, utterly determined to defeat Rockefeller, would change the pipeline path around Rockefeller's land. Then Rockefeller convinced the railroads not to sell the right-of-way for the pipeline to cross their tracks. Unable to cross a railroad track, the pipeline ended on one side of the track and started on the other side. Oil from the pipeline had to be loaded on wagons to cross the railroad tracks and then be put back into the pipeline. When this failed to stop the flow of oil, Rockefeller had the railroads park a train across track crossings to disrupt transfer operations.

Despite this towering wall of opposition, the independents managed to complete the pipeline. Pumping oil through a pipeline is far cheaper than shipping by railroad. The completion of the pipeline meant that Rockefeller had not only lost his strategic advantage over the independents, but that he now suffered from a strategic disadvantage. Once the independents could reach the east coast market cheaper than Rockefeller, Rockefeller did an about-face and became a pipeline builder. He eventually built 13,000 miles of pipelines connecting the Oil Regions with the east coast markets and took over nearly 90 percent of pipeline traffic, amply demonstrating what it was like to cross his path.

Rockefeller also manufactured kerosene lamps and sold them at cost to induce people to switch to kerosene. He pioneered in making lamps and stoves safer to lower the death rate of several thousand per year from kerosene fires and explosions. The hazardous nature of kerosene increased when unscrupulous refiners spiked their kerosene with more volatile naphtha rather than throw it away. Consumers could count on Rockefeller's "standard" kerosene product to be free of such dangerous adulteration. To expand his market beyond kerosene, Rockefeller spearheaded the development of other oil products including asphalt for road construction, special lubricants for railroad locomotives, and ingredients for paint, paint remover, and chewing gum. He made sure that Standard Oil stayed with the business it knew best: oil. Having established a horizontal monopoly in refining that stabilized the price of oil he then strove for a vertical monopoly by acquiring oil-producing properties. By 1879, Standard Oil's oil fields from Pennsylvania to Indiana pumped one-third of the nation's oil. This was also the year that Thomas Edison invented the electric light bulb, the start of a slow death for the kerosene lamp.

In the meantime, Rockefeller discovered that natural gas was frequently found along with oil and was flared or vented to the atmosphere. Because of his aversion to waste, Rockefeller started plowing his oil profits into developing a natural gas industry. Natural gas could fuel streetlights, buildings, and factories as a substitute for coal gas if there were a means to get natural gas from the oil wells to towns and cities already served with coal-gas pipelines. Standard Oil was active in building natural gas pipelines and obtaining municipal franchises to supply communities with natural gas, which was cleaner burning, cheaper, and had a higher heat content than coal gas. Sometimes he had to pay a bribe to get a municipal franchise and sometimes he resorted to corporate trickery. One municipality decided to split a franchise between two independent firms so that consumers could benefit from competition. While the two companies that won the split franchise seemed to be rivals, in reality both were subsidiaries of Standard Oil.

Is there anything that can be said in favor of the way Rockefeller conducted business? Actually there was one: When a competitor was crushed and had no choice but to sell to Standard Oil, Rockefeller would offer either cash or shares in Standard Oil, recommending the latter. Frequently sellers, after being beaten by Rockefeller into abject submission, took the cash just to avoid further entanglement with him. This was indeed unfortunate because the value of the stock would, in time, vastly exceed the value of cash.

Ida Tarbell, a journalist-author of a series of magazine articles starting in 1902 in *McClure's Magazine* entitled "The History of the Standard Oil Company," exposed the company's nefarious business practices. These articles turned public opinion against Rockefeller and fueled Theodore Roosevelt's aversion to monopolies. Ida was a perfect person to write such a series of articles. Her father was a joiner, or barrel maker, who profited in the early days of oil by being the first to make wooden tanks for storing oil, rather than pits dug in the ground. He built a house for his family by scavenging lumber from an abandoned hotel in Pit Hole. His days of prosperity ended abruptly with the advent of metal tanks. Throughout his life he was a strong advocate of American independents. He was allied with one that was eventually crushed by Standard Oil, a fate shared by Rockefeller's brother Frank.

By 1882 Standard Oil, a conglomerate of subsidiaries created by Rockefeller's numerous acquisitions, was becoming difficult to control. Rockefeller reorganized the company as the Standard Oil Trust, whereby control of forty-one companies was vested in nine trustees including Rockefeller, who operated out of Standard Oil's New York City office at 26 Broadway. As the years went by, Rockefeller controlled less of the company's operations and spent more time grooming his successors plus time in court fending off victims seeking restitution and at hearings fending off government inquiries into his business practices. Rockefeller was moving into the public spotlight and the public did not like what they saw. Rockefeller's business practices did not fit the picture of America as a land of opportunity for pioneers and family owned businesses. Forcing competitors to sell against their wishes, whether or not the price was fair, was not considered a fair business practice.

Rockefeller's business practices, while not technically illegal at the time, inspired legislation that made them illegal. The Interstate Commerce Act of 1887 required railroads, as common carriers, to charge the same rates for all customers and outlawed secret rebates and kickbacks and established the first federal regulatory watchdog agency, the Interstate Commerce Commission. In 1890 Congress passed the Sherman Antitrust Act, which banned trusts and combinations that restrained trade and sought to control pricing through conspiratorial means. In 1892 the Ohio Supreme Court ordered the local Standard Oil company to leave the Standard Oil Trust, but Rockefeller instead dissolved the Trust and set up a new corporate holding company, Standard Oil of New Jersey (New Jersey was selected for its lax corporate laws). Standard Oil Trust as a legal entity lasted only ten years, but its name would last forever. Independently of Standard Oil, Rockefeller also got involved in investments in mining iron and copper ores and banking, which, in the Rockefeller tradition, all made money. The banking investment turned out to be a predecessor bank to Chase Manhattan, which was eventually run by his grandson, David Rockefeller.

With all these successful achievements in business, Rockefeller had one more favor from God awaiting him: Theodore Roosevelt. Rockefeller the trust maker fought Roosevelt the trust buster for years before Roosevelt won in 1911 with the Supreme Court decision that forced Standard Oil to dissolve itself into thirty-four separate and distinct companies. Rockefeller, rather than holding shares in Standard Oil, now held the equivalent number of shares in thirty-four companies including what would become Exxon (Standard Oil of New Jersey), Mobil (Standard Oil of New York), Amoco (Standard Oil of Indiana), Sohio (Standard Oil of Ohio), Chevron (Standard Oil of California), ARCO (Atlantic Refining), Conoco (Continental Oil), Marathon (Ohio Oil), Pennzoil (South Penn Oil), and twenty-five others.

Before the divestiture, and with Rockefeller exerting less control, Standard Oil was becoming bureaucratic and lethargic, the twin banes of large successful organizations. After the breakup and an initial period of cooperation among the sister companies, each went their separate ways, opening up and exploiting new markets that Rockefeller had not envisioned. The net impact of splitting up Standard Oil was to invigorate the company with a host of new managements and multiply the stock value of Rockefeller's original holdings in Standard Oil many times over. Rockefeller, fully retired after the Standard Oil breakup, became far wealthier in retirement than when he was actively engaged in business. Rockefeller died in 1937 at age ninety-eight, two years shy of his goal and, by all accounts, well-pleased with the course of his life. He had given away all but $26 million of his money, although a nice chunk of change was in his son's hands. Whether God was pleased with him is unknown.

Enter Marcus Samuel

The story of the transition of a small trading house in seashells to a major oil company known as Shell Oil serves as a counterpoint to the story of Standard Oil. It has more twists and turns and impinges more on the affairs of other oil companies, which, one day, would become part of Big Oil. The story begins with Marcus Samuel, the father of the two brothers who would found Shell. The elder Marcus, a British Jew, purchased seashells and other objects from sailors who frequented the London waterfront. The shells were cleaned, polished, and, attached to shell boxes, sold in seaside towns and curio shops. By the 1860s the elder Marcus began to branch out into general merchandise purchased as it landed on the dockside in London. The elder Marcus saw the end of an era of shipping when a vessel left London with goods without any clear idea of what the goods would be sold for until the vessel arrived in Asia, and whose proceeds would purchase Asian goods whose value was unknown until the vessel docked in London. Trading was a real gamble in terms of the commercial risk: buying goods with no idea of what they would fetch if they survived the hazards of being carried aboard a vessel at sea. The opening of the Suez Canal, which shortened the voyage time between London and Asia, and the start of a regular mail service, which allowed buyers and sellers to communicate with each other, reduced the extent of operating in the dark although traders still had to contend with price changes during the weeks or months between buying and selling goods.

The elder Marcus's volume of business began to blossom as the British Empire expanded, first into India and then to British enclaves in Singapore, Hong Kong, Shanghai, and other Asian ports such as Bangkok, and finally the opening up of trade with Japan. Rather than buy and resell goods as they arrived in London, Marcus started working through agents in Bangkok, Singapore, and elsewhere to secure imports paid for by exports of British manufactured goods. The elder Marcus set up a trading house and conducted business through letters that took weeks to exchange, never visiting his agents. The agents learned to trust Marcus because he kept his word even if market conditions changed. This was a bit unusual in a world where reneging on deals was fairly common, particularly during times of financial distress when banks closed and trading houses collapsed. In 1870 the elder Marcus died, and the eldest son Joseph took charge of the family business, while the two younger sons, Marcus and Samuel, inherited only their father's reputation for sticking by his word.[10]

After spending some time at the family business, the younger Marcus—at twenty years of age—set out on his first voyage to Asia in 1873. Marcus discovered a famine while visiting his father's agent in India and surplus rice while visiting his father's agent in Bangkok. Marcus put together his first international deal with rice merchants and ship owners to relieve the famine in India, a deal that was both humanitarian and profitable. He returned home in 1874, shortly before his mother's death and, on his second voyage in 1877, made the acquaintance of the great trading families in Asia. At that time trade was either between Asia and Europe via the newly opened Suez Canal or within the borders of a nation. Marcus sold goods he acquired—not in England, as was expected—but to other Asian nations, "at the least possible distance." Strange as this must sound from a modern perspective, Marcus is credited with the start of intraregional trade among Asian nations as opposed to trade being confined within a nation's borders or with England.

Marcus reached Japan just as it was opening its borders to trade, and established an office to import English textile machines in exchange for Japanese wares such as rare seashells, china, and silk. As the years progressed, the two brothers, operating from their office in London, built up a substantial trading house working through trusted employees and third-party agents in Asia. By the 1880s they owned the largest foreign concern in Japan and were involved with all types of cargo including Japanese coal exports for fueling steamships and kerosene imports in tin containers, called case-oil, from the Black Sea port of Batum.

At that time Standard Oil was a leading force in the case-oil market, but it was not alone. The Russian czar permitted the development of Caucasus oil in 1873 by awarding a concession to the Nobel brothers, Robert and Ludwig; a third brother, Alfred, was the inventor of dynamite and originator of the Nobel Prizes. The two Nobel brothers developed the oil resources of the landlocked Caspian Sea, located in the Baku region of modern Azerbaijan. As in Titusville, oil seeped to the surface and was "mined" for centuries before the two Nobel brothers began drilling for it. Although we tend to think of the oil industry as strictly American, the Nobel brothers made several important contributions to drilling and refining oil and in shipping oil by pipeline and tanker. The Nobels led the effort to make Baku a major world supplier with Caspian oil, which at the beginning of the twentieth century accounted for over half of the world's supply of oil (11.5 million tons versus U.S. production of 9.1 million tons).

To get kerosene to Europe, the case-oil was shipped in barges from a Caspian Sea refinery through the Volga River and canal system, then transferred to the Russian railroad for transport to a Baltic Sea port, and then by water to Europe. The Nobels had high shipping costs and, once their case-oil arrived in Europe, they faced Standard Oil. Rockefeller moved into Europe early on, first moving kerosene in barrels to Europe on general cargo vessels and later in bulk on the world's first tankers. These early tankers proved to be dangerous. Fires and explosions often cut their lives short, a weak point that Marcus would eventually exploit.

The Nobels had learned a valuable lesson from Rockefeller's successful control of the railroads, which assured him a monopoly over American oil. The Nobels' version was to gain virtual control over water transportation up the Volga River. To beat this monopoly, the independent Russian producers started to build a railroad from Baku to Batum, on the Black Sea. If completed, Caucasus oil would be shipped by rail from Baku to Batum and then by tanker through the Black and Mediterranean seas to Europe. The oil would arrive in Europe cheaper than the Nobels transporting it to Europe via the Volga River and the Russian railway system to a Baltic port. This would place the Russian independents at a competitive advantage with the Nobel brothers in Europe. The Nobels, just as ruthless as Rockefeller, lowered the price of Russian oil and starved the Russian independents of the funds necessary to complete the railroad. Confident that they had crushed the Russian independents, the Nobels had inadvertently opened up the opportunity for the Paris Rothschilds, who entered the oil game by financing the completion of the railroad. The Rothschilds extracted an exclusive purchasing arrangement from the Russian independents as a price for financing the railroad. With a secure source of oil, the Rothschilds built a refinery at Batum and began to market kerosene in Europe in competition with the Nobels and Rockefeller.

The Nobels were in deep trouble because transporting oil by canal and railroad to the Baltic via the Volga River was more expensive than by rail to the Black Sea and then by tanker to Europe. Like Rockefeller, the Nobels were not easily beaten. They built a pipeline from the Caspian Sea to the Black Sea, using their brother's dynamite to clear the way. Now it was the Rothschilds' and the Russian independents' turn to "sweat" as it was cheaper to pipeline oil to the Black Sea than to transport it via railroad. Having lost their strategic advantage, the Rothschilds were in a weak bargaining position, locked in third place after Standard Oil and the Nobels in the race to supply kerosene to Europe.

In 1885 a London ship broker, Fred Lane, "Shady" Lane to his critics, was the London representative of the Paris Rothschilds. Lane approached Marcus with the idea of selling Rothschilds' oil in Asia. The Rothschilds were eager to diversify their market to counter their competitive disadvantage in Europe. But no matter where the Rothschilds attempted to sell kerosene, Standard Oil would step in, lower the price, and chase them away. Another approach was needed to establish the Rothschilds in Asia, and over the following years Lane and Marcus hatched a strategy to beat Standard Oil at its game.

First, the relatively expensive transportation of case-oil, including the cost of tin containers, would be replaced by bulk transport in newly built tankers from Batum to Asia via the Suez Canal. Storage depots would be built in the principal ports in Asia to receive the bulk oil shipments. The storage depots, where possible, would be connected to railroads or roads for bulk transport in railroad tank cars or horse-drawn wagons to inland destinations. To assure the success of the venture, the Rothschilds entered into a low-priced long-term supply contract for kerosene. As attractive as this sounded, it had one serious drawback. Bulk shipments of kerosene in tankers were not allowed to transit the Suez Canal because of their poor safety record. If Marcus could build tankers to a higher standard of safety and receive permission to transit the Suez Canal, then the Rothschilds would have a strategic advantage over Standard Oil.

The project faced enormous obstacles. The first obstacle was financing the tankers. Marcus became an alderman of the city of London, which, in addition to his being a successful businessman, would aid in garnering the necessary financing for the tankers, whose ultimate use was to be kept a secret from those providing the financing. The Rothschilds could not put up the financing as that would compromise the secrecy of the project. The second obstacle was that the Rothschilds had a hidden agenda: They intended to use the contract with Marcus as a means for putting together a more attractive deal to amalgamate their interests with Standard Oil. This made the Rothschilds an unreliable partner, although Marcus did not know it. The third obstacle was the Suez Canal Authority, who had no idea what tanker standards should be imposed to permit safe transits. Marcus was building tankers whose standards might or might not satisfy the Suez Canal Authority. The fourth obstacle was building storage terminals in Asia, for which Marcus had no experience, just as he had no experience with building tankers. The fifth obstacle was that Marcus, while a successful trader, had no background either in oil or in leading such a Herculean business enterprise, although he must have had the Rothschilds' confidence that he could successfully take on Standard Oil. The sixth obstacle was keeping Standard Oil from learning the entirety of the plan, in which case the project would face its full fury. The seventh obstacle was the two brothers themselves, who continually bickered with one another because they had different personalities, different approaches to business, and, most importantly, different perceptions of risk. The eighth obstacle was that the financial stake was of such a magnitude that, if it failed, Marcus would be disgraced. The ninth obstacle was that Marcus preferred to act through inexperienced blood relatives, two nephews in particular, rather than through those with experience, although operating through his nephews might have been necessary in Marcus's mind to preserve secrecy.

The tankers under construction for Marcus incorporated the lessons learned from fires and explosions on existing tankers. Kerosene would not be carried in the bow section of the ship, which would protect the cargo in case of a collision. Tanks were added to contain the thermal expansion of the cargo when the vessels passed through warm tropical waters. The individual cargo tanks were of limited capacity and airtight to enhance safety and would be thoroughly cleaned after discharging their cargo to prevent evaporating residues from forming an explosive gas mixture. The tankers would be registered with Lloyds Register's highest classification rating.

Two young nephews of the Samuel brothers were put in charge of building storage facilities in Asia, but they had no experience in acquiring property rights and building storage tanks. Port authorities opposed bulk storage facilities for oil products because they were considered potentially unsafe. Local business interests were against constructing storage tanks since change of a nature they did not understand could best be addressed by resisting it. The nephews were bombarded with micromanagement cables from their uncles that ran from close scrutiny of their expense accounts to attending to other aspects of the firm's trading activities. Their uncles' advice on building storage facilities was anything but helpful.

Owners of existing tankers who had been denied permission to pass through the Suez Canal were not keen to see a new class of tankers built that could. This would make their vessels obsolete, at least from the point of view of trading between Europe and Asia. Members of the Russian imperial family, who were large shareholders of a Black Sea fleet of tankers, were in a position to have the Russian government petition the Suez Canal Authority to deny permission to Marcus's new tankers. Other petitioners included tanker owners and tin-plate manufacturers of cases for holding oil, whose business would be threatened by bulk shipments of kerosene, plus a host of companies, many of which were not engaged in the case-oil trade or shipping. Standard Oil's name did not appear among those opposing Marcus's application. It would have been utterly out of character for Standard Oil to be absent from such proceedings, but for whatever reason Standard Oil preferred to pull the legal strings through other parties and remain hidden behind lawyer-client privilege.

In the end the Suez Canal Authority concluded that tanker transits would add to canal revenue and accepted Lloyd's highest classification rating as adequate criteria for safe passage. Despite all odds, including near-continuous interference from their uncles, the two young nephews succeeded in having storage tanks built in Bangkok and Singapore and were making progress in building tanks in Hong Kong and Kobe when, in 1892, the first tanker, the *Murex,* named after a seashell, passed through the Suez Canal with a cargo of Rothschild kerosene. The vessel unloaded its 4,000 tons of cargo at Bangkok and Singapore (actually at Freshwater Island, outside the jurisdiction of the Singapore port authority, which had denied permission to build an oil storage facility within Singapore). Ten more vessels were launched in 1893, creating a fleet of eleven vessels, all named after seashells as a tribute to the elder Marcus. By the end of 1895, sixty-nine Suez Canal tanker transits were made, of which all but four were tankers either owned or chartered by the Samuels. In 1906, Marcus shipped 90 percent of the 2 million tons of oil that passed through the Suez Canal. Marcus and the Rothschilds had beaten Standard Oil at its own game, a singular achievement, which by any measure must rank as a commercial miracle.

In 1892, after being told by a doctor that he was dying from cancer, Marcus organized the Tank Syndicate to carry on the tanker business after his death. The Tank Syndicate included family and friends such as merchants responsible for local distribution and individuals who had supplied storage tank facilities. The syndicate members were also responsible for garnering return cargoes for the tankers, which the Samuels sold in Europe. Trading transactions were done on the basis of a joint account for the syndicate members, all of whom became quite rich. When the doctor was proven wrong, the Tank Syndicate was reorganized as The Shell Transport and Trading Company in 1897.

All this was built on a house of cards. The Rothschilds were negotiating with Standard Oil and the Nobel brothers to form a world cartel, thus ending the intermittent price wars between the oligarchs. Standard Oil, sensing the importance of Marcus to the Rothschilds, opened negotiations to make Marcus part of Standard Oil. Marcus turned down a generous offer because he did not want to see a British firm become American or lose the Shell trademark and his identity as a businessman. In the game of King of the Hill, only one is left standing at the top, the primary reason why proposals for amalgamation among the Oil Kings failed. With the failure to come to terms with Standard Oil, Marcus was back skating on thin ice without a truly secure source of oil.

As fortune would have it, a company by the name of Royal Dutch in the Dutch East Indies produced oil, but was unable to transport and market its production. Royal Dutch was none of the things its name might imply. Its chief claim to fame was being the first oil company on record that relied on a government (the Dutch authorities in Dutch East Indies) to protect its oil holdings from insurgents. Royal Dutch had borrowed money to finance kerosene held in storage just as the price of kerosene crashed from Marcus's bulk-oil shipments and Standard Oil's campaign to chase American independents out of the Asian market. Royal Dutch approached Marcus about buying its Sumatra refinery output, but Marcus proved to be a tough negotiator, perhaps too tough. A subsequent rise in oil prices saved Royal Dutch and Marcus lost his first opportunity to obtain a secure source of oil and take over a company on terms that perfectly complemented his own. In the end, Royal Dutch would take over Shell on its terms.

In 1895 the cards turned on Marcus. Standard Oil, the Rothschilds, the Nobels, and the Russian independent producers reached a price agreement. The oligarchs controlled the entire world supply of oil except for that of the American independents. As oil prices stiffened, Marcus had to cut the shipping rate on his fleet to stay in business, although selling return cargoes of general merchandise carried on Shell tankers made up for the losses in shipping oil. As things were becoming more difficult for Marcus, fortune smiled and the Shell fleet profited from the Sino-Japanese war because different elements within Shell supplied both China and Japan. Shell would come out a winner no matter who won. Marcus represented that portion of Shell allied with Japan. Marcus was able to take commercial advantage of Japan's winning the war by becoming a merchant banker and floating the first Japanese sterling loan in London in 1896. Now a merchant banker, Marcus's star continued to ascend with his election as sheriff of London, which placed him in the direct line of succession to the highest civic office in Britain, lord mayor of London. With his newfound wealth, Marcus purchased a 500-acre estate bordering on the parsonage of Bearsted, marking the high point of his career when he was only forty-three years old.

The contract with the Rothschilds was half over and an alternative source of oil would have to be arranged if the contract were not renewed. As luck would have it, a Dutch East Indies mining engineer with an oil concession in Borneo showed up at Marcus's door in 1896. By this time Mark, the younger of the two nephews, was carrying quite a load. He was responsible for building tank storage facilities and inland distribution points, identifying new agents to handle distribution, ensuring proper discharge and cargo handling of the Shell tankers, and tending to a myriad of instructions from London on the firm's trading business plus continuing to explain every item on his personal expense account. He also covered his uncles' mistakes, such as how to get the kerosene from the company's tanks to users. Users could not accept bulk shipments; they bought kerosene in a tin. The uncles had not taken this last crucial step in the supply chain into consideration, thinking that the buyers would supply their own tins; they did not. The only tins were the blue Standard Oil tins, which had other uses that did not include buying Shell kerosene.

This, too, became Mark's responsibility. He was building storage facilities with no previous experience; now, with no previous experience, he had to build a factory for making Shell red tins that competed with the Standard Oil blue tins. Once the factory was set up, Mark was selected to do something else for which he had no experience: operate an oil field in Borneo. His preparation was a crash course consisting of a three-week visit to Baku, cut short to two weeks to hasten his return to Singapore. Mark's training proved inadequate for drilling for oil in the fever-ridden, rain-drenched, mosquito-infested, inaccessible jungle in Borneo at the Black Spot, a place where the soil was saturated with oil. Mark faced severe challenges in acquiring and getting the necessary equipment and workers to the site. Once on site, the equipment would break down and parts were difficult to obtain while tropical diseases decimated the workforce.

In retrospect, it would have been better for Marcus to make a deal with Royal Dutch, when it was having financial difficulties, to transport and market their refined oil rather than develop an oil field and build a refinery. Royal Dutch, with its headquarters in The Hague, Netherlands, had a successful oil concession in the Dutch East Indies and was knowledgeable about exploration, production, and refining. Royal Dutch was a perfect complement to Marcus: one company rich in exploration, production, and refining and poor in distribution and marketing, the other rich in distribution and marketing and poor in exploration, production, and refining. Marcus was betting on Borneo crude taking the place of Royal Dutch, but Borneo crude was not fit for making kerosene. It was more useful as a fuel oil substitute for coal to power factories and ships.

In 1898 Standard Oil decided to get control over oil production in the Dutch East Indies. To do so, Standard Oil let out a false rumor that its intent in taking control over Dutch oil producers was to stop production and replace Dutch oil with Standard Oil's American oil. The next step would be to get rid of the Russian oil coming in on Shell tankers and have the Asian market for itself. The rumor worked: shares in Dutch East Indies oil companies plummeted and Royal Dutch and Shell were again talking to one another. Since the original talks, Royal Dutch had not been sitting idle, depending on Shell for marketing and distribution. Deterding, a bookkeeper who by now was a rising star in Royal Dutch, strongly advocated

Royal Dutch having its own marketing department, if only to be able to play a tougher hand in the cat-and-mouse negotiations with Marcus. A cooperative arrangement between the two companies, signed in 1898, while flawed because agents of both companies continued to compete against one another, did prevent Standard Oil from carrying out its plans to bring the entire Asian market into its embracing tentacles.

That same year Marcus scored a publicity coup. The British warship *Victorious* went aground in the Suez Canal, much to the embarrassment of the British Navy. All attempts to free the vessel failed until Marcus showed up with the Shell-owned *Pectan,* the most powerful tug in the world. The tug freed the *Victorious* and Marcus deliberately did not submit a salvage claim, which he was entitled to, and in return received a knighthood from Queen Victoria. Not one to let a knighthood stand in the way of a commercial deal, and with Borneo oil being too heavy to make kerosene, but perfectly fit for burning as ships' fuel instead of coal, Marcus used the *Victorious* incident to establish a relationship with the British Navy. This was the opening shot of what would become nearly a fifteen-year campaign to induce the British Navy to shift from burning coal to oil, something that Marcus had already done with his tankers.

Marcus found strong support in a young naval officer who would one day be Lord Fisher, head of the British Navy. Coal smoke revealed the presence of a warship and oil burned with relatively little smoke. With a higher energy content, oil consumption would be less than coal, allowing warships to travel further without refueling. Refueling time would be considerably shortened since coal was at times carried on board in bags, whereas oil could be much more rapidly pumped aboard a vessel. Oil removed the necessity for stokers to shovel coal into the ship's boilers, reducing crew size. Converting space for holding coal to carrying ammunition increased the ship's battle endurance. However, to Fisher, the most important advantage of oil over coal was the greater speed that the British Navy had to have in order to stand up against the emerging German navy.

Marcus and Fisher, however, could not overcome the principal argument against converting to oil: coal was a domestic fuel whereas oil had to be imported from foreign sources. Thus, oil was less reliable and less secure than coal, a critical matter for warships. Although Shell had refueling stations for oil in the Pacific, they had none in the Atlantic. The lack of sufficient coverage to supply fuel oil was an obstacle to convincing ship owners and admirals to switch from coal to oil. Until this chicken-and-egg conundrum was resolved, the British Navy and ship owners who traded worldwide could not convert to oil. Nevertheless, ship owners trading within a region adequately covered by fuel oil bunkering stations could switch from coal to oil.

The fortunes of Shell and Royal Dutch oscillated like a pendulum on an overwound clock. In 1897 troubles hit Royal Dutch when its wells went dry. Royal Dutch then purchased Russian oil for sale through its marketing outlets in direct competition with Shell. A price war with Shell would have ended with the demise of Royal Dutch, but Marcus chose not to do so because he felt that the Asian market would grow to accommodate both companies. This was quite unusual thinking at a time when oil magnates did not hesitate to crush one another at the first opportunity. Unusual or not, this marked Marcus's second failure to acquire Royal Dutch.

In 1898 it was Shell's turn to face a sharp decline in its Borneo production. Shell's people in Borneo tried to keep the matter a secret from its competitors, but an agent in Singapore got wind of it and kept Standard Oil better informed of the situation than was Shell's London office. Declining production was just one of the worries on Marcus's shoulders. In addition to running a major oil company, he was trading goods that still included seashells, operating a merchant banking house for floating Japanese bonds in England, and participating in an active civic and social life. Marcus had little time to spend on the upcoming renewal of the Rothschild contract. He had to demonstrate that Shell, through its producing properties in Borneo, could live without the Rothschilds' contract in order to be able to renew the contract on favorable terms. Borneo crude would generate significant savings in shipping costs for Shell, but, perversely, would leave Shell tankers bereft of cargoes.

Marcus had to carry out this critical renegotiation in a business environment of continually shifting alliances among Standard Oil, the Rothschilds, the Nobels, and the Russian and American independents. One grouping of these companies would gang up against the others in one part of the world and another group would do the same somewhere else. Alliances came and went like liaisons in a brothel. How quickly the alliances could shift was clear when, in late 1899, Standard Oil broke its agreement with the Nobels and started a price war in Europe to get rid (again) of the American independents. The Nobels, caught by surprise and with a large inventory of high-priced kerosene, decided to join forces with the Rothschilds and the Russian independents to push Standard Oil out of Europe. Then Standard Oil decided to join the very group set up to ostracize it to exert a more formidable force against the American independents in Europe. The American independents could not compete against an alliance of Standard Oil, the Nobels, the Rothschilds, and the Russian independents. Shell was now in danger if this alliance were expanded to include Asia.

In response to this threat, Marcus started discussions with Dutch East Indies producing companies to secure an alternative source of oil, excluding Royal Dutch, which was still selling Russian oil in Asia in direct competition with Shell. In the midst of the Boer War, which strained relations between England and Holland, Marcus was able to strengthen his position with the Dutch independents in the East Indies, who found getting in bed with a British firm infinitely more tolerable than getting in bed with Standard Oil. Unfortunately, Marcus let an opportunity to fix long-term contracts with the Dutch independents, who resented Royal Dutch selling Russian oil in Asia in competition with their own, slip through his fingers.

Meanwhile Royal Dutch had obtained a new concession and was among the first to hire geologists to assist in identifying sites for exploratory drilling. The world was rapidly running out of sites where the surface soil was saturated with seep oil. Marcus was against hiring geologists because he thought that they were better able to tell where oil could not be found rather than where it could be found. He failed to realize that this negative information, if true, is valued intelligence. Despite his failed attempts to secure a long-term supply of oil, he was still making money, particularly when ship rates rose to replenish the British army during the Boer War. With shipping rates and oil prices escalating, Marcus, against his brother's objections, took long positions in kerosene to cover the period until Borneo would be producing kerosene in sufficient quantities to meet Shell's needs. Marcus had placed two bets: one on kerosene prices continuing to rise and another on Borneo producing kerosene in sufficient quantities to take the place of the Rothschild kerosene. He built more tankers (two, each at 9,000 tons, were the world's largest tankers at that time), expanded his storage facilities, and filled them with high-priced kerosene. He would lose both bets.

The year 1900 started well for Marcus. He reported record profits to his shareholders and called for a stock split to permit more shareholders to buy shares. He renewed his contract with the Rothschilds, but without the exclusive right to sell Rothschild oil in Asia. Marcus was not worried because the Rothschilds, without tankers, would not be able to sell their oil in Asia, an impediment that they would eventually find their way around. Only a few months later, Marcus's world began to collapse. It started with falling coal prices, which diminished the market for fuel oil as oil-fired ships reverted to coal. Then freight rates collapsed. Then the Russian economy slumped, further reducing demand for fuel oil. With less demand for fuel oil to run Russian factories, the Russian independents began producing more kerosene, creating a glut at Batum. As kerosene prices fell at Batum, Standard Oil dropped its prices, and the rest of the world followed suit. This left Marcus with a huge inventory of high-priced kerosene plus a slew of term contracts to continue buying kerosene at even higher prices. To make matters worse, the Boxer Rebellion broke out in China in 1900 and Shell's property was looted including 60,000 tons of kerosene along with the steel in the storage tanks. Troubles next spread to India, where Shell had more storage than all its competitors combined, also filled to the brim with high-priced kerosene. Shell competed against cheap kerosene from the Russian independents, the Nobels, Royal Dutch, and a new competitor, Burmah Oil.

Burma was the last place where oil was discovered by drilling into oil-saturated soil. Since Burma and India were British colonies, Burma could export oil to India without paying the import fees associated with oil from foreign sources such as the Dutch East Indies and Russia. The situation in China and India left Marcus's newly expanded fleet without cargoes at a time of low freight rates. To top this off, the Borneo oil field was producing a fraction of what was expected and the refinery built to process Borneo crude suffered severe operating problems.

Motorcars were just beginning to appear in England. With only a few thousand registrations, Marcus saw automobiles as another business opportunity, as did other oil magnates. Up to this time, naphtha produced from refineries was either burned or in some way discarded. Automobiles would be an ideal market for selling a waste product. Gasoline in England was already being sold in blue Standard Oil tins when Marcus began dreaming of bright red Shell tins. He had made the opening moves to sell gasoline in London by leasing storage space, overlooking the fact that gasoline was not a permitted cargo for transiting the Suez Canal in bulk tankers. Not yet having obtained permission from the Suez Canal Authority to use the canal, Marcus shipped a cargo of gasoline from his Borneo refinery around the Cape of Good Hope, a dangerous undertaking. Standard Oil was fully prepared for the arrival of Shell's first shipment of gasoline to England. It had forced every agent and distributor in Britain to enter into a contract not to sell any brand but Standard Oil. This knocked Shell out of the gasoline market in England. Then the carnivorous Standard Oil purchased a U.S. west coast fuel oil producer for the sole purpose of exporting fuel oil to Asia and put the final squeeze on Marcus. Caught in the Standard Oil juggernaut in England and Asia, "discussions" began between the two firms.

Meanwhile, Royal Dutch, which at times bordered on bankruptcy, aided by advice from geologists was now on the comeback trail with discoveries of new oil fields in the Dutch East Indies. Deterding, now president of Royal Dutch, had done all that he could in the past to prevent an amalgamation between Royal Dutch and Shell. Marcus had lost his strongest supporter at Royal Dutch with the death of Deterding's predecessor and now faced an individual who relished taking full advantage of Royal Dutch's ascendancy over an ailing Shell. Whereas in the past Marcus was absolutely determined that Shell would not play second fiddle to Royal Dutch, now the tables had turned and Deterding was just as adamant that Royal Dutch would not play second fiddle to Shell. To add insult to injury, Royal Dutch geologists discovered oil in the same location in Borneo where Shell, without geologists (at Marcus's insistence), had failed. Just as prospects for Marcus were almost pitch-black, a new twist entered his life.

Spindletop

Patillo Higgens left his hometown of Beaumont as a one-armed young man who could fight better than most Texans with two. He returned in the middle 1880s as a Baptist churchgoer and Sunday school instructor and made a living in real estate and timberland. He took his class to picnic on a large mound that rose fifteen feet above the flat prairie and covered thousands of acres. He punched a cane into the ground and lit the escaping gas to amuse the children. Higgens was intrigued by the sour smell, the square boxes that held blue, green, and yellow waters for bathing or drinking or passing livestock through to rid them of the mange, and St. Elmo's lights, which hovered over the mound at night. These were signs of something, but it was not until he paid a visit to the Oil Regions in Pennsylvania and elsewhere (trying to figure out how to get into the brick-making business) that he figured out the source of these mysterious signs.[11]

Without funds, he convinced others to purchase land on what would eventually be called Spindletop after a nearby town, and tried to keep himself in the picture. In 1892, Higgens formed a company called Gladys City to corral investors in what he saw as a future oil company (the stock certificates featured the portrait of a local young girl named Gladys, along with imaginary oil wells, tanks, and refineries). Higgens was convinced that oil would be discovered if a well were drilled to 1,000 feet, but time was against him. He had purchased options on some land parcels and was having difficulty raising the necessary funds.

Spindletop made life tough for drillers with its quicksand, gas pockets, and loose conglomerate. The first hole was drilled to a little over 400 feet before being abandoned by the driller. Higgens persisted. A second driller made it to 350 feet before Spindletop put a stop to his attempts to uncover its secrets. With a history of two dry wells, to beef up support for drilling a third well Higgens got a Texas state geologist's opinion about Spindletop. He did opine: Petroleum means rock oil and with no rocks in Spindletop, no oil. This is what Higgens did not want to hear. The geologist, utterly convinced of his findings, sent a letter to the local newspaper to warn the good people of Beaumont not to waste their money looking for oil. This letter convinced the local townspeople of what they already suspected: Higgens was losing his mind from the sour gas fumes coming from Spindletop. As a last act of desperation, Higgens advertised for investors. He received only one response, from a Captain Lucas.

Lucas was looking for sulfur, not oil, and had a theory about finding sulfur in salt domes. After listening to Higgens, it was an easy leap of faith to think that oil might also be found in salt domes. Higgens, short on cash, arranged for Lucas to obtain a lease on all of his Gladys City landholdings, for which Higgins ended up with a 10 percent share. Higgens was reduced to acting as an agent on commission for the company he had founded. Lucas brought in a rotary rig, not the traditional cable-tool rig used by the previous two drillers. Spindletop proved to be too much for Lucas's rig. Lucas ran out of money after drilling two dry holes. Now with four dry holes, and unable to raise funds locally, Lucas sought help from Standard Oil. After examining the property, Standard Oil's expert geologist opined that no one would ever find oil at Spindletop, as did another geologist, employed by the federal government.

Lucas then made contact with a team—Galey, a driller, and Guffey, a promoter—with close ties with the Mellons. Guffey demanded that Lucas get rights to all the land on Spindletop before doing any drilling and that Higgens be kept in the dark to keep their involvement a secret. With Mellon money backing Guffey, Lucas was able to get leases on 15,000 acres, except for what would turn out to be a critical omission, the many small tracts that ran across the top of Spindletop, which included the thirty-three-acre lot owned by Lucas. The new partnership left Lucas with a relatively small share, greatly diminishing Higgens's 10 percent interest as well.

Galey visited the property and drove a stake into the ground. Had he driven the stake fifty feet away, the well would have missed its target. Galey arranged for the Hamill brothers, who had their own rotary rig, to drill the well. When the Hamills hit the same quicksand that had stopped the other drillers, they found that using drilling fluid spiked with mud, obtained by driving a herd of cattle around a slush pit, would seal the sidewalls and keep the quicksand from filling up the well bore. This was the first use of drilling mud, now universally used in drilling. When they reached a point where the mud would not seal up the sidewalls, the Hamills devised a means of inserting a pipe casing that supported the walls of the well, allowing drilling to proceed. At about 650 feet the Hamills ran into gas pockets that made the circulating mud boil and flow up rather than down the drill pipe. The Hamills overcame this problem, along with others, when the drill struck the salt dome caprock at 880 feet.

The night of January 9, 1901, turned out to be the last great show of St. Elmo's fire, ghostly blue flames usually associated with an electrical discharge, ever seen on Spindletop. The next morning, while the Hamills were lowering drill pipe into the now 1,200-foot-deep drill hole, mud suddenly started to spurt high above the derrick. The crew ran for their lives as six tons of drill pipe blasted from the hole destroying the derrick. This was followed shortly thereafter by a cannon shot of gas followed by a one hundred-foot-high, 100,000-barrels-per-day geyser of oil, clearly visible from Beaumont and everywhere else within a twelve-mile radius, accompanied by a stupendous roar. Higgins found out about the oil geyser that afternoon when he rode into town. A few days later the Hamills would be the first to devise a way to cap an oil gusher, the local pronunciation of *geyser*.

Pandemonium reigned in Beaumont as in Pit Hole. In the months that followed, Beaumont grew from 9,000–50,000 inhabitants with six special trains running between Beaumont and Houston daily. Those who did not go back to Houston could share the same hotel room with twenty other people. The bars and the brothels never closed. Stock manipulators and scoundrels sold leases that either did not exist or turned out to be far from Spindletop, even far into the Gulf of Mexico. Stocks in companies without clear title to the land or without a promise to do anything were traded daily in an improvised stock exchange. It did not matter if the title to the land or a lease was bogus or compromised if it could be sold at a higher price. Fortunes were made on dubious securities and leases. Eventually lawyers would make even more money settling litigation over whom, exactly, possessed title to producing wells. Higgens was lost in all the pandemonium surrounding Spindletop. Like Drake, he would die without fame or fortune, but at least not a pauper.

After the discovery of Spindletop, Guffey lined up financial support from the Mellons. The Mellons, while primarily bankers, had previous experience in the oil patch. In 1889 the Mellons owned an oil field in Pennsylvania and had decided to fight rather than become Standard Oil property. In 1892 they succeeded in getting a contract with a French company to refine their oil. Immediately the Pennsylvania Railroad hiked its shipping rates to prohibitive levels and the Reading Railroad refused to carry Mellon oil. When the Mellons attempted to build a pipeline to the east coast, hired thugs of the Pennsylvania Railroad fought the pipe layers by day and ripped up laid pipe by night. The Mellons were forced to sell out to Standard Oil, but they did make a handsome $2.5 million for their troubles. (It was not Rockefeller's price that people objected to necessarily, but his forcing the sale against the sellers' wishes.)

There was one thing Texans, and the Texas legislature, was bent on doing: keeping Standard Oil out of Texas. They succeeded by passing antitrust legislation that made it virtually impossible for Standard Oil to establish a toehold in Texas. Spindletop gave birth to Gulf Oil, the successor company to Guffey Petroleum Company and Texaco, the successor to the Texas Fuel Oil Company. Moreover, the spate of oil exploration in the rest of Texas set off by Spindletop would create Sun Oil and Humble Oil, named after a town. Humble Oil would eventually become Standard Oil's entry into Texas oil fields when it acquired a half interest in the firm in 1917. Eventually, Humble Oil, the most misnamed company imaginable, would be absorbed into Exxon. Oil flowing from Spindletop, which would account for half the nation's production, broke the Standard Oil monopoly in America. Spindletop oil, heavy and better fit for burning as a fuel than for making kerosene, was immediately recognized as a replacement for coal.

Spindletop and Shell

The Mellons, back in the oil business by financially backing Guffey, wanted the oil sold to anyone but Standard Oil. Marcus realized that Spindletop crude was unfit for kerosene production, but was an ideal fuel oil. With Spindletop crude, Marcus could fulfill Fisher's dream of fuel oil being available in both hemispheres to supply the British Navy. In June of 1901, Marcus agreed to buy half of Guffey's production for twenty-one years at about twenty-five cents per barrel, plus a 50 percent share of the profit in the net sales of the oil with a minimum takeoff of 100,000 tons per year. This was the second major transaction for Marcus in 1901; the first was the sale of the company's seashell business to a relative.

In the game of oil, the positions of the chairs had again shifted. Standard Oil now saw Shell not as a competitor about to be crushed, but as a means of getting its hands on Spindletop oil. Rather than wiping out Shell, the objective now was to make Shell part of the Standard Oil family. Deterding knew that any alliance between Standard Oil and Shell would spell trouble for Royal Dutch, so Deterding entered the unholy alliance and the three companies divided the non-Russian world oil market among themselves. They actually reached an agreement on divvying up the world market by oil products, of which there were then five: kerosene, the mainstay of the business, lubricating oils, the emerging markets in gasoline, fuel oil, and what was called solar oil, a substitute for coal for manufactured gas to light municipal street lamps and buildings. There was also agreement on which geographic areas each firm would operate.

In the midst of these critical discussions in late 1901, Marcus took time out for the pomp and ceremony of becoming lord mayor of London. While Marcus was being showered with honors, the Rothschilds were attempting to unite themselves with the Nobels and the Russian independents into a single marketing entity to counter any Standard Oil-Shell-Royal Dutch combine. In early 1902, talks between Standard Oil, Shell, and Royal Dutch collapsed, despite their marked progress in carving up the world market. The cause of the failure was the same for the failure of every proposed amalgamation: the name of the game is King of the Hill, not Kings of the Hill. No one could agree on which oil company would head the combine other than their own.

The pleasantries exchanged during negotiations between Standard Oil, Shell, and Royal Dutch gave way to open commercial warfare. This rekindled negotiations between Deterding and Marcus, which led to the signing of the British-Dutch Agreement in mid-1902. After the signing, the Rothschilds wanted to join the two, which Marcus opposed and Deterding supported. In the end, Deterding won and the British-Dutch Agreement was amended to become the Asiatic Agreement, marking the birth of the Asiatic Petroleum Company. Because the agreement called for all three companies to participate in a joint venture for refining and marketing oil products in Asia, the Rothschilds had finally found a way around Marcus to market oil in Asia. The Rothschilds, as in the past, saw the agreement as a means to improve their negotiating strength with Standard Oil. Marcus saw the Asiatic Agreement as something temporary to keep Standard Oil at bay. Deterding saw the agreement as something permanent, leading to the final ascendancy of Royal Dutch over Shell. This was virtually assured when Marcus allowed Deterding to be in charge of Asiatic Petroleum's operations.

The gods turned against Marcus. The *Hannibal*, a British warship put on trials to test out Marcus's idea of burning oil, was enveloped in black smoke when the fuel was shifted from coal to oil. The experiment was a total failure because the wrong atomizers had been installed. This would delay the conversion of the British Navy from coal to oil for another ten years, much to the disappointment of Marcus and Fisher. The Port Arthur refinery built to process Spindletop oil was having serious operating problems, but this was nothing compared to the news that the production of the hodge-podge of oil wells at Spindletop, one nearly on top of the other, had gone into a sudden decline, particularly those owned by Guffey. A young nephew of the Mellons surveyed the scene and concluded that the refinery was unworkable, the oil was gone, and their investment was wasted. The only way to recoup the Mellon investment in Spindletop was to create a totally new integrated oil company with a massive capital infusion. Rockefeller came out of partial retirement to tell the Mellons personally, with some degree of relish, that there was no way Standard Oil would assist them.

Colonel Guffey was set aside and new management installed to allow the Mellons to reorganize Guffey Petroleum into what would become Gulf Oil. Honoring the Shell contract was impossible, not because oil production at Spindletop had essentially ceased, but that the price of oil was above twenty-five cents per barrel. The Mellons could not buy oil on the open market to honor the contract without taking an enormous financial loss. Unwilling to absorb such losses, the Shell contract was unilaterally canceled and Andrew Mellon inveigled Marcus to substitute a much less onerous contract, which in the end was not honored. Shell's tankers, built to carry Spindletop oil, were converted to cattle carriers.

Some think that Marcus should have sued the Mellons and saved Shell through litigation. This would not have been as easy as one might expect because the terms in the contract left something to be desired if exposed to the scrutiny of a court of law. Others thought that Marcus might have been thinking of the long-term implications of not suing the Mellons, perhaps hoping for a potentially profitable collaboration between Gulf and Shell in the future. The implication of future collaboration might have been a keen insight on the part of Marcus, but the short-term effect was disastrous.

If this was not bad enough, Marcus received word that Deterding was unhappy with the Asiatic Agreement and that adjustments would have to be made to the agreement, adjustments of a type that would not benefit Marcus. Although his investiture as lord mayor of London, with its pomp and ceremonies, was a great honor for Marcus,[12] the time consumed prevented his meeting with the representatives of the Rothschilds and the Nobels to deal with yet another problem, in Germany, where Shell was facing the full fury of Standard Oil. This placing of civic responsibilities ahead of business was to cost Marcus dearly.

In early 1903, Lane submitted a letter of resignation stating that he was unable to continue as a director of a company as poorly managed as Shell. He complained of Marcus's attention being diverted from the oil business to trading merchandise, running a merchant bank, participating in civic activities, placing inexperienced nephews in charge of major projects, and relying on a brother's opinion rather than a more formal approach to planning before making critical business decisions. Indeed, the head count in Shell's London office, the heart and soul of a major world oil enterprise, was just under fifty including clerks, typists, bookkeepers, and messengers.

Things were going from bad to worse with Deterding running Asiatic Petroleum. Deterding limited Shell's profits to freight paid for its tankers and rentals on its storage facilities. Money made in marketing and distributing kerosene in Asia ended up in the Royal Dutch accounts. By the sleight of hand of a very experienced and adept bookkeeper, Shell suffered declining profits while those of Royal Dutch rose. Moreover, Asiatic Petroleum was extremely late in issuing its financial reports, without which Shell could not issue its final financial statements. This proved to be something else that depressed the value of Shell shares. Deterding had placed Marcus in a desperate strait, having wrecked Shell's profits and the value of its shares. Exhausted from his year as lord mayor of London and disillusioned with those about him, Marcus was at the point of giving up, something Deterding had been striving for since taking charge of Asiatic Petroleum.

Before Shell fell under Royal Dutch rule, Marcus was given a last-minute reprieve in the form of a financial shot in the arm from the profits made by the Shell fleet's support of Japan in the 1904 Russo-Japanese War. This proved to be the incendiary that ignited the 1905 Russian Revolution when revolting oil workers set fire to the Baku oil installations, a dress rehearsal for 1917 and a training ground for Stalin. The pathetically slow progress of the coal-fueled Russian fleet as it sailed from the Baltic to its destruction off Japan provided impetus for the British Navy to switch to oil. When the British Navy did switch, Shell was no longer an independent company.

The emergence of the automobile age in the United States made gasoline a mainstay for

Standard Oil and kerosene a by-product. Standard Oil dumped its excess American kerosene in Europe and formed a joint marketing effort with the Rothschilds (Shell's partner in Asiatic Petroleum) and the Nobels to keep kerosene prices low. Shell, whose mainstay was still kerosene, had to face this combine alone. Everyone was losing money by selling kerosene in Europe, but Shell did not have the financial wherewithal to outlast the others. Like wolves gathering for the final kill, Shell was forced to sell six of its best tankers at a tremendous loss to recoup its investment in Germany. By 1906, beaten in Europe by Standard Oil and beaten in Asia by Deterding, Marcus had no choice but to appeal to Deterding for an amalgamation of the two companies.

Marcus went to Deterding's office. Deterding gave Marcus his first and final offer. If Marcus left without accepting the offer on the spot, the offer was dead and so was Shell. The offer was the formation of a holding company called Royal Dutch-Shell Group, of which Royal Dutch would own 60 percent and Shell 40 percent. While Marcus was nominally in control of the holding company, the King of the Hill was definitely Deterding. To further ensure Royal Dutch dominance, Deterding had Royal Dutch buy 25 percent of Shell's shares at thirty shillings per share when the price of the stock three years' previous had been three pounds. Deterding considered this a very generous offer under the circumstances, which it may have been. Maybe it was Deterding's way of thanking Marcus for passing up several opportunities to take over Royal Dutch and become King of the Hill himself.

Two new operating companies were formed, one British and one Dutch. The British company controlled transportation and storage and the Dutch company production and refining. Royal Dutch and Shell were then emptied of all assets and became holding companies in which each party held a 60–40 percent share in the two operating companies. Asiatic Petroleum continued to market products in Asia with two-thirds shareholding of this company reallocated 60 percent to Royal Dutch and 40 percent to Shell; the remaining third stayed in the hands of the Rothschilds. In 1907, when the Group was formally established, Marcus, though personally rich, considered himself an abject failure.

As with so much of his life, there was a new twist. Deterding, contrary to all the rules of the game, did not leave Marcus out in the cold. Deterding decided to operate out of Shell's London office, not out of Royal Dutch's Hague office. With Marcus sitting in the same office, Deterding found that he could be more effective if he kept Marcus informed of the latest developments and conferred with Marcus before making any major policy decisions. This consultative arrangement between Deterding and Marcus worked to their mutual advantage, and this unique method of managing a large firm survived the two individuals. After Deterding retired from Shell, all major decisions had to receive a favorable ruling from two committees, one representing Royal Dutch and the other Shell. The committees were made up of personnel with long-standing records of achievement who, rather than retire to a golf course in Scotland, met on a regular basis to confer on important matters and make recommendations based on their extensive experience. This consultative and collegial method of decision making, unique in the corporate world, has been adopted by the principal operating companies within the Royal Dutch-Shell Group.

Deterding, though a Dutchman, saw a greater commercial advantage if the newly formed Group was associated more closely with Britain than Holland to take advantage of operating within the

British Empire. In 1910 the British Navy finally switched to fuel oil, which was a great boon to the Shell Group. However the Group was considered non-British because its sources of oil did not lie within the British Empire and a Dutch company owned 60 percent of it. Shell still benefited by selling fuel oil obtained from foreign sources to qualified British companies, which, in turn, supplied the British Navy. Even though Deterding was "on top," Marcus was not idle. He turned his attention to Egypt, and, following up on rumors, insisted that the Group explore for oil because if found (and it was found) the Shell Group would have a source of oil on British colonial soil. This would permit the Shell Group to sell fuel oil directly to the British Navy. Deterding was no slacker either. He acquired oil properties in California that were later expanded to Oklahoma, allowing the Shell Group to confront Standard Oil on its home turf, plus getting involved with oil fields in Mexico and Venezuela. In 1912, with Lane in the middle, the Rothschilds exchanged their Russian holdings for stock in Royal Dutch-Shell, thereby becoming one of its largest shareholders. In light of what was to occur only a few years later, this exchange of oil properties in Russia for shareholding interests in Royal Dutch-Shell proved to be a most astute move because diversification mitigated the financial risk of having all one's eggs in a single basket.

Winston Churchill, as first lord of the admiralty, agreed with Fisher on converting warships from coal to oil, with one major reservation. Churchill believed that the British Navy should not rely exclusively on contracts from suppliers, but that the government should have its own oil fields to guarantee a supply of fuel for the navy. Marcus attempted to convince Churchill that the Shell Group, along with Standard Oil, could supply the British Navy in any location throughout the world. Marcus argued that the navy would be better served building storage facilities, not buying oil fields. Despite Marcus's pleas, the British government went ahead with Churchill's plan and purchased a 51 percent interest in Anglo-Persian Oil Company, an offshoot of Burmah Oil, in 1914, weeks before the start of World War I.[13]

Anglo-Persian Company was originally formed in 1908 when another oil explorer with the drive of Colonel Drake, ignoring a letter to cease looking for oil, found oil. A 130-mile pipeline, the first in the Middle East, was laid between the oil field and a refinery built in Abadan. Having only one outlet to the marketplace, through the Shell Group, Anglo-Persian Oil was in a weak bargaining position and vulnerable to a Shell Group takeover. Moreover, it was in a weak financial condition. The company wanted an investment by the British government to gain a major new client, the British Navy, and planned to expand its refinery with the proceeds of the investment become the largest in the world, diversify its markets, and serve those markets with its own tanker fleet. By owning an oil field, Churchill felt that he would not be at the mercy of oil companies with regard to price and supply. The British government did not interfere with the running of Anglo-Persian Oil and its members on the board of directors made sure that the company's operations did not conflict with the government's strategic objectives.

At the start of the war, the Shell Group chartered its entire fleet of over seventy tankers to the British Admiralty at prewar rates as a show of support for Britain. As a consequence, the company had to charter in other tankers at up to four times these rates to meet its needs. The Dutch side of the company transferred as much of their operations to London as possible. Marcus converted his mansion into a military hospital and his two sons and two sons-in-law served in the military. Only one survived.

The Shell Group became the sole supplier of aviation fuel and the principal source of motor vehicle fuel to the British Expeditionary Force. The toluene in the explosive TNT (tri-nitro-toluene) came from processing coal. While crude oil normally contains only trace amounts of toluene, Shell's Borneo crude was unusually rich in toluene (10 percent). To process Borneo crude for its toluene content, the Shell Group's refinery in Rotterdam was dismantled and "smuggled" to England. In addition to having the Shell Group invest in National War Loans, Marcus spearheaded the conversion of general cargo vessels into tankers and had others fitted with double bottoms for supplying fuel to the expeditionary forces in Europe. He was also active in introducing diesel propulsion to replace oil-fueled steam propulsion plants. In 1916, when it was clear that Romania would fall to German forces, Marcus and Deterding authorized company personnel to destroy the Shell Group's Romanian oil assets without any promise of restitution by the British government. Ironically, Shell Group gasoline was distributed in Britain before the war under a contract with British Petroleum, at that time a German-owned marketing company. British Petroleum shares were seized by the British government and turned over to the Anglo-Persian Oil Company, marking the official birth of BP.

The British government was the first government to have majority ownership of an oil company, but chose not to run it. The first government that actually ran an oil company was the Soviet Union after it expropriated the oil-producing properties of the Nobels, the Shell Group, and the Russian independents after the Russian civil war. But the Nobels did not leave empty-handed because Exxon bought their oil rights in Russia in 1920, on the remote chance that the Whites would win the Russian civil war. Although the Nobels received money for their oil properties, they were out of the oil business. The Rothschilds' loss could have been disastrous, aside from their one-third ownership of Asiatic Petroleum, had they not exchanged their Russian oil-producing properties for shares in Royal Dutch-Shell. The Russian independents were lucky to escape with their lives. Now the Soviet Union was in the oil business and depended on oil exports to build communism, in much the same way that Russian oil rebuilt the Russian economy after the fall of communism in 1991.

In 1920, despite all that Marcus had done in support of the British war effort during World War I, the public rose against him and accused him of greed in the face of rising petrol prices. Only a decade earlier the darling of London society, Marcus was now a pariah, accused of siphoning money out of everyone's pocket. It was his turn to endure the vituperation heaped on Rockefeller. Marcus's appeal to the harsh law of supply and demand for establishing the price for oil to clear the market did not endear him with the public. This display of public ill will might have played a role in his retiring as chairman and board member. In 1925 he became Lord Bearsted (Deterding was knighted in 1921 in recognition of his war services) and two years later both Marcus and his wife died within twenty-four hours of one another.

The Emergence of Oil as a Strategically Vital Commodity

Winston Churchill was the first government official to sense the strategic importance of oil. Oil had its beginnings in lighting, but kerosene lamps were giving way to electric light bulbs. Automobiles were toys for the rich at the beginning of the twentieth century, but when Henry Ford began to mass-produce Model Ts, the era of the horse and wagon ended. During the years prior to World War I, oil was becoming an integral part of national economies without anyone taking notice, but during World War I, when success in combat depended on a steady and reliable flow of oil to fuel military vehicles, tanks, and fighter planes, it was noticed. National survival placed a whole new emphasis on the importance of oil. Oil was no longer a consumer item, but a means to ensure military success.

World War II only reinforced the lessons learned in World War I. Oil followed only armaments and ammunition in importance for winning a war. Hitler, cognizant of Germany's lack of raw materials and energy, except for coal, built facilities in Germany that made gasoline from coal. Gasoline fueled the aircraft and tanks essential for the success of the blitzkrieg: a rapid deployment of armies to envelop an enemy before resistance could be organized. The Nazi army quickly invaded Romania after starting World War II to seize its oil fields. Hitler's thrust into the Soviet Union and Rommel's invasion of North Africa were to meet at the Baku oil fields, placing Middle East and Soviet oil under Axis control. Fortunately for the Allies, both suffered from severed supply lines. Hitler's army's replenishment lifeline was cut at Stalingrad as was Rommel's gasoline lifeline to North Africa. Likewise, in the Battle of the Atlantic, Hitler tried to cut the British lifeline of troops, armaments, ammunition, and oil flowing from the United States with submarine U-boats. Germany's capacity to wage modern warfare ended when the Allies finally won air supremacy and bombed Hitler's coal-to-gasoline production plants.

The war in the Pacific was likewise heavily influenced by oil and by attempts to interrupt its flow. In the months prior to Pearl Harbor, the United States imposed an embargo of scrap steel and oil to Japan as a sign of its disapproval of Japan's invasion of China. With the United States supplying 80 percent of Japanese oil, the embargo forced Japan to set its sights on the Dutch East Indies oil fields. The Japanese knew that the supply line of oil from the Dutch East Indies to Japan was long and vulnerable to naval interruption. Only one navy was powerful enough to interrupt Japan's oil lifeline; the oil embargo made Pearl Harbor inevitable. Severing the lifeline of raw materials and oil to Japan from its conquered territories in Southeast Asia was a major goal of the war in the Pacific.

The Era of the Seven Sisters

For over a half-century, between World War I and the oil crisis of 1973, the world oil business was conducted largely through the seven sisters: Exxon, Shell, British Petroleum (BP), Gulf, Texaco, Mobil, and Chevron, each ranking among the world's largest companies. Exxon, Mobil, and Chevron were the leftovers of the Standard Oil Trust breakup. Gulf and Texaco were the products of keeping Standard Oil out of Texas (Exxon eventually became a major player in Texas through its subsidiary, Humble Oil). BP, Churchill's brainchild, branched out far from its original purpose. Since the 1973 oil crisis, the seven sisters have been reduced to four: Exxon and Mobil have recombined, Chevron purchased Gulf Oil and combined with Texaco, and BP, while it did not combine with any of the other seven sisters, absorbed three leftovers of the Standard Oil breakup, Sohio, Amoco, and Arco. Apparently, the earlier conflicts that had plagued attempts to amalgamate had been overcome. As the decades passed, top executives with no links to the founders or their immediate successors, stepped aside, their hurt feelings assuaged by generous bonuses and retirement packages.

The seven sisters were fully integrated multinational companies that controlled every facet of the oil business. Upstream activities included exploring and developing oil fields. Downstream activities included refining crude oil and distributing refined products by pipelines, tankers, and tank trucks to gas stations and industrial, commercial, and residential end users. The oil companies felt that they owned an oil field, even if it was located in a foreign nation under a concession agreement. Every aspect of the oil business from exploring, drilling, production, refining, distribution, and marketing was not only controlled, but the assets in oil fields, pipelines, tankers, refineries, storage facilities, tank trucks and filling stations were owned by the oil companies. Price and production volumes were set with the oil companies sitting on one side of the table and oil producers on the other, with a generally one-way dialog between the two. This world collapsed in 1973, a pivotal year in the oil industry.[14]

The seven sisters both competed and cooperated. They competed with one another over market share and cooperated with one another in exploring and developing oil-producing properties. Oil industry leaders had to learn to deal with this dichotomy, but in a way they were groomed to both cooperate and compete from the beginning. An individual I know was starting to climb the corporate ladder as drilling manager in an isolated part of South America. Over the hill was another individual in charge of drilling for a competing oil company. When a drill bit broke in the middle of nowhere, the individual could order a replacement from the home office, but it would have taken weeks to receive it, which would have caused him to miss the scheduled completion date. Alternatively, he could walk over the hill and borrow one from his competitor. The competitor's drilling manager was more than willing to cooperate because he knew that he now also had a ready source of replacement parts across the hill that would allow him to complete his drilling program on time. Because the performance of both men would be judged in terms of the time and cost required to complete their respective drilling programs, both advanced their careers by walking over the hill when they needed spare parts.

Costly and risky oil exploration and development programs are often carried out by a syndicate of oil companies. The potentially enormous losses associated with exploration and oil-field development can be spread over the participating companies in a syndicate without having a single oil company bear the entire risk of loss. The risk of loss has not been reduced, but the extent of loss a single company must bear is limited to its share of the syndicate. To some degree the risk of loss is reduced since cooperation allows oil companies to share particular skill sets and technological expertise with others. Thus, not every company has to be an expert in every facet of exploration and development. In a well-structured syndicate, companies assume responsibility for specific functions they are particularly adept at fulfilling. Nevertheless, each participant keeps a wary eye on the others to ensure that no one takes advantage of a situation as Deterding did with Marcus.

Opening Up the Middle East

The opening up of the Middle East is synonymous with Calouste Gulbenkian. His father and uncle were petty merchants who rose to the position of being responsible for collecting revenues for the Sultan's privy purse in Mesopotamia. This gave them the opportunity to found a merchant bank to finance transactions between Constantinople and Baghdad.[15] As a reward for Gulbenkian's father's service to the Sublime Porte, he was given the governorship of Trebizond, where he became involved with kerosene imports from Baku on behalf of the Turkish government. Through contacts developed with the Baku oil exporters as a representative of the Turkish Crown, he enriched himself greatly as a private merchant handling kerosene imports into the Ottoman Empire.

His son Calouste was educated in Britain. As a young man in the 1890s, Gulbenkian was sent to the Caucasus to learn about oil, which began his lifelong interest in oil. He wrote a book about his experiences, including an assessment of the Baku oil industry, which attracted the attention of the Turkish Crown. Gulbenkian was commissioned to do a report on oil prospects in Mesopotamia (now Iraq). The book was a compilation of existing sources plus observations from railroad engineers who had been in Mesopotamia, a place Gulbenkian was never to visit. The book whetted the Sultan's appetite and induced him to transfer enormous land holdings from the government to the Crown.

Fleeing Turkey with this family during the Armenian massacres of 1896, Gulbenkian appeared on the world stage of oil as the London representative of Mantachoff, a leading Armenian Russian oil magnate. Gulbenkian worked with Frederick Lane, who he considered the father of the British oil industry, to bring Russian oil interests into the Royal Dutch-Shell Group. Gulbenkian's experience led him to believe in the importance of pooling oil resources, production, and marketing to achieve price stability, an idea shared by others responsible for creating the oil industry. Unlike his father, Gulbenkian had no interest in the business aspects of oil. He saw himself as a creative architect of oil business arrangements. His failure to seize upon an early opportunity to get involved with an oil concession in Persia, which became the basis for the Anglo-Persian Oil Company, led him to adopt his lifelong business obsession: Never give up an oil concession!

In 1908 an oil strike in Persia whetted Gulbenkian's interest in Mesopotamia. Gulbenkian convinced Deterding to open a Constantinople office with Gulbenkian in charge, although he continued to be a financial advisor to the Turkish embassies in Paris and London and to the Turkish government. However, others shared Gulbenkian's intuitive insight. The Germans were eager to build a railroad to Baghdad that would have oil rights for about thirteen miles on both sides of the track. The Anglo-Persian Oil Company saw Mesopotamia as an area with great oil potential. The Ottoman-American Development Corporation also had its eye on Mesopotamia. The British government, alarmed over growing German influence in Turkey, needed someone known to European oil interests who spoke the language, had the contacts and knowledge of the oil industry, plus possessed the business acumen, skills, and foresight to represent their interests in the Near East. Gulbenkian possessed all these traits and was in the right place at the right time; if he had not existed, the British government would have had to invent him.

In 1910, in addition to his other activities with Shell and the Turkish government, Gulbenkian became an adviser to British financial interests when they formed the National Bank of Turkey in order to make loans within the Ottoman Empire. Working under the auspices of the National Bank of Turkey, Gulbenkian formed the Turkish Petroleum Company in 1912; Deutsche Bank held 25 percent of the stock, Gulbenkian 40 percent, and the National Bank of Turkey 35 percent. To entice Shell into the deal, Gulbenkian gave Shell a 25 percent interest, reducing his to 15 percent. Neither the National Bank of Turkey nor the Turkish Petroleum Company had any Turkish investors.

When Anglo-Persian Oil pursued an oil concession in Mesopotamia, Gulbenkian rearranged the shareholding in the Turkish Petroleum Company to include Anglo-Persian Oil. (Gulbenkian believed that it is better to embrace rather than fight a potential competitor.) The ownership of the Turkish Petroleum Company was now split: Anglo-Persian Oil had a 47.5 percent share, Deutsche Bank a 25 percent share, Shell 22.5 percent, and Gulbenkian 5 percent, with the Turkish National Bank no longer sitting at the shareholders' table. In 1914, just before the outbreak of World War I, the Ottoman government wrote a letter to the British and German ambassadors in Constantinople acknowledging that the Turkish Petroleum Company had an oil concession in the provinces around Baghdad and Basra.

After World War I, Britain and France proceeded to carve up the Middle East as spoils of war, excluding the United States because it had not officially declared war on Turkey. The British government agreed with Gulbenkian's assertion that the concession granted to the Turkish Petroleum Company by the Ottoman government was still valid, even though the Ottoman Empire no longer existed. The British government wanted to turn Deutsche Bank's quarter share over to Anglo-Persian Oil. To avoid giving too much power to Anglo-Persian Oil, Gulbenkian inveigled the French government to take over the German quarter share interest in the Turkish Petroleum Company as a war prize. In 1922 the U.S. government, concerned over a possible shortage of crude oil, negotiated an interest in the Turkish Petroleum Company in the name of the Near East Development Corporation (again with Gulbenkian's support, based on his practice of embracing potential rivals rather than fighting them). The corporation did not specifically name any U.S. oil companies, but was eventually represented by six; these were reduced to two, Exxon and Mobil, by the end of World War II. After deliberations with Gulbenkian's involvement, the Turkish Petroleum Company was evenly split among the Near East Development Corporation, the French government, Shell, and Anglo-Persian, which accepted a halving of its share for a 10 percent overriding royalty. The new reorganization still contained Gulbenkian's 5 percent share. This would become a bone of contention from this point forward between Mr. Five Per Cent and his partners, even though there was not a single drop of known oil reserves. Without any activity in oil production and marketing, the oilmen, working out of luxury hotel suites, saw no value in Gulbenkian's creative architectural corporate designs, once completed. Gulbenkian noted this lack of gratitude by remarking that "oil friends are slippery!"

In 1925, the new nation of Iraq signed an agreement with the Turkish Petroleum Company, to be renamed the Iraq Petroleum Company, whereby the government of Iraq would receive a royalty on any oil produced, if any were discovered, until 2000. At some point in the discussions, the government of Iraq was promised 20 percent participation, but the participation was excluded from the final agreement. This would be a bitter source of contention between the Iraq Petroleum Company, owned by the oil companies, and the host government of Iraq for nearly half a century.

All this maneuvering was merely an academic exercise because the Iraq Petroleum Company was a scrap of paper until the 1927 discovery of one of the world's largest oil fields. Gulbenkian now insisted that the concession granted by the Ottoman Empire was not restricted to Iraq but included all the lands under the former empire. Gulbenkian took a map and drew a red line over what he thought was the former Ottoman Empire, which included all of the Middle East (Turkey, Jordan, Syria, Iraq, and Saudi Arabia) except

Kuwait and Iran. Although the Ottoman Empire did control the religious centers along the Red Sea (what was to become Saudi Arabia) its control over the vast emptiness of deserts inhabited by nomads was nominal, to say the least. No one, including Gulbenkian, foresaw the implications of having what was to become Saudi Arabia within the Red Line Agreement. Signed in 1928, the Red Line Agreement stipulated that no oil field within the red line could be developed unless there was equal participation by the companies owning the Iraq Petroleum Company, which, of course, included Gulbenkian's 5 percent share.

As the only oil company with operating experience in producing Middle East oil, BP initially handled Iraqi oil production. Exxon and Mobil soon became more actively involved as did Compagnie Francaise de Petroles (CFP), a national oil company organized by the French government in 1924, modeled after BP, to handle its share of the Iraq Petroleum Company. The world of oil now had three governments involved with oil: the British government's half interest in an independently run BP with a concession in Iran and Iraq, the French government's wholly owned interest in an independently run CFP, with a concession in Iraq, and the Soviet Union, which exercised absolute control over its oil resources.

In addition to opening up the Middle East and playing second fiddle to Frederick Lane in bringing in Russian oil interests to the newly formed Shell Group, Gulbenkian brought Shell into the Turkish Petroleum Company and helped raise money for Shell as an intermediary with New York investment bankers. He also arranged contracts for Shell to supply the French and Italian governments with petroleum products during World War I. In 1918 he orchestrated the Shell takeover of Mexican Eagle Oil Company, the start of Shell's activities in Mexico. To further cement his relationship with Shell, Gulbenkian arranged for his son, Nubar, to become the personal assistant to Deterding. It was rumored at the time that Nubar might be in line to succeed Deterding, but his son's career with Shell ended abruptly some years later when Gulbenkian yanked him away to become his personal assistant.

Gulbenkian was asked to act on behalf of British investors with an oil concession in Venezuela called, appropriately, Venezuela Oil Concessions (VOC). Gulbenkian brought this investment opportunity to Deterding's attention, which ended up with Shell owning two-thirds of VOC and Gulbenkian and other shareholders, including Venezuelans, with the remaining third. Deterding believed that any investment made by Shell was to be controlled and run in the best interests of Shell. Deterding practiced what he preached. As majority and controlling owner, Shell was in a position to determine the price of oil exported from Venezuela. It was in Shell's financial interests to set a low price for the exported oil, but not in the financial interests of the minority VOC shareholders. Gulbenkian's failure to reach an agreement with Deterding on his behalf and the behalf of other minority shareholders eventually led to a breach between the two.

After World War II, Exxon and Mobil, the two remaining U.S. shareholders in the Iraq Petroleum Company, took on the decades-old task of squeezing out Gulbenkian's 5 percent share. During these discussions, Teagle (Exxon's president) referred to Gulbenkian as an oil merchant, to which Gulbenkian angrily responded that he was a business architect, not an oil merchant, a perfect description of his role in oil. He once pointed to a strange-looking ship in a harbor and asked what it was and had to be told that the ship was a tanker that might be carrying his oil! While oilmen had a jaundiced view of Gulbenkian, Gulbenkian's view of oilmen as cats in the night—"by their sound no one can tell if they're fighting or making love!" —was equally negative.

Finally Gulbenkian was informed that the 1928 Red Line Agreement was void because it violated American antitrust legislation, an interesting tactic on the part of oil companies, which were occasionally threatened by Congress for violating the same legislation. For Gulbenkian's alleged violation, Exxon and Mobil stated that they were no longer bound by the 1928 Agreement. The revised 1948 Agreement left Exxon and Mobil free to develop Saudi Arabian oil reserves on their own. Between 1948 and 1954, Gulbenkian negotiated a replacement for the Red Line Agreement from his various hotel suites. The 1954 agreement not only reaffirmed his 5 percent interest in the Iraq Petroleum Company, but he was also reimbursed for previous unpaid receivables.

Gulbenkian's annual succession of seventeen- and eighteen-year-old mistresses ended with his death in 1955 at the age of eighty-six. After his death, he bequeathed the bulk of his wealth and future revenue to a foundation based in Lisbon. In the end, the oilmen won. Gulbenkian's 5 percent share was wiped out with the nationalization of the Iraq Oil Company in the 1970s; but so too were theirs. Nevertheless, the Gulbenkian foundation has continued to prosper, with income from his shareholdings in an oil company with interests in the Middle East and elsewhere.

Early Attempts at Oil Price Controls

Rockefeller, of course, was the first to attempt to control prices, and he pretty much succeeded when he achieved 90 percent control over the U.S. refinery industry. His idea of an acceptable price for kerosene was the price that would not encourage outsiders to build refineries. Too high a price would only create more problems for Rockefeller by providing an incentive for others to get into the refining business. This idea is still alive. OPEC realizes that an oil price that is too high financially underwrites the development of high-cost non-OPEC oil fields that will eventually erode OPEC's market share.

The first to attempt to bring order to the oil industry on a global scale was the oil power brokers of the day, Teagle, of Exxon (a distant relative of Maurice Clark, Rockefeller's first partner) and Deterding, of Shell. In 1922 they stood together, along with others, to present a united front in dealing with oil sales by the Soviet Union, which they viewed as buying back stolen property. While the two power brokers were shaking hands and expressing mutual dismay over Soviet duplicity in expropriating oil properties without compensation, Deterding secretly purchased a large quantity of Soviet oil at less than the agreed price with Exxon, which he promptly dumped in the Far East. Subsequent attempts by Teagle and Deterding to restore some semblance of order sometimes worked and sometimes did not, but in 1927 Deterding abandoned any further pretext of cooperating with Exxon over the matter of Soviet oil. This time the reason was not related to oil, but to his second marriage to a White Russian. Cross-accusations between Teagle and Deterding eventually induced Deterding to start what turned out to be a disastrous price war. The Soviets thought that they had succeeded in creating chaos in the world oil patch by successfully playing one oil company off another, perhaps bringing back memories of the Nobels and the Rothschilds. Soviet satisfaction over spreading confusion in the capitalistic world of oil stemmed not so much from their conspiratorial plans, or Deterding's ill-fated venture into a price war, but from a world flooded with crude from the Soviet Union, Mexico, and Venezuela.

The 1920s started with a feeling that oil would be in short supply, so the U.S. government forced Exxon and Mobil to get involved with Middle East oil through its interest in the Turkish Petroleum Company. By the late 1920s, and continuing on through the global depression of the 1930s, the world was awash in oil. Something had to be done. Oil companies had made massive investments on the basis of a certain projected price of crude oil; as crude prices sank, so did the return on these investments. In 1928, in a Scottish castle, Deterding held a social affair that happened to include Teagle from Exxon and Mellon from Gulf Oil and other oil magnates, including the head of BP. This social affair led to a pooling arrangement to control price through cooperation in production and in sharing incremental demand among the cartel of supposedly competing oil companies. The reference price would be American oil in the U.S. Gulf, with adjustments to take into account freight from the U.S. Gulf.

Once this system was set up, other oil companies joined. If a participating oil company purchased oil in the Middle East and sold it in France, the selling price would not be the FOB price in the Middle East plus freight from the Middle East to France, but the price of oil in the U.S. Gulf plus freight from the U.S. Gulf to France. This system stabilized the price at a healthy level for the oil companies as long as others joined, which they did. With a mechanism in place for allocating incremental production to meet growing demand among the participating oil companies, the global oil business, with the exception of Soviet oil, was under the control of a cartel of oil companies. Of course, for those U.S. oil companies involved in this arrangement to fix pricing and production, this was in direct violation of the Sherman Anti-Trust Act. The Rockefeller dream of world control over oil, for the most part, had finally come true, but not with domination vested in the hands of an individual, but a small group of executives who, in the aggregate, controlled most of the world oil. The success of this agreement hinged on all the individuals continuing to cooperate, something rarely seen in the world of oil.

In 1930, only two years after the system was set up, price stability was threatened by yet another mammoth oil discovery. Like Drake and Higgens, an old wildcatter, Dad Joiner, persisted where others had given up. Joiner did not drill on land that had promising geologic characteristics, but on land owned by promising widows who might invest in Joiner's ventures. Joiner must have had a way with the widows for they were all financially disappointed with Joiner's ventures; except for one, on whose east Texas farm in Kilgore Joiner brought in a gusher. Joiner had proved the oil geologists wrong and Kilgore became another Pit Hole and Spindletop all rolled into one, with oil derricks almost on top of one another pumping with all their might. This strike would lead to the discovery of other oil fields in east Texas much larger than anyone imagined. Unfortunately, Joiner was in financial straits from his past ventures with widows and could not hold onto his holdings. Forced to sell out to H.L. Hunt, who made billions on Joiner's and other east Texas properties, Joiner was to die as poor as Drake and Higgens.

The east Texas oil boom, coming at the time of the Great Depression, created a glut and oil prices collapsed locally to ten cents a barrel. Teagle and Deterding were powerless because they did not control the east Texas oil fields. The Texas "independents" demanded federal and state intervention. The state governments of Texas and Oklahoma obliged and declared martial law on the basis that the independents were squandering a valuable natural resource, particularly at ten cents a barrel. Using conservation to justify the states' actions, and the local militia to enforce their will, oil production was stopped. Then the Texas Railroad Commission was authorized to set up a rationing system to control production. Although individual producers cheated whenever they could, the Texas Railroad Commission eventually got the upper hand over the producers and was able to ration production of individual wells and prices rose. This government action to protect and conserve a natural resource, which today would be viewed as environmentally desirable, served the interests of the global oil cartel as well. Thus, capitalism and conservation joined hands with a common objective, but different goals. Deterding's pooling arrangement and the Texas Railroad Commission's rationing of production stabilized the world price of oil and both were valuable lessons for OPEC when it gained control over oil prices and production in the 1970s.

Enter Saudi Arabia and Kuwait

With the price of oil reestablished by controlling east Texas production, the last thing the oil companies wanted was another east Texas discovery. Another oil rogue, New Zealander Frank Holmes, believed that oil was waiting to be discovered in Arabia. Gulbenkian's Red Line Agreement prohibited exploration in Arabia without the joint cooperation of the signatories. Socal, the name of Chevron at that time, was not a signatory of the Red Line Agreement, and for $50,000 bought Holmes's concession in Bahrain, an island nation off of Saudi Arabia, and in 1931 struck oil. While Bahrain would never become a major oil producer, it indicated that Holmes might be right about Arabia.

In 1927 the desert king Ibn Saud subdued his rivals along the Red Sea coastline and named his new kingdom after his clan. In 1930, desperate for money, King Saud inveigled Socal to buy a concession in Saudi Arabia. The major oil companies, bound by the Red Line Agreement and in no mood to discover more oil, passed up the opportunity to make a deal with King Saud. Socal did some exploration, which turned out to be promising; but short on capital in the event that oil were discovered, the company teamed up with Texaco, another nonsignatory to the Red Line Agreement. Texaco bought a half share of Socal's interests in Bahrain and Saudi Arabia. Eventually oil was discovered in Saudi Arabia and in 1939 King Saud opened up a valve and oil began to flow into an awaiting tanker. The king was so pleased that he increased Socal's and Texaco's concession to an area as large as Texas, Louisiana, Oklahoma, and New Mexico combined.

Frank Holmes was also involved with opening up Kuwait, which was also outside of the Red Line Agreement. Eventually BP and Gulf set up a joint venture after a fair degree of behind-the-scenes maneuvering by the British and U.S. governments. In 1938 oil was discovered. Although Frank Holmes was instrumental in opening up oil exploration in Bahrain, Saudi Arabia, and Kuwait, all successful finds, he made no fortune from the enormous wealth that he was instrumental in creating for the oil companies and producers. Originating and transforming a good idea to reality does not necessarily translate into personal wealth. This is the lesson of Drake, Higgens, Joiner, and Holmes; something else was needed.

Exit the Key Players

Hitler inadvertently took down three leading oil company executives. The first to fall was Deterding, who was showing signs of mental imbalance (megalomania) as his management style became increasingly dictatorial. In his memoirs, composed in 1934 in the midst of the Great Depression, when tens of millions of idle workers were desperately seeking work, he wrote that all idlers should be shot on sight. Upset over the loss of Shell properties in Russia after the revolution, Deterding's position against communism hardened with his second marriage to a White Russian and his third to a German. Deterding became a Nazi sympathizer because of their determination to rip communism out root and branch. Deterding would not be the only industrialist, statesman, monarchist, or church leader to support the Nazis for this reason. The board of directors removed Deterding from his position in 1936 by forcing him to retire, and he died six months before World War II started. Shell's penchant for collegiality and corroboration in the decision-making process might be partly in reaction to Deterding's last years of rule.

The second to fall was Rieber, the head of Texaco. In 1937 Rieber diverted Texaco tankers taking oil to Belgium to support Franco in Spain, and in 1940 he got around a British oil embargo against Germany by shipping oil to Germany from neutral ports. Unable to take money out of Germany, Rieber worked out a barter agreement whereby he accepted German-built tankers in exchange for oil. Rieber was forced to resign in 1940 in the wake of a British intelligence revelation that a Texaco employee was sending information to Germany about American war preparations.

The third to fall was Teagle, who had entered into an agreement before the rise of Hitler with I.G. Farben, a German chemical company. Farben was to research and develop synthetic rubber for Exxon in exchange for Exxon's patents for tetraethyl lead, a vital ingredient in aviation fuel. Teagle was unable to see the military implications of this arrangement even after Hitler's rise to power and after the Japanese had overrun the rubber plantations in Southeast Asia. Teagle refused to break what he considered first and foremost a business deal, which remained in force until revelations by the U.S. Justice Department led to his resignation in 1942.

All three were counterpoints to Marcus Samuel, who put civic duties and patriotism above business. Deterding, Teagle, and Rieber put business above all else. Buy for a little less here, sell for a little more there, was their key to success. Business plans were to fit the immutable laws of supply and demand. The name of the game is making money. Politicians come and go and have little use other than passing laws and establishing regulations that protect business interests or guarantee their success. Governments rise and fall, but business remains forever; it is the great constant.

Shareholders and Stakeholders

The modern corporation is based on the premise that its mission is maximizing shareholder wealth. One way to do this is to spawn new products and expand market reach to millions of individuals as Rockefeller did. Another way to maximize shareholder wealth is to widen the spread between the price received for a product and its cost of production, also a Rockefeller practice. While maximizing wealth for a corporation's shareholders is what the game is all about, there are other constituencies, or stakeholders, affected by the operation of a private corporation. For instance, an oil company has some degree of latitude concerning where profits are assigned. Profits can be shifted between upstream activities (crude oil production) or downstream activities (refining and marketing) through internal transfer prices. If an oil company has its oil fields, refineries, distribution system, and market within the borders of a single nation, such as the United States, it does not matter how profit is assigned internally when a company consolidates its financial statements and tax returns. The federal government collects the same in income taxes regardless of how internal transfer prices are set, although internal transfer prices can affect state income taxes. When an oil company is buying crude oil from one nation, processing it in a second, and selling in a third, the internal assignment of profits through transfer pricing can heavily influence taxes and royalties paid by oil companies to host governments. This in turn affects the well-being of the people of oil-exporting nations, who are, in every sense of the word, stakeholders in a company that is exploiting their nation's natural resources.

Deterding noted the importance of the triangle linking the mutual interests of an oil company with the people and with the host government in which all three should benefit from developing a nation's oil resources.[16] Although Shell operated in Mexico, the government and the people felt they were getting a raw deal from the oil companies and, in 1938, nationalized the industry. The oil companies struck back by refusing to buy Mexican oil until they received restitution, which Pemex, the newly formed national oil company of Mexico, was forced to pay in order to gain access to foreign markets. Now two nations directly controlled their oil resources: the Soviet Union and Mexico. Yet the oil companies did not learn the essential lesson of Mexico: A one-sided relationship in which an oil company exploited the oil resources of a nation with limited benefit to the people or the government was not in the best long-term interests of the oil company. No one viewed Mexico as a harbinger of more to come when new oil discoveries in Venezuela diverted oil company attention from Mexico.

Development of Saudi Arabia's Oil Fields

Saudi Arabia was the answer to Washington's worry, one that had first bothered Theodore Roosevelt and would come back now and then to haunt government energy policymakers: The world was going to run out of oil. Socal and Texaco operated in Saudi Arabia under the corporate umbrella of Aramco, the Arabian-American Oil Company. Socal and Texaco advanced the idea during the early years of World War II of the U.S. government setting up a Petroleum Reserve Corporation to buy a controlling interest in Aramco and constructing a refinery on the Persian Gulf. The idea was well received by Franklin D. Roosevelt, who, like Churchill, was attracted by the idea of government ownership of a foreign oil field. However, the oil companies abruptly broke off negotiations in 1943. Only in hindsight can one see the timing between the success of Rommel in North Africa and the proposal for the Petroleum Reserve Corporation and Rommel's defeat in 1943 with the proposal's demise. Obviously, oil company investments in the Middle East would be in danger if Rommel succeeded in his master plan to link his army in North Africa with Hitler's in Baku. Oil companies generally oppose government intervention in their operations unless, of course, such intervention promotes their agenda.

The U.S. government then proposed constructing a thousand-mile pipeline to carry Saudi crude to the Mediterranean and the oil companies would guarantee a 20 percent interest in the oil fields as a naval reserve. The Trans-Arabian Pipeline (Tapline) pipeline was completed, without U.S. government involvement, in 1950 when Saudi crude was loaded on a tanker in Sidon, Lebanon. The pipeline, passing through Saudi Arabia, Syria, and Lebanon, was shut down in 1975 during a time of turmoil in Lebanon. However, the pipeline's capability of carrying oil cheaply to Europe when in operation meant a great deal to Socal and Texaco.

Having achieved such success in Saudi Arabia, Socal and Texaco passed up an opportunity to become dominant players in the oil business by not wanting to challenge the other major oil companies. They felt that involvement of the other major oil companies was necessary for access to oil markets, capital to develop Saudi oil resources, and garnering diplomatic support if there were an unfriendly successor to King Saud. Admitting Exxon and Mobil and excluding the other signatory oil companies violated the Red Line Agreement. Using American antitrust legislation as a lame excuse, Exxon and Mobil walked away from the Red Line Agreement and joined Aramco, thereby locking BP, Shell, and CFP out of Saudi Arabia.

Aramco proved to be a model for a company operating in a host nation. Its employees had their own town and concentrated on the business of finding, developing, and operating the oil fields and building and running refineries, pipelines, and terminals. By any measure, Aramco was considered a "good corporate citizen." Aramco permitted the United States to have two allies diametrically opposed to one another. The state department dealt directly with Israel and, when necessary, used Aramco as a go-between in its dealings with Saudi Arabia. In the twenty-first century the company is known as Saudi Aramco, with 54,000 employees of whom 86 percent are Saudis. The company prides itself on its ability to manage Saudi energy resources and contribute to the nation's development.

The Shoes Begin to Fall

It is one matter when producers supply 10 percent of the world's oil, which can easily be replaced by other sources. This keeps the producers in a weak bargaining position as they learned in Mexico. It is another matter when their share grows to 30–40 percent, which no longer can be replaced; then their bargaining position is not quite so weak. The oil companies failed to realize the growing bargaining strength of the oil producers that accompanied the growing world dependence on foreign oil. The next shoe to fall after the Mexican nationalization of its oil industry came in 1948, when Venezuela passed a law for a 50:50 sharing of profits, an idea of Juan Pablo Perez Alfonso, the Venezuelan oil minister and chief architect of OPEC. The idea was not total anathema to the oil companies if sharing profits meant forestalling nationalization as had occurred in Mexico (better to have half than none). Moreover, the oil companies had the power to define profitability by how they allocated profits through internal transfer pricing.

King Saud, whose huge family's lifestyle had become incredibly expensive, joined the fray and demanded a share of the profits. Aramco turned to the U.S. government for support, and the government, fearing a communist takeover in the Middle East, agreed to have the Aramco partners treat the additional payments to Saudi Arabia as a foreign income tax. This was a great boon to the Aramco partners because this meant, under rules on double taxation, that taxes paid to the U.S. government would decrease one dollar for every extra dollar in taxes paid to Saudi Arabia. In other words, the U.S. government, hence U.S. taxpayers, was subsidizing the extra cost of oil. Such a ruling could not be restricted to some oil companies, equal treatment demanded that this apply for all. The upshot of this ruling was that it became more profitable for oil companies to develop oil properties overseas than domestically. The oil companies could shift a part of what they were paying foreign suppliers in the form of taxes to reducing their U.S. taxes, something that would not apply to a U.S. source of supply. Another tax bonanza for the oil companies was applying the oil depletion allowance to foreign as well as domestic sources of oil. These two tax rulings placed oil companies in a quasi tax-free environment at that time, which is not true today.

The Next Shoe to Fall

BP, still half-owned by the British government, had expanded into activities far beyond those envisioned by Churchill. While its principal source of oil was still Iran, BP had a major position in Iraq and Kuwait and had developed a worldwide marketing network served by its fleet of tankers. In 1951, a new Iranian leader appeared on the scene, Mohammad Mossadegh, who called for nationalization of Iranian oil fields after BP's refusal to adopt a deal similar to that between Aramco and Saudi Arabia. The Iranian prime minister, who opposed Mossadegh, stated that he would not allow Iran to repudiate its concession with BP. That remark caused his assassination, opening the way for Mossadegh to become prime minister and nationalize BP's oil fields. The Labor Party, then in power in Britain, was hardly in a position to enforce this legacy of colonialism. With no help from the British government, BP took legal action, not in Iran, but in every nation where a cargo of Iranian oil landed. This lasted two years. By then the civil unrest that resulted from the loss of revenue led to a coup, encouraged by the CIA, which reinstated the son of a previous shah. In 1954 an agreement was hammered out whereby the National Iranian Oil Company, formed by Mossadegh, would remain owner of the oil fields along with the Abadan refinery. However, the oil would be sold through a consortium in which BP had a 40 percent share, Shell 14 percent, with the rest divided among CFP and the five remaining American sisters. In other words, the seven sisters, eight counting CFP, had total market control over Iranian oil production. The agreement taught the oil companies their first lesson: ownership of an oil field is not nearly as critical as access to its oil.

Later on five smaller U.S. oil companies inveigled a 5 percent share. Among these were Getty Oil and Tidewater, both owned by Jean Paul Getty. Getty was the son of a lawyer who struck it rich in oil in Oklahoma. The son was just as talented, if not more, as his father. Getty became a billionaire, partly as a result of his flying with an oil geologist over the Neutral Zone between Saudi Arabia and Kuwait. The Kuwait side of the Neutral Zone was already producing oil. The geologist noted from the air that a certain sector of the Neutral Zone in Saudi Arabia had a geology similar to that of the oil-producing sector in Kuwait. Getty immediately started negotiating with Ibn Saud for a concession. Drilling revealed a huge oil field, big enough to make Getty a billionaire and for the geologist to be reimbursed for his travel expenses.

Besides Getty there was Hunt, another billionaire not given to sharing with those responsible for his wealth (Higgins comes to mind), and Armand Hammer. Hammer had received a medical degree, but did not practice medicine, as his father had, who had befriended Lenin. Hammer took advantage of his father's relationship with Lenin to make commercial deals in the Soviet Union, including setting up a pencil factory and purchasing Russian art treasures for pennies on the dollar. Hammer, at an age when many contemplate retiring, got interested in oil and eventually took over a small oil company called Occidental Petroleum. By dint of his determination and driving force, Hammer transformed Occidental Petroleum

Table 5.1

Shareholders' Percentage

	Iran Consortium	Iraq IPC	Saudi Arabia Aramco	Kuwait KOC	Abu Dhabi Petroleum
BP	40	23.750	—	50	23.750
Shell	14	23.750	—	—	23.750
Exxon	7	11.875	30	—	11.875
Mobil	7	11.875	10	—	11.875
Gulf	7	—	—	50	—
Texaco	7	—	30	—	—
Socal	7	—	30	—	—
CFP	6	23.750	—	—	23.750
Others	5	—	—	—	—

into an international oil company with the discovery of three major oil fields in Libya. Hammer would play a pivotal role in the oil crisis of 1973.

Another thorn in the side of the seven sisters was Enrico Mattei, head of the Italian State Oil Company, who was able to prick the seven sisters by negotiating an independent concession with the Iranian National Oil Company (NIOC) in 1957 and making a private deal with Khrushchev for cheap Soviet oil, much as Deterding before him had done. The seven sisters then had to contend with CFP's discovery of oil in Algeria. New discoveries of supply remained ahead of rapidly growing demand. Despite the best efforts of the seven sisters to keep production matched with demand to sustain prices, there was a glut of oil on the market and oil prices remained cheap. Unbeknownst to the Iranian government, the oil companies in the consortium that purchased the output of the NIOC made a secret side-agreement to reduce Iranian sales in order to avoid a global glut of oil. Neither the shah nor the NIOC knew about this agreement, which effectively made Iran a swing producer to maintain world oil prices.

This perhaps marked the zenith of oil company power. The oil companies had reinstated their position in Iran even though their properties had been nationalized by preventing access to the world market, the same trick they had used in Mexico. Mossadegh's political demise served as a warning to other interlopers. Notwithstanding the success of Hunt, Getty, Hammer, and Mattei, there were limited opportunities for third parties to reach the market unless they went through one or more of the seven sisters. The seven sisters exerted the power of Rockefeller's horizontal monopoly on a global scale. Table 5.1 lists the shareholders of the various Middle East oil concessions in play up to the eve of the 1973 oil crisis.

Nasser's 1956 takeover of the Suez Canal did not affect the oil companies as much as it created fortunes for tanker owners. Because it took longer to get the oil around South Africa, Humble Oil, the Texas subsidiary of Exxon, took advantage of the temporary shortage of oil in Europe and raised crude prices by thirty-five cents per barrel. This incurred the wrath of Congress, which from a contemporary perspective appears ludicrous when price changes of thirty-five cents per barrel are hardly noticed. Of course, thirty-five cents per barrel of oil when it cost around $2 per barrel was a large change in percentage. What this showed was a major consuming government's keen interest in keeping a lid on oil prices; in fact, one might conclude that consuming governments depended on oil companies to keep a lid on oil prices. Keeping communists out of the oil-producing nations and keeping oil prices low for consumers were the reasons why the U.S. government never seriously pursued antitrust actions against the American oil majors, which clearly violated the Sherman Anti-Trust Act when they cooperated with competitors to fix prices and limit production. The British government took a far more pragmatic view of the situation and did not share the U.S. government's vexation when oil companies attempted to stabilize something as critical to the world economy as oil.

The Birth of OPEC

By the late 1950s cheap Soviet crude was cutting into the seven sisters' markets in Italy, India, and Japan. The seven sisters had to lower their prices in these nations to maintain their market presence, which, of course, meant lower profit margins. In 1959, Exxon resolved that it must cut posted prices to oil producers to preserve its profit margin. When the other oil companies followed suit, the Arab oil producers organized the first meeting of the Arab Petroleum Congress, the fruit of private talks between the oil ministers of Venezuela and Saudi Arabia. A second round of Exxon-inspired cuts provoked a stronger surge of unity among the oil producers. Another meeting in 1960 of the oil ministers of Saudi Arabia, Iran, Iraq, Kuwait, and Venezuela gave birth to the Organization of Petroleum Exporting Countries (OPEC). The purpose of OPEC was not to raise oil prices, but to prevent further reductions in posted prices. The original unity of purpose was gone by the second OPEC meeting in 1961, when a rough and tumble battle broke out among OPEC members as each sought to garner a larger export volume at the expense of others. OPEC was behaving no differently than the earliest oil drillers in Pit Hole; it was every man for himself.

By no measure could OPEC be considered a success during the 1960s. There was little coordination among the members and politics kept getting in the way of negotiations. Meanwhile, new sources were coming onstream, such as Nigeria, putting more pressure on OPEC's approach to maximizing revenue by maximizing production, another reminder of Pit Hole. In 1965, OPEC failed at an attempt to gain control over future increases in production just as it failed to gain control over current production. The seven sisters meanwhile were trying to restrain production to prevent further declines in oil prices. The irony is that in only ten years, OPEC would take over the oil companies' role of restraining production to control prices. The role reversal would not be complete as the OPEC idea of price in the 1970s would be radically different than that of the oil companies in the 1960s.

The 1967 Six-Day War between Israel and Egypt sparked the first Arab boycott. The war was over before the boycott had any effect, which was doomed anyway when Venezuela and Iran refused to join. The formation of the Organization of Arab Petroleum Exporting Countries (OAPEC) within OPEC in 1970 did not succeed in strengthening the resolve of OPEC to bring order to the oil market. Order, of course, meant maximizing the respective production volume of each member to maximize revenue. Oil company attempts to rein in production to maintain prices, which varied for each member of OPEC, irritated the oil producers, who now had to contend with new oil production from Qatar, Dubai, Oman, and Abu Dhabi.

In 1970 the Alyeska Pipeline Company was formed to handle the 1968 oil discovery by Arco (then Atlantic Richfield) in Prudhoe Bay on the north slope of Alaska. Compared to the Middle East exporters, this is expensive oil. Arco, short on crude, viewed the development of the North Slope field as vital to its survival. Two other major participants were Exxon and BP, the latter having acquired Sohio to gain greater access to the U.S. market. These two companies, with more cheap Middle East oil than they wanted, did not need expensive North Slope oil. At first the environmentalists were successful in blocking the building of an 800-mile pipeline to Valdez. Congress set an interesting precedent by overriding environmental concerns in the wake of the 1973 oil crisis and authorized the construction of the pipeline. Alaskan oil began flowing in 1977.

Another source of high-cost oil was the 1969 discovery of the Ekofisk oil field in the Norwegian sector of the North Sea by Phillips Petroleum. This was followed a year later by the BP discovery of the Forties field north of Aberdeen and the following year by the Shell and Exxon discoveries of the Brent field off the Shetland Islands. The involvement of Exxon, BP, and Shell in oil fields far more costly to develop than buying Middle East crude, intentionally or unintentionally, could be interpreted as manifesting their concern over the rising dependence on Middle East oil.

The 1973 oil crisis was not caused by a shortage of oil. Indeed, the greatest worry right up to the eve of the crisis was how to keep new production from flooding the market and further weakening oil prices. The producers were worried about anything that would shrink their export volumes. The shah of Iran wanted to increase export volumes in order to expand Iran's military power and rapidly develop its economy, and saw his role as a guarantor of stability of the Middle East, for which he had received President Richard Nixon's blessing.

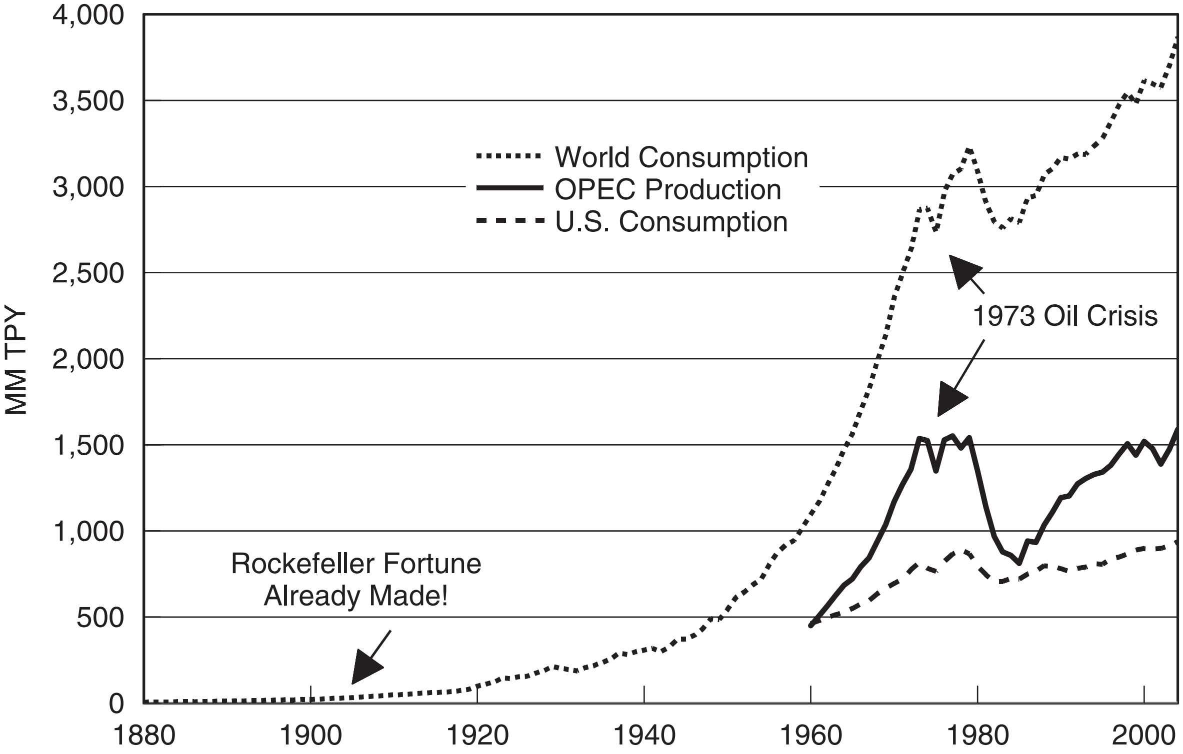
The 1973 Oil Crisis

Figure 5.1 shows the growth of world oil consumption from the beginning of the oil age, with U.S. oil consumption and OPEC production since 1960.[17]

From the birth of the automobile age around 1900, oil consumption began to double about every decade. Even the Great Depression did not dampen growth in oil consumption, but the age of oil did not begin in earnest until after World War II, when successive doublings really started to kick in (one penny doubled is two pennies, two pennies doubled is four, doubled again eight, doubled again sixteen, doubled again thirty-two). The slopes of the curves for both world oil consumption and OPEC production appear to be about the same from 1960 to 1973, which implies that nearly all incremental oil demand was coming from the OPEC nations. A closer examination of the figures, however, reveals that in 1960 OPEC was supplying 38 percent of world oil consumption, 47 percent in 1965, and 56 percent in 1973, meaning that OPEC exports were growing faster than world oil demand. The slope of U.S. consumption is less than world consumption, which implies that oil consumption was growing faster elsewhere in the world before the 1973 crisis, which was true in Western Europe and Japan. While on the surface it appears that the United States was not a major contributor to the 1973 oil crisis, the fact is that the nation was heavily responsible as it made the transition from being the world's largest oil exporter to the world's largest oil importer, as shown in Figure 5.2.

The early 1970s was a period of rapidly rising U.S. oil imports, of which a greater portion was coming from the Middle East. The oil crisis halted growth in U.S. consumption and twenty years were to pass before U.S. consumption would surpass its 1978 peak. Middle East exports around 2000 were just about back to where they were in

Figure 5.1 **Growth in World and U.S. Oil Consumption and OPEC Production**

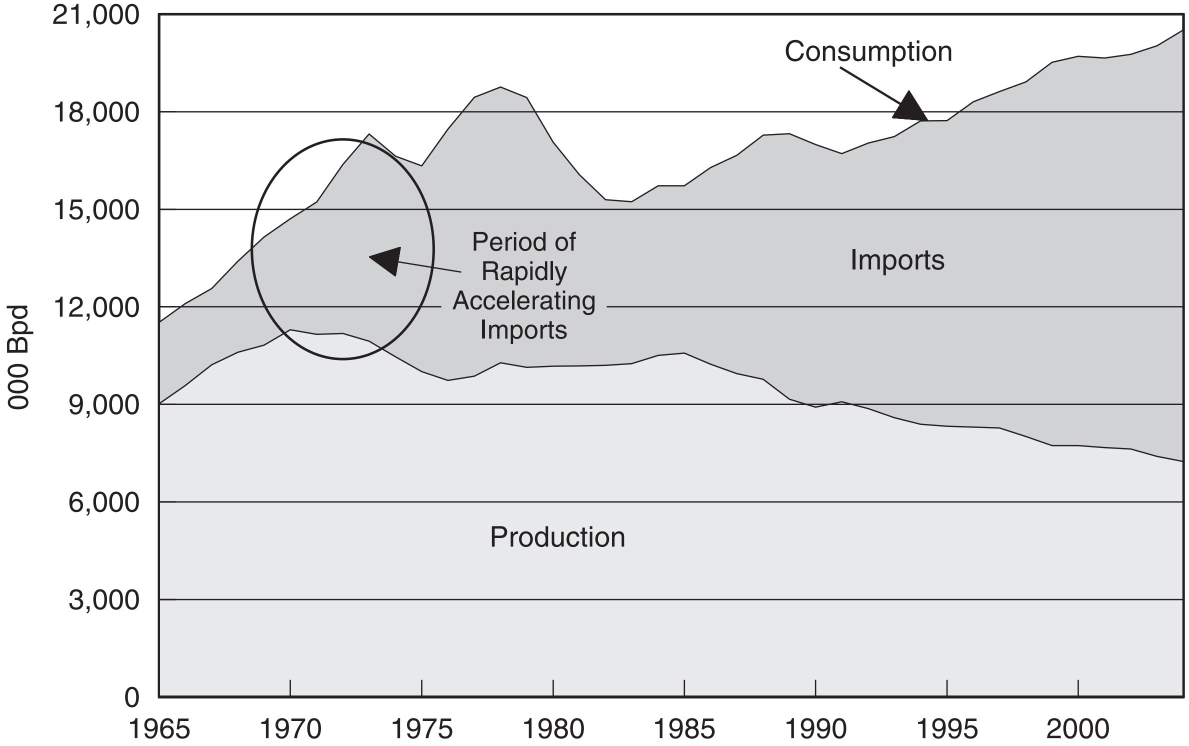


the late 1970s. While the United States is criticized as the energy hog of the world, Figure 5.1 suggests that it had been much worse in the past. The U.S. portion of world oil consumption was 42 percent in 1960, which had declined to 25 percent by 2004. Obviously, incremental growth in oil consumption has been concentrated elsewhere.

The high point of oil company ascendancy over national powers was the BP-inspired embargo against Mossadegh that led to his fall from power in 1953 and brought Iran to heel. Between then and the 1973 oil crisis, there was a shift from a buyers' to a sellers' market that occurred without public fanfare. The question raised by Figure 5.2 is: Why did it take so long? Another way of putting it, from the consumers' perspective, would be that the oil companies should be congratulated because they had kept oil prices low for as long as they did. Yet, there had to be an underlying malaise with respect to the situation. Why else would major oil companies start searching for oil in such high-cost areas as the North Slope and the North Sea?

The underlying shift from a buyers' to a sellers' market needed a precipitating event to make it manifest. Actually, it was a series of events that started with Colonel Gadhafi's successful military coup in Libya in 1969. At this time, Libya was supplying

Figure 5.2 **U.S. Oil Consumption and Production**



about one-quarter of Europe's needs with high-quality and low-sulfur crude. Moreover, Libya is located on the right side of the Suez Canal. The Canal was closed in 1956, then reopened in 1957 when Nasser nationalized it, then closed again in 1967 during the Israeli-Arab War and not reopened until 1975. Libya received no premium for its oil, considering its quality and nearness to market. Gadhafi was not to be cowed by the major oil companies' resistance to any price change. In 1970 Gadhafi struck at the weakest link in the supply chain: the independents, particularly those dependent on Libyan crude. Of these, the most dependent was Occidental Petroleum. Gadhafi chose his target wisely.

Hammer pleaded with the majors to sell him replacement oil at the same price he was paying for Libyan oil. In their shortsightedness, they offered Hammer higher-priced oil. Facing a disastrous interruption to supply, Occidental gave in to Gadhafi's new price and tax demands, which were relatively modest from today's perspective. Flushed with victory, Gadhafi went after the majors. To everyone's surprise, the majors did not embargo Libyan crude and replace it from other sources as they had with Mexico and Iran. Instead, they capitulated to Gadhafi's demands, a stiff price to pay for not coming to Hammer's aid. The producers now sensed that a fundamental change had taken place in the market. The world was shifting from a buyers' to a sellers' market.

As a consequence of Gadhafi's success, a hastily convened OPEC meeting in Caracas in late 1970 agreed to higher minimum taxes and higher posted prices that, when announced, only made Gadhafi leapfrog with even greater demands, followed by Venezuela. This infuriated the shah because it challenged his leadership. To shore up the resistance of the independents to further OPEC demands, the majors agreed that appropriately priced replacement oil would be provided to the independents to prevent them from caving in to producer demands. It was too late.

With U.S. government support, the oil companies attempted to get the oil producers to agree to common terms and to moderate their demands, that is, to get control over Gadhafi. A meeting was held in Tehran in 1971 attended by delegates from the oil-producing nations, the oil companies, and the U.S. State Department. The shah insisted that Libya and Venezuela not attend. The majors hoped that the presence of the State Department would aid in their negotiations, but it proved to be a weak straw. The State Department wanted to avoid a confrontation between the oil companies and producers because it depended on Iran and Saudi Arabia to act as regional police to suppress radicals who espoused communism. The State Department and the oil majors were not on the same page. Similarly, government representatives of several European nations and Japan proved equally unable to influence the outcome. Without strong government backing, and considering the importance of OPEC oil in the general scheme of things, the oil companies made no new demands and shifted their approach from confrontation to a call for moderation. It was now a matter of damage control.

The capitulation of the oil companies to the oil producers was the final piece of evidence that convinced the oil producers that the market had shifted in their favor. One top oil executive publicly quipped that the buyers' market was over. The agreed price increase in February of 1971 was an extra thirty cents per barrel on top of the posted price, escalating to fifty cents per barrel in 1975. This price adjustment held for the Gulf producers; now a meeting was necessary with Libya. A separate Tripoli agreement, signed six weeks after the Tehran agreement, called for a higher price without Libya providing a similar guarantee on future prices. The shah was infuriated by Gadhafi's leapfrogging over what he had agreed to.

Whereas the 1960s were years of worry over looming oil gluts, the early 1970s were years of a growing concern over a potential shortage, a reversal of the change in perception that occurred between the early and late 1920s. This change in sentiment spurred the oil producers to increase their demands for part ownership of their natural resources in the two-year hiatus between the Tehran agreement and the oil crisis of 1973. The oil producers felt that the original concessions granted to oil companies belonged to a bygone age of colonialism and imperialism. They wanted to move into the modern era and control their national resources through joint ownership rather than merely collecting taxes on their exports. The producers favored joint ownership with the oil companies over nationalization because nationalization removed the oil companies' incentive for making money in the upstream, or production, side of the business. By limiting their profits to the downstream side of refining and marketing, oil companies would only be interested in buying crude at the cheapest price and the producers would be back to undercutting one another as the only way to attract an oil company's attention.

Joint ownership turned out to be an idle thought. The British withdrawal of their military presence from the Middle East in 1971 created a power vacuum that allowed Iran to seize some small islands near the Strait of Hormuz. Gadhafi used Iranian aggression as an excuse to nationalize all of BP's holdings in Libya, along with Bunker Hunt's concession, and then 51 percent of the remaining concessions, including Hammer's. Algeria and Iraq joined in the frenzy of nationalizing oil assets. In early 1973 the shah announced his intention not to have the NIOC renew its operating agreement with the oil companies when it expired in 1979, and to transform NIOC from a domestic oil producer into a major global oil company.

By making separate deals with oil companies, the oil producers were fast learning how to play one of the seven sisters off another just as effectively as the seven sisters used to play one producer off another. The oil companies were beside themselves as their oil fields and physical assets were transferred from their books unto the books of the oil producers. They were at loggerheads over an approach that would minimize their loss of power and enable them to obtain restitution. Their appeals to the U.S. government for help were interpreted as a sign of weakness. Then the independent oil companies broke ranks with the seven sisters and began a bidding war to assure their oil supplies, another sign of weakness. The imposing facade of oil company power was exposed for what it was: an imposing facade.

With governments standing helplessly aside, the oil companies prepared to meet with the OPEC producers in Vienna in October 1973. The meeting took place just as Syria and Egypt invaded Israel, hardly an auspicious omen. The meeting broke down when the oil producers demanded a price hike to $5 per barrel. The oil companies played a weak hand and tried to refer the matter to their respective governments before making a formal reply. Oil companies had never appealed to their governments for permission before, so why now unless they were in desperate straits? Shortly after, in mid-October, King Faisal delivered an ultimatum to Nixon: immediate cessation of U.S. military aid to Israel or face an embargo. The ultimatum arrived just as the U.S. Senate had overwhelmingly voted to send reinforcements to Israel.

Events were now entirely out of the hands of the oil companies and consuming nations. In quick response to continued U.S. military support of Israel, the members of OPEC meeting in Kuwait unilaterally raised the price of a barrel of oil from $3 to $5, accompanied by a 5 percent cutback in production. The oil weapon, mentioned in the past, was now taken out of its sheath for the first time. The production cut was intended to sway the United States not to continue supporting Israel. Then, three days later Saudi Arabia announced a 10 percent cutback in production plus an embargo of oil to the United States and the Netherlands. This embargo had to be carried out by the oil companies themselves, even though a majority of them were U.S. companies. Of course, Saudi Arabia could not stop the oil companies from supplying oil to the Netherlands and the United States from other sources. Nevertheless, the embargo created a hiatus in oil moving into the United States, resulting in long lines at gasoline stations in November only a month later. The irony was that on October 21, when the embargo went into effect, Israel agreed to a cease-fire. But the Humpty Dumpy of the old world could not be put back together again. The oil companies made fruitless attempts to regain control over market prices. The first oil shock reached its apogee in December, when independents panicked over oil supplies and Iran conducted an auction with the highest bid coming in at $17 per barrel.

One argument advanced for raising oil prices by the oil producers was the fact that European governments collected more in taxes on a barrel of crude than what they received for selling a finite and depleting resource (as will be seen at the end of this chapter, this relationship still holds). Another was that when oil displaced coal, it proved that oil was under priced with respect to coal. Hence, it was in the long-term interests of energy consumers to reinstate coal as a source of energy, which could be accomplished, according to the shah, if crude were priced at $11.65 per barrel, the price necessary to make oil products from coal and shale oil at that time. The benefit to consumers was that a higher price of oil would cut oil consumption and postpone the time when the world would run out of oil.

The shah was absolutely right. If the oil crisis had not happened, there presumably would have been three more doublings between 1973 and 2003. World oil consumption was 2,750 million tons in 1973; three doublings is a projected consumption of 11,000 million tons compared to the 3,637 million tons consumed in 2003. An oil crisis was inevitable at some point because there is no way for production in 2003 to triple to accommodate a continued doubling of consumption every decade.

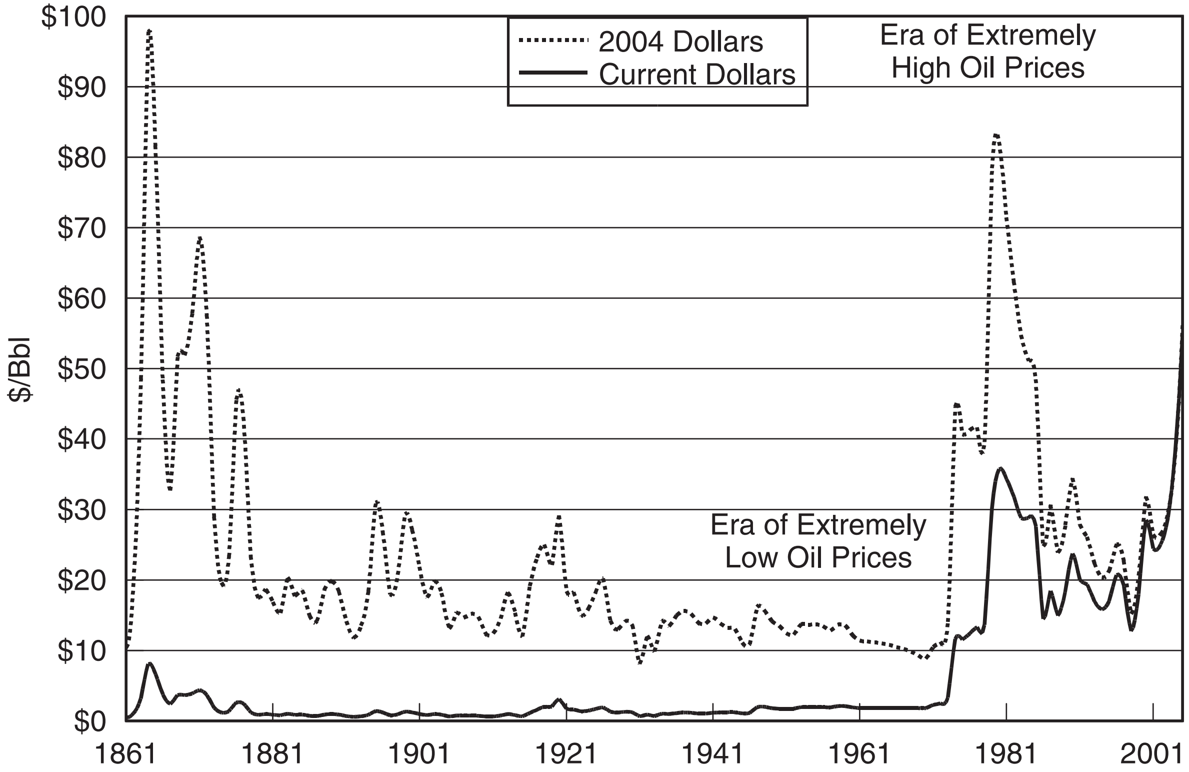
As the shah was justifying why oil prices had to be increased, an oil auction held in Nigeria fetched a whopping $23 per barrel, although the winner did not show up to take delivery. At the end of 1973, with an OPEC meeting to determine the appropriate price for a barrel of oil, the shah of Iran unilaterally announced a price of $11.65 per barrel, much to the chagrin of the other producers.[18] Even though the shah would be accused of moderation in a sea of immoderation, his price still represented a doubling of the then-posted price and a quadrupling of the posted price only a few months earlier. He accompanied his announcement of the new price with the warning that Western living styles would have to change and everyone would have to learn to work harder. The world no longer had to face a cartel of oil companies, but a cartel of oil-producing states. The greatest transfer of wealth in history was about to occur.

The Era of High Oil Prices

Figure 5.3 shows the history of oil prices in constant and current dollars. Current dollars are the dollars of the day; the actual price of oil paid at a point in time. Constant dollars reflect the purchasing power of 2004 dollars. Crude prices expressed in constant 2004 dollars are higher than in current dollars, reflecting the loss of purchasing power of dollars through inflation.[19]

In terms of constant 2004 dollars, the all-time peak in oil prices occurred in 1864, when prices expressed in constant 2004 dollars were just below $100 per barrel. This explains a lot about Pit Hole. With Rockefeller in control by 1880, oil prices in terms of 2004 dollars averaged $18 per barrel, and varied between $12–$30 per barrel until 1910. From 1910–1930 the average was $17 per barrel with the same range. Thus, oil prices were fairly stable for a half century. The Great Depression of the 1930s saw average prices decline to $13 per barrel, ranging between $8–$16 per barrel, again in 2004 dollars. The 1940s and 1950s were a continuation of depression prices, averaging $13 per barrel with a narrow range between $11–$16 per barrel. This is another thirty years of essentially constant prices that were close to those of the Depression. The 1960s up to 1972 was the absolutely worst period for oil producers, with an average price of $10.60 per barrel, ranging between

Figure 5.3 **History of Middle East Crude Oil Prices**



$9–$12 per barrel. It is ironic that the oil producers were facing the lowest prices in the history of oil while export volumes were virtually exploding. As long as exploding export volumes stayed ahead of exploding import volumes, the oil companies could maintain the upper hand. As soon as exploding import volumes got ahead of exploding export volumes, which happened when Saudi Arabia imposed its embargo, all hell broke loose.

After the 1973 price hikes, the shah now had the means to make Iran the military powerhouse of the Middle East, transform its economy to that of a modern state, and make NIOC a global oil powerhouse. Rather than giving him the means to pursue his grandiose dreams, all he got for his financial bonanza was exile (he went on a vacation from which he never returned in early 1979). The Iranian Revolution, which broke out in 1978 as national strikes, ended in 1979 with the ascendancy of Khomeini, a cleric with a decidedly anti-Western bent. The Iranian Revolution marked the second oil shock when the cessation of Iranian crude exports of over 5 million barrels per day (bpd) caused oil prices to climb precipitously to over $40 per barrel on a spot basis, prices that would not be seen again until 2003, in terms of current dollars, but not in constant dollars. Even the price peak of above $70 per barrel in 2005 did not exceed the peak $82 per barrel price in 1980 in constant dollars.

The cessation of Iranian production and the accompanying panic buying and hoarding brought about a reoccurrence of long lines of automobiles at gasoline filling stations. As Khomeini was finding his way around Tehran, Saddam Hussein staged a coup and made himself dictator of Iraq. Two years later, in 1981, Saddam cast his eye on Khomeini's army, whose weapons were no longer being supplied by the United States, and whose officers, commissioned by the shah, had been purged and replaced by loyal, but untrained, revolutionaries. Saddam decided that Khomeini's army, unlike the shah's, was no match for Iraq's army, newly equipped by the Soviet Union, and invaded Iraq.[20]

While the Iranians and Iraqis were waging war and Saudi Arabians were having problems digesting their newfound wealth, changes in the world of energy were at work that would come back to haunt the oil producers. Among these was a worldwide economic decline that reduced overall energy demand. High oil prices instigated a desperate search for alternative sources to oil, leading to a resurgence of coal, an accelerated pace in building nuclear power plants, a greater reliance on natural gas and anything else not called oil, including wood-burning electricity-generating plants. There were great gains in energy efficiency where cooling a refrigerator, heating a home, running an automobile, truck, locomotive, marine, or jet engine could be achieved with significantly less energy. Conservation of energy took the form of keeping indoor temperatures higher in summer and lower in winter, driving the family car fewer miles, and recycling energy-intensive products such as glass, aluminum, and paper. Companies set up energy managers to scrutinize every aspect of energy use in order to identify ways to reduce consumption.

In addition to slashing demand, high-priced oil caused an explosion in non-OPEC crude supplies, best exemplified in the North Slope of Alaska and in the North Sea. The North Slope of Alaska is an inhospitable place to develop and operate an oil field and necessitated the construction of an 800-mile-long pipeline to the port of Valdez over mountain ranges and tundra. North Slope production peaked at 2 million bpd a few years after the pipeline started operating in 1977. The North Sea was an even greater challenge with its hundred-knot gales and hundred-foot seas. Floating oil-drilling platforms explored for oil in waters a thousand feet deep. "Oceanscrapers," structures higher than the Empire State Building, were built on land, floated out to sea, and flooded (carefully) to come to rest on the bottom as production platforms. North Sea oil started with 45,000 bpd of output in 1974 and grew to over 500,000 bpd in 1975, to 1 million bpd in 1977, to 2 million bpd in 1979, to 3 million bpd in 1983, eventually peaking at 6 million bpd in the mid-1990s. Every barrel from the North Slope and North Sea was one barrel less from the Middle East.

Oil exporters dictated prices after the 1973 oil crisis, but continually changing prices implied that OPEC could not control the price as well as the oil companies had. When oil prices fluctuate widely, no one knows, including the oil producers, what will be tomorrow's price. This provides speculative opportunities for traders who try to outwit or outguess oil producers. All they needed was a place where they could place their bets. Once the traders started placing bets, buyers and sellers of oil had an opportunity to hedge their investments against adverse price changes.

Future and forward contracts of commodities with wide price swings were already traded, providing buyers with a means to hedge against the risk of a rising price and sellers a means to hedge against the risk of a falling price. The first futures were traded in grain in the nineteenth century. Grain growers could then short the futures market and lock in their revenue whereas bakers could buy futures and lock in their costs. Futures then spread to other agricultural products and industrial metals to stabilize prices, provide a means of hedging against price swings, and function as chips in a gambling casino for speculators, whose buying and selling add depth to the market. There was no reason to have a futures contract in gold, interest, and currency exchange rates when these were essentially fixed by government fiat. As governments lost control over gold prices, interest, and currency exchange rates during the 1970s, future contracts were developed to help buyers and sellers deal with the risk of price and rate volatility.

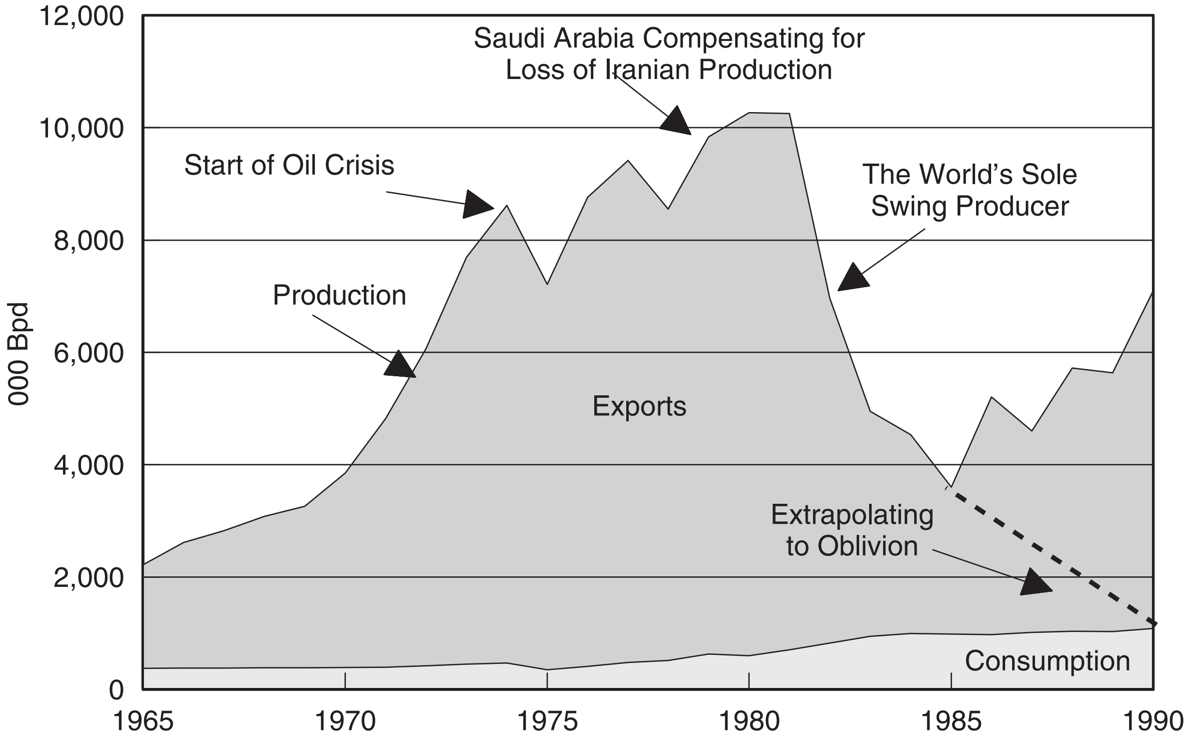
When oil companies controlled oil prices within a narrow range, there was no point in having futures. When they lost control over pricing, and with oil prices gyrating widely from a combination of oil producer greed, political instability, and Middle East conflicts, it was only a matter of time before someone would create a futures contract in oil. The New York Mercantile Exchange (NYMEX), with a long history in butter, cheese, and eggs, and later potatoes, needed a new trading commodity to keep its doors open. In the early 1980s, NYMEX started trading futures in heating oil, then gasoline, and finally crude oil. First attracting primarily speculators, soon oil companies as buyers and oil producers as sellers started trading. The development of a cash and futures market, with contracts that could be settled in monetary, or physical, terms and with marker crudes expanding from West Texas Intermediate to a variety of specific crudes in the Middle East, West Africa, and the North Sea eroded the oil producers' control over price. Since the early-1980s, the primary determinant of oil prices has been the relationship between supply and demand. The oil producers (OPEC) attempt to influence price by cutting back or expanding production, and in this indirect way affect the price of oil. But they no longer dictate price as they had in the years immediately following the 1973 oil crisis.

The End of the Era of High Oil Prices

With consumers doing everything they could to reduce oil consumption, and with every OPEC and non-OPEC producer operating full out, taking advantage of the price bonanza to maximize revenue, it was becoming increasingly difficult to maintain price. There had to be a swing producer to maintain a balance between supply and demand to keep prices high and, as Figure 5.4 shows, the swing producer was Saudi Arabia.

Saudi Arabia's production was initially boosted as replacement crude during the Iranian Revolution in 1978 and 1979 and during the early years of the Iran-Iraq war. After production in Iran and Iraq was restored, Saudi Arabia had to cut back sharply to maintain price. Those holding huge inventories in anticipation of further price increases had a change of heart when some semblance of order was restored and prices began to decline. Liquidating excess inventories caused OPEC oil demand to slump just as panic buying and hoarding caused a jump. With OPEC members producing full out, Saudi Arabia had to cut production sharply to keep prices from eroding further. Saudi Arabia was now playing the same historical role played by the United States. The Texas Railroad Commission had the authority to control oil production to maintain oil prices. The United States ceased being a swing producer in 1971 when the Commission authorized 100 percent production for all wells under its jurisdiction.

Figure 5.4 **Saudi Arabia Oil Production and Consumption**



From the perspective of 1985, with cessation of exports just over the horizon, Saudi Arabia was at the end of the line of playing the role of swing producer. Something had to be done. In 1985 Saudi Arabia unsheathed the oil weapon, not against the consuming nations but against its fellow OPEC members. Saudi Arabia opened the oil spigot and flooded the market with oil, causing oil prices to collapse below $10 per barrel. Threatening to wipe out OPEC financially, Saudi Arabia forced its fellow producers to sit around a table and come to an agreement on production quotas and a mechanism for sharing production cutbacks whereby Saudi Arabia would cease to be the sole swing producer. The cartel would now act as a cartel.

The Era of Moderate Oil Prices

Thus began the era of moderate oil prices, shown in Figure 5.3. Immediately world and U.S. consumption began to increase (see Figures 5.1 and 5.2) along with OPEC (Figure 5.1) and Saudi Arabian (Figure 5.4) production. What happened to energy conservation and efficiency? By the mid-1980s most of the mechanisms to achieve energy conservation and efficiency were already in place. Energy conservation and efficiency are noble undertakings; wasting a nonreplenishable resource cannot be justified. The dark side of energy conservation is that it only works when prices are high. If energy conservation and efficiency succeed in decreasing demand to the point where prices fall, then it becomes a different ball game. Suppose an individual buys a fuel-efficient car when the price of gasoline is high. The individual is using less gasoline. If repeated over millions of individuals, reduced consumption may be sufficient for the price of gasoline to fall. Once gasoline is cheaper, there is a temptation to take an additional vacation trip, perhaps as a reward for having a fuel-efficient automobile, which increases gasoline consumption.

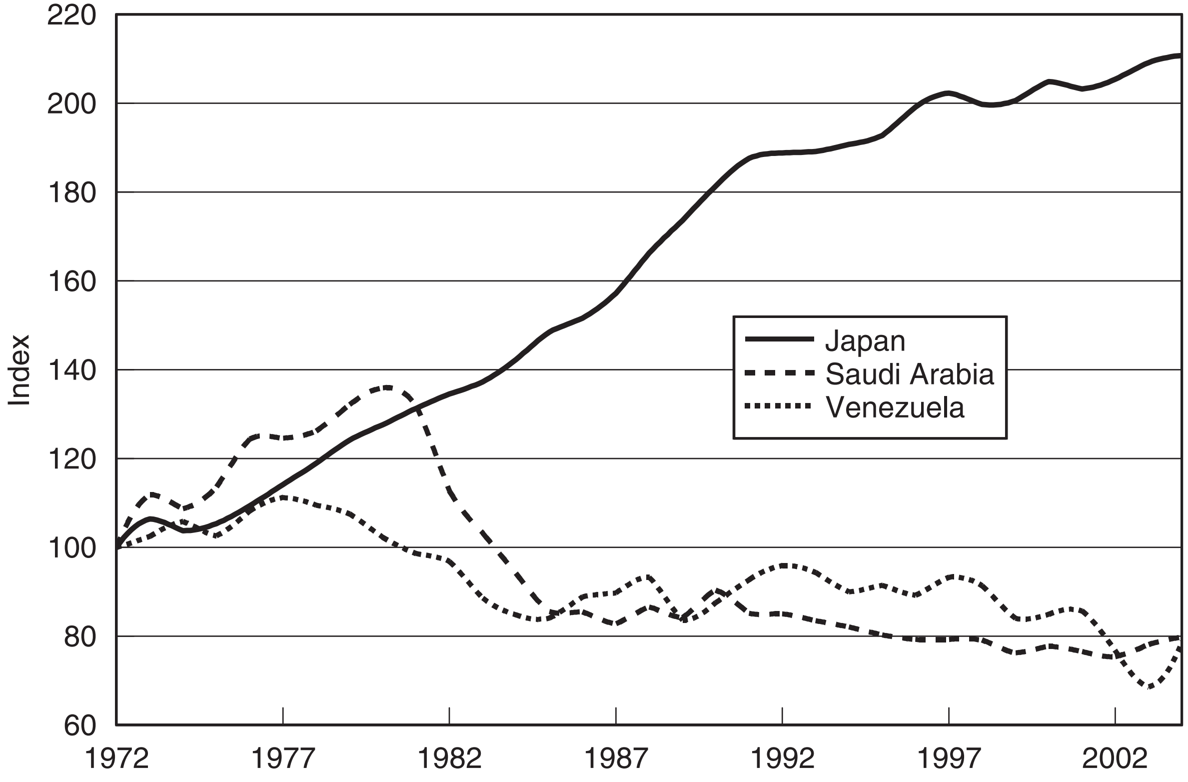
A house has been insulated and the temperature is lowered to use less heating oil in winter to cut heating oil consumption. If repeated in millions of homes, the cut in consumption may be sufficient to cause the price of heating oil to decline. When this occurs, the temptation is to increase the indoor temperature for greater comfort, causing heating oil consumption to rise. Fuel-efficient jet engines cut jet fuel consumption. If the airline industry converts to fuel-efficient jet aircraft, reduced consumption eventually cuts the price of jet fuel. Suppose that fuel-efficient jet aircraft are underemployed from a lack of passenger traffic. As jet fuel prices fall, the temptation is to use the savings in jet fuel to underwrite a cut in the price of passenger tickets to attract more business. Cheaper tickets encourage more flights, increasing jet fuel consumption. Thus, if conservation and efficiency succeed in cutting demand to the point where energy prices decline, then cheaper energy will restore consumption, closing the gap between current usage and what energy consumption was before conservation and efficiency measures were put into effect. This phenomenon is clearly seen in Figures 5.1, 5.2, and 5.4. Ultimately, conservation and efficiency are self-defeating, which does not mean that energy conservation and efficiency should be discarded. It has to be recognized that high prices have to be sustained in order to maintain the benefits of conservation and efficiency.[21]

It Is Not Oil Prices Anymore, It Is Oil Security

The British pulled out of the Middle East in 1971, leaving a power vacuum that contributed to the unfolding of events that led to the 1973 oil crisis. Before the 1973 oil crisis, U.S. military presence was limited to providing military weapons and advice to Saudi Arabia and Iran either for cash or as part of a military aid package. After the oil crisis, the U.S. military presence and involvement ballooned. It started under Carter in 1979 when forces loyal to Khomeini held U.S. embassy personnel hostages for 444 days, a situation worsened by a failed rescue mission. The U.S. Navy was charged with keeping the Arabian Gulf sea lanes open during the long Iran-Iraq War from 1981–1989. Failing to vanquish Iran, and desperate for money to repay loans for military equipment, mainly from the Soviet Union, Saddam cast his eye south to another neighboring state, Kuwait. Furious that Kuwait had refused to cancel Iraq's debts as he had requested and short on funds, the temptation to take over Kuwait's enormous oil fields proved more than he could resist.

The United States led the coalition forces in Gulf War I of 1990–1991. The retreating Iraqi forces set fire to Kuwait's oil fields, creating an environmental disaster in their wake. U.S. military presence in the Middle East remained strong between Gulf War I and Gulf War II in 2003 when it was Iraq's turn to be devastated. Now, with occupying troops in Iraq, the United States and the United Kingdom are de facto OPEC members. President George Bush's policy of transforming Iraq to an island of democracy in a sea of autocracy, led by a stable government capable of keeping the terrorist elements at bay, has to succeed. There is

Figure 5.5 **Per Capita GDP in Saudi Arabia, Venezuela, and Japan Since the Oil Crisis** (Index: 1972=100)



no alternative. If this policy fails and terrorists seize control of Iraqi, there will be no oil security anywhere in the Middle East.

When examining the current situation in the Middle East, one begins to consider the lasting benefits of the oil producers finally receiving just compensation for their oil exports. While living standards have increased in Kuwait and the smaller Gulf producers, can the same be said of Iraq, Iran, and Nigeria? Wars and corruption have taken their toll. Figure 5.5 is the per capita gross domestic product (GDP) of Saudi Arabia and Venezuela, the founding nations of OPEC, and Japan, a leading oil importer, indexed at 100 in 1972. Admittedly, this may not be the best way to determine whether the population has benefited from higher oil prices, but it is at least an indication of the pace of internal development of a nation's capacity to produce goods and services.[22]

The chart shows that per capita GDP contracted in Japan immediately following the price hikes in 1973, then resumed its upward course despite the high price of oil. Ironically, the Japanese benefited from the oil crisis even though they import all their energy needs. In the early 1970s the Japanese were producing higher-quality, more fuel-efficient automobiles than the mediocre gas guzzlers being sold at that time in the United States. In the wake of the oil crisis, the Japanese succeeded in capturing a significant share of the U.S. automobile market, which they managed to keep after Detroit began producing higher-quality automobiles with better gas mileage. Per capita GDP for other nations in Asia, particularly the Industrial Tigers, which include South Korea, Taiwan, Singapore, and Hong Kong, would show an even more dramatic increase.

For the oil exporters Saudi Arabia and Venezuela, per capita GDP expanded in the years immediately following the 1973 oil crisis, particularly for the former. However, these gains began to evaporate during the era of high oil prices and continued to erode during the era of moderate oil prices. What is surprising is that the decline in per capita GDP has fallen below 100. This implies that these nations are producing fewer goods and services now on a per capita basis than they were before the oil crisis. However, this does not necessarily mean a lower standard of living because per capita GDP may not entirely reflect the portion of petroleum revenue that is distributed to the people in the form of social, educational, and medical services. The falling per capita GDP does imply an increasing dependence on oil revenue to sustain living standards. This can be blamed on an understandable, though not a constructive, attitude that people in oil-exporting nations should not have to work. Perhaps the oil producers should heed the shah's advice that Westerners should learn to work harder.

Another contributing factor for the declining per capita GDP in Venezuela and Saudi Arabia is their booming populations. The population of Venezuela has more than doubled, from 11 million in 1972 to a little over 26 million in 2005, while the population of Saudi Arabia has more than tripled, from 6.6 million in 1972 to 24 million in 2005. GDP would have to double for Venezuela or triple for Saudi Arabia simply to keep per capita GDP constant. Part of Venezuela's decline in per capita GDP in the early 2000s was a consequence of civil unrest. Perez Alfonso, principal architect of OPEC and Venezuelan oil minister in the 1970s, wrote that oil was the "devil's excrement" and would eventually ruin Venezuela. Maybe he is being proven right.

Oil and Diplomacy

Oilmen, most of whom have engineering backgrounds, often end up playing a diplomatic role to protect their investments in foreign nations. An excellent example before the 1973 oil crisis occurred when Great Britain imposed a trade embargo against the then-existing nation of Rhodesia (now Zimbabwe). South Africa deemed such an embargo was illegal for the oil companies operating in South Africa. No matter what BP and Shell did, as British companies operating in South Africa they were breaking someone's law. If they continued to trade with Rhodesia, they violated British law. If they stopped trading, they violated South African law. During this period of apartheid, Shell was despised by the general public for dealing with South Africa and, perversely, reprimanded by the South African government for its practice of hiring, training, and giving blacks positions of responsibility and authority in violation of apartheid. As a result, Shell found itself breaking the law in both Britain or South Africa and, simultaneously, being criticized by outsiders for having investments in South Africa, and by insiders (the South African government) for not upholding the spirit of apartheid.[23]

Another time when oil companies found themselves on the proverbial horns of a political dilemma was during the Yom Kippur War in 1973, when U.S. oil companies had to enforce an embargo of Saudi crude to America. The challenges of oil companies operating in foreign nations remain no less daunting. Helping Russia to reopen its oil resources have put oil companies at loggerheads with the government over its practice of unilaterally changing tax laws and terms of contractual agreements, with no means of judicial appeal, and restrictions on the rights of minority shareholders. Oil company executives and Russian government officials have had to work together to come to some sort of compromise that affects Russian laws on taxation, the nature of contracts, and judicial appeal, and the rights of minority shareholders before major investments could be made. The laying of pipelines in the Caspian Sea region brings oil companies in contact with governments hostile to one another through which the pipelines must pass. Tariff structures, security measures, and ways for resolving disputes have to be just as carefully planned as selecting the pipeline route and engineering its construction. Resolving the conflicting interests of different peoples and governments to determine a fair share of the benefits of oil exports for the people and their governments, along with the participating oil companies, still poses enormous challenges.

Oil and Environmentalists

In more recent times, oil executives have had to learn to deal with environmental groups that have learned ways to pursue their agendas other than public communication media and demonstrations. Shell's plans to dispose of an abandoned North Sea oil platform were changed by environmental groups lobbying for a government ruling that resulted in a different and far more costly means of disposal. In addition to being active in sponsoring environmental laws, environmental groups have learned to gain their objectives through loan covenants, conditions that have to be satisfied before funds can be advanced. In response to environmental group lobbying, the World Bank imposed environmental conditions as loan covenants that affected the construction of two pipeline projects; one for moving oil from an oil field in Chad to a port in Cameroon and the other for moving oil in Ecuador from an oil field in the Amazon over the Andes to an exporting port. These loan covenants ensured that a portion of the oil revenues would be paid directly to indigenous peoples, along with changes in pipeline routing to deal with environmental concerns. Both the government oil companies and those building and operating the pipelines had to agree to comply with these loan covenants to obtain World Bank financing.

The environmentalists point to oil as being primarily responsible for pollution, along with coal. Pollution-emission regulations, a concept that no one can really oppose, can pose significant operational challenges for the oil companies. Yet these seemingly insurmountable barriers to their continued existence are surmountable. The oil companies have learned to cope, if not thrive, in this changing business environment. The "Beyond Petroleum" of BP, which must sound sacrilegious to the ears of oilmen of yore, is recognition that oil companies must operate in an environmentally friendly way and consider issues beyond their focus on oil.

Of course, this world of environmental concern exists mostly in North America, Europe, and Japan. The rest of the world is more interested in making economic progress without any overdue sensitivity about the adverse environmental consequences, as epitomized by the nonreaction of most Asian nations to a brown cloud of pollution that hangs above a large portion of the continent. Nevertheless, oil companies must respond to environmental challenges in the quality of their product and in the nature of their operations, or face a daunting public relations challenge.

The Role of Oil Companies After the Oil Crisis

Although the oil companies were literally thrown out of producing nations after the oil crisis, there has been a return to the situation that existed prior to the oil crisis. Oil companies are, to widely varying degrees, involved with oil production with nearly all the producers that had previously nationalized their oil fields and facilities. The major difference is that the accounting entries for oil reserves in OPEC nations have been obliterated from oil companies' books. These were always fictional because the oil fields were located outside the nations of domicile of the oil companies. The oil companies never had any legal recourse to protect their property rights against actions taken by host governments. This makes ownership a spurious claim, to say the least.

Nationalized oil companies operate under encumbrances that Big Oil does not have. Nationalized oil companies are limited in their activities to exploiting a nation's wealth of oil and natural gas and, by government fiat, do not and cannot look outside the box. Most nationalized companies are not run as government bodies even though they are wholly owned by their respective governments. Normally, they operate as quasi-independent oil companies. While some have managed their nations' energy resources and infrastructures quite well, others have a less than sterling record. Some oil-exporting nations are looking at privatization, at least in part, as an alternative to a nationalized oil company or to introduce a taste of competition to reinvigorate a moribund organization. Lest we forget, *privatization* would not be a word had it not been for the failure of government-owned companies to deliver the goods and services that they were set up to provide.

While nationalized oil companies do not worry about making a profit or surviving in an extremely competitive world, their financial life is not one of idle comfort. Nationalized oil companies are the chief revenue generators for many oil-producing nations and they have to fight over every dollar with their exclusive shareholders, their national governments. Nationalized oil companies sometimes come up short in the struggle over whether a dollar of revenue should be spent supporting a social program, funding a government expenditure, or being plowed back into the oil infrastructure. Some nationalized oil companies are short on funds needed to maintain oil productivity, others lack technical expertise to expand their oil infrastructure or have limited access to markets. Having chased the oil companies out with a broom, oil companies are back under a variety of contractual arrangements to assist nationalized oil companies with capital infusions, technical expertise, and market access. Since these are the same functions oil companies provided before their oil reserves and properties were nationalized, the circle has been closed.

Oil companies have learned that what counts is not who owns the oil fields but who has access to the oil. Access is provided under a variety of joint venture and production-sharing agreements with the producers. These agreements would not be necessary if the nationalized oil companies could fully replicate the oil companies' contributions in capital, technology, and market access. Having said that, some oil producers have been successful in becoming more integrated by acquiring refineries and service stations in consuming nations and tankers to ship their oil. Kuwait purchased Gulf Oil's refinery and distribution system in northern Europe. Venezuela purchased refineries in the United States as well as offshore, and owns a chain of gas stations in the United States. Both nations have tanker fleets to transport a portion of their oil exports. These investments assure Kuwait and Venezuela of outlets for their oil exports and secure transportation. These moves by producing nations to become integrated oil companies have not diminished the role for the major oil companies and hundreds of independents in the global oil business. Just as Marcus Samuel remarked, there is room for both Shell and Royal Dutch to succeed in an expanding Asian oil market (although he may have lived to regret that remark), so too is there room for government and privately owned oil companies to succeed in an expanding world oil market.

A Changing World

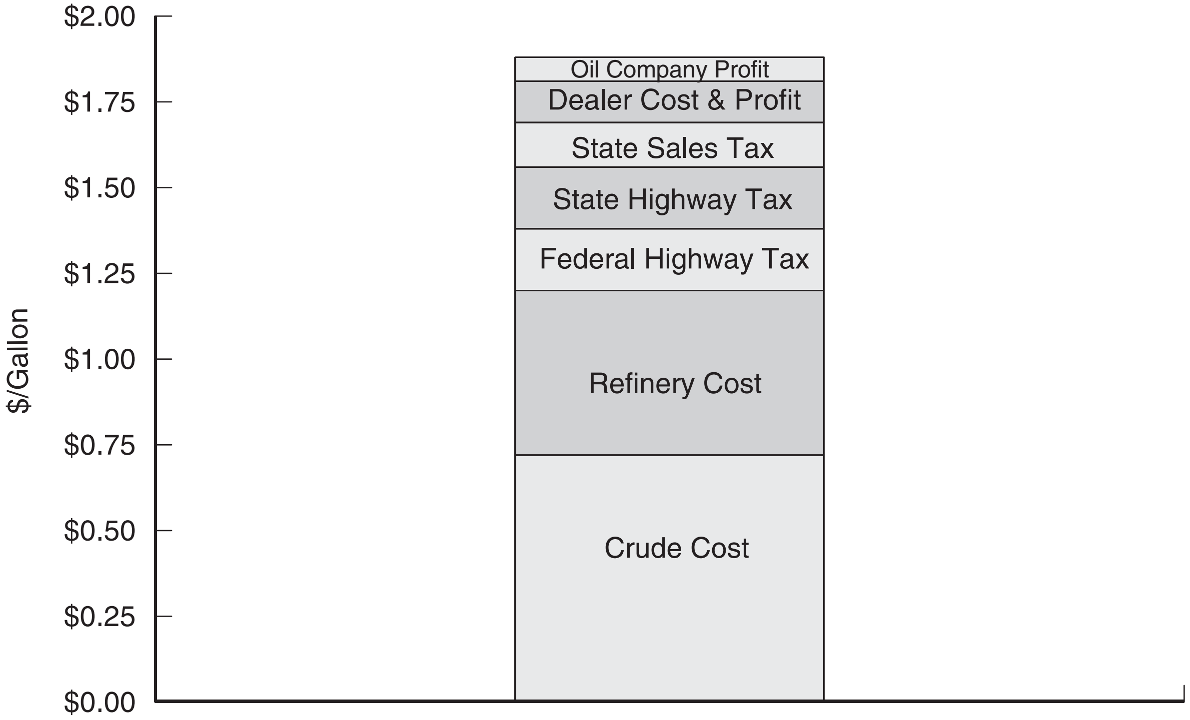
The world of the twenty-first century is different for the oil companies, but in one respect it is easier: They need not worry about pricing. That is no longer in their hands, or not nearly as much in the hands of the oil producers as they would like; that role has been taken over by the immutable laws of supply and demand. Moreover, they are no longer concerned about ownership of oil in foreign lands as long as they have access to that oil, which does not seem to be a problem at this time. Before the oil crisis, the goal of the oil companies was to reduce costs, that being the price of crude oil. The irony is that profits are not based on costs, but on the spread between the price of oil products and the cost of crude oil. It does not matter what the cost of oil is as long as the margin can be maintained. In addition, an oil company can enter into a variety of contracts and financial derivatives such as swaps, futures, and forwards to hedge against the risk of an adverse change in oil prices.

One can expand this concept to the environmental cost of doing business. It does not matter what incremental costs are placed on an oil company's operations to safeguard the environment as long as every oil company is bearing the same cost. Then it becomes just another cost of doing business, such as the cost of crude oil or the obligation to pay taxes, all of which are simply passed on to consumers as higher prices. Ultimately, it is the consumer who pays for higher-priced crude, increased environmental costs, and additional tax burdens. As long as these are approximately equally borne by all oil companies and can be passed on to consumers, why should the oil companies care? All they have to focus on is maintaining their margins, which in the last analysis means covering their costs. Furthermore, they really do not have to be overly concerned about security of supply. Before the 1973 oil crisis, it was not a significant concern and, since the crisis, the responsibility has been assumed by the taxpayers who foot the bill for an American and British military presence in the Middle East.

Are Oil Companies' Margins All That Great?

Some politicians accuse oil companies of making unconscionable profits. Oil companies do make a lot of money. As an example, Shell Oil reported $12.5 billion in net income for 2003, of which $12 billion is attributable to oil and natural gas. Shell sold 12.2 million bpd of oil, split roughly in half for motor vehicle fuels (gasoline and diesel fuel) and crude oil and residual fuel. If we estimate that two-thirds of the $12 billion in net income is from motor vehicle fuel sales, then the profitability of the Shell Group is equivalent to 8.5 cents per gallon. As a side note, the Shell Group reports an income tax rate of 51 percent, which implies that they paid out over $12 billion in income taxes to different tax authorities. Amerada Hess, an American independent, had profits of $643 million in 2003. If $116 million of net gains in asset sales is netted out and the remainder divided by 153 million barrels of oil product sales of all types, neglecting profitability from its natural gas sales,

Figure 5.6 **Cost Factors of a U.S. Gallon of Regular Gasoline in 2000**



then the company makes a profit of 6.6 cents per gallon of sales.

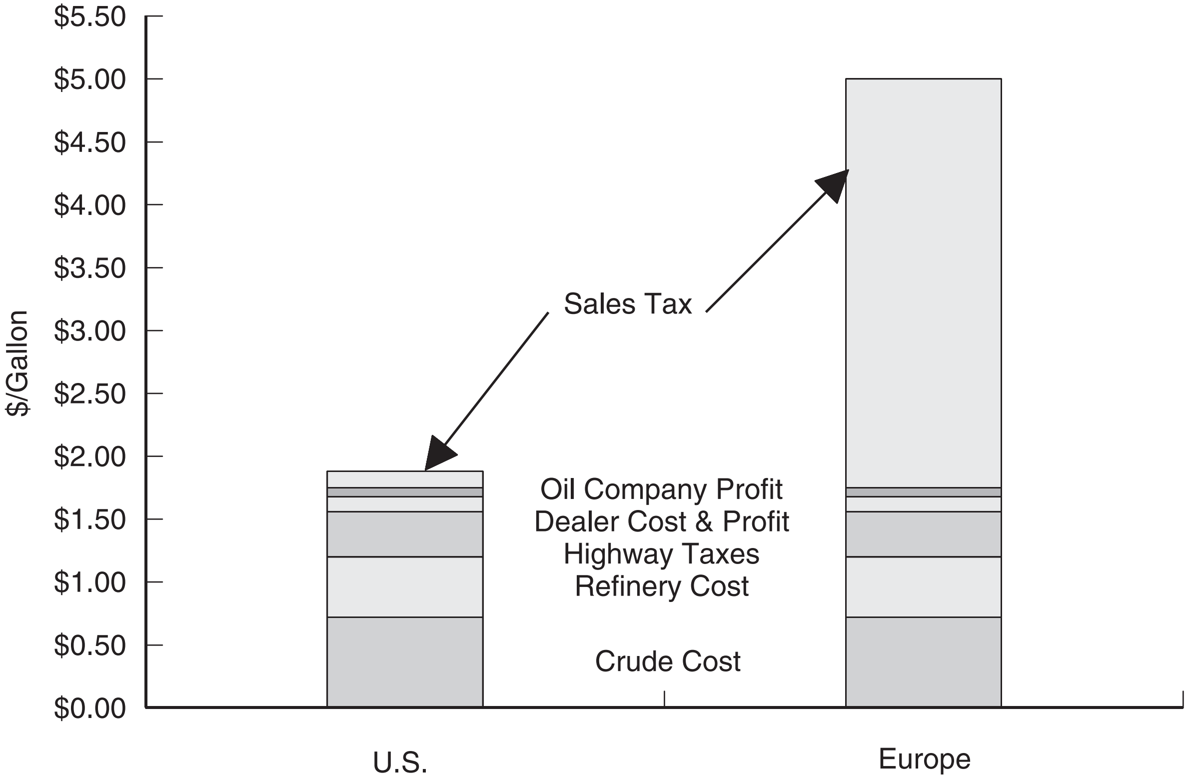
Both of these calculations are approximate and, if anything, overstate the amount of oil company profits inherent in the sale of a gallon of gasoline or diesel fuel because the calculations neglect, or do not fully take into account, profits from natural gas sales. This profit margin is not sacrosanct; it changes from year to year. The same calculations for 2002 yield a profitability of seven cents per gallon for Shell. Since Amerada Hess had a loss in 2002, it paid, at least in an accounting sense, for the privilege of distributing energy products to the consuming public.

The oil companies earned more on a cents-per-gallon basis in 2005, a record-breaking year, but 2005 is not a representative year from a historical perspective. Figure 5.6 compares oil company profits of an assumed seven cents per gallon for motor vehicle fuels with the other cost factors for a gallon of gasoline in California, serving as a typical case for the United States.[24]

The state and federal highway taxes are sufficient to maintain, upgrade, and build roads and highways. Figure 5.7 represents a rough cost factor comparison between European and American prices, assuming that the same highway taxes, refinery and dealer costs, and oil company profit margins apply to Europe, where gasoline sold for about $5 per gallon when U.S. prices were close to $2.

European governments still make far more in tax revenue from selling a gallon of gasoline or diesel

Figure 5.7 **Comparison of Gasoline Prices in Europe and the United States**



fuel than the oil producers get for selling a gallon of a depleting natural resource; one of the justifications of the price hikes in 1973. As a point of comparison, three liters of Poland Spring™ mineral water was purchased on sale for $1.99 during the summer of 2004. This works out to $2.50 per gallon and would be significantly greater for an upscale mineral water not on sale. At the time, gasoline was selling for less than $2 per gallon. When one compares the effort to explore, develop, refine, distribute, and market one gallon of gasoline with the cost of bottling one gallon of mineral water, plus taking into consideration the amount paid to governments in the form of sales and highway taxes and paid to the producing nations for the crude, the oil industry has got to be one of the most efficiently run operations on Earth.

The Future Role of Oil Companies

Most oil companies specialize in some facet of the oil business. They have neither the capital nor the technical expertise nor the desire to explore energy alternatives outside the oil box. Big Oil is a relatively small group of publicly traded corporations that play a paramount role in finding and developing large oil fields and in refining and marketing oil. Unlike smaller oil companies, they have an eye on the future of energy with or without oil. Big Oil is aware that the energy business goes beyond getting crude out of the ground and gasoline into a tank. They realize that the era of oil may draw to a close much as the era of biomass in the nineteenth century and the era of coal in the twentieth. This is not to say that oil will disappear any more than biomass and coal have disappeared. The major oil companies are investing in the development of alternative fuels to oil. If not enthusiastic endorsers of developing alternative energy technologies, they are certainly cognizant of their own well-being. Oil is but one facet of the energy industry, and if the role of oil changes, Big Oil wants to be part of that change. This is the only way they can ensure their survival as major players in the energy game.

Big Oil's ability to adjust to a changing business environment has been amply demonstrated. They have survived the greatest assault imaginable on their privileged position by losing control over vast oil resources once considered their own. They have also lost the ability to determine the price of oil. Such losses could have led to their demise, yet they are prospering more now than ever before. Once unceremoniously thrown out of oil-producing nations, they have since been invited back by the nationalized oil companies that had taken over their oil fields and distribution and refining assets.

There are some who say that oil is too important to be left in the hands of businesspeople bent on making a buck. Oil should be in the hands of a benign government body that knows best how to serve the wide interests of the people rather than the narrow interests of the shareholders. As alluring as this sounds, the privatization of the ex-Soviet oil industry revealed the outmoded technology, managerial ineptness, and the disregard for the environment of a government-owned and -operated oil company.

Unfortunately, profit has a bad name. For many, all profit means is the right of unscrupulous individuals or companies to gouge the public when the opportunity arises, as was exercised by certain energy traders and merchants who supplied electricity to California during the 2000 energy crisis. Allegations have been made of certain supplying companies holding back on generating electricity to create an artificial shortage that hiked electricity rates. In fairness, the California state regulatory body that established a flawed energy policy and provided poor oversight must bear some of the blame. Nevertheless, this crisis—along with the exposure of executive compensation for certain companies of hundreds of millions at the expense of corporate liquidity and stakeholder value—reinforced the public's generally negative image of corporate executives as profit-gouging, irresponsible, selfish, self-serving gluttons.

Profit means that revenue covers costs. No one can seriously argue against the concept of a public or private undertaking covering its costs, that is, having enough money in the bank to pay its bills. The only objection that can be raised against profits is the degree of coverage. As has been shown, profits made by two oil companies expressed in cents per gallon are quite modest in comparison to what consuming governments receive in the form of taxes and what producing governments receive in the form of revenue. The key question is whether we are getting value for what the oil companies charge for their services. By focusing on making money, oil companies have been able to bridge the gap between consumers and suppliers, acting as a neutral third party serving the widely divergent interests of both.

The oil companies' possession of engineering technology, capital resources, and market access cannot be duplicated. If we interpret profits as some excess over costs, which by any measure cannot be considered excessive, then the oil companies should continue to play their historical role of a neutral buffer between consumers and suppliers. If we believe what they report in their annual reports, major oil companies view themselves as energy companies with a particular focus on oil, with that focus subject to change as conditions warrant. If tar deposits in Canada and Venezuela and oil shale deposits around the world become technologically and economically feasible, the oil companies will be there. If another fuel replaces gasoline as the fuel of choice, the oil companies will be part of the transition. Their survival as major global companies hinges on their ability to adapt to changing times. As this chapter readily shows, they have proven their adaptability in the past and there is every reason to expect that they will do so in the future.

Notes

1. Mir Yusif Mir-Babayev, "Azerbaijan's Oil History: A Chronology Leading up to the Soviet Era," *Azerbaijan International* 10 (Summer 2002), www.azer.com.

2. "Abraham Gesner: Lighting up the Nineteenth Century, 1797–1864," Volume IV, *Pathfinders: Canadian Tributes*, Canada Heirloom Series of Canada's Digital Collections, http://collections.ic.gc.ca/heirloom_series/volume4/136-139.htm.

3. "The History of the Oil Industry: Oil Through the Ages, 347 BC to 1859," www.sjgs.com/history.html.

4. Early Petroleum History Web site at http://little-mountain.com/oilwell/.

5. J.T. Henry, *The Early and Later History of Petroleum* (New York: Augustus M. Kelley Publishers, 1970). First published 1878 by Jas. B. Rodgers Co., Philadelphia.

6. James S. Robbins, "How Capitalism Saved the Whales," www.alts.net/ns1625/gesner.html.

7. Unless otherwise noted, my primary source for the life of John D. Rockefeller is Grant Segall, *Anointed with Oil* (New York: Oxford University Press, 2001).

8. John D. Rockefeller, Jr. philanthropic works are described in the June 2004 issue of the *Smithsonian Magazine,* Dorie McCullough Lawson, "Who Wants to Be a Billionaire?"

9. Rockefeller considered Flagler the brains behind Standard Oil. After 1885, and while still a director of Standard Oil, Flagler became a prominent Florida developer in St. Augustine, Palm Beach, and Miami. He also organized the Florida East Coast Railroad, which ran the length of the state from Jacksonville to Miami. He extended the railroad to Key West, which was considered an engineering feat of the day (see www.flagler.org).

10. Unless otherwise noted, my primary source for the life of Marcus Samuel is Robert D.Q. Henriques, *Bearsted: A Biography of Marcus Samuel, First Viscount Bearsted and Founder of "Shell" Transport and Trading Co.* (New York: Viking Press, 1960).

11. James A. Clark and Michel T. Halbouty, *Spindletop* (Houston, TX: Gulf Publishing, 2004). First published 1952 by Random House.

12. As Lord Mayor, Marcus did what he could to counter pogroms in Romania and Russia, warning the Russian ambassador that continued violence against the Jews would lead to the undoing of the Czar.

13. As mentioned previously in the text, Burmah Oil had the unique status of being, in the eyes of the government, the only true British oil company. This mantle was taken over by BP when it became involved with North Sea oil. Burmah Oil would eventually be without oil-producing properties when the Burmese oil fields were exhausted. As a result of speculative investments, Burmah Oil went bankrupt in 1974 and its 20 percent holding in BP was taken over by the Bank of England.

14. Unless otherwise noted, my primary source on the oil business up to the 1973 oil crisis is Anthony Sampson, *The Seven Sisters: The Great Oil Companies and the World They Made* (New York: Viking Press, 1975).

15. Ralph Hewins, *Mr. Five Percent: The Story of Calouste Gulbenkian* (New York: Rinehart & Co., 1958); Nubar Gulbenkian, *Portrait in Oil: An Autobiography of Nubar Gulbenkian* (New York: Simon and Schuster, 1965).

16. Stephen Howarth, *A Century in Oil: The "Shell" Transport and Trading Company, 1897–1997* (London: Weidenfeld & Nicolson, 1997).

17. The statistics reported in Figure 5.1 can be found in *BP Energy Statistics* (London: British Petroleum, 2005).

18. A detailed chronology of events prepared by National Public Radio (NPR), "Gas & Oil Prices: A

Chronology," is available at www.npr.org/news/specials/oil/gasprices.chronology.html.

19. British Petroleum, *Statistical Review of World Energy 2005* (London*:* Author, 2005).

20. The primary source for the post-1975 history recounted here can be found in Daniel Yergin, *The Prize: The Epic Quest for Oil, Money, and Power* (New York: Simon & Schuster, 1991).

21. Herbert Inhaber, *Why Energy Conservation Fails* (Westport, CT: Quorum Books, 2002).

22. International Monetary Fund, *International Financial Statistics Yearbook* (Washington, DC, 2005); International Monetary Fund, *World Economic Outlook, A Survey by the Staff of the International Monetary Fund* (Washington, DC, 2005).

23. Stephen Howarth, *A Century in Oil: The "Shell" Transport and Trading Company, 1897–1997* (London: Weidenfeld & Nicolson, 1997).

24. See the California Energy Commission at www.energy.ca.gov for oil prices of $30 per barrel. The oil company profit of $0.07 per gallon has been separated from the California Energy Commission estimates for combined refinery cost and oil company profit margin.

6

Oil

This chapter describes the journey oil takes from deep in the earth until it reaches consumers in a wide range of products from plastics to motor vehicle fuels to fertilizers and pesticides. The sojourn starts with exploration and the development of oil wells onshore and offshore, the refining and transportation of oil products, and the use of enhanced recovery methods to get the most out of oil fields. The adequacy of oil reserves to continue to fuel our economy, the potential of nonconventional oil sources, and alternative motor vehicle fuel substitutes are covered. The chapter ends with a discussion of the geopolitical aspects of oil, with a call to internalize an externality called oil security.

The Earth as an Oil Manufacturer

Terra firma it is not; the daily chronicle of earthquakes and volcanoes attests otherwise. The earth, with a radius of less than 4,000 miles, has a center core with a radius of about 2,000 miles made up mostly of an alloy of nickel and iron. The center of the core is solid with a liquid outer portion. If stripped of the mantle and crust, the core would shine as brightly as the sun, heated by the weight of the overburden and radioactive decay. Between the core and the outer crust is a nearly 2,000-mile-thick mantle of semiliquid rock and metals called magma. Magma is less dense than the core of pure metal, but denser than the crust, which is made largely of rock. Magma is a viscous fluid with upward convection flows of hotter magma from near the core balanced by downward flows of cooler magma from near the crust. There may be an internal structure to magma consisting of gigantic plumes hundreds of miles high. The upper 50- to 150-mile portion of the mantle is called the asthenosphere, which has a chemical composition closer to that of the crust than the underlying magma.

Literally floating on top of the asthenosphere is a relatively thin, brittle crust made up of mainly less dense stone mixed with metals called the lithosphere. The lithosphere is only between four to seven miles deep beneath the oceans and up to sixty miles deep beneath mountain ranges. Its average depth of nearly twenty miles is only 0.5 percent of the radius. Oceanic crust is mostly relatively heavy basalt while the continental crust is mostly relatively light granite. The crust is broken into major segments called tectonic plates including the Eurasian, North and South American, African, Pacific, Antarctic, Australian, Arabian, and Indian plates. There are also smaller plates such as the Philippine plate, the Juan de Fuca plate (off west coast Canada and United States), the Caribbean plate, the Cocos plate (west of Central America), the Nazca plate (west of South America), and the Scotia plate (between the Antarctic and South American plates). These plates separate, collide, or slip by one another in response to underlying flows of magma in the asthenosphere. We recognize the continents and oceans as major geological features, but so too is the mid-ocean ridge, largely unseen except where it protrudes above the ocean in Iceland and the Azores. Made of two mountain chains separated by a rift valley, the mid-ocean ridge can be up to two miles in height above the ocean floor and as much as 1,000 miles wide. The ridge is 35,000 miles in length and encircles the world like the stitching on a baseball. Tectonic plates along the mid-ocean ridge are separating about as fast as a fingernail grows with new crust formed by upwelling magma from volcanoes and fumaroles. The East African rift marks a plate separation that may one day be a new body of water like the Red Sea; others include Lake Baikal, Rio Grande, and Rhine Graben.

Subduction occurs when two plates collide and one overrides the other, forcing the lower plate back into the mantle to become new magma. Subduction zones are located at ocean trenches such as the Chilean trench and the Marianas trench in Indonesia and are marked by volcanoes. If two tectonic plates collide, but neither is massive enough to cause the other to submerge, a mountain chain may then emerge (e.g., the Caucasus and Himalayas). The collision between two plates in the Middle East was not sufficient to create a mountain range or a subduction zone, but folds in the rock capable of trapping the earth's greatest concentration of oil and natural gas. A fault is created when plates slip by one another laterally such as the San Andreas fault in California. Faults and folded rock are critical in the formation of oil reservoirs.[1]

The generally accepted theory of the origin of oil in the Western world is that it comes from dead animal matter, but a cemetery, where bodies decay and turn to dust, is not a future oil field. For the earth to manufacture a fossil fuel, the decaying process must be interrupted either by dead plant and animal matter falling into oxygen-starved waters or rapid burial. It was once thought that oil came from dinosaurs, the symbol of Sinclair Oil. This theory has been discredited because the earth could not possibly have sustained a population of dinosaurs, even over millions of years, large enough to create so much oil, even if one ignores the special circumstances that must accompany death for dinosaurs to become a fossil fuel. In the twenty-first century the accepted theory postulates that ocean plankton, algae, and other forms of simple marine life die and sink into oxygen-starved waters that prevent further decay. Sediments from rivers mix with the partially decayed matter to form an organically rich concoction. Continued burying by more layers of sediment squeezes out the water, and when buried by a mile or more of new sediment over millions of years, the original sediment is transformed to sedimentary rock and the organic matter to oil.

Based on this hypothesis, a favorite place to explore for oil is near river mouths such as the Mississippi and Niger rivers, where shifting deltas and rising and falling sea levels over millions of years have created widespread oil and gas deposits at various ocean depths. Geologists look for what appear to be buried river mouths in sedimentary rock for likely places to drill. A location where oil and natural gas may be in the making is the Bosporus, where Mediterranean water enters the Black Sea. Dead and dying marine organisms sink to the bottom and mix with the sediments in oxygen-starved waters. If buried by a mile or more of overburden during the next many millions of years (a scenario that requires the Black Sea to drop by more than a mile or the land around the Black Sea to rise by more than a mile, or some combination of the two) then the sediment will be transformed to rock and the entrapped organic matter to oil. Then a new budding oil field can await discovery by a petroleum geologist yet to be born far, far in the future.

The crust has three general types of rocks. Upwelling magma, when cooled, becomes igneous (fire-formed rock) such as granite and basalt. Deeply buried igneous or sedimentary rocks are subjected to enormous heat and pressure and are transformed into metamorphic rocks such as marble and slate. From a Western petroleum geologist's point of view, the presence of igneous and metamorphic basement rock underlying sedimentary rock is good reason to stop drilling because oil and gas are found only in sedimentary rocks.

Two of the three types of sedimentary rocks are created by erosion, which would level the earth were it not for emergence of new mountain chains from colliding tectonic plates (the Himalayas are still rising as the Indian plate continues to plow into the Eurasian plate). Wind, rain, and flowing water are the principal agents of erosion. When water seeps into the cracks and crevices of rock and freezes, it expands, fracturing the rock and making it more vulnerable to erosion. The debris of erosion suspended in flowing water is another powerful force of erosion that can cut deep gorges into solid rock. When carbon dioxide in the atmosphere mixes with water, it forms dilute carbonic acid that eats away at limestone and forms underground caverns. Finally, the debris of erosion, as gravel, sand, and clay, is deposited in river deltas. Sources of sediment that are not from the result of erosion are shells of dying plankton deposited on the ocean floor and the precipitation of dissolved calcium carbonate from evaporating seawater in shallow lagoons. Under sufficient pressure from overburden, and with calcium carbonate and silica acting as cementing agents, gravel is transformed into a conglomerate, sand to sandstone, clay to shale, lime mud to a gray to black limestone, and microscopic seashells to white limestone called chalk, such as the White Cliffs of Dover. The most common form of sedimentary rock is shale and the most common form of petroleum reservoir rock is sandstone.

The compressive force of plate movements near subduction zones creates folds in hot, plastic sedimentary rocks. An upward fold is called a syncline and a downward fold is called an anticline. Anticlines are shaped like an upside down or inverted bowl and faults (breaks in sedimentary rock layers caused by tectonic plate slippage) of nonpermeable rock (or caprock) form traps to prevent oil and gas from completing their migratory journey to the earth's surface. If they are not trapped, natural gas and the lighter components in oil evaporate as they migrate toward the earth's surface, finally emerging as a viscous crude or thick tar called seep oil.

Different types of sedimentary rock layered one on top of the other are on display at the Grand Canyon, each with a unique geological origin. A typical cross-section of the earth contains a mile of sedimentary rock with about 100 layers of various types of sedimentary rocks underlain by a basement of igneous or metamorphic rock. Sedimentary rock layers on the ocean floor are only about one-half mile thick on average, with the thinnest at the mid-ocean ridge, becoming thicker as the ocean floor approaches the continental shelf. The thickest sediments (about ten miles in depth) are located on continental shelves, but can also be found inland. Colorado is one such place where sedimentary rocks were formed when that area of the world was covered by a shallow sea. Glaciers can strip basement rock of their sedimentary rock cover as occurred in eastern Canada and what is now New York's Central Park. Drilling for oil in Central Park would be useless: no sedimentary rock, no oil. Even with the earth covered by a mile of sedimentary rock, the presence of sedimentary rock does not mean the presence of oil. The secret to successful oil drilling is identifying traps overlain with nonpermeable caprock and underlain by sedimentary rock whose pores, or spaces between the grains, are saturated with oil and natural gas.

Sedimentary rocks buried deep in the earth were originally made from debris from eroding mountains or the deposition of shells from marine life. These rocks may one day be uplifted by colliding tectonic plates and again become mountains vulnerable to erosion. The rough jagged peaks of the Alps, Rockies, Andes, and Himalayas are geologically new compared to the far older rounded Urals and Appalachians. As new and old mountains erode, their sediments are deposited in river deltas, which if buried by a mile of overburden, become new sedimentary rock, which over time may be thrust up again as a recycled mountain range. The Himalayas are such a range where, three to five miles above sea level, marine fossils can be found in sedimentary rock.

Formation of Oil

Dead organic matter must lie in either stagnant, oxygen-free waters at the bottom of the sea until buried or be buried quickly after death and achieve a concentration of 1 to 3 percent by weight to become a future oil reservoir, although this concentration can be as high as 10 percent. The next step is burying the organically rich sediment deep enough to generate the temperature and pressure necessary to transform organic matter to oil. With 7,000 feet of overburden, the pressure is sufficient to raise the sediment's temperature to 150°F, the minimum to produce a heavy and generally undesirable grade of crude oil. Preferred light crudes are produced as one approaches 18,000 feet and 300°F. Beyond 18,000 feet, the temperature and pressure are sufficient to transform oil to graphite (carbon) and natural gas. The oil window is 7,000–18,000 feet below the surface of the earth, meaning that sediments at river mouths must be buried between 1.5–3.5 miles of debris to produce oil by either the ocean bottom sinking or the surrounding land mass rising or a combination of both. The properties of the oil depend on the type of organism, its concentration, depth of burial, and the nature of the surrounding sediment. Oil properties vary from one field to another and no two oil fields have exactly the same properties. Commercial grades of crude are really a mix of oil from different oil fields in the same region that have similar properties. A few are from different oil fields with dissimilar properties such as Urals, a specified mix of light sweet crude from western Siberia and heavy sour crude from the Ural region of Russia.

Once formed in source rock, oil and natural gas, being lighter than water, begin to migrate laterally and vertically through migratory rock. Oil and gas pass through the pore space within the sedimentary rock structure and through fractures in rock layers. This migration may extend as far as 200 miles from source rock. The rate of migration depends on the porosity and permeability of the migratory rock. Porosity is a measure of the spaces (pores) within the rock that can be filled with fluids (oil, gas, and water) and permeability is a measure of the ease with which a fluid can pass from one pore to the next. Both are critical in determining the flow of hydrocarbons (and water) into a well; generally speaking, the greater the porosity, the greater the permeability. Oil and gas migration continues until interrupted by an intervening rock formation shaped like an inverted bowl or a fault made of a well-cemented rock with no spaces between the grains. Once migrating oil and gas are trapped in reservoir rock, natural gas, the lightest, rises to the top of the reservoir and forms a gas cap; saltwater, the heaviest, sinks to the bottom, leaving oil in the middle. In some reservoirs, a small concentration of natural gas may remain mixed with crude oil without forming a gas cap; in still others, there is no associated natural gas. The subsurface water that makes up the water table is fresh, produced by rain percolating through the soil; but the water beneath the water table is more or less as saline as ocean water.

Contrary to a popular conception that originated with the dawn of the oil age, an oil reservoir does not consist of a void space filled with a pool of oil; rather, it is migratory rock turned reservoir rock, saturated with oil and gas, that has been prevented from continuing its journey to the earth's surface. The geometry of a trap is one determinant of the size of an oil field; the larger the dome or fault of caprock and the greater the distance from the top of the trap to the spill point (where oil and gas can flow around the caprock and continue migrating to the surface) the larger the size of the potential oil field. Other determinants are porosity, which determines the quantity of oil and gas contained in the reservoir rock, its permeability, which determines the flow of the oil and gas to a well and its potential recoverability, and, of course, the concentration of oil and natural gas in the reservoir rock. Sandstone has the largest pores for the greatest porosity and permeability, followed by limestones, and then shales, which have the smallest pores. Most reservoir rocks are sandstones and limestones, but even here tight sands and dense limestones have low degrees of porosity and permeability.

Salt domes are another mechanism that can trap oil. Salt can be deposited hundreds or thousands of feet deep when ocean water in shallow lagoons, which were periodically connected and disconnected from the ocean, evaporates to form salt pans. These accumulations of salt pans, which can reach depths of thousands of feet, are then buried by a mile or more of overburden. The less dense salt does not begin to flow through the overburden until the overburden has reached a depth where the lighter density salt can exert sufficient thrust to begin flowing through weak spots in what has now become a mile or more of rock. The plug of salt works its way toward the surface, fracturing rock layers along the way and forming potential traps for oil. The top layer of the salt plug becomes a dome of nonpermeable gypsum, limestone, dolomite, or other rock residue left after salt has been leeched out by subsurface waters. Salt dome caprock can be 100–1,000 feet thick and a salt plug can be from one-half to six miles across and extend as much as four miles below the surface. There are hundreds of salt domes in the Gulf of Mexico and the coastal plains of Texas, Louisiana, and Mississippi; most do not trap oil. Spindletop, an exception—with oil trapped below its salt dome of dolomite—was drained after only one year of unrestrained production. It became a new site for oil production when, twenty years later, oil was found trapped in fractured rock along its flanks.

The amount of partially decayed organic or biotic matter that would have to be contained in ocean sediment to create all the known reserves of oil and natural gas might strain one's imagination. In the 1950s Russian petroleum geologists proposed an alternative theory for the origin of oil, but it was generally discredited by Western geologists. But the theory was revived in the 1990s when an oil well named Eugene Island 330 suddenly began producing more oil. The well had originally begun production in 1972, and peaked at 15,000 barrels per day. By 1989 production had declined to 4,000 barrels per day. Then, suddenly, production rose to 13,000 barrels per day and estimates of its reserves were revised upward from 60 to 400 million barrels. The well was located near a huge towerlike structure with deep fissures and fractures and the new oil was from an earlier geological age. Seismic evidence seemed to suggest that the new oil was flowing up from one of the deep fissures.

The inorganic or abiotic origin of oil theorizes that there are vast deposits of natural gas in the earth's mantle. Natural gas penetrates the crust and is transformed into crude oil as it rises toward the surface with its properties determined by surrounding rock. As it ascends to the surface, the oil picks up organic matter in the sedimentary rock, which, according to those who support the abiotic theory of oil's origin, explains the presence of organic matter in oil. If natural gas rises close to a volcano, it is transformed into carbon dioxide and steam, gases commonly emitted by volcanoes. The presence of helium in oil and gas, but not in sedimentary rock or organic matter, is put forth as further evidence of an inorganic origin. The most compelling argument against the abiotic origin made by Western petroleum geologists is that they have successfully discovered oil and natural gas on the basis of a biotic or organic origin. Unfortunately, this is the same argument advanced by Russian petroleum geologists who have discovered oil in the base rock beneath sedimentary rock on the assumption that oil has an abiotic origin.

If the abiotic explanation is true, as some earth scientists maintain, then oil and gas may become sustainable forms of energy if the earth produces oil and gas as fast as we consume them. This would have an enormous impact on energy policy if oil and gas were being replenished by the earth or if oil and gas reserves are underestimated by a factor of 100 as suggested by some advocates of an abiogenic origin.[2]

Oil Exploration

In the early years of oil, drillers imagined that they were drilling for a pool or an underground river of oil using oil seeps as a guide for where to drill. Such exploration was successful if the surface oil came from an oil reservoir directly beneath the seep. Many seeps offered little reward to the driller as they merely marked the spot where the migratory rock breeched the earth's surface. Drilling straight down missed oil embedded in a layer of migratory rock slanted at an angle to the surface. Since the first oil was found near a creek, early oil drillers followed creek beds, thinking that oil flowed beneath running water. Once oil was discovered, production wells were placed as close as possible to one another to ensure commercial success. Spacing wells increased the chance of drilling a dry hole beyond the perimeter of an oil reservoir. This practice—having the greatest possible concentration of wells furiously pumping oil—rapidly exhausted an oil reservoir and another search for seep oil began. Once sites marked by seep oil were exhausted, and the creek theory debunked, oil drillers turned to geologists for advice on prospective sites. Geologists examined the land for hints of the presence of three necessary conditions for oil: (1) source rock to generate petroleum, (2) migratory rock through which petroleum moves toward the earth's surface, and (3) reservoir rock where there is an impediment preventing further migration. Whether sedimentary rock is source, migratory, or reservoir rock is a matter of circumstance.

Early geologists became geophysicists when they started using gravity meters and magnetometers to search for oil. A gravity meter is sensitive to the density of rocks below the surface. A mile of sedimentary rocks on top of basement rock is dense compared to a salt dome or a layer of porous reef or lighter rocks, which are detectable as anomalies or variations in gravity. Gravity meters were particularly useful in discovering salt domes in Louisiana and Texas during the early 1900s and in the 1948 discovery of Ghawar, the world's largest oil field, in Saudi Arabia. Magnetometers, because they are sensitive to anomalies or variations in the earth's magnetic field generated by magnetite in basement rock, are useful for estimating the thickness of overlying sedimentary rock. Both gravity and magnetic anomalies may indicate an anticline or fault that holds an oil reservoir.

Seismic analysis measures the time interval between creating a sound burst and the return of its echo from a subsurface geological structure embedded within sedimentary rock capable of reflecting sound. Seismic analysis is very useful in identifying potential traps. Dynamite was first exploded in shot holes dug through the surface soil to solid rock. The shot holes were laid out in geometric patterns to get a better idea of the subsurface structures. Later, explosive cord was used in a trench about one foot deep or suspended in air. Nowadays seismic work on land may utilize a truck that lowers a pad to sustain most of its weight. Hydraulic motors in the truck vibrate the ground, creating sound waves whose echoes can be analyzed for subsurface structures.

Seismic surveys on land are often conducted in difficult, inaccessible, inhospitable, and often unhealthy terrain such as jungles, deserts, mountains, or tundra. Seismic surveys are easier to conduct at sea. An array of pressurized air guns towed by a seismic boat is fired and the returning echoes are recorded by hydrophones, also under tow. The geometry of the array of air guns, their size, and sequence of firing are arranged to obtain a high signal-to-noise ratio of any returning echoes to more easily identify subsurface structures. While the first successful use of a seismic survey occurred in 1928 when an oil field in Oklahoma was discovered, seismic analysis did not reach its full potential until after World War II with the advent of computers capable of processing the enormous volume of data contained on a digital magnetic tape.

The first seismic pictures were two-dimensional (2-D) vertical views of what was beneath the ground. While this was valuable information in itself, a better picture of the size of a potential oil field would emerge if its horizontal dimensions were known. This could be obtained by taking a series of 2-D seismics; but by the 1980s computer processing speed, data storage capacity, and software programs had advanced to the point of being able to digest and analyze the mountain of seismic data necessary to obtain a three-dimensional (3-D) view of a subsurface structure. A 3-D seismic on land involves parallel receiver cables with shot points laid out perpendicular to the receiver cables. At sea a vessel, or several vessels sailing in parallel formation, towing several lines of air guns and hydrophone receivers collects the requisite data. Once processed, an underwater subsurface structure can be rotated on a computer screen in order to assess its shape in terms of length, width, and depth from any angle. While 3-D seismics are costly, they are also cost-effective because they lower the probability of drilling unsuccessful exploratory oil wells, reduce the need for development wells to determine the size of an oil field and make it possible to plan the placement of production wells to effectively drain a reservoir. Three-D seismics can identify natural gas reservoirs by the unique sound reflections of natural gas in rock. Four-D seismics are a series of 3-D seismics taken over time to assess the remaining reserves of a producing field.

Drilling Rights

The United States and Canada are unique in permitting individuals and companies to own both the surface and the subsurface rights of land. All other nations consider subsurface minerals the property of the state regardless of who owns the surface land. In the United States and Canada, surface rights to build a house or farm the land can be separated from the subsurface rights to explore and develop mineral finds. If separated, a lease agreement has to be reached between the owners of the surface and subsurface rights with regard to access to the land, the conditions for exploration and the development of any discovered minerals, including oil and gas. Lease agreements usually contain a bonus payment on signing and a royalty payment to be paid to the owner of the surface rights if minerals, including oil and gas, are found and stipulate a time limit for the start of exploration. If exploration has not started by the time established in the lease, the lease becomes null and void and the subsurface mineral rights revert to the owner of the surface rights. Leases can also be farmed out to third parties who conduct exploration, and working interests can be sold to third parties to raise funds to develop an oil or mineral find.

Large portions of the United States and Canada are not owned by individuals, but by the federal governments. In the United States, the federal government holds auctions for mineral rights on its land holdings and offshore waters. Rights to drill on blocks on the continental shelf, whose depth is within the capability of offshore drilling rigs, are offered periodically in a closed-bid auction. The highest bidder has a five- or ten-year period, depending on the depth of the water, to begin exploration or the mineral rights revert back to the federal government. The U.S. government receives a one-sixth royalty if oil is found. Canada has different rules that vary among the provinces. In addition, if oil is discovered on land, there are government regulations on the spacing of production wells to avoid overproduction, the fruit of the bitter lessons learned from the early exhaustion of oil fields in western Pennsylvania, Spindletop, and elsewhere. Of course, providing a long-term optimal return on a costly investment is also a strong guiding force for oil field managers in determining the spacing between producing wells.

For the rest of the world, governments own the mineral rights regardless of who owns the surface land. Oil companies normally enter into individual or collective contracts with governments or their national oil companies for the three phases of oil and gas operations: exploration, development, and production. The type of contract most commonly used before the oil crisis in 1973 was an exclusive concession granted to an oil company for a defined geographic area. The oil company bore all risks and costs and the host government received some combination of bonuses, taxes, and royalties if oil and natural gas were discovered. Since the oil crisis, the most common form of contract has been the production sharing contract that was first written in Indonesia in 1966. Again the oil company bears all risks and costs. The oil company explores for oil, and, if it is successful, develops the oil field for production. A large share of the initial oil and gas revenue is dedicated to the recovery of exploration and development costs. After these costs, with a stated rate of return, have been recouped, the oil company and the host government, usually through its national oil company, share the remaining revenue. Some contracts have the host government's national oil company bear a portion of the costs and risks of exploration and the responsibilities of development. Service contracts are payments to oil companies for services rendered in exploration, development, or production, with the profits of production residing solely with the host government. However, a host government may provide incentives to the oil company for meeting certain goals, and, from this perspective, the oil company shares in the profits. Of course, some national oil companies prefer to go it alone without assistance from Western oil companies.

Drilling Equipment

There are three types of wells: exploratory or wildcat wells, appraisal or development wells, and production wells. Wildcat wells are drilled a significant distance from known oil fields in search of new oil. This is the pure gambling aspect of oil. Less risky exploratory wells can be drilled near existing oil fields in search of extensions to an existing field or a neighboring field. If oil is discovered, then the development phase of an oil field begins with the drilling of appraisal, or step-out, wells to measure the extent of an oil field and determine the number and placement of production wells. Appraisal wells are normally abandoned after an oil field has been evaluated. Drilling onshore or offshore is similar, with the major difference being the nature of the drilling rig and in having from several hundred feet to two miles of drill pipe between the rig and the well bore.

Cable-tool rigs have a long history for drilling for freshwater or brine, which was evaporated for its salt content. By the time of "Colonel" Drake, the founding father of the oil industry, a cable-tool rig was a four-legged wooden tower, called a derrick, between seventy-two and eighty-seven feet high. A steam engine drove a wooden walking beam mounted on a Samson post that created an up-and-down motion. A chisel-pointed steel cylinder, or bit, about four feet in length, attached by a cable or rope to the walking beam, pulverized the rock at the bottom of the well. Every three to eight feet, the bit was raised and a bailer lowered into the well to remove the rock chips. Drake was the first to use a cable-tool rig to drill, not for freshwater or brine, but for oil. He was also the first to install a casing of large diameter pipe in the well to keep water from filling the well and to prevent the sides of the well from caving in, a practice still in use today.

The advantage of the cable-tool rig was its simplicity of design and operation, requiring only two or three men (a driller, a tool dresser, and maybe a helper). The disadvantage was that it was slow, averaging about twenty-five feet per day depending on the type of rock. While rotary drilling bits had been used to drill for water as far back as the early 1820s, the cable-tool rig was exclusively used to drill for oil after Drake's discovery. Captain Lucas, developer of Spindletop, is credited with the first use of a rotary rig for oil exploration, where cable-tool rigs could not drill to the desired depth in the soft, sandy soil. The Hamill brothers, employed by Lucas, were major innovators in rotary rig operations. The rig employed at Spindletop, while primitive by contemporary standards, possessed the essential elements of a modern rotary rig. Early rotaries required five or more people to operate and were much more efficient in drilling a hole than a cable-tool rig.

The modern tricone rotary drilling bit, exclusively in use today, is capable of drilling hundreds or a few thousand feet per day depending on the type of rig and the nature of the rock. A rotary drilling bit is a fixed attachment at the end of a drill string, rotated by rotating the entire drill string. The invention of the tricone rotary drilling bit in 1908 by Howard Hughes, Sr., founder of Hughes Tool Company, with its greater productivity should have spelled the instant death of the cable-tool rig, but it did not. Drilling with a rotary drilling bit requires more costly equipment and a larger and more knowledgeable crew, making it more expensive and complex to operate. Despite its higher productivity, as late as 1950 half of the drilling rigs in the United States were still cable-tool rigs. Desperately in need of replacement after being idle during the Great Depression of the 1930s, and worn out operating without spare parts during World War II, cable-tool rigs quickly passed from the scene by a massive conversion to rotaries. This greatly benefited

Hughes Tool and the son of its founder, Howard Hughes, Jr., the infamous billionaire Hollywood movie producer, aircraft designer, mining mogul, casino owner, front for a government secret mission to raise a sunken Soviet submarine, tax evader, and ladies' man, who died a bitter and mentally disturbed recluse in 1976. The lesson with Hughes is that one does not have to discover oil to become an oil billionaire. Maybe there is another lesson. . .

When drilling on land, the ground is first prepared to support a rig and a cellar is dug into the ground in preparation for the conductor casing. For shallow wells of up to 3,000 feet, the rig can be mounted on the back of a truck. For deeper wells, the rig is broken down into segments, transported by truck, and reassembled at the site. The deeper the well, the stronger the rig has to be to support and pull out the drill string. Drilling a well starts with drilling, called spudding, a hole twenty to one hundred feet deep to cement in a conductor casing of up to twenty inches in diameter. The conductor casing stabilizes the top of the well and provides an anchorage for the blowout preventer, which, as the name suggests, is a surefire way to seal a well against a blowout of natural gas and oil.

A cable passing through the topmost crown block of a rig is connected to a kelly (a very strong four- or six-sided molybdenum steel pipe forty or fifty-four feet in length) by a swivel. The sides of the kelly are gripped by a rotary table turned by electric motors powered by diesel engines of 1,000–3,000 horsepower. Attached to the kelly is drill pipe in lengths from eighteen to forty-five feet, but most commonly thirty feet. Every thirty feet, drilling is stopped to add another length of drill pipe below the kelly. The outer diameter of drill pipe varies between three and six inches. Drill pipe nearest the bit is heaviest in gauge to provide additional weight to control drilling and prevent the drill string from kinking and breaking.

The tricone drill bit is a solid fixed cone at the bottom of the drill pipe with three counter-rotating sets of teeth of steel, high-grade tungsten carbide steel, or industrial diamonds, depending on the type of rock and the speed of drilling. The well bore is larger in diameter than the drill pipe to allow the drill pipe to rotate and slide up and down in the well. The drill string, driven by the kelly, rotates fifty to one hundred turns per minute, enabling the teeth of the drill bit to pulverize the underlying rock. The teeth on the drill bit wear out after 40–60 hours of use on average, but can last as long as 200 hours, depending on the type of rock and the type of teeth. The success of the tricone drill bit was that it lasted much longer than previous bits, sharply reducing the number of trips that had to be taken to replace the drill bit. The increased drilling productivity allowed Hughes to charge a premium price for his bits and, thereby, amass a large fortune.

Tripping out refers to the process of pulling the drill string and unscrewing each length of pipe for stacking in the derrick. For an offshore drilling rig that is floating on mile-deep water and drilling a well two miles into the earth's crust, tripping out requires pulling up and disconnecting three miles of drill pipe. For deeper waters and wells of greater depth, this may mean up to eight miles of pipe, a length at the current limits of drilling technology. After a new bit is attached, the reverse process of *tripping in*, or connecting from three to eight miles of pipe, is performed, so that during an entire trip six to sixteen miles of pipe must be handled. Taking a trip is dirty, tough, and dangerous work; no wonder the workers are called roughnecks. Drilling is a twenty-four-hour, seven-day (24–7) operation requiring three shifts of workers working eight-hour shifts, or two shifts of workers working twelve-hour shifts plus spare shifts to give the workers time off.

Drilling mud, which carries away the debris of pulverized rock from the bottom of the well, is forced down the center of the drill pipe and passes through the middle of the tricone bit. The mixture of mud and pulverized rock is forced to flow up the annulus, the space between the drill pipe and the walls of the drill hole or well bore. Separators on the surface remove rock chips so that the mud can be recycled. Drilling mud, first introduced by the Hamill brothers when they drove cows through a mud pit at Spindletop, is now a science. Mud consists of a mixture of clay, weighting material, and chemicals mixed with water or diesel oil. Bentonite clay remains suspended in water for a long time after agitation and adding barite or galena controls drilling mud's viscosity and weight. Viscosity affects how fast the mud can pass through the tricone bit and the weight of the mud, along with the weight of the drill pipe, must exceed the pressure of oil, gas, or water in the well to prevent a blowout. Depending on the circumstances, it may be necessary to add bactericides, defoamers, emulsifiers, flocculants, filtrate reducers, foaming agents, or a compound to control alkalinity. When drilling through soft or porous rock, rising mud can penetrate the surrounding rock to strengthen the sides of the well and form a seal, preventing subsurface fluids from flowing into the well.

Nothing about drilling is easy. One of the many challenges facing a driller is the possibility that the drill string might bend and bind itself to the well wall or break as a result of metal embrittlement, which occurs when hydrogen sulfide enters a well. Nothing is more risky. Safe operations to reduce the risk of an accident are of paramount importance. An unexpected release of high-pressure gas into the well, whose expansion in the mud lowers its density, may lead to a blowout. A blowout can shoot drill pipe out of a well like cannon shot, wrecking a rig and killing or maiming the drillers. If a blowout is about to occur, pipe rams and other means within the blowout preventer seal the well. Drillers are sensitive to the dangers that threaten drilling by relying both on instrument readings and sight, sound, and smell for warnings of potential trouble.

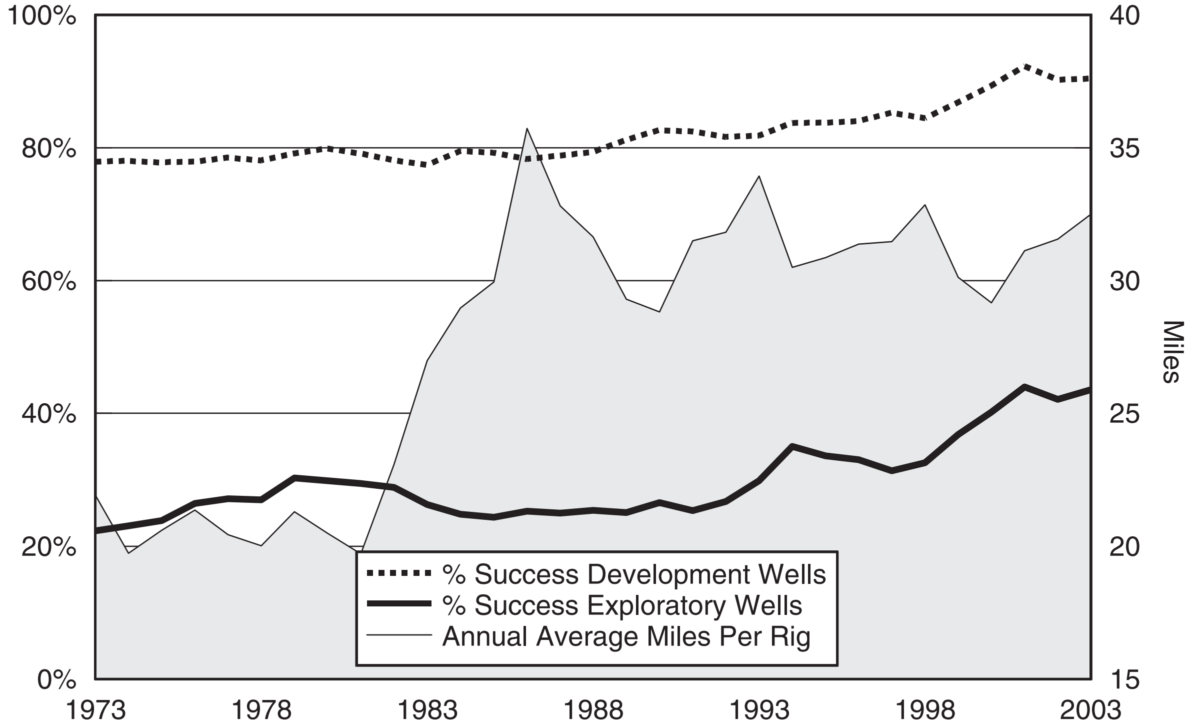
Figure 6.1 shows the probability of success for the exploration and development of natural gas and oil wells in the United States, along with a measure of drilling productivity.[3]

In 1973, one out of five exploratory wells was successful, but this increased to two out of five wells by 2003, primarily as the result of improvements in computer software, processing speeds, and storage capacity to perform 3-D seismic analysis. However, these exploratory wells were not true wildcat wells in that most of them were drilled in the vicinity of known fields. True wildcat wells, drilled far from any known source of oil, have much lower success ratios.

Four out of five production wells were successfully completed in 1973, and this improved to nine out of ten in 2003. Failure in drilling production wells can be caused by the well hole missing the oil reservoir (3-D seismic pictures can reduce the chances of this happening) or by twisting, binding, or breaking of the drill string, shattering of the drilling bit, or the damage incurred when a tool is dropped down the well. When the drill string or drilling bit breaks, various methods of fishing out the broken pipe or bit have been devised, including the use of powerful magnets and explosives. If these fail, the well may have to be abandoned.

Figure 6.1 also shows the improvement in drilling productivity. In 1973, the average rig drilled wells whose total depth amounted to twenty miles each year. If the average depth per well is two miles, then this is equivalent to drilling ten wells a year. During the early 1980s, when oil prices hit their highest levels in modern times, drilling activity blossomed from 2,000 rotary rigs in operation in

Figure 6.1 **Success Ratios and Productivity of U.S. Oil and Gas Wells**



1977 to nearly 4,000 in 1981. These new, technologically improved rigs increased average total depth drilled by 50 percent, to over thirty miles. After oil prices fell, so did the number of operating rotary rigs, to about 1,000, a number that has remained more or less at this level up to the present time. Even with a run-up in oil prices in 2004, land rig activity remained flat in the United States as oil exploration continued to move into more promising areas overseas and in offshore waters.

Directional Drilling

One would think that a drill string made up of thirty-foot lengths of steel pipe would be rigid, but this is not at all true when the drill string is measured in miles. It is rather flexible, sometimes described as similar to spaghetti, and, like a rubber band, actually twists several times when being rotated before the bit begins to turn. Drilling a vertical hole requires constant attention. The bit, turning clockwise, tends to introduce a clockwise corkscrew pattern to what one would suppose would be a straight hole. Moreover nonhorizontal layers of hard rock can cause a deviation or drift from the vertical. Drift measurements are necessary every several hundred feet of drilling, or when a bit is being changed, to verify the direction of the well hole.

Wells on land are drilled vertically, and a producing oil field is served by many individual wells. Historically, the only slanted wells were those that tapped a neighbor's oil field! Vertical wells spaced in a geometric pattern would be far too expensive in deep offshore waters, where it is more economical to have one production site served by several directional wells. A directional well is first drilled vertically and then cased. A hole or window is cut into the casing and an installed whipstock bends the drill string at a preset angle. A pilot hole of ten to fifteen feet in length is drilled, and when it is certified to be at the right direction and angle, which can be as great as 60 degrees, a turbo drill can be used. With a turbo drill, the drill pipe no longer rotates to drive a fixed drill bit. Mud driven down the center of a stationary drill pipe turns a rotary bit attached to the end of the drill string where it exits the drill pipe and enters the annulus. A magnetic compass and a gyroscope control the orientation and the degree of deviation of the well. The degree of deviation can be changed at various points along the way until the well is horizontal. Generally speaking, horizontal wells are more effective in draining an oil reservoir than vertical wells, particularly those with low permeability. A recent technological innovation is slim-hole drilling, which reduces the time required to drill wells and lowers the cost because of less tubular steel and cement to case a well. Another innovation that can further reduce drilling costs is coiled-tubing drilling, which eliminates the need for screwing and unscrewing drill pipe during trips to replace drill bits or for other reasons. High-pressure drilling mud powers a hydraulic motor that rotates the drill bit as it passes through the tubing.

Offshore Drilling Rigs

Seventy percent of the earth's surface is covered by water, a tempting prospect for oilmen when the promise of discovering large onshore oil fields dimmed. In the 1890s a slanted well was drilled onshore to tap an offshore oil deposit and oil rigs were built on wharves extending out to waters thirty feet deep in the Santa Barbara channel. In 1910 a gas well was completed one mile from shore in Lake Erie. In 1937 a grounded barge in shallow waters off Louisiana was the first submersible drilling rig, which does not mean that operations were conducted underwater. The submerged barge formed a barrier to keep the water out and exposed the bottom for drilling as though it were on land.

The birth of the modern offshore drilling industry occurred after World War II. The Depression of the 1930s dampened demand for developing new oil fields and World War II dedicated the nation's resources to the production of ammunition and armaments, making it necessary for onshore oil fields to be worked with old equipment. Steel was not available for building rigs, for manufacturing spare parts, or, for that matter, for producing consumer goods and automobiles. The end of the war also marked the end of postponing the good life. Americans proceeded to build homes, manufacture consumer goods to fill the homes, and build roads and manufacture automobiles to fill the roads. The postwar demand for oil exploded, but the chance of discovering major or giant new fields onshore was not overly bright because much of the land surface had already been explored. The gentle sloping Gulf of Mexico and its relatively shallow waters, bordered by onshore oil fields, provided a promising new region for exploration.

In very shallow waters, submersible drilling barges were sunk to rest on the bottom and keep water away from the well. This allowed drilling to proceed as though it were on dry land, a practice still in use. Floating drilling barges were and still are used in waters up to about twenty feet in depth. The drilling platform is mounted in the center of the barge and the drill pipe runs through a moon pool, a hole in the bottom of the barge surrounded by a cylinder that keeps water from flooding the barge. Kerr-McGee was among the first companies to convert excess barges and landing craft built during World War II to floating offshore rigs and was a technological innovator in developing a new industry. In 1947 the first offshore production rig was built on steel pilings driven into the sea floor by a pile-driver that repeatedly drops a heavy weight on a piling. The rig was in waters eighteen feet deep, twelve miles off the coast of Louisiana.

Legislation was passed in 1953 that granted state control over mineral rights in coastal waters and past that point to the federal government. Sales of leases are an important source of revenue for states that permit offshore drilling (Louisiana and Texas) and the federal government. Most states have banned offshore drilling in their coastal waters or new exploratory activity if offshore operations already exist (California). States have also succeeded in prohibiting exploration in offshore waters outside their jurisdiction such as the Hudson River canyon, where some believe an oil field may exist. In 1998 President Clinton signed an executive order extending a moratorium on leasing oil-drilling sites on the outer continental shelf until 2012. This moratorium includes nearly all the coastlines of California, Washington, Oregon, the Aleutian Islands, New England, the north and mid-Atlantic, and the eastern Gulf of Mexico, which includes an extensive area off the coast of southwest Florida. The Middle East is not the only place where oil companies face geopolitical risk.

In 1954 a submersible called "Mr. Charlie" was built. It consisted of a lower barge with a moon pool on which cylindrical columns were built to support a drilling platform. The barge was sunk in waters up to forty feet deep by flooding the lower barge, leaving the drilling platform above the water surface. When the well was completed, water was pumped out to refloat the barge for towing to a new location. "Mr. Charlie" drilled hundreds of wells before retiring in 1986 and was a springboard in the development of offshore drilling technology.[4] For deeper waters, jack-up rigs were developed. These consisted of three legs and an upper and lower hollow hull. In transit, the lower and upper hulls are together with the legs sticking up in the air. At the drill site, the lower hull is flooded and the legs are jacked down until the lower hull is firmly on the ocean floor. For hard ocean bottoms, cylinders built on the legs, rather than a lower hull, support the rig. With the lower hull or cylinders firmly on the bottom, the jack-up rig continues to jack down the legs, raising the upper hull with the drilling platform off the surface of the water until it is 100 feet above the ocean surface. The drilling platform must be high enough above the water surface to prevent waves from striking and possibly capsizing the rig. The first jack-ups were used in waters up to 80 feet deep, and are now capable of operating in waters up to 400 feet deep. In order for the drilling platform to be 100 feet above the water surface, the legs have to be 500 feet long. When being towed to a new site, the jacked-up legs stick up 500 feet into the air (equivalent to a thirty-story building), presenting a rather ungainly sight with transits limited to fair weather.

Drilling in deeper waters required another major step in technological development, which began with the government-sponsored Project Mohole (1958–1966), which envisioned drilling a well 25,000 feet deep in 15,000 feet of water. The purpose of the project was not to advance offshore drilling, but to enhance the understanding of the earth's geology by drilling through the crust in a place where it was thin enough to obtain a sample of the earth's mantle. While the project itself was never completed, it helped advance the technology of deep-water drilling. Five holes, one 600 feet deep, were drilled in 11,700 feet of water through a moon pool in the bottom of the drilling vessel. A system of swivel-mounted propellers was invented that imitated the method lobstermen use to keep their boats on station while they retrieve lobster pots. The system was a precursor of computer-controlled dynamic positioning.[5]

Drill ships and semisubmersibles are employed for deep-water drilling. A semisubmersible is a floating drilling platform mounted on top of a set of columns connected to large pontoons. The pontoons are empty when in transit, then flooded when on-station to lower the rig to a semisubmerged state with the pontoons thirty to fifty feet below the ocean surface. When in a semisubmerged condition with the drilling platform above the surface, the rig is quite stable even in stormy seas such as in the North Sea, where waves may strike the drilling platform. An anchor system was sufficient to keep the first generation of semisubmersibles on-station for drilling in water up to 2,000 feet deep. Modern ultra-deep fifth-generation semisubmersibles drill in waters up to 10,000 feet deep, and are kept on-station by a dynamic positioning system. Either sound impulses from transmitters located on the ocean floor or global positioning navigational satellites keep track of the rig's position and computer-controlled thruster engines maintain the rig on-station.

Drill ships have a drilling rig mounted at the center of the ship where the drill pipe passes through a moon pool. Depending on the depth of water, a drill ship remains on-station utilizing either an anchoring system or dynamic positioning. While top-of-the-line drill ships can also drill in waters 10,000 feet deep, drill ships lack the inherent stability of semisubmersibles. The vertical motion of a drill ship from wave motion would place an enormous strain on the drill string. A motion-compensated crown block keeps the drill string steady even though the ship is not. The *Discoverer Enterprise,* built by Transocean, can drill in water up to 10,000 feet deep, with well depths (combined horizontal and vertical length) of 35,000 feet, a figure not that far from the overall goal of the Mohole Project. At its extreme limits of operational capability, the drilling rig is rotating over eight miles of drill pipe. To reduce the number of connections that have to be made when making a sixteen-mile trip, this drill ship uses 135-foot pipe lengths rather than the more conventional 90-foot lengths usually used for offshore drilling.[6]

Similar to drilling on land, offshore drilling starts with a 100- to 250-foot, 30-inch diameter conductor pipe to which the blowout preventer is bolted. Remotely operated vehicles (ROVs) do the necessary work at the bottom of the ocean with lighting systems, television cameras, and remotely operated "arms" and "wrists" doing work at pressures that would crush a human body. A flexible, metal hollow tube called the marine riser connects the drilling rig with the borehole. Transponders and underwater television identify the location of the borehole and a jet-assist device at the bottom of the marine riser lets the driller guide the riser into the borehole. The drill string passes through the marine riser forming a closed system for circulating drilling mud. Once it is set up, drilling progresses much as it does on land, except for longer drill pipes and the additional risks of drilling at sea. Crew safety takes on a greater sense of urgency because the crew on an offshore drilling rig has limited means of escape if the rig catches fire, explodes, or capsizes during a storm. An additional risk is drilling into a large pocket of natural gas, which, if released in massive quantities, surrounds the rig and causes a loss of buoyancy that may result in its capsizing or sinking.

Drilling offshore wells requires a crew that varies with the size of the operation, but basically consists of a driller, who oversees operations, plus an assortment of derrickmen, motormen, diesel-engine operators, pump operators, mud men, and crane operators, plus roughnecks and roustabouts. Roughnecks handle connecting and disconnecting drill pipe and roustabouts handle supplies, including bringing drill pipe onboard the rig. Others are employed in maintenance and repair of machinery and equipment, keeping the rig on-station, and maintaining a livable environment. Schedules vary, but basically there are four crews, two aboard the rig working twelve-hour shifts and two on shore waiting their turn. The length of their stays onboard the rig depends on the rig's location: the more remote the location, the longer the stay. Crew transfers can be accomplished either by fast boats, for rigs relatively close to shore, or by helicopters.

The growing number of offshore-drilling rigs capable of drilling in ever deeper waters is an indication of the growing challenge of finding new oil fields; an indication that we may be running out of places to look for oil. The world fleet of offshore rigs totaled 586 in mid-2004, of which 383 were jack-up rigs, 160 were semisubmersibles, 36 were drill ships, and 7 were submersibles. Of the 383 jack-up rigs, 174 are capable of drilling in waters over 300 feet in depth. The 160 semisubmersibles have gone through five generations of technological advancement; none of the first generation rigs remain. Eighty-four of the second generation, thirty-six of the third, twenty-eight of the fourth, and twelve of the fifth-generation rigs are still in operation. High-end semisubmersibles and drill ships, used for drilling exploration and development wells, can cost up to $350 million each. Depending on the operational capacity of the rigs and the state of the market, the rate to employ a semisubmersible or drill ship can range from $30,000 to over $200,000 per day. Although a rule of thumb is that each million-dollar increment in construction cost increases the daily rate for employment by $1,000 per day, wildly fluctuating rates are not so much in tune with what it costs to build a rig as much as in fluctuations in the underlying demand with respect to supply. Net demand for jack-up and semisubmersible rigs in 2004, in terms of numbers, was about the same as in 1990, but newer rigs have greater operational capabilities and higher productivity than older rigs, thus masking rising demand. Table 6.1 shows the principal areas of the world employing jack-ups and semisubmersibles.[7]

Of these areas, the North Sea is by far the most challenging, where drilling proceeds 24–7 in freezing weather with frequent, strong, and long-lasting storms. These rigs have to be capable of operating in hundred-knot winds and hundred-foot waves. However, rigs operating in the relatively calm waters of the Gulf of Mexico must be able to withstand hurricane force winds and waves. As seen in Table 6.1, employment of semisubmersibles and drill ships in deep water drilling is concentrated in the Gulf of Mexico (off the United States and Mexico), the North Sea, and the offshore waters of Brazil, Nigeria, Angola, and other West African nations. Brazil has introduced many technological developments: It was the first nation to successfully drill in water more than 8,000 feet deep, and, in 2003, a Transocean drill ship operated at a record depth of 10,000 feet in the Gulf of Mexico. The company also held the record for the deepest subsea completion in waters of nearly 7,600 feet. "Other" in Table 6.1 includes rigs in-transit between areas, awaiting employment, and about forty technologically or operationally outmoded rigs in mothball status (another indication of the increasingly demanding environment for oil exploration and development).

Table 6.1

Employment Patterns of Jack-ups and Semisubmersibles (Mid-2004)

	# Jack-ups	# Semisubmersibles	# Drillships & Submersibles	Total
North America	89	21	13	123
South America	48	31	7	86
North Sea	26	29	1	56
Middle East	53	—	—	53
Far East/India	41	4	6	51
Africa	18	17	3	38
SE Asia/Australia	25	12	1	38
Mediterranean	22	5	—	27
Russia	14	7	—	21
Other	47	34	12	93

Once an oil field has been discovered and appraised, a permanent offshore production platform is installed. In shallower waters, a production platform is built on steel piles. For deeper water (hundreds of feet), a bottom-supported steel structure with hollow chambers to hold water is constructed on its side on land. Upon completion, the structure is placed on a barge and towed to the site. The barge is then partially flooded to launch the structure, which initially floats horizontally on the surface. By flooding designated chambers, the structure is slowly brought to an upright position; further flooding sets the structure on the sea bottom. Steel piles that pass through the structure's jacket are driven into the sea bottom. Once the structure is fixed to the sea bottom, equipment modules are loaded and assembled on the production platform.

Pile drivers cannot operate in the thousand-foot ocean depths of the North Sea. There, a gravity-based platform was built on land with a massive steel and concrete bottom with steel and concrete legs connected to a platform that extended above the ocean surface. The structure, ranking among the tallest on Earth, was built on its side, maneuvered onboard a barge, and towed to deep water. There the barge was partially flooded to offload the platform, which was then partially flooded to an upright position. After being towed to its final site, other hollow chambers in the base were flooded to allow the platform to sink vertically until its massive base penetrated deep into the ocean bottom. The remaining hollow cylinders connecting the base with the platform could be used for more seawater ballast, for storage of the diesel fuel needed for its operation, or crude oil. Gravity-based platforms were extremely expensive and are no longer built.

Nowadays, the tension-leg buoyant platform is used for producing oil and natural gas in deep waters. The platform floats above an offshore field; hollow steel tubes called tendons connect the floating production platform with heavy weights on the sea floor. These tendons are under tension and pull the platform down into the water to prevent it from rising and falling with waves and tides. Tension production platforms are very stable and have been successfully employed in the Gulf of Mexico and elsewhere in waters thousands of feet deep. Although tension production platforms are built to survive extreme weather, in 2005 hurricane

Katrina proved too much for one of them and turned it upside down. Like offshore exploratory rigs, production rigs operate 24–7.

Deep-sea production platforms are usually connected to a shoreside terminal by underwater pipelines, except in isolated regions of the North Sea, the offshore waters of Canada (Hibernia), West Africa, and Brazil, where shuttle tankers move oil from storage tanks within the production platform to a terminal. Besides isolation, another reason for not pipelining oil to a shore location is security of supply in locations where terminal operations are threatened by civil disturbances. When an oil field is exhausted, a production platform becomes obsolete and has to be removed and disposed of at considerable cost. In 1995 the environmental group Greenpeace aroused sufficient public opposition to Shell Oil's plan to move an obsolete production platform to deep water for sinking that Shell had to opt for the far more expensive method of towing the platform to a shoreside facility for dismantling.

A Floating Production, Storage, and Offshore Loading vessel (FPSO), often a conversion of an older but still seaworthy large crude carrier, has a production platform incorporated on the vessel's hull above a moon pool. The vessel provides storage for the oil and has an offloading arm for pumping crude from its tanks to shuttle tankers for transport to shoreside terminals. The advantage of a FPSO over a fixed production platform is that it is far less costly to build and install and its storage capacity eliminates the need for an underwater pipeline. The FPSO, because it relies on an anchoring system to remain on-station, cannot serve deep oil fields that require dynamic positioning. FPSOs can, however, exploit offshore oil fields that are too small to economically justify building a fixed production platform and laying a pipeline to shore. Once an oil field is exhausted, a FPSO sails to another oil field, and the company avoids the cost of dismantling a fixed platform and building a new one.

Evaluating a Well

Completing a production well, whether on- or offshore, is more costly than drilling an exploration or appraisal well. A careful evaluation of various logs obtained during the course of drilling an exploratory or appraisal well has to be completed prior to making a decision on whether to drill a production well. The lithographic, or sample, log records the nature of the coarser samples of rock chips separated from the drilling mud as to the type of rock, texture, grain size, porosity, microfossil content, and oil stains. Oil stains are examined in ultraviolet light to assess their nature and quality. The drilling-time log records the rate of penetration through subsurface rocks; a change in the rate of penetration indicates a change in the type of rock. The mud log records the chemical analysis of drilling mud for traces of subsurface gas and oil at various depths. The wireline well log, first introduced by Conrad Schlumberger in France during the 1920s, was, as with Hughes Tool, the basis for another fortune not directly related to owning oil-producing properties. The wireline well log was obtained by removing the drill string and inserting a sonde, a torpedo-shaped device laden with instruments. The first instrument was an electrical log to measure the resistance of the rocks to electricity. Changes in resistance indicate the degree of saturation of water, oil, and gas. Later additions included a natural gamma ray log to read the background radioactivity of rocks in the well. Since shale is the only sedimentary rock that emits radiation from radioactive potassium, the gamma ray log identifies the presence of shale rock or the degree of shale in mixed rock. A gamma-emitting radioactive source in the sonde creates a density or gamma-gamma log from returning gammas to measure porosity. The neutron porosity log records the results of bombarding rock adjacent to the well bore with neutrons. The intensity of returning neutrons from collisions indicates the presence of hydrogen, which is found in oil, gas, and water. The comparative results of the neutron porosity and gamma-gamma density logs identify the presence of natural gas. The caliper log measures the diameter of the well bore, which can widen when soft rocks slough off from the upward flow of mud in the annulus. This information is needed to calculate the amount of cement needed to case the well. The acoustic velocity, or sonic, log measures the speed of sound through a rock layer, which, for a known type of rock, indicates its porosity and the presence of fractures. The dip log measures the orientation of rock layers, or slant, from the horizontal.

Originally the sonde, with its various sensors, required pulling the drill string and removing the drill bit; since 1980, it is located just above the drill bit to provide real-time log analysis. Results of these logs are interpreted at each increment of depth by experts as to the likely productivity of the well; a key factor in deciding whether to complete a production well. If the experts decide to abandon a well, its conductor casing is pulled for salvage and the well is cemented at appropriate levels to prevent saltwater and oil seepage from rising and polluting surface waters.

Completing a Well

After the decision has been made to complete a well, the process starts with preparing a well for casing. Casing stabilizes a well, preventing the sides from caving in and protecting freshwater aquifers near the surface that might be polluted with oil, gas, and saltwater. If the casing is to be installed in a single operation after the well is completed, drill pipe is lowered with a used bit to circulate mud and remove any remaining cuttings from the bottom of the well. Wall scratchers remove mud from the sides of the well. Casing is thin-walled steel pipe, usually in thirty-foot lengths sized to fit inside the well bore. After the well is prepared for casing, casing pipe is screwed together and lowered into the well. A guide shoe guides the casing down the well and centralizers position the casing string in the center of the well. A float collar near the bottom of the casing string acts as a check valve to prevent mud in the well from flowing up the casing pipe. After the casing is in place, portland cement is mixed with additives to control its density and the timing required for the cement to set. Cement is pumped down the center of the casing through the float collar, forcing its way through the bottom plug out of the casing pipe and up the annulus between the outer casing wall and the well bore. Then a top plug is added and mud is pumped down the casing, which forces the remaining cement in the casing into the annulus. The driller has to ensure that an adequate amount of cement is injected into the annulus to complete the cementing of the casing string. When the top plug meets the bottom plug, cementing is complete and the wiper plugs, guide shoe, and cement at the bottom of the well are drilled out and the mud is removed from the casing pipe.

A variation of this method is to case a well in segments. After the well is drilled, the lowest section of the casing is cemented first, then a plug is installed at the top of the casing. More casing is added with holes drilled in the coupling with the installed casing to force cement out of the bottom to fill the annulus. The plug is removed and the process is repeated until the entire casing is installed. Some wells have three or four concentric casing strings installed in segments as the well is being drilled, with the largest diameter casing installed at the top of the well. After the casing is installed, the well is drilled deeper and another, narrower casing is added. This process continues until the narrowest casing string is added at the bottom of the well.

If the well ends in a producing zone, the bottom of the well is opened and filled with gravel. Smaller diameter liner pipe is run down the casing to the bottom of the well. Then the casing pipe is perforated and fractures are created in the rock and in any impregnated mud to ease the flow of oil and gas to the perforated casing wall and then into the liner pipe. Perforation was first accomplished in 1932 using a bullet-gun, a device lowered to the bottom of the well that fires bullets similar to ball bearings in all directions. The bullet gun is still in use and has been successfully employed in horizontal wells. Bullets are reduced to fine particles after firing. Shaped charges are also used, along with hydraulic injection of large volumes of diesel oil, nitrogen foam, water, or water with acid under high pressure for limestone reservoirs (the acid contains an inhibitor to prevent corrosion of the steel casing and tubing). Working over a well, which must be done several times over its lifetime, includes not only fixing mechanical problems and cleaning out the bottom of the well, but also taking measures to enhance the permeability of the surrounding rock.

The annulus between the liner pipe and the inner casing wall is sealed to prevent oil and gas from coming in contact with the casing pipe, which would corrode and weaken the casing. If the well passes through several producing zones, each has its own tubing and packing to ensure that the output of each zone is segregated in order to identify the output of each zone. Normally, a well will not have more than three producing zones. Shaped charges, or firing bullets, perforate the casing and fracture the surrounding rock at each producing zone. Seals are installed to ensure that oil and gas enter the liner pipe, not the casing pipe.

A "Christmas tree," normally made from a single block of metal, is mounted on top of the casing with a master valve that can shut off a well under emergency conditions. Other valves control the pressure and flow from each producing liner pipe or tubing string within the well with associated gauges that measure the tubing pressure. A new well usually has sufficient reservoir pressure to cause the oil to flow naturally from the top of the liner pipe or tubing string. If the reservoir pressure declines to a point at which oil no longer flows from the well, the most common form of lifting device is the sucker rod pump. A motor powered by electricity or natural gas from the well drives a walking beam mounted on a Samson post to obtain a vertical up-and-down motion to drive a pump. On the downward stroke, a ball unseats from a seal, letting oil flow into the pump. On the upward stroke, the ball seats force the oil up while the space below the pump fills up with more oil. The pumping rate has to be less than the fill rate for the pump to operate properly. A gas lock can form in the pump if natural gas is present. A sucker rod pump may have to be used in natural gas fields that release a lot of water. A gas lift system, which injects some of the natural gas produced by the well into the annulus between the tubing and the casing, can be installed for wells producing a mixture of saltwater, oil, and gas. Gas lift valves installed along the tubing string allow the gas to enter the tubing. The expanding bubbles in the liquid force the mixture of water, oil, and natural gas up the tubing. Gas lift systems are simple and inexpensive to operate, but are only effective for relatively shallow wells. Alternatively, an electrically or hydraulically driven submersible pump can be installed at the bottom of a well.

Moving the Oil to the Well

The primary force that causes oil and gas to flow through pores in the reservoir rock toward the bottom of the well is the pressure differential between the oil and gas within the reservoir rock and the pressure at the bottom of a well. The driving force in the reservoir can be provided by dissolved natural gas in the oil or by a natural gas cap on top of the oil that expands as oil is removed from a reservoir. Natural gas cannot maintain the same initial reservoir pressure as it expands, which causes oil production to decline with time. Subsurface water entering an oil reservoir from its bottom or sides as a primary driving force can maintain nearly constant reservoir pressure and oil production. An oil well goes "dry" when natural gas or water reaches the bottom of the well. Gravity can also be an effective drive mechanism for wells drilled into the bottom of steeply inclined reservoirs. Most oil reservoirs have more than one of these four primary driving forces.

Natural gas reservoirs are either driven by expanding gas or by water. Natural gas wells do not go dry in the sense that oil does, but their pressure may decline to somewhere between 700–1,000 psi, the lowest pressure acceptable for a gas pipeline. A compressor can extend the life of a gas well. If production from an oil well falls below 10 barrels per day, it is known as a stripper well. Stripper wells number about half a million in the United States, with many producing as little as 2 or 3 barrels per day. They are kept in production or reactivated if shut-in as long as revenue exceeds the costs of operation and reactivation.

Maintaining Reservoir Pressure

The recovery factor—the portion of the oil and gas removed from a reservoir—depends on the driving force. The recovery factor is lowest for oil reservoirs driven by natural gas in solution with the oil or by gravity, higher if driven by a natural gas cap, and higher yet if driven by water. The overall average recovery factor for oil fields is only about one-third (natural gas fields have higher recovery factors). Thus, when a well that relies on the natural drive of the reservoir goes "dry," about two-thirds of the oil is still in the ground.

Secondary methods to maintain reservoir pressure and promote oil recovery normally involve injecting water or natural gas. Injection wells, either specifically drilled or converted from abandoned producing wells, are placed to enhance the flow of oil in the direction of the producing wells. Water injection is the most common method for maintaining the pressure of an oil reservoir and is an environmentally acceptable way of getting rid of any brine produced by the well to avoid contaminating the freshwater table. If brine cannot be pumped into subsurface rock below the freshwater table, it must be disposed of in an acceptable manner. Brine may be placed in open tanks to let evaporation get rid of most of the water before disposal.

Depending on the type of reservoir rock, an alkaline chemical such as sodium hydroxide is mixed with the injected water to enhance recovery. Injected water must be compatible with the type of reservoir rock to ensure that a potential chemical reaction does not decrease its permeability. Pores in the reservoir rock can be plugged by injecting suspended solids in the water or by slimes feeding on injected bacteria and organic matter. Natural gas from an oil well, called associated natural gas, is normally sold, but for isolated wells far from natural gas pipelines, it is often reinjected into the oil field to maintain reservoir pressure. However, natural gas is not as effective as water in enhancing oil recovery.

Secondary methods can raise recovery to 40 percent on average from the one-third average recovery of primary methods. To reach 50 percent, tertiary or enhanced recovery methods must be employed. The price of oil plays a critical role in determining whether more costly tertiary recovery methods should be employed. Thermal recovery is utilized when the remaining oil is heavy and viscous. "Huff and puff" burns crude oil at the surface of the well to produce steam that is injected down a well. The well is shut in to allow steam to heat up the surrounding crude to reduce its viscosity, enhancing its flow through the rock. Then the well is put back into operation to extract the heated crude. Steam flooding is a continuous process in which injected steam maintains pressure on previously injected condensed steam to drive heated crude toward the producing wells. Placement of the steam injection wells is critical to ensure that the oil flows in the right direction. Thermal recovery is effective as long as crude production exceeds the amount burned to produce steam.

A fireflood is setting subsurface oil on fire and keeping it burning by forcing large quantities of air down an injection well, with or without water to create steam. The heat reduces the viscosity of the crude while increasing the pressure within the reservoir rock to enhance the flow of oil toward the producing wells. The amount of air has to be limited to avoid burning all the oil in the reservoir. Firefloods cannot be used if there is any appreciable sulfur in the oil because of the formation of sulfuric acid that eats away the liner pipe. While simple in concept, firefloods are difficult in practice.

A chemical flood involves inserting detergent into injected water to form tiny droplets of oil to aid in their migration to a producing well. As long as water is not present, miscible floods of natural gas liquids such as butane and propane act as solvents and wash the oil out of the reservoir rock. This is one of the most effective tertiary methods of oil recovery, but it is very expensive unless the butane and propane can be recovered for recycling. Carbon dioxide floods involve either carbon dioxide as a gas or dissolved in water. Soluble in oil, carbon dioxide promotes migration to the producing wells by increasing the volume of oil and reducing its viscosity. Injected carbon dioxide can be separated from the oil at the surface of the producing well for recycling. This is not sequestration of carbon dioxide as it returns to the surface dissolved in the oil. Carbon dioxide is brought to the well in a liquefied state in tanks or is piped in from wells that produce large amounts of carbon dioxide or as a waste by-product from nearby power, chemical, and fertilizer plants.

Tertiary recovery methods do not always succeed and require high-priced crude oil to justify their cost, but they do reduce the need to find new oil fields. With tertiary recovery, about half of the oil can be removed from an oil reservoir on average, although, as with any average, there are higher and lower recovery factors. Nevertheless, tertiary recovery methods still leave about half of the oil entrapped within the pores of reservoir rock after an oil field has gone "dry."

Getting the Oil to a Refinery

Most wells produce a mixture of oil and saltwater with or without associated natural gas. The output from a well enters a gas oil separator unit shaped like a cylinder where natural gas, if present, rises to the top and water sinks to the bottom, leaving oil in between the two. A heater or demulsifier may be necessary to break down an emulsion of oil and water. A certain retention time is necessary to allow the two to separate. Natural gas, if present, is diverted to a natural gas pipeline gathering system. Once separated from water, oil is pumped to a staging area that serves a number of wells and then through collecting pipelines to larger capacity pipelines that eventually connect to refineries.

Pipelines provide the lowest cost means of moving crude oil and oil products on land. Crude oil pipelines are not built unless there are sufficient reserves to guarantee pipeline throughput and provide an adequate financial return. Technological advances made in building the "Big Inch," a twenty-four-inch pipeline, and the "Little Inch," a twenty-inch pipeline from the U.S. Gulf region to the northeast during World War II set off an explosion in pipeline construction. Modern trunk lines are up to forty-eight inches in diameter and have a throughput capacity of 1–2 million barrels per day, depending on pumping capacity. Additives to make oil more "slippery" by reducing the friction or turbulence at the boundary layer between the oil and steel pipe can improve pipeline throughput capacity. The speed of oil in a pipeline is not very impressive, about that of a fast walk, but over twenty-four hours a pipeline with a diameter of four feet can move a lot of oil. The pipeline industry in the United States and Canada is regulated as a common carrier. Tariffs are set to limit earnings on investment with assurances that all shippers have equal access and pay the same basic rate.

Oil pipelines are like blood vessels in a living being, with the United States having hundreds of thousands of miles of gathering and collecting pipelines connecting countless producing wells to refineries. Most offshore oil fields such as those in the Gulf of Mexico and the North Sea are connected to land by underwater pipelines, although more remote fields use shuttle tankers. The Louisiana Offshore Oil Port (LOOP), located about twenty miles off the Mississippi River mouth in deep water, is a system of three single buoy moorings that serves large crude carriers carrying oil from the world's exporting oil nations. A discharging crude carrier pumps cargo from its tanks through a hose to the floating buoy. The floating buoy is connected via an underwater pipeline to an offshore marine pumping station. The pumping station moves the crude to onshore salt caverns for storage and connection to other crude oil pipeline systems that serve two-thirds of U.S. refinery capacity from the Gulf Coast to as far north as Chicago and as far east as the Middle Atlantic states. Russia also has an extensive crude oil pipeline system to handle domestic distribution and exports to Europe. Crude oil pipelines have been built to ship landlocked Caspian crude to Black Sea ports, and a major pipeline has been built to ship Caspian crude to a Mediterranean port in Turkey. Major projects under consideration involve pipelining Siberian oil to China and Japan and Russian oil to Murmansk, a year-round ice-free Arctic port, for export to Europe and the United States.

In addition to crude oil pipelines, product pipelines take the output from refineries to oil distribution terminals near population centers. Large product pipelines move refined products from the U.S. Gulf Coast refineries to the Atlantic and northeast markets and from Russian refineries to their markets in Europe. Tank trucks complete the movement from storage tanks at pipeline distribution terminals to wholesalers and retailers. In a few nations, railroads still move crude and oil products where the volume is insufficient to justify building a pipeline.

Tankers and Barges

Water transport is an even lower-cost alternative than pipelines because the "highway" is free, although investments have to be made in ports, terminals, and ships. Tankers and barges move about half of the oil produced either as crude from exporting terminals to refineries or as refined oil products from refineries to distribution terminals and customers. All the OPEC producers export oil by tanker, although pipelines can shorten the tanker voyage. The first Middle East export pipelines, now inoperative, carried Saudi crude to ports in Lebanon and Syria, eliminating the tanker movement from the Arabian Gulf to the Mediterranean via the Suez Canal. A portion of Iraqi crude is pipelined to a Mediterranean port in Turkey and some Saudi crude is pipelined to a Red Sea port for transfer to tankers for transit to the southern terminal of the Sumed pipeline that parallels the Suez Canal. Oil is shipped in tankers from the northern terminal of the Sumed pipeline in the Mediterranean to ports in southern and northern Europe. The Sumed pipeline allows the use of very large tankers that cannot transit the Suez Canal fully loaded to move Middle East crude to Europe. However, in about ten years' time the Suez Canal will be widened and deepened enough to accommodate most of the world's largest tankers fully loaded.

Refineries on or near the coastline distribute oil products locally by small tankers and barges. Barges distribute the output of Rotterdam refineries up the Rhine River into central Europe and along the northern European seaboard and from refineries in the United States Gulf up the Mississippi River and along the Atlantic seaboard. Product carriers move cargoes from export-refining centers in the Caribbean, Mediterranean (southern Italy), Middle East, and Singapore to nearby and far-off markets. Price differentials arise between regions when planned production and distribution do not exactly match demand. Traders take advantage of price differentials once they exceed shipping costs to arrange a shipment from a low-priced to a high-priced market. Arbitrage trading completes the balancing of global refinery supply with global consumer demand.

Standard Oil was the first company to export oil. The initial shipments of kerosene from the United States to Europe were carried in barrels on general cargo sailing vessels, some of which were lost at sea when leaking fumes came in contact with an open flame in the ship's galley. The first tanker, the *Gluckauf,* built in Germany in 1886, was compartmentalized into several cargo tanks whose outer tank surface was the hull itself, now called a single-hull tanker. The vessel's deadweight ton capacity (dwt) was 3,000 tons. As a rough rule, the cargo capacity of a tanker is about 95 percent of its dwt. Shell Trading was a major impetus in building larger and safer tankers in the early part of the twentieth century to ship Black Sea kerosene to Asia through the Suez Canal. By the end of World War II, the standard tanker, which had been built in large numbers for the war effort, was 16,000 dwt. As world oil movements increased in volume in the postwar era, tankers grew in carrying capacity to take advantage of their inherent economies of scale. The same size crew is required regardless of the size of the ship and the cost of building a vessel does not rise proportionately with its carrying capacity. Thus, the larger the tanker, the less its operating and capital costs in terms of cents per ton-mile of transported cargo. While there was talk of mammoth tankers of 750,000 and 1 million dwt in the early 1970s, the 1973 oil crisis cut short the development of these behemoths. Indeed, the fall in oil exports after the 1973 crisis brought on the most devastating and long-lasting tanker depression in history.

Very few tankers over 500,000 dwt were built (the largest are just over 550,000 dwt) as they proved to be too unwieldy to serve most of the world's terminals and ended their days as storage vessels. Water depth in ports, channels and alongside terminals, terminal storage capacity, cargo availability, and the annual throughput volume determine the optimally sized tanker for each trade. The largest tankers, called Very Large

Crude Carriers (Vlccs), range between 200,000–350,000 dwt. These vessels, which in 2004 numbered 430, dominate Middle East exports. Seventy percent of Middle East cargoes are destined for Asia and 30 percent around South Africa primarily to North America and the rest to northern Europe. This is the opposite of the split in cargo destinations in the early 1970s, when these tankers made their debut, and reflects the growing importance of Asia for Middle East exports. While these vessels were originally built to serve Middle East crude exports exclusively, nowadays Middle East exports provide about 70 percent of Vlcc employment. The remaining 30 percent hauls primarily West African crude to the United States and Europe and as backhaul cargoes to Asia. Other backhaul cargoes to Asia are North Sea crude, fuel oil from Europe and the U.S. Gulf, and orimulsion (a mixture of 70 percent bitumen and 30 percent water, which is burned as a substitute for coal) from Venezuela. A few Vlccs move Saudi crude from the Red Sea pipeline terminal to the southern Sumed pipeline terminal and from the northern Sumed terminal to northern Europe.

The next size category, Suezmax tankers between 120,000–200,000 dwt, numbered 290 vessels in 2004, and are primarily employed handling crude exports from West and North Africa and the North and Black Seas. Tankers smaller than Suezmaxes have more diverse trading patterns. Yet, despite there being about 4,300 tankers above 20,000 dwt, the 430 Vlccs, which represent 10 percent of the world fleet in number, make up 40 percent in carrying capacity. Clean or refined oil products are usually transported in carriers of less than 50,000 dwt, although naphtha shipments between the Middle East and Japan are carried in product carriers as large as 100,000 dwt. Clean products tankers are smaller than crude carriers, reflecting terminal capacity and water depth restrictions and lower throughput volume of clean products versus crude oil trades. They are also more sophisticated than crude carriers, with coated tanks and segregated cargo-handling systems to ensure cargo integrity. There are thousands of tankers and barges below 20,000 dwt, but these vessels are normally involved with intraregional distribution of oil products, not interregional trading.

Oil Spills

Although larger sized tankers reduce the number of tankers needed to transport oil, and, hence, the number of collisions, the environmental consequences of large tankers breaking up in open waters is worsened considerably by the greater quantity of oil that can be spilled. Tankers sinking far out at sea barely get mentioned in the press, but an oil spill that reaches land is another matter. Two of the first large oil spills were the *Torrey Canyon* in the English Channel in 1967 and the grounding of the *Amoco Cadiz* on the French coast in 1978. This sharpened environmental opposition to tankers, which came to a head in the 1989 *Exxon Valdez* spill in Alaska. Although only 15 percent of the vessel's cargo entered the environment (the rest was safely off-loaded on barges), it was enough to foul nearly a thousand miles of pristine coastline. The uproar over this spill was responsible for the passage of the U.S. Oil Pollution Act of 1990, which greatly increased the limits of liability of oil spills and required a gradual phase-in of double-hull tankers calling on U.S. ports. This was followed by amendments to international conventions that required double-hull construction for all tankers delivered after July 1996, along with a mandatory phaseout schedule of single-hull tankers based on age.

Double-hull tankers have a space between two hulls, where the inner hull is the exterior surface of the cargo tanks. Thus a grounding, or a collision, must be of sufficient force to pierce both hulls before oil can be spilled into the environment. The space between the outer and inner hulls holds ballast water to maintain a tanker's stability when it is empty and returning to a load port, and is empty when the tanker is carrying a cargo. In single-hull tankers, ballast water had to be carried in the cargo tanks. Although these tanks were cleaned prior to taking on ballast water, there was still some contamination of ballast water from oily residues. Ballast water in double hull tankers is free of oil pollution. However, this does not prevent the migration of sea life from one part of the world to another when ballast water is pumped out of the vessel at the load port.

The sinking of the *Erika* in 1999 polluted the French shoreline and the sinking of the *Prestige* in 2002 polluted the Spanish and Portuguese shorelines with fuel oil. The lighter ends of crude oil tend to evaporate when released, somewhat reducing environmental damage. Fuel oil is the residue of the refining process after the lighter end products have been removed. This makes fuel oil a worse pollutant than crude oil. The environmental damage wrought by these two spills reinforced public determination for "oil-spill-proof" tankers. Like the unilateral action taken by the United States after the *Exxon Valdez* incident, the European Union unilaterally shortened the phase-in of double-hull standards in European waters without bothering to obtain international approval or cooperation.

No one makes money in an oil spill other than those involved in cleanup operations and in handling lawsuits stemming from real or perceived damage. Certainly tanker owners and oil companies do not profit from an oil spill. The 1989 *Exxon Valdez* spill has cost Exxon $3.5 billion in cleanup costs and compensatory damage claims, and the company is still in court appealing a $4.5 billion punitive damage judgment. If Exxon fails in its appeal, the final bill for the spill will be $8 billion, plus interest accrued during the appeal process.

Tanker owners and oil companies have taken positive and costly steps to ensure the safe delivery of cargo. The record for tanker spills has improved markedly since the 1970s, with less crude spilled in absolute (total tons) and in relative terms (percentage of oil carried). But this record of achievement, never accepted in the public's mind as a manifestation of good intentions, evaporated as soon as the first drop of fuel oil from the *Erika* and the *Prestige* reached the shoreline.

Most people take great solace in the double hull being the magic cure for tanker spills. Actually, spills are the result of human error, the root cause of collisions, groundings, floundering on reefs, shoals, and rocks, poor design, shoddiness of construction, lack of thoroughness in tanker inspections, and in not maintaining a vessel fit for service at sea. It is true that double-hull construction prevents oil spills from low-energy collisions or groundings in which only the outer hull is breeched. This is not true for high-energy collisions or groundings in which both hulls are breeched. The *Exxon Valdez,* a single-hull tanker, floundered on an underwater rock that breached its single hull. The crude cargo, being less dense than water, kept the vessel afloat, permitting barges to come alongside and remove 85 percent of the cargo. Had the vessel been double hulled, the floundering would most probably have breeched both hulls. Water entering the space between the two hulls would have sunk the vessel, making it more difficult to off-load the cargo, and perhaps resulting in greater oil spillage. Try selling that concept to members of Congress reacting to public outrage!

Refining

There are approximately 40,000 oil fields in the world, which means there are 40,000 grades of crude oil because no two crude oils from different oil fields are exactly the same. However, oil from different oil fields in the same geographic region, with more or less common characteristics, share the same gathering and collecting systems that blend the slight differences into a common commercial oil such as West Texas Intermediate, Brent Blend, and so forth. Each commercial grade of crude oil has unique properties that determine its value with respect to others.

American Petroleum Institute (API) degree ratings measure the density of crude oil. Light crudes have a lower density than heavy crudes and are between 30–50 API degrees. Condensates, extra-light forms of crude oil found in natural gas fields, are as high as 65 degrees. Medium crudes are 22–30 degrees and heavy crudes vary between a very viscous 7–22 degrees. Sweet crudes are under 0.5 percent sulfur and sour crudes are over 1 percent sulfur, with intermediate crudes between the two.[8] Crude oils are also classed as naphthenic or paraffinic. Naphthenic crudes are more highly valued because they produce more naphtha, the principal ingredient in gasoline and the principal driver of the entire oil industry. Paraffinic crudes are waxy, an undesirable trait. Some extra-heavy waxy crudes are unfit for refining and are burned directly as a fuel. Waxy crudes require heating coils in the cargo tanks to keep the oil warm enough to be pumped in cold weather. There have been a few instances of heating coils failing during cold weather transits, resulting in the cargo congealing into one enormous ship-shaped candle.

The most highly valued crudes are naphthenic, light, sweet crude such as West Texas Intermediate. The output product slate of a refinery using light sweet crude is skewed to gasoline and other valuable light-end products. Arab Light is a paraffinic light sour crude oil, less desirable and less light than West Texas Intermediate. A heavy crude has an output product slate skewed to gasoil and fuel oil such as Duri, an Indonesian heavy sweet crude and Bachaquero 17, a Venezuelan heavy sour crude. The output product slate of a particular crude oil depends on the design of the refinery and its mode of operation. Some refineries are rather simple in design and restricted to light sweet crudes. Others are designed to run on a single type of crude oil with little ability to vary the output. If the output is too great for one product and not enough of another, the refinery operator may export one and import the other to balance supply and demand. Often the residues of simpler refineries, called straight run, are sold to more sophisticated refineries capable of cracking straight run into more useful products. More sophisticated refineries, so-called merchant refineries, can take a variety of crudes and process them with different modes of operation for different product slates. A mathematical modeling technique called linear programming selects the type of crude based on delivered cost, the output slate based on product prices, and the refinery mode of operation that maximizes profitability.

There is no such thing as a generic or plain vanilla oil product. Each oil product has several grades, each with a specific slate of characteristics or requirements to meet the demands of different markets. Motor gasoline has different specifications or limitations on octane rating, vapor pressure, sulfur, lead, phosphorus, gum, and corrosive impurities in addition to volatility standards (the degree of evaporation at specified temperatures). Specifications of gasoline sold in Europe are different than those in the United States. The U.S. gasoline market is particularly fragmented, with different specifications for different states that complicate life for refiners.

Jet fuels have specifications on acidity, aromatics, olefins, sulfur and mercaptans (a malodorous form of organic sulfur), flash point, gravity, vapor pressure, freezing point, viscosity, combustion and corrosion properties, and thermal stability. Although the same as jet fuel, kerosene for lighting and heating has, as one might expect, fewer and less demanding specifications. Gasoil for home heating and diesel fuel have standards that vary in terms of flash, pour, and cloud points, carbon, ash, viscosity, specific gravity, cetane (analogous to octane) rating, sulfur, and corrosive impurities. Diesel fuels have another set of specifications, depending on the type of diesel engine. Even heavy fuel oil, the bottom of the barrel, the residue of the refining process, has various specifications with regard to flash and pour points, water, sediment and ash content, sulfur and viscosity, depending on its end use, that is, whether it is to be burned in industrial plants or as bunkers for marine engines.

Refining is a bit of a misnomer since refining suggests purification. Refining is not so much purifying crude oil, but transforming it into different products by separating, altering, and blending various hydrocarbon molecules. The refinery process starts with preheating crude oil and adding chemicals and water. The mixture sits in a desalting unit where gravity separates the oil and water, washing out inorganic salts and trace metals that can corrode refining equipment and poison catalysts. Atmospheric distillation first heats crude oil above 720°F and the resulting vapors enter a distillation column or fractionating tower stacked with perforated trays. Hydrocarbon vapors rise and condense to a liquid on the trays and are transformed back into a vapor by heat exchange with other upwelling hot vapors. The vaporized hydrocarbons rise to a higher tray, condense, and are turned back to a vapor and rise again. Eventually a particular hydrocarbon vapor reaches a tray where it condenses to a liquid, but cannot collect enough energy from passing hydrocarbons to change back to a vapor. This continuous exchange of heat between liquid and vapor allows hydrocarbon molecules of a similar nature to collect on the same tray. The sorted liquid hydrocarbons are drawn off through outlets placed at different heights on the distillation column. The lightest hydrocarbons with the lowest boiling points or temperatures of condensation are drawn off at the top of the fractionating tower and the heaviest hydrocarbons with the highest boiling points or temperatures of condensation at the bottom.

Starting at the top of the fractionating tower, methane in the oil escapes without condensing and is collected and used in the refining process. Flaring of unwanted gases, while common in the past, now means a loss of revenue. The lightest hydrocarbons of butane, propane, and ethane condense below 90°F. A refinery does not just produce simple butane and propane, but also more complex forms such as butylene and propylene. To provide a brief taste of the complexity of the refining process, an alkylation unit with either a sulfuric or hydrofluoric acid catalyst (a catalyst promotes a chemical reaction without being part of it) can transform butylene to alkylate, a high-octane ingredient for motor gasoline or aviation fuel, plus other light end by-products, butane and isobutane.

Light end products of the refining process can become part of the gasoline pool or end up as petrochemical feedstock to create the wonderful world of plastics. Walk around a house and look at all the objects made from plastic. One would be surprised at the extent of plastic in automobiles or the use of plastic in medical facilities (tubing and plastic bags for intravenous feeding and a host of other uses, blood sample vials, gowns for patients and medical personnel, bedding, gloves, and even body parts). This amazing world of plastics comes from the light ends of the distillation process that are feedstock for steam crackers that produce ethylene plus a whole array of other petrochemicals such as propylene, butadiene, butylene, benzene, toluene, xylene, and raffinate. Ethylene can be changed into other petrochemicals such as polyethylene, ethylene oxide, dichloride, and others to become plastic packaging, trash bags, plastic containers, antifreeze, flooring, paints, adhesives, polyester for textiles, and upholstery for furniture. Propylene goes through its intermediary transformations to end up as polyurethane foams, polyester resins, protective coatings, film, and adhesive for plywood. Butadiene ends up in tires, rubber goods, nylon, and high-impact plastic products. Benzene becomes polystyrene, which is found in insulation and disposable dinnerware, while other forms of benzene become detergents, fiberglass, herbicides, and pesticides. Toluene and xylene can end up in the motor gasoline pool or in paints, coatings, and in polyurethane and polyester products, depending on their respective value in the gasoline pool or as paints and plastics.

The next level down in a fractionating tower produces light naphthas that condense between 90°F–175°F and become part of the gasoline pool. Heavy naphthas condense between 175°F–350°F and are fed into a catalytic reformer to produce a mix of reformate for high-octane gasoline and BTX (benzene, toluene, xylene). The mix of reformate and BTX from a catalytic reformer can be varied according to their respective prices in the gasoline pool or as petrochemicals. The butane and isobutene by-products of naphtha reforming are sold or used elsewhere in the refining process and the hydrogen by-product is consumed in a refinery's hydrotreating and hydrocracking units.

Kerosene condenses between 350°F–450°F and can be sold as kerosene or jet fuel with or without a run through a hydrotreater. A hydrotreater uses hydrogen from the naphtha reformer and a catalyst to purify kerosene and gasoil to improve combustion performance and remove sulfur. Sulfur comes out as hydrogen sulfide and is then reduced to pure sulfur for sale to industrial users and fertilizer manufacturers. Light gasoil condenses between 450°F–650°F and is sold as heating oil and diesel fuel. Heavy gasoil condenses between 650°F–720°F. Catalytic cracking splits the long hydrocarbon chains of heavy gasoil into shorter chains by breaking carbon-carbon bonds with a special silicon dust catalyst. The resulting free carbon sticks to the silicon dust, which inhibits its effectiveness until it is burned away in a regenerator. The output of the cat cracker is primarily naphtha and gasoil; the mix is adjustable to make more gasoline during the summer or more heating oil during the winter. Heavy cycle oil produced by the cat cracker is either recycled or becomes part of the residual fuel pool.

In addition to catalytic cracking, hydrocracking is another method used to break long hydrocarbon chains into shorter chains of more valuable naphtha, jet fuel, and light gasoil. Hydrocracking employs high temperatures (650°F–800°F) and hydrogen from the naphtha reformer under high pressure (1,500–4,000 psi) in the presence of a catalyst to split hydrocarbon chains. Refiners prefer to consume hydrogen by-product from naphtha reformers rather than purchase it or strip it from methane. Refinery operators have a long history of the safe production and distribution of hydrogen within a refinery, which may come in handy someday if society begins the slow shift to a hydrogen economy.

What is left at the bottom of the distillation column is called atmospheric (or atmos) or straight

Table 6.2

Historical Development of Refining Processes

	Process	Purpose	By-Product
1862	Atmospheric distillation	Produce kerosene	Naphtha, tar
1870	Vacuum distillation	Lubricants	Asphalt, resids
1913	Thermal cracking	Gasoline	Resids
1916	Sweetening	Reduce sulfur	Sulfur
1930	Thermal reforming	Improve octane	Resids
1932	Hydrogenation	Remove sulfur	Sulfur
1932	Coking	Gasoline base stocks	Coke
1933	Solvent extraction	Improve lubes	Aromatics
1935	Solvent dewaxing	Improve pour point	Waxes
1935	Catalytic polymerization	Improve gasoline yield and octane	Petrochemical feedstocks
1937	Catalytic cracking	Improve gasoline octane	Petrochemical feedstocks
1939	Visbreaking	Reduce viscosity	Distillates, tar
1940	Isomerization	Alkylate feedstock	Naphtha
1942	Fluid catalytic cracking	Improve gasoline yield and octane	Petrochemical feedstocks
1950	Deasphalting	Increase cracking feedstock	Asphalt
1952	Catalytic reforming	Upgrade low-quality naphtha	Aromatics
1954	Hydrodesulfurization	Remove sulfur	Sulfur
1956	Inhibitor sweetening	Remove mercaptans	Disulfides
1957	Catalytic isomerization	Convert to high octane molecules	Alkylation feedstocks
1960	Hydrocracking	Improve quality and reduce sulfur	Alkylation feedstocks
1974	Catalytic dewaxing	Improve pour point	Waxes
1975	Residual hydrocracking	Increase gasoline yield from resids	Heavy resids

run resid. Simpler designed refineries that cannot further process straight run resid normally sell it to more sophisticated refineries that can. Vacuum distillation heats straight run to nearly 1100°F, then injects a blast of steam in a vacuum to create light and heavy vacuum gasoil. The heavy vacuum gasoil can be fed to a cat cracker to further break down the hydrocarbon chains into lighter end products. What is left is called flasher bottoms, a heavy fuel oil burned as an industrial and utility fuel, as bunkers for marine engines, or made into lubricating oils. Viscosity breakers, or visbreakers, also break up long molecular chains of hydrocarbons to recover more gasoline and gasoil from resids. Cokers crack heavy refinery streams into light products, leaving nearly solid carbon, called petroleum coke, which looks like charcoal briquettes and is burned like coal. Petroleum coke and asphalt are the very bottom of the bottom of the barrel. Considering the nature of asphalt and petroleum coke, one can conclude that refinery operators have learned to squeeze the last light hydrocarbon molecule out of crude oil. All this did not happen overnight. Table 6.2 shows the historical development of refinery processes (note how many are associated with increasing gasoline yield).[9]

Oil Reserves

Oil resources are the totality of oil in the ground. Half of this is irretrievable, even with the most costly recovery methods. Oil that is retrievable is called reserves. Reserves of an oil and gas field are not known with certainty until the last well is dry. Reserves are an estimate of the amount of oil and gas that can be removed from a reservoir under current oil prices employing current extraction technology, not the amount of oil resources actually in the ground. Thus, an improvement in the price of oil that can support more costly recovery methods, or an advance in oil-extraction technology, can change the amount of proven reserves. Oil resources are fixed by what is in the ground whereas reserves are a variable dependent on oil prices and extraction technologies. Proven oil reserves can be considered working inventory, but not an inventory that appears on the balance sheets of oil companies. Proven oil reserves are reported as a footnote in an annual report. The reported book value of a share of oil company stock based on its balance sheet does not include the value of the company's proven reserves. Proven reserves are, however, acceptable as collateral for bank loans.

Proven reserves are reserves that can be calculated with reasonable accuracy based on field production and the results of appraisal or development wells that measure the potential size of an oil field. The calculation of proven reserves is based on the volume of the pay zone, the porosity and permeability of the reservoir, the degree of oil saturation, and the recovery factor. Porosity is obtained from well logs or cores and oil saturation from a resistivity well log. The recovery factor is estimated by the reservoir drive, nature of the oil, and permeability of the reservoir rock. Another method of estimating proven reserves is based on the decline curve, the falloff in production over time. The materials balance method is another mathematical approach that correlates the volume of oil, water, and gas produced with the change in reservoir pressure.

Proven reserves are either developed (within reach of existing wells) or undeveloped (new wells would have to be drilled to access the oil). Probable and possible reserves are calculated in a fashion similar to proven reserves, but their lower classification reflects the greater degree of uncertainty associated with the underlying data. Rule 4.10(a) of Regulation S-X under the U.S. Securities Act of 1933 was promulgated to protect investors from being fleeced by unscrupulous speculators selling east Texas oil properties. The required methodology for calculating proven reserves is based on actual production. In 2004 the U.S. Securities Exchange Commission (SEC) ordered Shell Oil to remove over 4 billion barrels of oil, equivalent to 20 percent of its reserves, from proven reserves because Shell had not followed the prescribed methodology. Shell had categorized certain deep-water reserves as proven based on the results of exploratory wells and 3-D seismic analysis of their reservoir structures. Shell retorted that the SEC was using a dated methodology applicable to onshore reservoirs, not deep-water offshore reservoirs. The SEC response was that its rules are clear: An assessment of proven reserves must be based on actual production from existing wells using an analytical approach that can substantiate at least a 90 percent chance of recoverability. Without following the SEC script for determining reserves, this portion of Shell's reserves could not be considered proven, but could be considered probable if a 50 percent chance of recoverability could be demonstrated or, lacking that, the reserves could be considered possible. Thus, while Shell's total proven, probable, and possible reserves remained unchanged, the fraction considered proven took a significant hit.

More ominous was the *Petroleum Intelligence Weekly* (*PIW*) report in January 2006 that Kuwait's assessment of proven reserves of 99 billion barrels, representing 10 percent of known world reserves, might be overstated by as much as four times. If true, then Kuwait's proven reserves are only 25 billion barrels. *PIW* estimated that proven and unproven reserves may total 48 billion barrels, about half the official estimate. If true, writing off

5–7.5 percent of the world's known petroleum reserves in one blow cannot be lightly dismissed.

Are We on the Downward Slippery Slope?

Colombia, Egypt, Indonesia, the United Kingdom, and the United States have two things in common: They are all suffering from declining oil production and were once oil exporters and are either now importers or will soon become importers. Oil production in Colombia and Egypt is falling because they are not finding sufficient new oil to replace their declining reserves. Faced with rising consumption, these oil-exporting nations are expected to make the transition to oil importers during the next few years. Indonesia, an exporting nation since the dawn of the oil industry, is in the same boat, with its oil fields maturing with little in the way of new finds. Indonesia will be the first OPEC member to become a net oil importer in the near future. This will bring a new perspective to the bargaining table when OPEC sits down to discuss production quotas to maintain a desired range for oil prices. However, Indonesia will remain a net energy exporter as it continues to exploit its vast natural gas and coal resources. The United Kingdom, with its relatively low oil reserves, which are not being replenished by new discoveries, crossed the line to become a net importer in 2004.

The United States, once the world's largest oil exporter, is now the world's largest oil importer. Despite discoveries of oil in the North Slope of Alaska and in the Gulf of Mexico, oil production in the United States has been in a slow decline from exhaustion of Lower 48 oil fields and the decades-long prohibition of exploration in the Arctic National Wildlife Reserve and in offshore waters other than Louisiana and Texas. The prohibition of oil drilling in offshore Florida waters was strongly supported by yacht owners who did not want their ocean vistas ruined by the drilling rigs that provide oil for their fuel-guzzling yachts—an obvious disconnect between desire and reality. Those who own gas-guzzling SUVs for their daily trips to the shopping mall, yet oppose anything the oil industry proposes, are guilty of the same disconnect.

As Figure 6.2 shows, oil production for this group of nations peaked in 1985 at 15.7 million barrels per day (bpd), and in 2005 produced 11.0 million bpd, down 4.7 million bpd.

World oil reserves are 1,188 billion barrels (1.188 trillion barrels), a figure that includes 270.5 billion barrels of OPEC write-ups, shown in Table 6.3.[10] These write-ups are held in suspicion as they were not accompanied by new discoveries. While it is true that existing reserves could have been recalculated to the higher totals, it is also true that, during this time, OPEC was setting production quotas based on proven reserves. A warranted or unwarranted write-up of proven reserves would have resulted in a higher oil production quota and higher revenue.

Adjusting the published proven reserves by the amount of the write-ups in Table 6.3 yields 941.2 billion barrels in proven reserves, a number reflected in Figure 6.3.

Figure 6.4 shows the ratio of the published and adjusted reserves and oil production for each year since 1980. The ratio of reserves to production based on published reserves indicates that current production will exhaust published proven reserves in forty years, and in thirty years if based on adjusted proven reserves. The ratio is in a slight decline by both measures. Care has to be exercised because the ratio of thirty or forty years assumes that production does not increase and no new discoveries are made, both weak points in determining how long it will take for the world to run out of oil.

Table 6.4 lists the world's largest oil fields. The cumulative percentage is based on the adjusted

Figure 6.2 **Nations with Declining Oil Production**

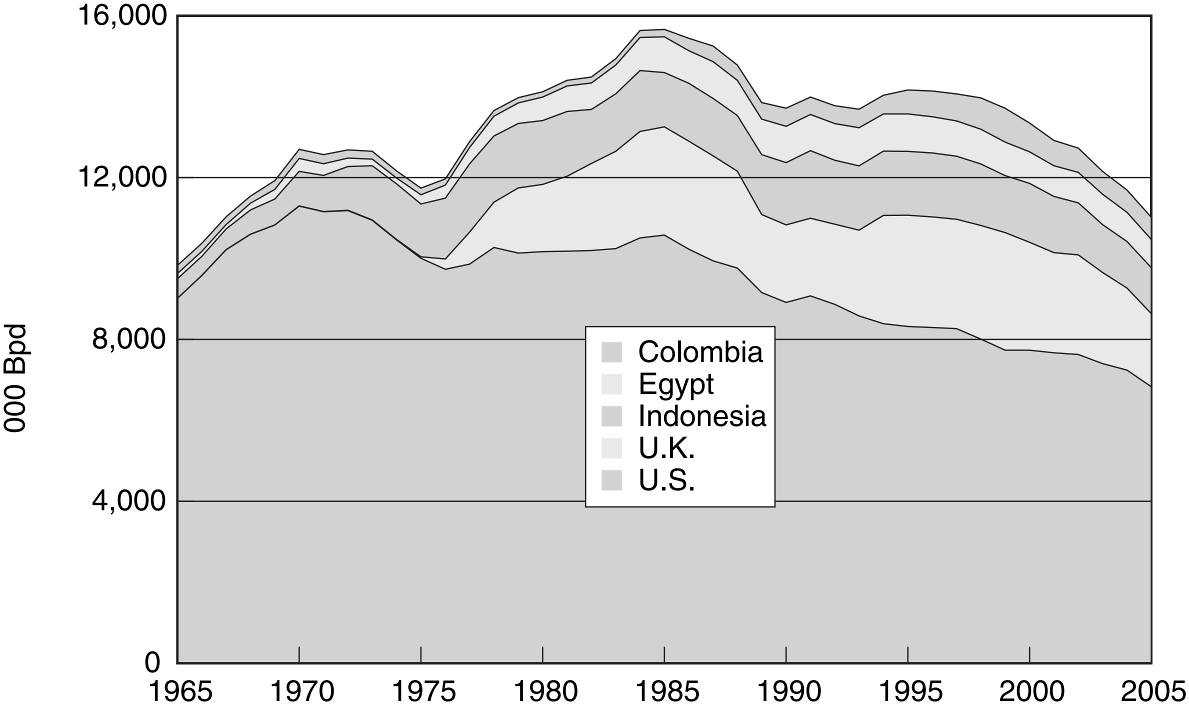


Table 6.3

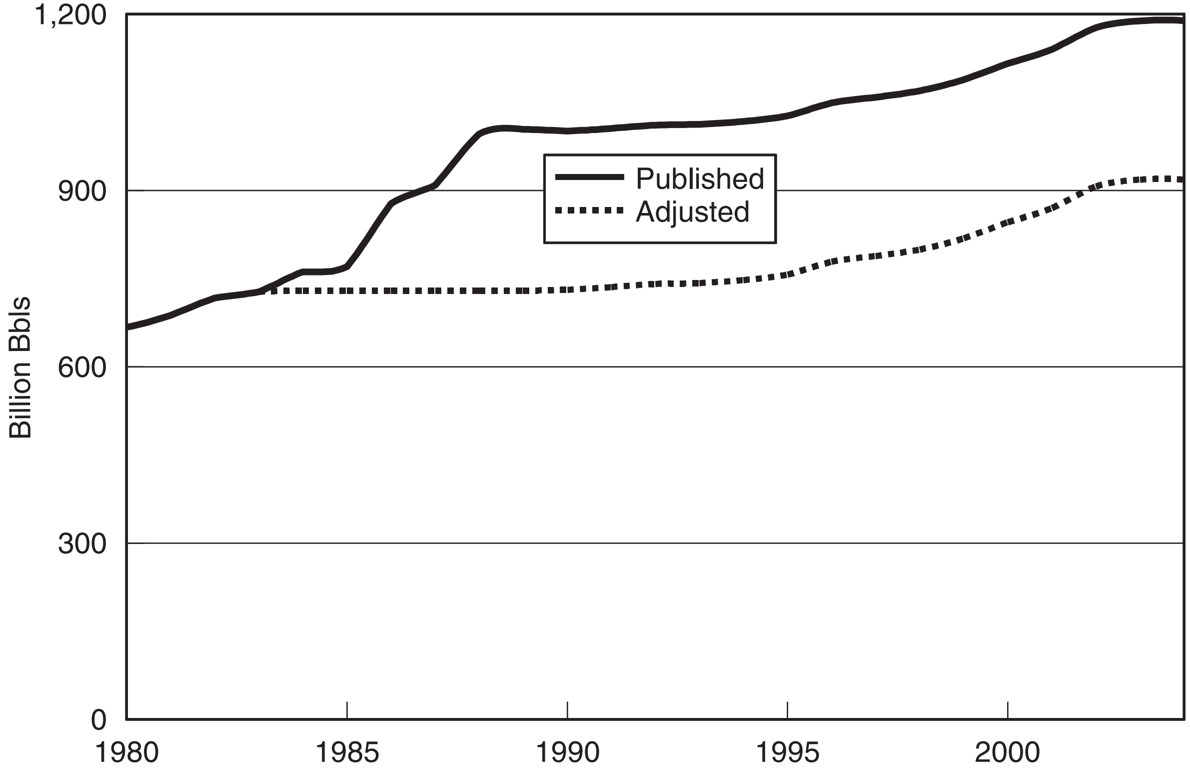
Write-up of OPEC Reserves

Nation	Year	Write-up of Reserves Billion Barrels
Kuwait	1983	25.7
Venezuela	1985	26.5
Iran	1986	33.9
Iraq	1986–87	35.0
United Arab Emirates	1986	64.2
Saudi Arabia	1988	85.4

proven reserves of a little over 900 billion barrels. The Ghawar field represents 7 percent of the world's proven resources. The total of the Ghawar and the Greater Burgan fields represent 10 percent of the world's proven reserves, and so forth.[11]

These eighteen supergiant fields account for one-third of the world's known proven reserves in 40,000 oil fields. Two-thirds of these were discovered in and prior to 1960, forty-five years ago. All but three are in the Middle East. Figure 6.5 shows the distribution of the world's oil producers. As one

Figure 6.3 **World's Proven Reserves**



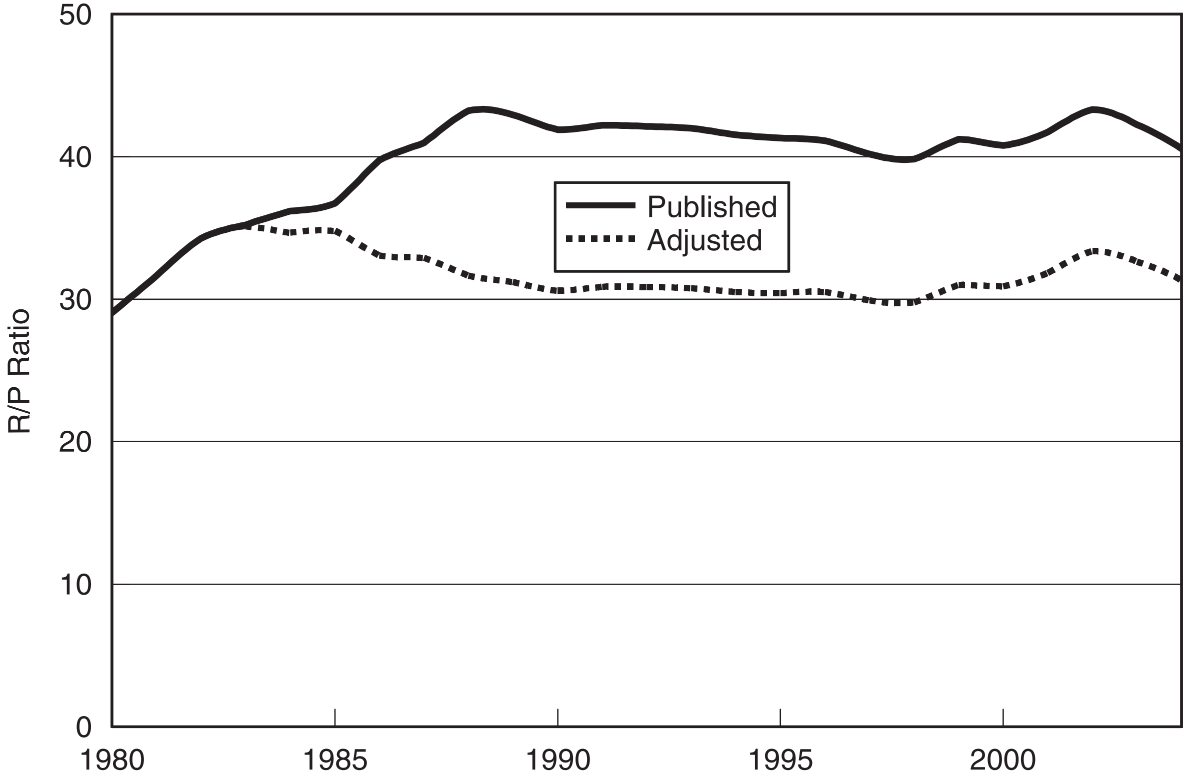
can see, OPEC dominates with a 41-percent share of the world's oil production.

Figure 6.6 shows the major world oil consumers, with North America and Asia responsible for nearly 60 percent of consumption. Clearly Asia and North America are both dependent on imports, but Asia is much more dependent on Middle East imports than North America. The location of major world oil reserves illustrates with great clarity the geopolitics of oil—the world is utterly incapable of extricating itself from reliance on Middle East oil. If reserves are adjusted, on the basis of unadjusted reserves, the Middle East share climbs to 62 percent.

In 1956 M. King Hubbert, a geophysicist with a background in exploration for Shell Oil, postulated that U.S. oil production would peak in the early 1970s based on an assessment of discoverable oil (known oil reserves plus that yet-to-be discovered). Scorned by his contemporaries, he turned out to be basically right. Hubbert was off a bit on the actual timing of the peak in production because, since he made his original prediction, more oil was discovered in Alaska and in the Gulf of Mexico than he anticipated. But he was not off by much.

Modern-day followers of Hubbert assess the quantity of ultimately discoverable oil and compare that to cumulative production on a global scale.

Figure 6.4 **World Ratio of Reserves to Production**



Oil production peaks when cumulative production has consumed half of the ultimately discoverable reserves. Ultimately discoverable reserves consist of known reserves, including enhanced production from played-out fields through tertiary recovery methods, and an assessment of what has not yet been discovered. When on the downhill slope of a bell-shaped curve, exploration and extraction become more expensive as fewer and smaller oil fields are discovered in more remote areas and more costly methods have to be employed to maintain production in aging oil fields. Furthermore, oil becomes more viscous as a field ages, a fact that increases refining costs.

When applying Hubbert's thinking on a global scale, one still has to deal with the challenge of assessing ultimately discoverable oil. As with discoveries in Alaska and the Gulf of Mexico, any discovery that increases the assessment on ultimately discoverable oil postpones the peaking of production. Right now the favorite assessment for ultimately discoverable oil is between 2–3 trillion barrels. The lower estimate comes from followers of Hubbert,[12] while the higher estimate is from the U.S. Geological Survey.[13]

If we take the lower estimate of ultimately discoverable and recoverable oil of 2 trillion barrels and assume that 1 trillion barrels have already

Table 6.4

World's Largest Oil Fields

Ultimate Recovery Oil Millions Bbls	Country	Field Name	Discovery Year	Cumulative Percentage
66,058	Saudi Arabia	Ghawar	1948	7
31,795	Kuwait	Greater Burgan	1938	10
22,000	Iraq	Rumaila North & South	1953	13
21,145	Saudi Arabia	Safaniya	1951	15
17,223	Abu Dhabi	Zakum	1964	17
17,000	Iraq	Kirkuk	1927	19
16,820	Saudi Arabia	Manifa	1957	20
13,390	Venezuela	Tia Juana	1928	22
13,350	Iran	Ahwaz	1958	23
13,010	USA-Alaska	Prudhoe Bay	1967	25
13,000	Kazakhstan	Kashagan	2000	26
12,631	Iran	Marun	1964	27
12,237	Saudi Arabia	Zuluf	1965	29
12,000	Iraq	Majnoon	1977	30
11,800	Iran	Gachsaran	1928	31
10,276	Abu Dhabi	Murban Bab	1954	32
10,265	Saudi Arabia	Abqaiq	1940	33
10,000	Iran	Fereidoon	1960	34

been consumed, we are either peaking now or shortly will be. If the higher estimate of 3 trillion barrels is valid, then we have some breathing room. The higher estimate places us 0.5 trillion barrels away from peaking if we have already consumed 1 trillion barrels. At consumption levels of 80 million barrels per day, we will consume the 0.5 trillion barrels separating us from peak production in seventeen-and-one-half years. If consumption grows by 1 million barrels per day, which is less than historical growth, peaking occurs in sixteen years. At an annual growth of 1.5 percent per year, peaking occurs in fifteen years. Whether peaking occurs at fifteen or twenty years is not critical; what is critical is that, even if peaking occurs in twenty years, we will end this century with no oil. Many of us will be dead by then, but what of our progeny?

The times of trouble do not begin when the oil is gone, but after production has peaked. This is not a prescription for cheap oil, but expensive oil. Oil extraction costs are already rising as we explore in more inhospitable and remote locations for oil. The historical survey of water-depth capacity and level of sophistication of offshore-drilling rigs attests vividly to the increasing challenge of finding new oil fields. Higher-priced oil slows or depresses economic activity in industrialized nations and sends developing nations with little indigenous supplies of oil and a perennial negative trade balance into an economic tailspin. Every unsuccessful exploratory well decreases the overall chance of finding another megagiant oil field. Huge finds are necessary to increase reserves in a world of rising consumption. Decreasing oil reserves after peak production will create greater stress among nations in an ever more evanescent search for security of supply. We may be at the beginning of the times of trouble with our military involvement in the Middle East, which started in 1990 with protecting Kuwait and escalated sharply with the invasion of Iraq in 2003.

Figure 6.5 **Major World Oil Producers**

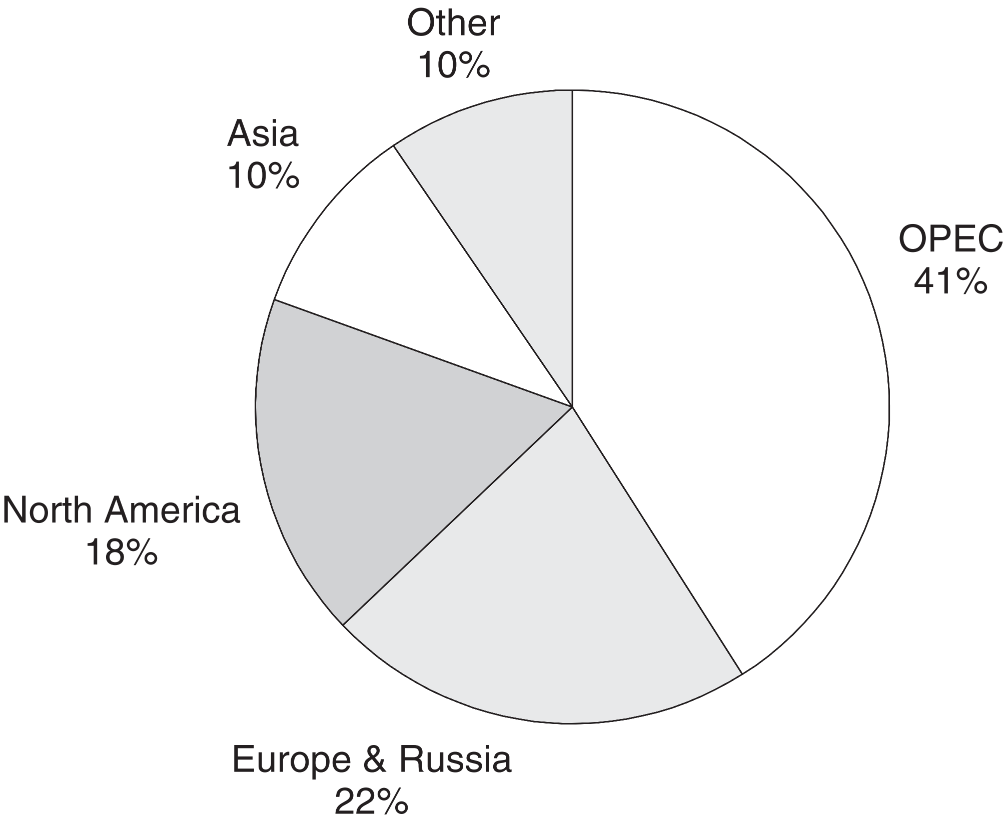
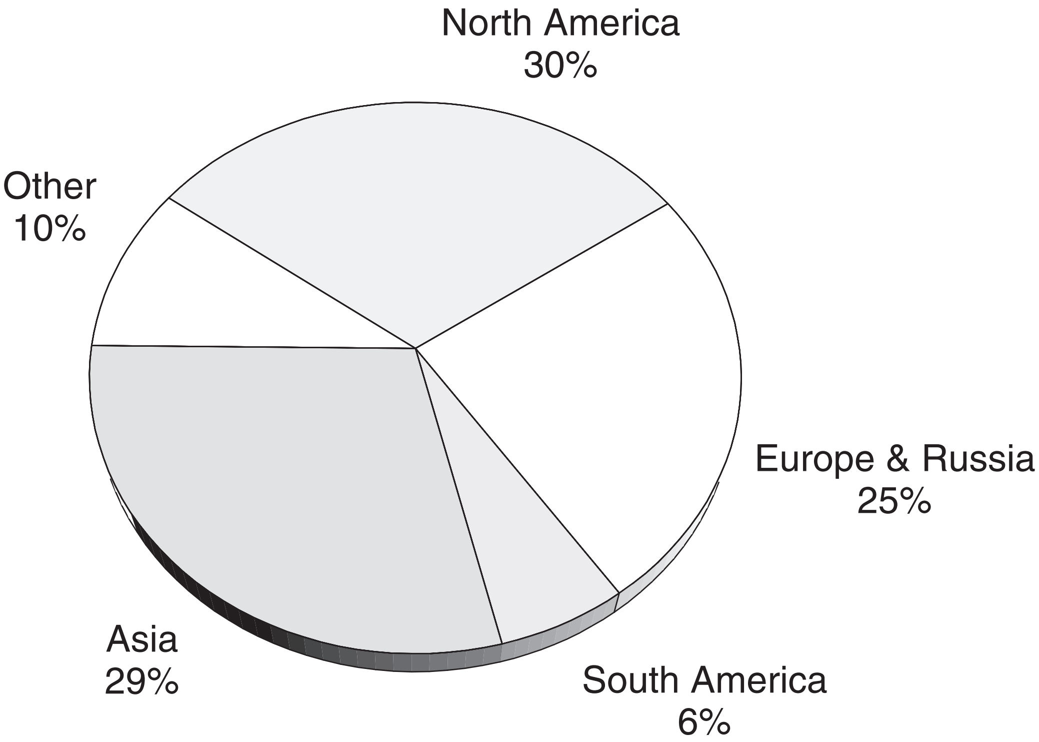


Figure 6.8 shows the historical record for discovering giant oil fields of greater than 500 million barrels. Clearly the peak of discovery has passed. While Figure 6.8 is the number of giant oil fields that have been discovered, Figure 6.9 is the amount of proven oil reserves in these fields. The rate of addition to proven oil reserves is compared to consumption. Prior to 1968, the problem faced by oil executives was how to control production to maintain price in the face of mounting discoveries. Since 1968, with the exception of two or three years, discoveries have not kept up with consumption by a significant margin. Current estimates are that discoveries compensate for only half of consumption; a surefire prescription for running out of oil.

In 1980, remaining proven oil reserves were about 670 billion barrels, compared to the current estimate of 1.1 trillion barrels (unadjusted). One may wonder how reserves can be getting larger if the rate of discovery of new fields lags behind consumption. Part of the answer is that Figures 6.8 and 6.9 only measure large finds of over 500 million barrels; smaller fields are not being counted. Part of the answer also lies in the fact that proven reserves of a new field, once established, may not

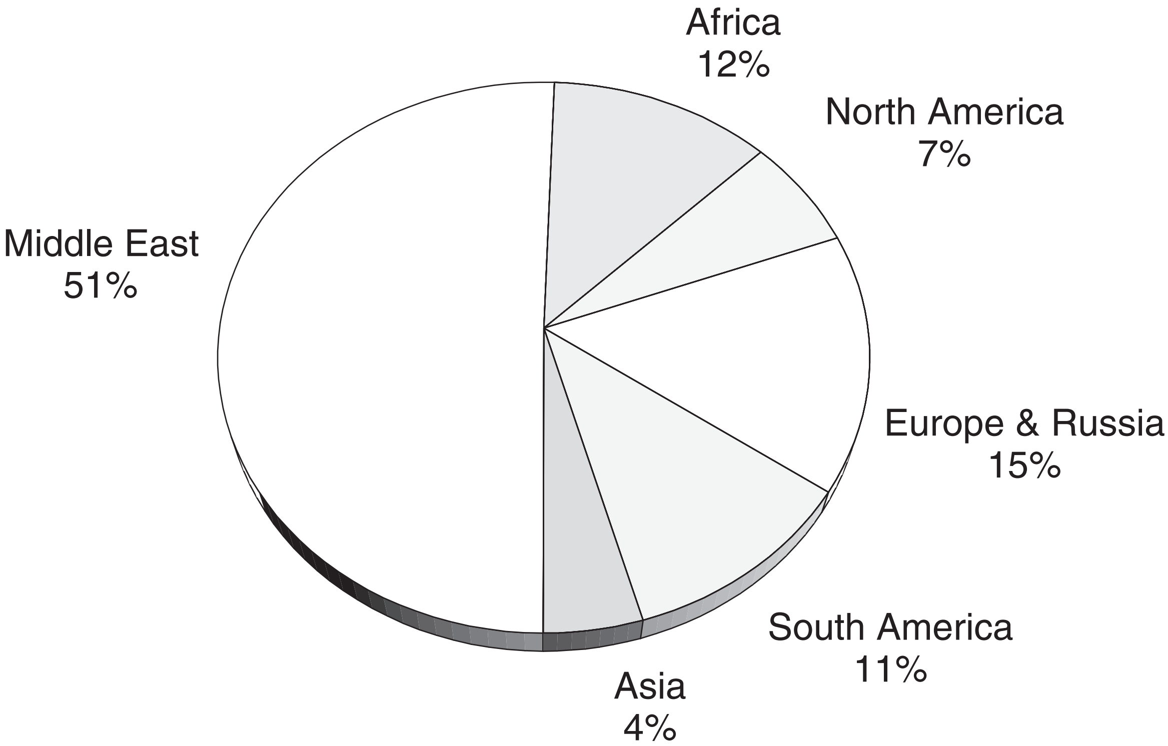
Figure 6.6 **Major World Oil Consumers**



always be written down as it is being depleted. Proven reserves of major exporting nations with no discoveries of note remain the same year after year despite significant production. This cannot be. Reserves should take into consideration both new discoveries and the depletion of existing reserves.

Since the beginning of the oil age, predictions of the world running out of oil have been made and all have been proven wrong. In 1879 the U.S. Geological Survey was formed, in part because an oil shortage was feared. In 1882 the Institute of Mining Engineers estimated that there were 95 million barrels left, an amount that would be exhausted in four years at the then-present consumption rate of 25 million barrels per year. In the early 1900s Theodore Roosevelt opined that there were about twenty years of reserves left and hearings were held in Washington on the adequacy of supply. In 1919 the *Scientific American* warned that there were only twenty years of oil left in the ground and made a plea for automobile engines to be designed for greater energy efficiency (déjà vu?). In 1920 the U.S. Geological Survey estimated that U.S. reserves were only 6.7 billion barrels, including what was known and remaining to be discovered (current reserves are 31 billion barrels after eighty-odd years of production).[14] In the 1920s the U.S. government, worried over the adequacy of oil supplies, secured an interest in the Turkish Petroleum Company and had to almost

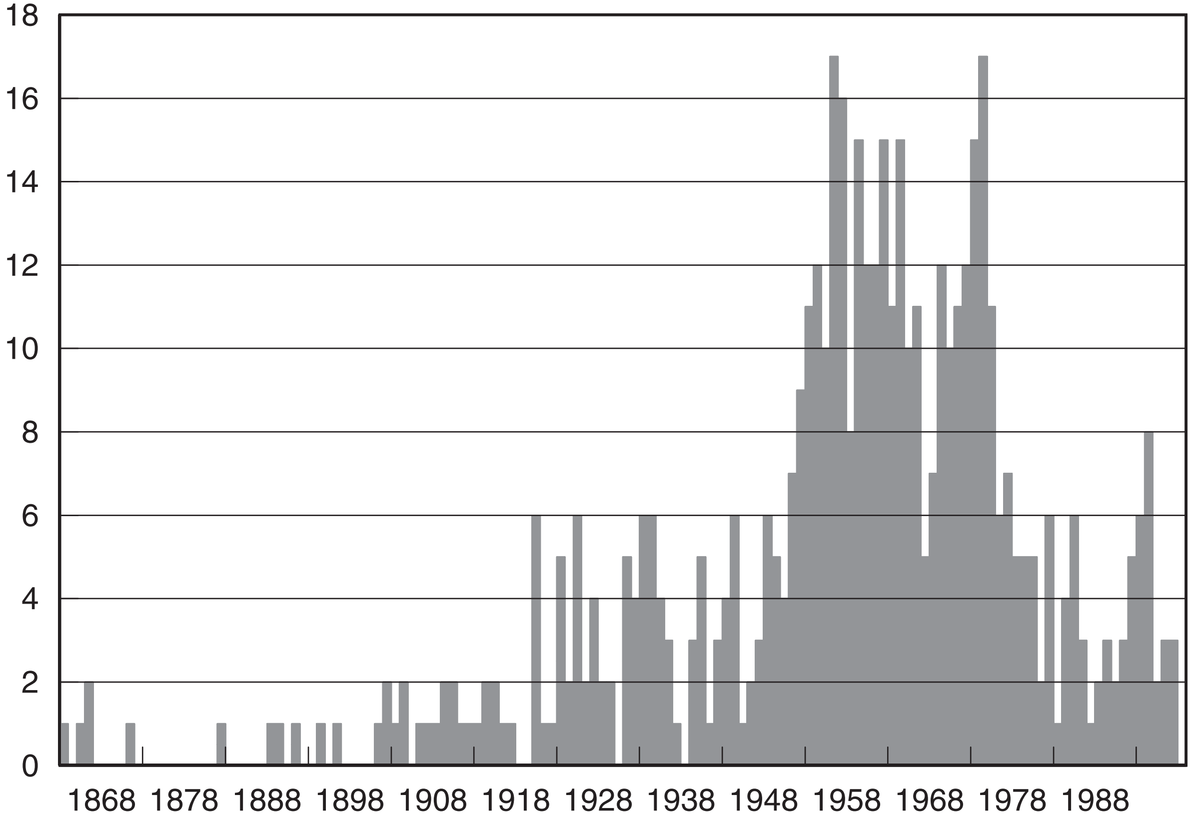
Figure 6.7 **Major World Oil Reserves** (Adjusted)



coerce reluctant U.S. oil companies to get involved with Middle East oil.

All these forecasts have been proven wrong, but that does not mean that the current spate of dire forecasts is necessarily wrong. The big difference between past and present forecasts is a lack of large oil discoveries. A thirty-five- or forty-year dry spell should not be ignored. If oil cannot be replenished as fast as it is being consumed, then it is a wasting asset. It is not a question of whether, but when, we run out of oil. Yet, all this hand-wringing is based on proven reserves. There is something unsettling about considering only proven reserves, which is a variable based on current oil prices and current extraction technology. Increase price or improve extraction technology, and presumably proven reserves will increase from the reclassification of probable reserves to proven and, perhaps, possible reserves upgraded to probable. Oil reserve statistics are also subject to manipulation for political or commercial reasons. Perhaps some OPEC nations exaggerated their reserves to get a higher production quota. Perhaps others do not want the world to know the true amount of their reserves in order to sustain oil prices. A geographic area may be a very strong candidate for harboring enormous oil reserves, but exploration might be postponed on the theory that what may be in the ground will be worth more if discovered tomorrow than if discovered today.

Figure 6.8 **Frequency of Discovery of Giant Oil Fields**



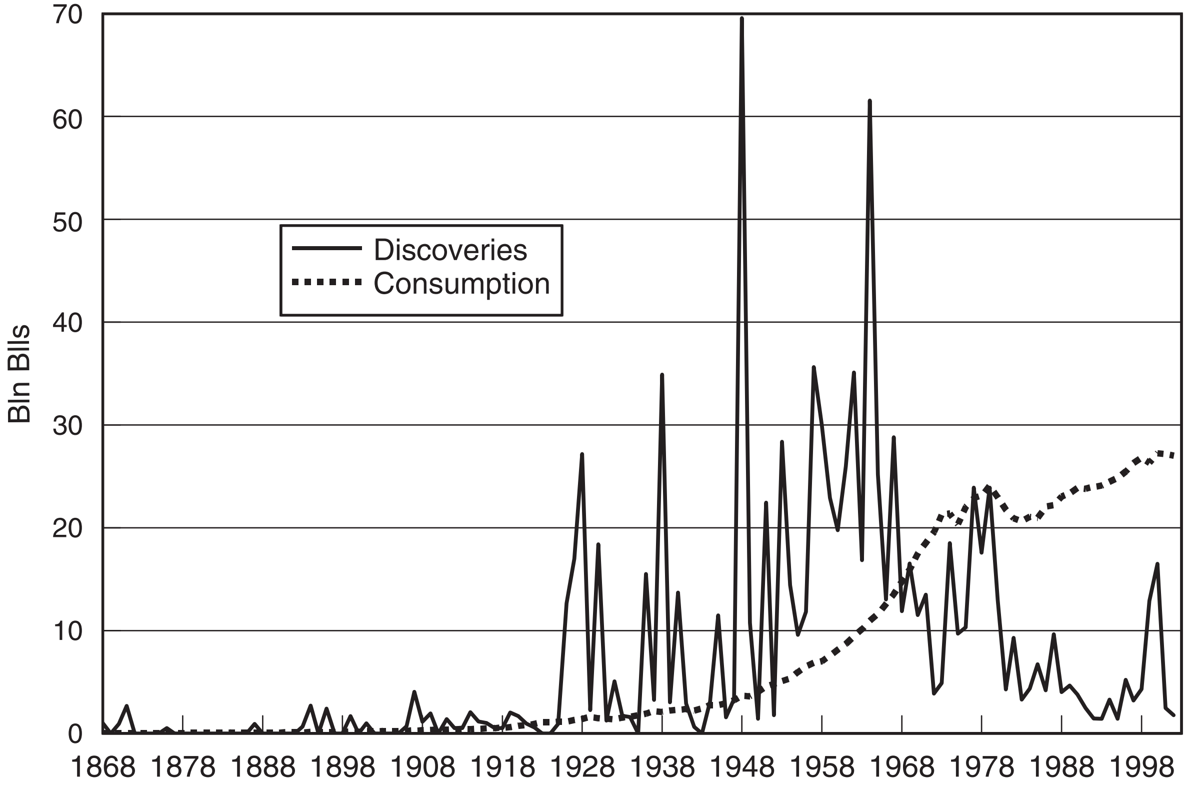
Probable and possible reserves and unconventional sources of oil could make a difference. As things stand, unless the world experiences the thrill of a discovery of supergiant fields with some degree of regularity, it is clear from Figures 6.8 and 6.9 that:

- the frequency of discovering major oil fields is dropping;
- the size of newly discovered oil fields is falling; and
- consumption is getting ahead of additions to proven reserves.

Was 2004 a Watershed Year?

Two prominent government projections indicating a doubling to a near tripling of Middle East oil exports by 2025 and 2030 to keep up with world growth in demand may not be tenable.[15] The past presumption that Saudi Arabia has the spare capacity to export whatever is required may no longer be valid. The reservoir pressure in Ghawar, the world's largest oil field and responsible for 60 percent of Saudi output, is maintained by pumping in seawater. Over time seawater mixes with

Figure 6.9 **Comparison of Major Discoveries versus Consumption**



crude oil, with an increasing concentration of seawater in oil coming from producing wells. The end of Ghawar will be marked when only saltwater is coming out of its oil wells. While Saudi Arabia vociferously denies the rumor that Ghawar is showing signs of aging by an increasing presence of seawater, other supergiant oil fields in Russia, Mexico, Venezuela, the United States, and Indonesia do show indisputable signs of aging.

While once Saudi Arabia was thought capable of ramping up its production when another OPEC member such as Nigeria and Venezuela stopped exporting oil, this is no longer considered possible other than for short disruptions. Events in 2004 indicated that Saudi Arabia was not capable of pumping sufficient quantities of light to medium sweet grades of crude to satisfy demand by the world's refinery operators, and that is why the price of sweet crude rose to the mid-$50-per-barrel range. Although Saudi Arabia was able to pump heavy sour grades in large quantities, the capacity of the world's refineries was insufficient to process this crude, causing a record price spread between these two grades.

The world was up against two constraints in 2004: the capacity of the oil producers to meet demand for light sweet crude and the capacity of the oil refiners to process heavy sour crude. The decline in OPEC's theoretical maximum production capacity from 39 million bpd in 1979 to 32 million bpd in 2004 has to be treated as another warning that the Middle East cannot be considered an infinite source of oil.[16] The matter is worse when viewed in terms of spare capacity. In both 1979 and 2004, OPEC production was 31 million bpd. This implies a spare capacity of 8 million bpd in 1979 and only 1 million bpd in 2004. There is little margin for OPEC as a group to satisfy world demand if a single OPEC nation ceases to export for other than a short period of time.

Even Russia, with its large reserves and America's hope for the future, may not be entirely dependable. The government attack on Yukos, Russia's leading oil producer that accounts for 1.5 percent of world oil production, in 2004 for tax evasion amounting to $15 billion, led to its corporate dismemberment. This is another example of Russia's penchant for unilaterally redefining tax liabilities for its own political agenda. Changing the rules of the game does not make Russia a safe place to invest. In early 2006, Russia interrupted natural gas shipments to the Ukraine over a price dispute that also interrupted natural gas flows to Europe. This posed a serious question regarding the dependability of Russian energy exports.

The United States was once the world's swing producer when oil production was controlled by the Texas Railroad Commission. The Texas Railroad Commission curtailed production to maintain price in order not to waste a natural resource. Of course, maintaining price was also in the interests of Big Oil, but not necessarily for conservation purposes. In 1971, the Texas Railroad Commission authorized 100 percent production for all wells under its jurisdiction, thus ending the days of the United States being a swing producer. Since the oil crisis of 1973, the mantle of swing producer has been worn by Saudi Arabia. Saudi Arabia, by all accounts, has been a fairly responsible swing producer, seeking a price that was not too low to support its social programs in medical care, housing, and education for its rapidly growing population as well as providing funds to build a value-added infrastructure of refineries and petrochemical plants. Saudi Arabia also realizes that too high a price dampens world economic activity, subsidizes the development of high-cost oil fields elsewhere, and promotes alternative sources of energy. This is the lesson Saudi Arabia learned to its disadvantage during the late 1970s and early 1980s.

OPEC maintains what it deems an acceptable range of oil prices by raising production quotas when oil prices are too high and lowering them when prices are too low. Since most oil producers within OPEC operate at or near their maximum sustainable rates, the nation with the greatest capacity to increase production and the strongest will to reduce production is Saudi Arabia. In 2004, Saudi Arabia made several announcements of its intention to increase production to dampen oil prices. Increased Saudi production did not cool oil prices as expected. Contrary to the naysayers who maintained that Saudi Arabia was bluffing, tanker rates soared, proof of higher Saudi export volumes. What this showed was the lack of sufficient spare capacity to keep oil prices from getting out of control.

The primary source for incremental oil demand that taxes OPEC's capacity to produce oil in 2004 was China, which rose in rank to become the world's second largest consumer. Its 9 percent annual economic growth was not only taxing world oil production capacity, but also a number of commodities such as copper, tin, zinc, platinum, steel and iron ore, aluminum, lead, nickel, and so forth, driving up prices for all these commodities to record levels.[17] Capital goods exports from Europe and Japan for machinery and electricity-generating equipment surged in response to China's rapid industrialization. Unless there is a significant cooling of economic activity in China, we may have already entered a world of real constraints in which we cannot produce sufficient quantities of desirable grades of crudes or refine sufficient quantities of undesirable grades of crude, with the distinct possibility that other commodities may also come into short supply. From the point of view of oil, the world is on the razor's edge with regard to runaway oil prices, dependent on mild weather to reduce heating oil demand, no delays in the scheduled additions of new non-OPEC and OPEC oil supplies, and no long-lasting disruptions among oil exporters with restive populations. This is not the most comfortable position to be in.

The precarious state of our position in the world of oil can be seen in the Baku-Tbilisi-Ceyhan (BTC) Oil Pipeline, built by British Petroleum, which came onstream in 2005. Its fully rated capacity, to be attained near the end of 2006, is 1 million bpd of Caspian crude that will be exported from the Turkish port of Ceyhan. That may seem like a lot of oil by any measure. But is it? This crude is needed as replacement of declining North Sea oil production, which in 2005 was already down 1.2 million bpd since 2000, and is projected to continue falling. Thus, this new source of crude does nothing but partially fill the gap of declining North Sea oil and in no way can satisfy the incremental growth in global oil consumption of over 1 million bpd per year.

China is well aware of its inability to expand its domestic oil production in significant volumes, yet it is unwilling to reduce its growth in oil demand by sacrificing economic development. This leaves the nation vulnerable because it relies on the Middle East as a major and growing source of oil. In order to reduce its reliance on Middle East oil, China has been actively pursuing diversification of oil supplies by encouraging Russia to build an oil pipeline to ship Siberian oil to China, by investing or buying oil properties in Indonesia, Europe, and Canada, and by taking an active role in the development of oil projects in other nations, such as a major oil export project in Sudan. With China shopping for oil in the Atlantic basin, it is sure to come in conflict with the United States over an increasingly scarce, and vital, commodity.

Events in 2004 raised the question of whether refinery capacity in the United States was adequate. The well-publicized fact that no grass-roots refinery has been built in the United States since 1976 ignores another fact: that there was excess U.S. refinery capacity between the mid-1970s and mid-1990s, which provided little in the way of an economic incentive to build new ones. However, there was an ongoing program of upgrading and debottlenecking existing facilities that increased refinery throughput capacity, called refinery creep. Luckily for the United States, its refiners invested in facilities to handle heavy grades of crude oils that remained in plentiful supply during 2004. Moreover, the switch by Europeans from gasoline- to diesel-driven automobiles freed up gasoline refining capacity to help meet growing U.S. gasoline import needs. Once this spare capacity is consumed, presumably more refineries will have to be built; if not in the United States then somewhere in the Atlantic basin, and if not there, then in the Middle East, which will only increase our reliance on Middle East oil.

One might think that high oil prices would be a strong incentive for building refinery capacity, but that is not how the system works. It is the spread between the price of oil products and the cost of crude oil that determines refinery profitability.

A high price of crude oil does not automatically translate to high refinery profits unless the spread widens. Even a widening price spread between crude and refined oil products may not be sufficient to induce refinery construction. A refinery operator who is making a great deal of money on a refinery with a cost base of $200 million may not have the financial wherewithal to construct an equivalent-sized refinery that would cost $2 billion unless the spread widens further.

Of course, none of this need happen. This logjam of constraints can be broken by building more refinery capacity to handle lower grades of crude oil and by the discovery of a supergiant oil field or two. One potential area is the South China Sea, where territorial claims by six littoral nations have inhibited exploration. Another potential area for discovery of a supergiant oil field is, surprisingly, Iraq, which has large areas of unexplored terrain.

Synthetic Crude

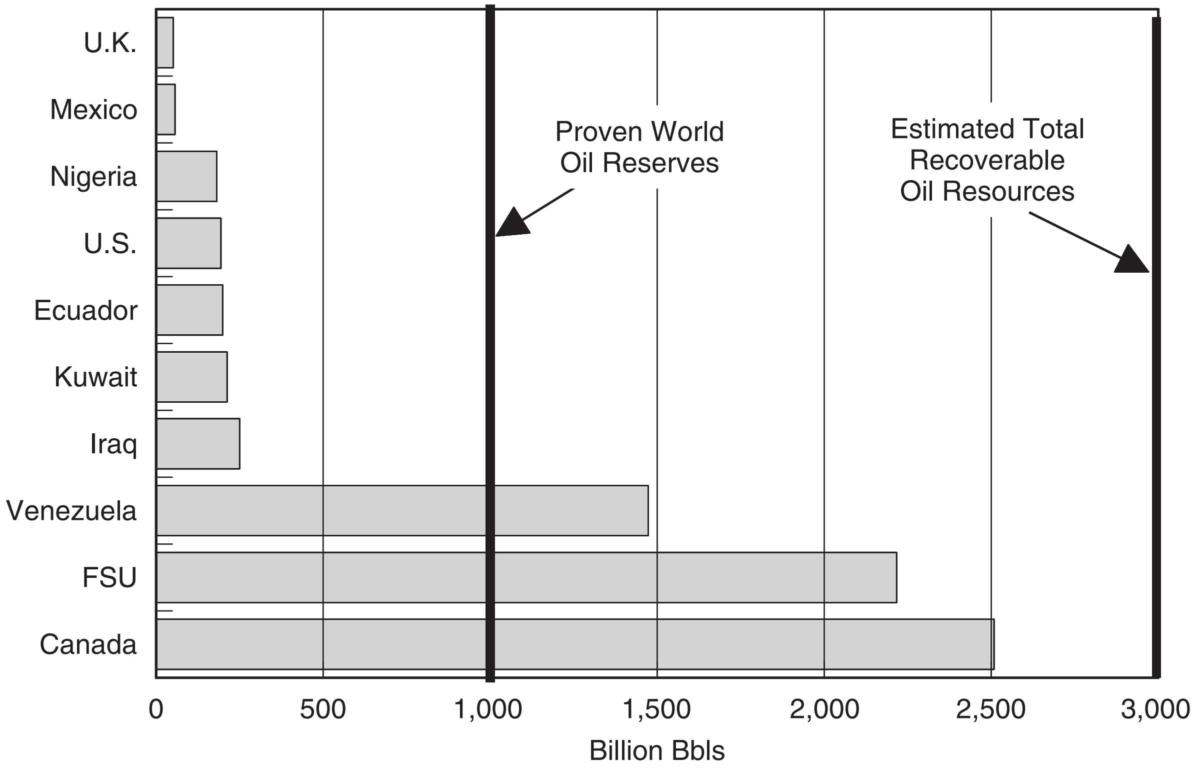
Synthetic crude, or syncrude, is a nonconventional source of oil that must be processed to produce an acceptable grade of crude oil before it can be fed into conventional refineries. Major nonconventional sources of synthetic crude are bitumen deposits in Canada, Venezuela, and Russia, plus oil shale, which is found in various parts of the world. Bitumen is a thick, sticky form of crude oil, sometimes called extra-heavy oil, with the consistency of molasses; it is too viscous to flow in a pipeline unless it has been mixed with a light petroleum liquid. Bitumen also has a high sulfur and metals content that complicates upgrading to a syncrude fit for a conventional refinery. These undesirable traits are offset by huge deposits of bitumen, as shown in Figure 6.10. With the exception of a portion of heavy crude in Venezuela's proven oil reserves, these deposits are not included in official oil reserve statistics.

In Canada, huge volumes of oil migrated horizontally and vertically through more than seventy miles of rock without entrapment by anticlines or faults. As the seep oil emerged on the surface, it mixed with sand and some clay and became a feast for microorganisms that transformed the oil into bitumen. It is estimated that the bacteria consumed between two and three times the present volume of bitumen, an incredible amount of oil when one considers that Canadian tar sands hold about 2.5 trillion barrels of bitumen. This implies that the original source rock generated 5 to 7.5 trillion barrels, compared to 2 or 3 trillion barrels of ultimately discoverable global oil reserves, including all that has been consumed since Drake. That is a whole lot of ocean plankton, algae, and other forms of simple marine life to die and settle in oxygen-starved sediment in a single province of Canada.

As a recoverable resource using present technology, Canadian tar sand is equivalent to 300 billion barrels of crude oil, slightly more than the 263 billion barrels of proven oil reserves in Saudi Arabia and about 40 percent of the Middle East's proven oil reserves. The Athabasca deposit in Alberta is the world's largest tar sand deposit, containing about two-thirds of Canadian bitumen resources. Tar sand with about 12 percent bitumen by weight is mined similar to the way coal is surface-mined. The overburden is removed and stockpiled for reclamation after mining operations cease. Natural gas is burned at the mining site to thaw the frozen ground and free up tar sand. Giant mining shovels working 24–7 fill huge trucks with 360–380 tons of tar sand for transport to an upgrading plant.[18]

The first step is an extraction plant where the tar sand is crushed and mixed with hot water. It is then sent to a large separation vessel where sand

Figure 6.10 **World Heavy Oil and Bitumen Resources**



falls to the bottom and bitumen, trapped in tiny air bubbles, rises to the top of the water as froth. The froth is skimmed off, mixed with a solvent, and spun in a centrifuge to remove the remaining water and sand. Water and sand residue, called tailings, are placed in a settling pond where any remaining bitumen is skimmed off the surface. Sand is mixed with water and returned to the mine site by pipeline to fill in mined-out areas, and the water in the settling pond is recycled for the next batch of tar sand. This method minimizes undesirable environmental consequences and recovers over 90 percent of the bitumen in the sand. For deeper deposits of tar sands, wells are drilled and high-pressure steam is injected into the well to soften up the surrounding bitumen. Bitumen flows into the well and is then pumped to the surface. When production slows, a cycle of "huff and puff" softens up another batch of bitumen. This method recovers a relatively small portion of the bitumen.

After the bitumen is extracted, it is ready for upgrading, which converts it into a synthetic, or processed, crude with a density and viscosity similar to conventional crude oils. Upgrading involves removing carbon and sulfur and adding hydrogen. Coking removes carbon atoms from the large, carbon-rich hydrocarbon chains, breaking them up into shorter chains. Hydrotreating removes sulfur, which is sold to fertilizer manufacturers. Hydrocracking adds hydrogen to hydrocarbon chains to increase the yield of light-end products when the syncrude is refined in a conventional refinery. The processed syncrude is mixed with condensate, a very light oil associated with natural gas production, for pipelining to a refinery.

The process for making syncrude requires a lot of natural gas (also needed as a source of hydrogen for hydrocracking and hydrotreating) to heat the tar sand for loading in trucks and the water to extract the bitumen from the sand. A great deal of water is used for separating the bitumen from the sand and in pumping the spent sand back to the mine site. Thus, the enormous reserves of bitumen are ultimately dependent on the availability of water, a naturally replenished resource, and natural gas, a wasting resource. Without a pipeline to ship the local supplies of natural gas to market, syncrude production creates value for stranded gas. With a pipeline that can ship natural gas to markets in Canada and the United States, the issue would become whether to consume natural gas locally for syncrude production, where its value is determined by the price of crude oil, or to sell it as commercial pipeline natural gas to the Lower 48. In terms of the environment, natural gas consumed in mining bitumen and producing syncrude, plus the energy consumed in transporting bitumen from the mine to the syncrude plant, add to carbon emissions when syncrude is substituted for an equivalent amount of crude oil. Syncrude production in Canada was 800,000 bpd in 2003 and is expected to increase to 2 million bpd by 2012.

Bitumen deposits in Venezuela are located in the Orinoco region and essentially lie on the surface. Unlike Canada, the bitumen is not intermingled with the soil. Bitumen mixed with 30 percent water and other chemicals is sold as a coal substitute called Orimulsion. Bitumen for syncrude manufacture is mixed with naphtha to reduce its viscosity for pipelining to a syncrude plant. There the naphtha is recovered and pipelined back to the bitumen deposit for recycling. The costs of getting bitumen to a syncrude plant and preparing it for processing are far less in Venezuela than in Canada. Consequently, syncrude facilities are more costly in Canada. As a point of comparison, increased production of 145,000 bpd for one project in Canada cost $4 billion and increased production by 155,000 bpd for another project cost $3.5 billion. Incremental production in the amount of 300,000 bpd cost $7.5 billion, for a unit cost of $25,000 per bpd of capacity.

Venezuela entered into four grass-roots joint ventures with foreign oil companies. All are in operation and are expected to produce 600,000 bpd of syncrude in 2005. The total capital costs of about $12 billion work out to $20,000 per bpd of capacity. While probably not the best way to do a cost comparison between Venezuela and Canada, the lower unit cost in Venezuela reflects no investment in mining and extracting facilities, only in a pipeline. With a five-year hiatus between plant approval and operation and with no current plans to expand capacity, the earliest year Venezuela could increase its syncrude production would be after 2010.[19] Syncrude projects in Venezuela and Canada require oil prices of about $30 per barrel to serve the large capital requirements of syncrude plants and pay the currently high market value for the large quantities of natural gas consumed in making syncrude.

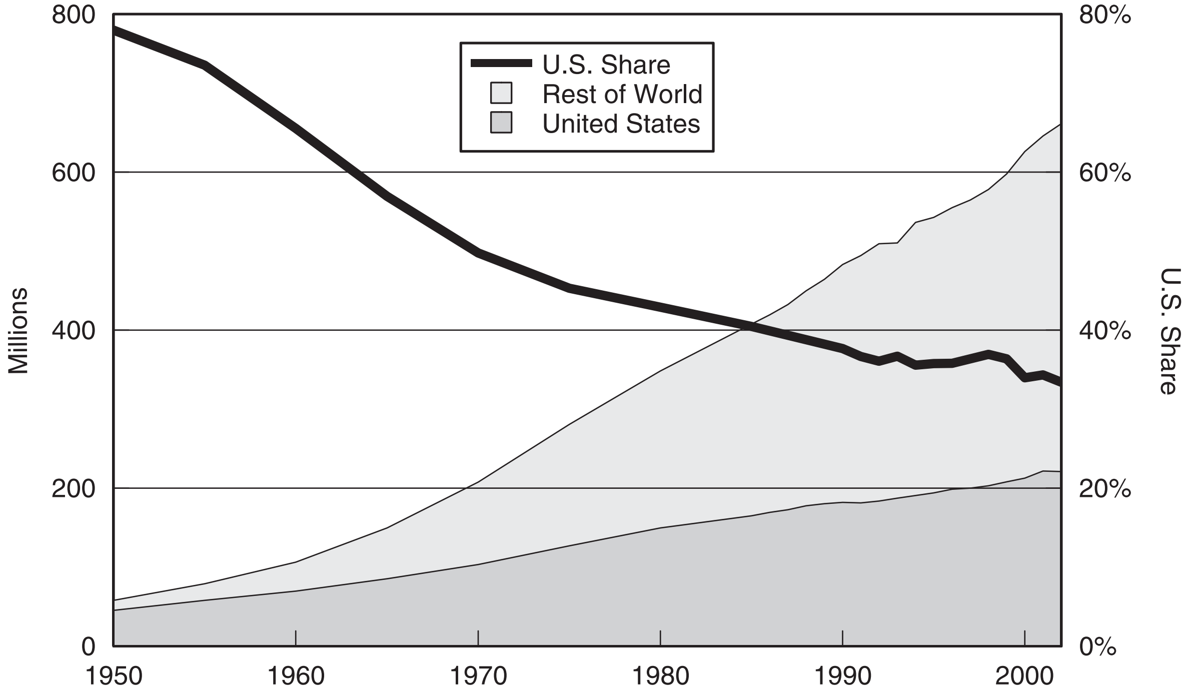
Syncrude may not be a significant substitute for conventional crude oil, considering its capital and natural gas requirements, but it could be effective in reducing U.S. reliance on Middle East crude oil imports. Middle East oil imports to North America have been relatively flat at 2.7 million bpd since 1999 (some Middle East crude is exported to east coast Canadian refineries whose output is both consumed domestically and exported to the United States). With Canadian synthetic crude production projected to grow by 1.2 million bpd in 2012, then another half dozen or so plants of 200,000 bpd capacity built in Venezuela would be sufficient to replace current levels of Middle East imports. Although one may argue over the geopolitical risk associated with substituting syncrude from Venezuela for crude from the Middle East, the point is that syncrude can reduce U.S. dependence on Middle East imports. Venezuela's attitude toward the United States could change if Venezuela viewed a large step-up in syncrude output as a means of creating and sustaining a "Greater Venezuela." However, syncrude is not a practical solution if we are passing the point of peak production of conventional oil.

Syncrude can also be made from oil shale. Like so many things, "oil shale" is a misnomer. The oil in the rock is not oil, but an organic material called kerogen that has not been heated to the requisite temperature to become oil. Hence, the process of making oil from oil shale involves the application of heat to complete the process. Nor is it necessary for the rock to be shale, it can be any kind of rock that contains kerogen, although normally it is a relatively hard rock called *mari*. About 72 percent of the world's oil shale resources are in the United States. Of this, about 70 percent lies in a 16,000-square-mile area, mostly in Colorado, with extensions into eastern Utah and southern Wyoming, called the Green River formation. The Green River formation is estimated to have as much as 2 trillion equivalent barrels of oil. Other nations with large oil shale resources are China, Brazil, and Morocco, each with 5 percent of world reserves, Jordan, with 4 percent, and the remaining 9 percent in Australia, Thailand, Israel, Ukraine, and Estonia. Estonia burns oil shale for power generation.[20]

Oil shale has to be mined, transported, crushed, and heated to a high temperature (450°C) in the presence of hydrogen to produce a low-quality crude oil. The process requires a great deal of water, a commodity in short supply in the western United States, plus natural gas as a source of energy to heat the oil shale and for hydrogen. The crushed rock residue takes up more volume than the original rock, presenting a significant disposal problem in the scenic Rockies. The United States invested a great deal of money during the oil crisis in the 1970s to commercially develop oil shale, but to no avail. As promising as oil shale might appear, in a practical sense mining of shale does not offer a viable solution to a shortfall in conventional oil production.

This negative outlook may change. Shell Oil's Mahogany Project in northwest Colorado took a different approach to oil shale. Rather than mining the oil shale and then extracting the oil, heating elements were embedded one-half mile into the ground to heat the shale rock *in situ* to 700°F over four years. During this time, heated shale produced and released natural gas and a light high-grade crude suitable for refining gasoline. To prevent hydrocarbons from getting into groundwater, a wall of impermeable ice between twenty and thirty feet thick had to surround the heated oil shale. The ice wall required a great deal of electricity and water, the latter in short supply. While three barrels of water are consumed for every barrel of hydrocarbon output, the energy in the hydrocarbon output is estimated to be about 3.5 times the energy consumed to heat the shale and form the ice wall. Denser oil shale formations may be capable of producing up to a billion barrels of oil per square mile. If these estimates of oil shale resources are accurate, and if this method proves to be commercially feasible, the Green River Formation will have about eight times the proven oil reserves of Saudi Arabia.

Figure 6.11 **U.S. versus World Automobile Population**

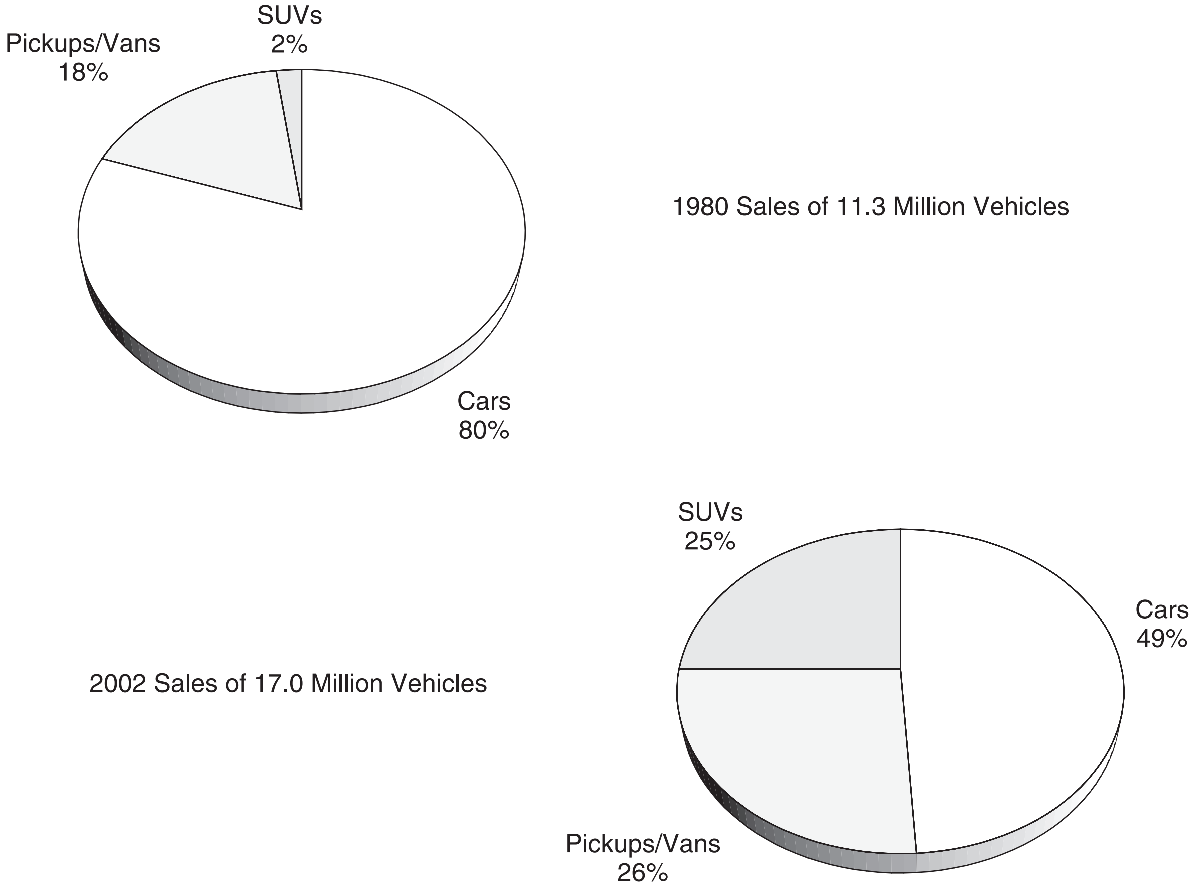


Challenge of Oil: The Automobile

Oil is the fuel of choice for automobiles, trucks, buses, railroad locomotives, aircraft, and ships. Ship's bunkers are the waste or residue of the refining process. Trucks, buses, and railroads consume diesel fuel and airlines jet fuel. For the most part, automobiles run on gasoline, although Europe has succeeded in inducing a switch from gasoline- to diesel-fueled automobiles through tax incentives. The United States is unique in that vans and pickups are used for personal and commercial use, whereas other nations tend to use these vehicles exclusively for commercial purposes. Figure 6.11 shows the world and the U.S. populations of automobiles (SUVs, vans, and pickups included only for the United States).[21]

Henry Ford pioneered the motor vehicle industry in the United States in the early 1900s. Even as late as 1950, 70 percent of the world's motor vehicles were registered in the United States. Since then the U.S. share has declined to 33 percent (as of 2002) with a tenfold growth of the world population of motor vehicles from 70 million to nearly 670 million. The initial rapid growth in the population of motor vehicles after World War II was concentrated first in Europe (1960–1985), then Japan (1970–1985), afterward to the emerging economies of the Industrial Tigers (Korea, Taiwan, Singapore, and Hong Kong), and from there to

Figure 6.12 **Composition of New Car Sales in the United States**



1980 Sales of 11.3 Million Vehicles

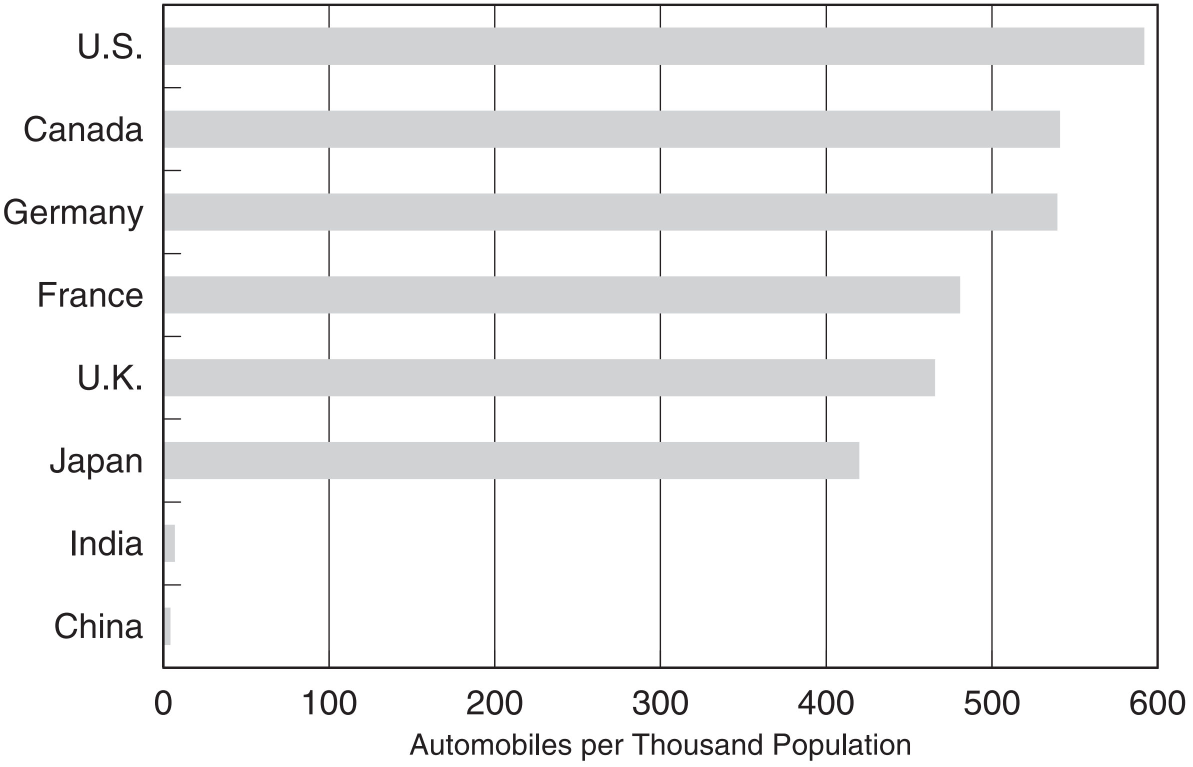
2002 Sales of 17.0 Million Vehicles

South America, eastern Europe, and other areas of the world. The center of motor vehicle population growth is now shifting to India and China.

The United States's love affair with sports utility vehicles (SUVs) is illustrated as a percentage of new car sales in Figure 6.12. Figure 6.13 shows the number of automobiles owned per 1,000 people (U.S. automobile figures includes SUVs, pickups, and vans).

The greatest potential growth in automobile ownership is in China and India. If the number of automobiles in India and China were to increase to 100 per thousand people, the population of automobiles would jump by over 200 million on top of a present population of nearly 700 million. With 51 million bpd of motor vehicle fuel consumed in 2004, an increase of 28 percent in the motor vehicle population would translate to 15 million bpd of incremental motor vehicle fuels. Assuming a refinery yield of 70 percent for motor vehicle fuels, incremental demand for gasoline and diesel fuel would increase demand for crude oil and refinery

Figure 6.13 **Automobiles per One Thousand Population** (2002)



capacity by about 20 million bpd. For this to happen, there would have to be significant expansion of the global capacity to produce and refine crude oil, which, from today's perspective, looks highly unlikely.

Vehicles That Use Alternative Fuels

One way to attack oil demand for motor vehicles is to find an alternative to gasoline. The U.S. Department of Energy defines alternative fuels as substantially nonpetroleum methods that enhance energy security and the environment. The list includes methanol and ethanol fuels of at least 70 percent alcohol, compressed or liquefied natural gas, liquefied petroleum gas (LPG), hydrogen, coal-derived liquid fuels, biofuels, and electricity, including solar power. All are currently more costly than gasoline and diesel oil and, more importantly, lack an infrastructure for serving customers. As Table 6.5 shows, the energy content of alternative fuels is lower than gasoline and diesel oil, which means that motor vehicles that use alternative fuels will get lower mileage (miles per gallon) than those running on gasoline diesel fuel.[22]

The antigasoline public sentiment is fueled by the desire to improve air quality and the concern over oil security. California suffers from the worst polluted air in the nation and the state energy authorities are acutely aware of declining California

Table 6.5

Energy Content of Motor Vehicle Fuels

Fuel	Btu/Gallon
Diesel	129,000
Gasoline	111,400
E85(3)	105,545
Propane	84,000
Ethanol (E100)	75,000
Methanol (M100)	65,350
Liquid hydrogen	34,000
CNG at 3,000 psi	29,000
Hydrogen at 3,000 psi	9,667
Electricity	
Biomass	

and Alaska oil production and its impact on oil security. The state leads the nation in initiating legislation to promote the demise of the conventional gasoline engine. Many states look to California as a model for air pollution legislation.

Ethanol and biodiesel, both biomass fuels, were discussed in Chapter 3. While biodiesel can be readily used as a substitute for diesel fuel, availability is a serious problem because there are only 142 biodiesel fueling sites in the United States, of which 22 are in North Carolina, 17 are in California, and 14 are in the state of Washington. Gasohol, a mixture of 5–10 percent ethanol, known as E-5 or E-10, is widely sold in the Midwest corn-growing region. No engine modifications are necessary for this low concentration of ethanol, which in 2006, was introduced into the Northeast market. E-85 and E-95, mixed with 15 percent and 5 percent unleaded gasoline, respectively, are considered alternative fuels. There are only 188 sites in the United States that sell E-85: 87 are in Minnesota and 13 are in Illinois, the second-ranking state. Major American car manufacturers sell E-85 flexible-fueled automobiles that can be fueled by either E-85, if it is available, or gasoline.

Methanol is an alcohol fuel made from natural gas, but it can also be produced from coal and biomass. The primary methanol fuel is M-85 (85 percent methanol and 15 percent unleaded gasoline). In order to use methanol as a fuel, engine parts made of magnesium, copper, lead, zinc, and aluminum must be replaced to prevent corrosion. Methanol cannot be handled in the same distribution system as petroleum products, and the necessity of building a new distribution system limits methanol's potential use as an automobile fuel. There are no commercial outlets in the United States selling methanol, and there is little point to change to methanol from an environmental point of view because emissions from M-85 are not significantly lower than those from gasoline.

Interest in natural gas (mostly methane) as an alternative fuel stems from its clean-burning qualities and its availability through a well-developed pipeline distribution system. But engines must be modified in order to accommodate natural gas, which is stored in tanks either as compressed (CNG) or liquefied (LNG) natural gas. CNG-fueled vehicles require compression stations either at distribution centers or homes served by natural gas. Natural gas distribution companies commonly use CNG-fueled vehicles. Major American car manufacturers have models that run on CNG exclusively or are bi-fueled to run on either CNG or gasoline. Most CNG-fueled vehicles are restricted to fleet buyers such as natural gas producers or distributors who have ready access to the fuel. LNG-fueled vehicles must have some way of keeping natural gas in a liquefied state unless their tanks can withstand the pressure created when the liquid gasifies. There are sixty-two sites where LNG can be purchased: thirty-five are in California; Texas ranks second with six. CNG is sold at 1,035 sites of which 194 are in California followed by 65 in Utah. The run-up in natural gas prices and the growing need to import natural gas are inhibiting factors for a significant conversion of automobiles from gasoline to natural gas.

LPG is propane or butane alone or as a mix. LPG is a by-product of natural gas processing and petroleum refining. As a vehicle fuel in the United States, LPG is mainly propane and has been in use for over sixty years. Propane is gaseous at normal temperatures and must be pressurized to remain in a liquid state. Emissions from the combustion of propane, however, are significantly lower than those produced by gasoline and diesel fuel, making propane the fuel of choice for forklift trucks and other vehicles that must operate in closed spaces such as warehouses and terminals. There are a few fleets of municipal taxis, school buses, and police cars of propane-fueled vehicles, with nearly 4,000 refueling sites nationwide, of which nearly one-fourth are in Texas. However, because the quantity of LPG is fixed by refinery operations and domestic production of natural gas, switching from gasoline to LPG on a large scale is unlikely because of its limited availability.

Electricity can run a vehicle via a battery or a fuel cell. Batteries store electricity, while a fuel cell generates electricity. The cost and weight of batteries have discouraged their use in the past, but progress has been made for battery-powered vehicles to become technologically and economically feasible. Nissan produces two automobiles that run on electricity for sale in California. Some electricity-generating companies are thinking about electric automobiles, but this would involve running another electric cable to a garage with its own meter to measure the amount and time of electricity usage. Recharging batteries overnight and on weekends would be far less costly because those are the times when electricity-generating utilities have spare, low-cost generator capacity. There are 830 commercial electricity "refueling" sites in the United States, with California again in the lead with 514 sites, Georgia in second place with 87, and Massachusetts third with 41. A large number of battery-powered vehicles would have significant repercussions on electricity generation, transmission, and distribution capacity.

The assertion that electric vehicles are pollution-free is true only if viewed in isolation. Electricity stored in motor vehicle batteries is not pollution-free if it is generated from burning fossil fuels. Nor is the electric vehicle energy-efficient when the inefficiencies of electricity generation and transmission are taken into account. From this viewpoint, electric vehicles are neither pollution-free nor energy-efficient. However, if the electricity to power an electric vehicle comes from wind or solar energy, including solar-powered cars, then the argument that electric vehicles are pollution-free is valid.

Vehicles fueled by hydrogen are another possibility. The chief advantage of hydrogen is that the only emission produced by combustion is water. California has set a goal of a minimum level of hydrogen-fueled vehicles, but with no means of enforcing compliance. There are only seven hydrogen-fueling sites in the United States, of which five are in California. As with CNG, compressed hydrogen can be burned directly in a modified conventional engine. Whereas CNG is pressurized at 3,000 psi, hydrogen must be pressurized at 10,000 psi in order to store enough fuel in a normal-sized tank for 300 miles of travel. Piping not only has to withstand this pressure, but seals have to be specially designed to keep small atoms of hydrogen under high pressure from leaking and forming a potentially explosive mixture. Opponents of hydrogen point to the Hindenburg as a good reason not to have a tank of high-pressure hydrogen in a car, yet they are willing to ignore the danger of having a highly volatile tank of gasoline a few feet away from passengers.

Hydrogen is the preferred fuel for a fuel cell as it emits only water. Hydrogen cannot be considered pollution-free if it comes from the electrolysis of water where the source of electricity is burning fossil fuels. If the source is hydroelectric or electricity from solar cells or wind, then the hydrogen-fueled vehicle as a system, including the hydrogen source, can be considered pollution-free. Fuel cells are discussed under Hydrogen Economy in Chapter 10; but for now the prognosis for large numbers of automobiles fueled by hydrogen, either for direct combustion in a conventional engine or feedstock for a fuel cell, is quite dim.

Despite intense government and private efforts to support the technological development of vehicles run by alternative fuels, including the development of fuel cells, over the past three decades (all of this started in the aftermath of the 1973 oil crisis), most of the world's fleet of automobiles and trucks is still powered by gasoline and diesel fuel. In the United States, the total population of automobiles, vans, pickups, SUVs, trucks of all sizes, and buses is 221 million; with a population of people close to 300 million, there are a little over 700 motor vehicles of all types for every 1,000 Americans. Nearly all are fueled by gasoline and diesel. Vehicles fueled by LPG (including forklifts) number 200,000. CNG-fueled vehicles number nearly 145,000, and an equal number are fueled by E85 (85 percent ethanol and 15 percent gasoline), excluding flexible-fuel vehicles that run on either E85 or gasoline. Vehicles run by electricity number 55,000, with those run on methanol (M85) and LNG coming in last at 5,000 and 3,000, respectively.

It is possible that long-term oil consumption could be affected by technological breakthroughs that bring the cost of alternative fuels closer to that of gasoline and diesel fuel. Although not a challenge to crude oil yet, the potential conversion of natural gas to motor vehicle fuels (discussed in Chapter 7), currently being spearheaded by Shell Oil, could have some impact on crude oil demand as a motor vehicle fuel, but not in the near future.

One thing is certain; the lack of significant progress for motor vehicles fueled by alternative means is not being hampered by the automobile industry. Automobile manufacturers share the general public sentiment that the conventional gasoline engine is becoming archaic and needs to be replaced. Carmakers are not wedded to the oil industry, but they are wedded to a technology that works. Until there is an alternative fuel that works, which also means that it is widely available to the public, oil will remain the preferred fuel. There will be no easy divorce from oil.

Enhancing Engine Efficiency

There has been substantial private and governmental support to enhance engine efficiency. Increased efficiency has two benefits: better mileage with less pollution. Doubling mileage cuts both fuel consumption and pollution emissions in half.

The U.S. Department of Energy, through the National Renewable Energy Laboratory, has funded development costs for hybrid electric vehicles (HEVs) with high fuel economy and low emissions. GM and Ford have announced their intention to manufacture hybrid vehicles although Honda and Toyota have had models available for a number of years (another testament to the decline of American leadership in technology).

An HEV obtains higher mileage by converting the energy lost during deceleration to electrical energy that can be stored in a battery. HEVs can be designed in a series or parallel configuration. In a series configuration, the primary engine drives a generator that powers electric motors to drive the vehicle and charge the battery. The vehicle is driven solely by the electric motors. HEVs now on the market have a parallel configuration in which the car is driven directly by a gasoline-fueled engine augmented by electric motors. The electric motors are run by electricity stored in a battery that supplements the power from the gasoline engine, and cut in when the car needs extra power. The nickel metal hydride battery is recharged during deceleration, when regenerative braking captures the energy normally passed to the environment as waste heat, and also during normal motor operation if a charge is necessary.

Supplemental acceleration using electric motors means that automobiles can be built with gasoline or diesel engines of lower horsepower, which consume less fuel. Fuel efficiency is further enhanced by "cutting out" the firing of a cylinder at cruising speed when on level ground. In addition, an HEV has less weight because its engine components are made from lighter-weight aluminum, magnesium, and plastic. The vehicle's body, also made of lightweight aluminum, is aerodynamically designed to reduce wind resistance. Depending on the style of HEV, mileage can range from thirty to more than sixty miles per gallon. An HEV's mileage performance, compared to that of a conventional automobile, is more impressive in stop-and-go traffic than in steady highway driving. HEVs' engine emissions meet California's stringent ultra-low vehicle emission standards. The most striking aspect of HEVs is that their sales jumped when gasoline prices spiked during the summer of 2004, sending SUV sales into a slump. This lesson is vital in coming to terms with the future.

Internalizing an Externality

The government supports the oil industry by ensuring security of supply. This is not a subsidy to the oil companies because oil companies are not in the business of military interventions. Government participation in the civilian economy is common. The automobile industry would have been truncated (to say the least) if town, county, state, and federal governments did not build roads. It would be just as unfair for automobile companies to be responsible for building roads as it would be for oil companies to ensure oil security. The big difference is in the method of payment. The cost of building and maintaining roads falls on the user in the form of a gasoline tax, whereas the cost of oil security falls on the taxpayer as a government expenditure. It is high time that those who benefit from oil bear the full cost of oil through an oil security tax.

The United States consumes 6 million bpd of gasoil as both heating oil and diesel fuel for equipment, machinery, motor vehicles (mostly trucks), and locomotives. Railroads are actually a fuel-efficient alternative to trucking. The advantage of trucks over railroads is their flexibility: they can go anywhere. The advantage of railroads is their inherent efficiency: A train crew of only three members can haul several hundred containers or truck trailers on flatbed railcars, a number that would require an equal number of truck drivers. Aside from this labor savings, railroads with steel wheels on steel tracks are far more energy-efficient than trucks with rubber wheels on concrete or asphalt roads. An optimal blend of both modes of transport is intermodal transport (piggyback) of combining trucks for short-distance delivery with rail for long-distance hauling. A higher price for diesel fuel would provide an economic incentive to get some of the trucks off the road and their payloads on a train. This would reduce oil consumption and highway congestion.

In 2002, estimates of how much gasoline and diesel fuel motor vehicles consume in the United States were 112 billion and 34.8 billion gallons,

Table 6.6

How Far Can a Gasoline Tax Go?

	Cost in $ billions	$/Gallon Tax to Cover Cost
2004 U.S. Defense Budget	$405	$2.20
2004 Federal Budget Deficit	$375	$2.04
2004 Federal Budget Deficit not counting Social Security Surplus	$536	$2.91

respectively. This works out to 7.2 million bpd for gasoline and 2.3 million bpd for diesel oil, for a total of 10 million bpd (rounded). Looking into the future, let us suppose that automobiles and trucks consume 12 million bpd of gasoline and diesel fuel.[23] How much revenue would a $1 per gallon oil security tax on motor vehicle fuels generate? Twelve million barrels per day multiplied by 42 gallons per day at $1 per gallon would produce $500 million per day, or a little over $180 billion per year. Table 6.6 shows how much of the budget deficit could be covered by a motor vehicle tax.[24]

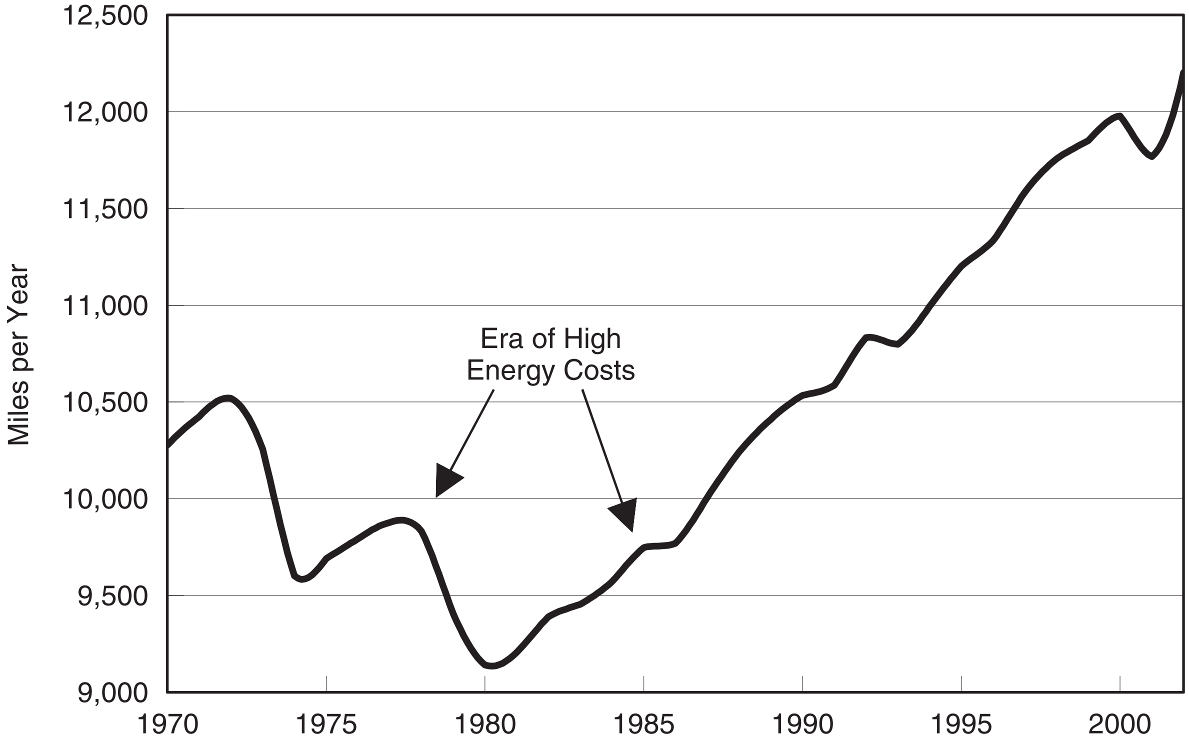
A $2–$3 hike in the price of gasoline would make U.S. prices comparable to what Europeans pay. European governments rely on motor vehicle fuel taxes as a significant portion of their revenue, and maybe we should consider this possibility. A high price on gasoline would have two beneficial impacts: It would provide an incentive to purchase more fuel-efficient automobiles and make motorists think about their driving habits. Figure 6.14 shows the impact of high-priced gasoline on the average number of miles driven by U.S. automobiles.

In round terms, the average miles driven per year fell about 10 percent during the era of high energy costs during the late 1970s and early 1980s. People were not deprived of the driving experience, they just thought more carefully about how they drove their cars. Carpooling, taking the bus or train to work, not visiting the shopping mall every day, letting the kids take the bus from school rather than picking them up, and shortening the distance traveled for family vacations can have a significant impact on the number of miles driven in a year.

Let us ignore vehicles that run on diesel fuel because it is used mainly by trucks, buses, and locomotives. Any discretionary travel is done driving gasoline-fueled automobiles including vans, pickups, and SUVs. During the era of high energy costs, the average automobile was driven about 9,500 miles per year compared to the 12,200 miles per year reported for 2002. If we could reduce the average annual miles driven by 10 percent, to 11,000 miles, which is significantly above the 9,500 miles typical of the early 1980s, this would represent a savings in gasoline consumption of 10 percent. Suppose that a few years of individuals buying more fuel-efficient cars, not necessarily HEVs, and fewer SUVs could create fuel savings of another 5 percent, for a combined savings of 15 percent in gasoline consumption.

What does that figure mean? With a consumption of 10 million bpd in 2004, throwing in a little for diesel-powered automobiles, a 15 percent reduction would mean a reduction of 1.5 million bpd, a bit more than half of our Middle East imports of 2.7 million bpd. Adding in the incremental crude from tar sands production in Canada, and perhaps some additional syncrude plants to be built in Venezuela, we are not far from eliminating Middle East imports in North America. This means we might be able to extricate ourselves from what really is a government subsidy, not to oil companies, not to automobile owners, but to other Middle East oil consumers. Figure 6.15 reflects the relative dependence of the United States on Arabian Gulf exports compared to the rest of the world.[25]

Figure 6.14 **Average Miles Driven per Year for Passenger Cars**

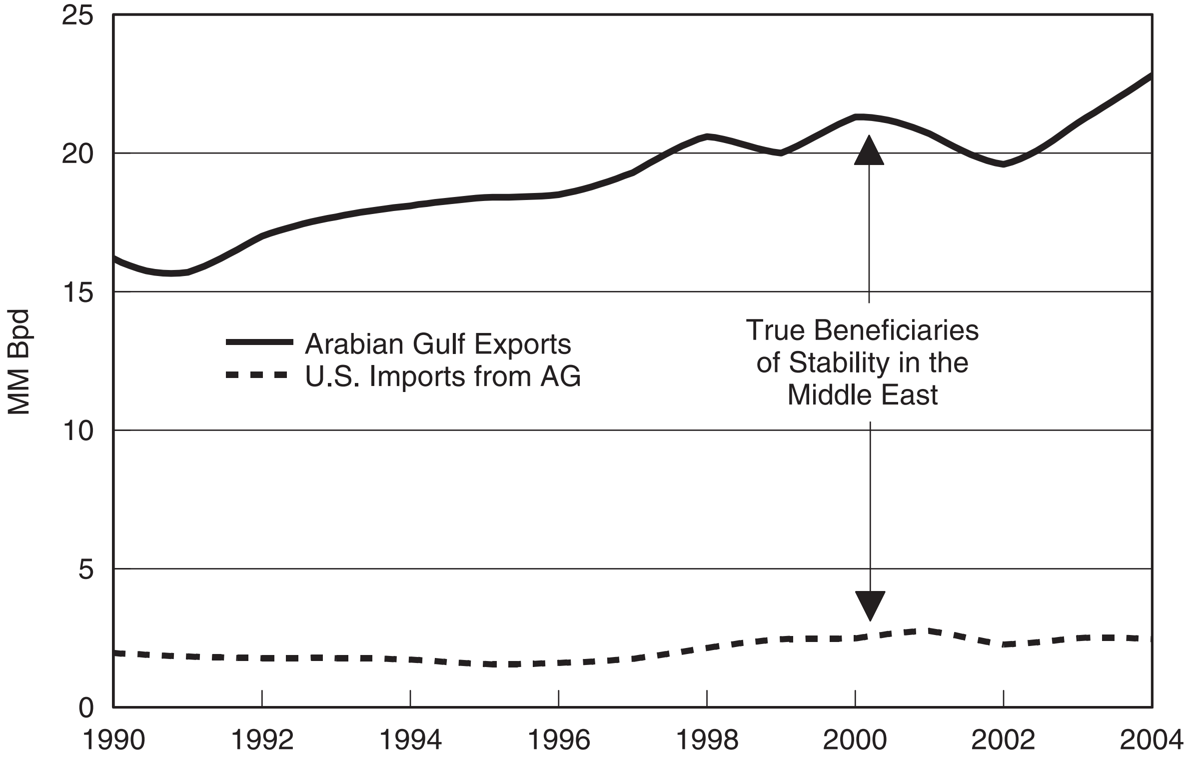


Is This Politically Acceptable?

Of course reducing our dependence on oil from the Middle East by hiking gasoline prices is not politically acceptable, but what relevance is that? If we do nothing, which is how our political system seems to work with regard to energy, we will be doing our part to ensure that the world remains dangerously close to supply not being able to satisfy demand. The problem with being on the razor's edge is that one can easily fall off. In the world of oil, an extended supply disruption in Nigeria or Venezuela or Iran, or any number of other possibilities can reduce the oil supply below demand. Once this happens, there is no upper limit on oil prices. Then we will be paying $4–$5 per gallon or more with the incremental proceeds flowing to the oil producers. On the other hand, if we charge ourselves $4–$5 per gallon by imposing an oil security tax, then we can extricate ourselves from the Middle East and save billions of dollars and many lives, increase the availability of oil for the rest of the world, and in so doing, ease the strain on global oil supplies.

Stated in the most simplistic way, we have a choice of paying $4–$5 per gallon or paying $4–$5 per gallon. The choice is only differentiated by when and where the money flows: to the U.S.

Figure 6.15 **Arabian Gulf Oil Exports and U.S. Imports from the Arabian Gulf**



government to ease the budget deficit or to the coffers of the oil exporters. In either case, there will be an economic incentive for the development of alternative fuels.

It is high time for doing something to at least reduce the growth in demand and our dependence on Middle East oil and becoming more serious about developing alternative motor vehicle fuels. The old saw that alternative fuels cannot be developed because of our enormous vested interest in the infrastructure of oil production facilities, refineries, and associated processing plants, pipelines, ships, storage tanks, and distribution facilities is not going to carry much weight when that infrastructure goes dry.

Notes

1. The primary sources of information throughout this section and my discussion of refining are: Charles F. Conaway, *The Petroleum Industry: A Nontechnical Guide* (Tulsa, OK: PennWell Books, 1999); Norman J. Hyne, *Nontechnical Guide to Petroleum Geology, Exploration, Drilling, and Production* (Tulsa, OK: PennWell Books, 2001); Robert O. Anderson, *Fundamentals of the Petroleum Industry* (Norman, OK: University of Oklahoma Press, 1984).

2. Aspects of the abiogenic gas debate can be found in J.F. Kenney, Vladimir A. Kutcherov, Nikolai A. Bendeliani, and Vladimir A. Alekseev, "The Evolution of Multicomponent Systems at High Pressures: VI. The Thermodynamic Stability of the Hydrogen–Carbon System: The Genesis of Hydrocarbons and the Origin of Petroleum," *Proceedings of the National Academy of*

Sciences 99, 17 (August 2002), pp. 10976–10981, and at the American Association of Petroleum Geologists Web site at www.aapg.org/explorer/2002/11nov/ abiogenic. cfmg.

3. The information presented in Figure 6.1 can be found in the *Monthly Energy Review* (Section 5), published by the Energy Information Administration, Washington, DC (2004).

4. For more information on Mr. Charlie visit www.rigmuseum.com.

5. For more information on the Mohole Project see www.nas.edu/history/mohole.

6. More information on deep-water drilling is available at the Transocean Company's Web site at www.deepwater.com.

7. The statistics on offshore drilling rigs presented in Table 6.1 are from *Offshore Drilling Monthly* (May/June 2004), published by the Oil Service Group of Jefferies & Company, a New York investment firm (www.jefco.com).

8. From *Oil Literacy* (New York: Poten & Partners, 1988).

9. The information provided in Table 6.2 is from the *OSHA Technical Manual,* Chapter 2, Section IV (Washington, DC: U.S. Department of Labor, 1999); also available online at www.osha-slc.gov.

10. The information shown in Table 6.3 on world oil reserves comes from *BP Energy Statistics* (London: British Petroleum, 2005).

11. Michel T. Halbouty, ed., *Giant Oil and Gas Fields of the Decade (1990–1999)* (Tulsa, OK: American Association of Petroleum Geologists, 2003).

12. "Deepwater Oil Discovery Rate May Have Peaked, Production Peak May Follow in Ten Years," *Oil & Gas Journal* (July 26, 2004), one of a six-part series on Hubbert Revisited, Tulsa, OK; the Association for the Study of Peak Oil and Gas (www.asponews.org and www.peakoilnet); lecture by C.J. Campbell, "Peak Oil," available online at www.geologie.tu-clausthal.de/Campbell/lecture.html; *The Wolf at the Door: Beginner's Guide to Oil Depletion*, available online at www.wolfatthedoor.org.uk; "Future World Oil Supplies," see www.dieoff.org; Hubbert Peak of Oil Production (www.hubbertpeak.com).

13. U.S. Geological Survey, *World Petroleum Assessment 2000* (Washington, DC: Author, 2000).

14. Bill Kovarik, *The Oil Reserve Fallacy: Proven Reserves Are Not a Measure of Future Supply,* published online at www.radford.edu/~wkovarik/oil/oilreservehistory.html.

15. The *International Energy Outlook 2004* (Washington, DC: Energy Information Administration, 2004) calls for a 200 percent increase in OPEC exports by 2025, whereas the *World Energy Outlook* (Paris: International Energy Agency, 2004) calls for a 270 percent increase.

16. "OPEC Adds Less than Meets the Eye," *Petroleum Intelligence Weekly* (June 28, 2004).

17. Asian Development Bank Annual Economic Report for 2004 Web site www.adb.org .

18. Petroleum Communication Foundation, *Canada's Oil Sands and Heavy Oil* (Calgary, Alberta, Canada: Author, 2000). The Petroleum Communication Foundation's Web site is www.centreforenergy.com/.

19. U.S. Department of Energy Country Briefs for Venezuela, available online at www.eia.doe.gov/emeu/international/venez.html.

20. Julie Cart, "U.S. Backs Squeezing Oil from a Stone," *Los Angeles Times* November 30, 2004; World Energy Council (www.worldenergy.org).

21. The figures presented in Figure 6.11 are taken from *Transportation Energy Data Book* (Washington, DC: U.S. Department of Energy, 2004).

22. U.S. Department of Energy, *Transportation Energy Data Book* (Washington, DC: Author, 2004). This report was based on the National Alternative Fuels Data Center of the U.S. Department of Energy, www.afdc.doe. gov.

23. The *Transportation Energy Data Book,* published by the U.S. Department of Energy (2004) is the source for the statistics on gasoline and diesel fuel consumption in 2002. A reasonable estimate of near-term consumption of gasoline and diesel fuel in motor vehicles would be 12 million bpd.

24. The budget data presented in Table 6.6 is from the Congressional Budget Office, available online at www.cbo.gov.

25. The statistical information was provided by the U.S. Energy Administration (www.eia.doe.gov).

7

Natural Gas

With natural gas reserves double that of oil and with plenty of potential for expansion, the "oil and gas industry" may one day be dubbed the "gas and oil industry." This chapter covers the history of natural gas from its beginning as a manufactured gas made from coal. It is the most regulated of fossil fuels because only one natural gas pipeline can be connected to a house, just as a house can have only one electrical cable. Like electricity, natural gas is in the midst of deregulation (liberalization). How natural gas travels from the earth to the points of consumption is discussed, along with the growth of the international trade in natural gas and the possibility of developing nonconventional sources of methane.

Background

Natural gas is made up of primarily methane, a carbon atom surrounded by four hydrogen atoms. It is the cleanest-burning fossil fuel with only water and carbon dioxide as products of combustion. Carbon monoxide emissions, if any, are caused by insufficient oxygen to support combustion. Nitrous oxides stem from nitrogen in the air that reacts with the heat of a flame. Natural gas produces far less nitrous oxides than oil and coal, which contain nitrogen within their molecular structures. Burning natural gas produces virtually no sulfur oxides and no particulate or metallic emissions. A greater ratio of hydrogen to carbon atoms means less carbon dioxide emitted per unit of energy released than coal and oil. Moreover, the technology of electricity generation using natural gas has a higher thermal efficiency than coal and oil, further lowering carbon dioxide emissions for the same output of electricity.

Natural gas fields have about double the reservoir recovery (70–80 percent) than oil (30–40 percent), which lessens the need to continually find new gas fields. Unlike oil, natural gas requires relatively little processing to become "pipeline-quality." On the minus side, natural gas has always been a logistical challenge. In the beginning decades of the oil age, much of the natural gas produced in association with crude oil was flared (burned) or vented to the atmosphere. The primitive state of pipeline technology restricted natural gas to the local market. Large amounts of natural gas associated with oil production were available with the development of oil and gas discoveries in the U.S. Southwest. With no nearby markets to consume the gas and no means to get the gas to distant markets, vast quantities of natural gas associated with crude oil production were vented to the atmosphere. This waste of a "free" energy source and the waning of natural gas fields in Appalachia provided a strong incentive to improve pipeline technology to connect suppliers with consumers over long distances in a safe, reliable, and cost-effective manner.

Another drawback is that leaking natural gas can asphyxiate the occupants of a building or trigger a fire or an explosion that can level a building or, on occasion, a city block. Fires fueled by broken gas mains in the aftermath of earthquakes, such as occurred in San Francisco in 1906 and Kobe in

1995, exacerbated the damage and suffering. Unlike liquid petroleum products, consumers have no way of storing natural gas. Natural gas delivery systems must be designed to handle extreme vagaries in demand.

Most pipelines are largely confined to a single nation or region, such as North America, where they connect producers and consumers in Canada, the United States, and Mexico. The United States alone has 300,000 miles of transmission pipelines and about 1 million miles each of gathering and distribution pipelines. Russia has a well-developed natural gas pipeline system to serve its domestic needs and also exports large volumes to the European pipeline grid. The grid crosses national borders, much as pipelines cross state borders in the United States, connecting European consumers to gas fields in Russia, the Netherlands, the North Sea, and Algeria via two undersea trans-Mediterranean pipelines. One pipeline from Algeria crosses Morocco and the Mediterranean Sea to Spain (near Gibraltar) and the other crosses Tunisia, the Mediterranean, Sicily, and the Strait of Messina to mainland Italy.

There are limits to pipeline transmission that leave enormous reserves of stranded gas beyond the reach of consumers. With limited domestic consumption, natural gas associated with oil production in the Middle East, Southeast Asia, and West Africa was either flared or vented to the atmosphere or reinjected into oil fields. Flaring and venting are a horrendous waste of energy, equivalent to burning money in this world of high energy prices. Reinjection maintains the pressure of oil fields and preserves a valuable energy resource. In recent decades a new method of shipping natural gas in a liquefied state aboard highly specialized tankers has emerged. Liquefied natural gas (LNG) export plants, coupled with specialized import terminals and LNG carriers, have monetized these reserves of stranded gas by making them available to consumers. Since LNG cargoes can be shipped from exporting terminals in North and West Africa, Latin America, the Middle East, Southeast Asia, and Australia to receiving terminals in the United States, Europe, and Asia, natural gas is being transformed from a regional to a globally traded commodity like oil. Furthermore, technological progress has been made in converting natural gas to liquid motor vehicle fuels, giving natural gas access to the same delivery system that serves petroleum.

A long-standing and complex relationship exists between natural gas and electricity as the proliferation of electric and gas utilities suggests. Gas both supplies fuel to generate electricity and competes with electricity to supply consumers with a means to cook, heat water and living spaces, and run appliances. Both became federally regulated commodities in the 1930s as a result of interstate transmission. Later on, natural gas regulation was expanded to include natural gas suppliers of regulated interstate transmission pipelines. This experiment in total regulation of an industry turned into a bureaucratic quagmire, and internal contradictions and undesired consequences plagued the regulators. The final solution to the problems induced by regulation was deregulation of natural gas production, transmission, and distribution, beginning in the late 1970s. Again, the link between electricity and natural gas can be seen in the parallel deregulation of electricity generation, transmission, and distribution. While the breakdown of the monopoly status of natural gas and electricity is quite advanced in the United States and the United Kingdom, it is still an ongoing process. Deregulation (liberalization) is actively being pursued elsewhere in the world.

Natural gas as an energy source looks extremely promising for the coming decades and has a number of advantages working in its favor. However, natural gas, like oil, will eventually become another depleting resource, a fact that the world will eventually have to contend with. Some say that day, while admittedly farther away than that for oil, may not be that far in the future.

Early History of Coal Gas

The beginning of the natural gas industry was not natural gas from the first oil wells in western Pennsylvania, but manufactured gas from coal—a case of a synthetic or manufactured fuel preceding the use of a natural fuel. In 1609 a Belgian physician and chemist reduced sixty-two pounds of coal to one pound of ash and speculated about what had happened to the missing sixty-one pounds, the first published account of coal gas. While burning coal in the presence of air reduces it to ash, heating coal in a closed environment, without a fresh supply of air, produces coke, tar, and gas. Coke, primarily carbon, is burned as a fuel or consumed in steel production. Coal tar was originally a waste product, dumped willy-nilly in streams, rivers, ponds, and on land adjacent to manufactured gas plants. "Free" coal tar, as a potential raw material for useful products, became the cornerstone of the chemical industry by first being transformed to creosote, tar, pitch, wood preservatives, mothballs, and carbon black. Later on the chemical industry learned to extract benzene, toluene, and xylene (as gasoline components or feedstock for the petrochemical industry) plus phenol and polynuclear, aromatic hydrocarbons found in synthetic fibers, epoxies, resins, dyes, plastics, disinfectants, germicides, fungicides, pesticides, and pharmaceuticals. Unfortunately, a large amount of the early coal tar was not processed. When natural gas replaced manufactured gas, the shift left a legacy of thousands of abandoned manufactured gas plants, now classified as hazardous and toxic waste dumpsites under the Superfund Program of the U.S. Environmental Protection Agency (another example of the consumer not paying the full cost for a service).

The purpose of manufactured gas plants was not to make coke or coal tar, but coal gas, a mixture of hydrogen, carbon monoxide, carbon dioxide, and methane. The heat content of coal gas, made up partially of noncombustible carbon dioxide, is half that of natural gas. The first demonstration of coal gas as an energy source occurred in 1683 when an English clergyman stored coal gas in an ox bladder; when he was ready to use it, he pricked the bladder and lit the outgoing gas. The first demonstration of coal gas as a means of illumination occurred a century later, in 1785, when a professor of natural philosophy lit his classroom by burning coal gas in a lamp. In 1801, a French engineer used coal gas to light and heat a Parisian hotel. William Murdoch, an engineer working for Boulton and Watt, the manufacturer of James Watt's steam engines, produced coal gas that passed through seventy feet of copper and tin pipe to light a room in his house in 1792 and, in 1802, to light a foundry. He experimented with various types of coal heated to different temperatures for varying lengths of time, which led to the first major commercial use of coal gas: to light the Manchester cotton mills for round-the-clock operation. For his pioneering work, Murdoch was called the father of the gas industry.[1]

Other advances included purifying coal gas by passing it through limewater and devising meters to measure its usage. Friedrich Albrecht Winzer, a German entrepreneur, proposed the first centralized gas works where gas would be made in large quantities and pipelined to customers for lighting and heating. Germany was not ready for the idea, so he anglicized his name to Fredrick Albert Winsor and sold the concept to the Prince of Wales, a fellow German of the house of Hanover, who had gaslights installed for celebrating King George III's birthday in 1805. In 1812 the Westminster Gas Light and Coke Company was chartered by parliament, and by 1815 the company was supplying London from a centralized coal gas producing plant via twenty-six miles of gas mains of the same three-quarters-inch pipe used to make rifle barrels. This placed England in the forefront of a new industry and a font of technological know-how for the introduction of gas lighting in Europe and America.

Two sons of Charles Peale, a well-known portrait painter of Revolutionary War heroes (including fourteen of George Washington), played important roles in what would lead to the formation of the Gas Light Company of Baltimore in 1816. In 1817 the company received a franchise from Baltimore to provide gas lighting. Progress was slow, and only two miles of gas mains were supplying 3,000 private and 100 public lamps by 1833. The company's activities spread into manufacturing and repair of gas meters, along with producing chandeliers, pipes, and fittings in order to be able to sell coal gas for illumination. The spread of manufactured gas in major cities for lighting was not particularly rapid, beginning with Baltimore in 1817, New York City in 1825 (the Great White Way of Broadway was first lit with gas, not electricity), Philadelphia in 1836 (the first municipally owned gas works), Cincinnati, St. Louis, and Chicago in the 1840s, San Francisco in 1854, Kansas City and Los Angeles in 1867, and Minneapolis and Seattle in the 1870s. The advantage of coal gas was easy shipment of coal to manufactured gas-producing plants strategically located in the center of their markets, which minimized the cost of laying pipes, thus causing coal gas plants to proliferate. The slow adoption of manufactured gas for lighting was its cost, which ranged between $2.50–$3.50 per thousand cubic feet (Mcf). In twenty-first-century dollars, this would be equivalent to about $60 per Mcf compared to $6.75 per Mcf for the average price of natural gas delivered to residential consumers during the 1990s.

Consumers have a choice of oil companies when buying liquid petroleum products, the hallmark of a competitive market. Manufactured gas began as a natural monopoly because laying multiple gas mains to give consumers a choice of supplier was not deemed cost-effective. Moreover, manufactured gas companies required municipal assistance, support, and cooperation to get into business including a franchise to be sole supplier, permits to lay gas mains under city streets, and a contract to light city streets. All these were necessary for a manufactured gas company to assure potential investors of sufficient revenue for a return on their investments. Once the gas mains were laid for city lighting, it was a relatively simple matter to connect to residences and businesses. While municipal authorities recognized that a single company could provide gas at a lower cost than two competing companies with twice the investment in facilities and pipelines, they also recognized that a single company, once ensconced in a market as a natural monopoly, would not be cheaper. Thus, a franchise that granted a monopoly also specified municipal oversight on entry, expansion, exit, safety, and rates to protect the public interest.

Local or municipal regulation worked well with manufactured gas providers whose plant and distribution system were within the legal jurisdiction of a municipality. Things changed when natural gas began to displace manufactured gas because natural gas fields were normally outside of a municipality's legal jurisdiction. This complicated regulation for the municipalities, a problem that was solved when state governments replaced municipal regulation. Since a natural gas field generally served several municipalities, the natural gas industry opted for statewide rather than municipal regulation, a situation that promised greater consistency of rules and rates and reduced the number of regulators to be dealt with (or influenced).

History of Natural Gas

Sacred fires in Persia and elsewhere were natural gas seeps that may have been ignited by lightning. The temple of Delphi was built around a "burning spring." Around 400 BCE the Chinese discovered natural gas bubbling through brine, which they separated and burned to distill salt. Around 200 CE the Chinese learned to tap natural gas deposits and route the gas through bamboo pipes to distill salt from seawater and cook food. The earliest reference to natural gas in the United States was in the 1600s when explorers noted certain Indian tribes burning gaseous emissions from the earth. In 1821, a more organized approach to capturing escaping or seep gas started in Fredonia, New York, when a gunsmith piped seep gas to nearby buildings for lighting. In 1827, another source of naturally occurring seep gas was harnessed to supply a lighthouse on Lake Erie. In 1840, the first industrial use of natural gas occurred in Pennsylvania, where gas was burned to heat brine to distill salt, the same thing the Chinese had done more than two millennia earlier.

While natural gas provided the lift for Drake's well, for the most part, natural gas found along with oil was vented to the atmosphere. Drilling for oil and discovering natural gas was equivalent to a dry hole. Natural gas was normally out of reach of municipalities and was unable to compete with manufactured gas protected by municipal franchises. In 1872 the Rochester Natural Gas Light Company was formed to provide natural gas to Rochester, New York, from a field twenty-five miles away. Pipes made of two- to eight-foot segments of hollowed-out Canadian white pine logs reflected the primitive state of pipeline technology. The problems associated with a rotting and leaking wooden pipeline eventually led to the company's demise. In the same year a five-and-a-half-mile, two-inch-wide wrought-iron pipeline was successfully constructed to carry waste gas from oil wells near Titusville to 250 townspeople.

But cast- and wrought-iron pipelines were plagued by breaks and leaking connections held together by screws. Before the day of compressors, transmission distance was limited by gas well pressure. In 1870, Pittsburgh became the first city to start consuming natural gas as a substitute for coal to clean up its smoke-laden atmosphere. The Natural Gas Act, passed in 1885 by the Pennsylvania legislature, permitted natural gas to compete with manufactured gas. This proved to be the driving wedge that enabled natural gas to penetrate the manufactured coal gas business and resulted in the formation of Peoples Natural Gas, which by 1887 was serving 35,000 households in Pittsburgh. Another Pittsburgh natural gas distributor, Chartiers Valley Gas, was the first company to telescope pipe from an initial eight to ten and finally twelve inches in diameter to reduce gas pressure before it entered a home, business, or industrial plant. By this time screws had given way to threaded pipe to hold pipe segments together. Dresser and Company, formed in 1880, specialized in pipe couplings, and in 1887 received a patent for a leak-proof coupling that incorporated a rubber ring in the pipe joints; an invention that would dominate the market until the 1920s.

George Westinghouse, inventor of the compressed-air railroad brake, became interested in natural gas and decided to drill for natural gas. He selected, of all places, his backyard and, lo and behold, he struck natural gas as one might expect for the rich to get richer. He became one of the largest gas distributors in Pittsburgh, and relied on the natural gas produced from one hundred wells in and around Pittsburgh, including his backyard. Westinghouse was well versed in the dangers associated with natural gas such as gas users not turning off their gas appliances (lamps, stoves, heaters) when natural gas pipelines were shut down for repair of breaks and leaks. When pipeline service was restored, a nearly odorless and colorless gas seeped into homes and shops, threatening to kill those within from asphyxiation, fire, or explosion. Westinghouse put his experience with compressed air to good use and originated a number of patents for enclosing main gas lines in residential areas with a conducting pipe to contain gas leaks, introducing pressure regulators to reduce gas pressure before it entered residences and commercial establishments, and cutoff values to prevent any further flow of gas once gas pressure fell below a set point.

These improvements made Pittsburgh the center of the natural gas industry by the late 1880s, with 500 miles of pipeline to transport natural gas from surrounding wells to the city and another 230 miles of pipeline within the city limits. Andrew Carnegie, the steel magnate, promoted the use of natural gas in steelmaking. Natural gas became the fuel of choice not only for steel mills, but also glassmaking plants, breweries, businesses, homes, and a crematorium. Hundreds of natural gas companies were formed to sell gas to municipalities in Pennsylvania, West Virginia, Ohio, and Indiana with a local supply of natural gas. Some of these gas fields were rapidly depleted, forcing a switch back to manufactured gas. Early customers were simply charged a monthly rate for a hookup without a means to measure the amount of gas consumed. When meters were eventually installed, a new business sprang up: renting "gas dogs" to greet meter readers on their days of visitation.

John D. Rockefeller entered the natural gas business in 1881. True to form, through mergers with existing pipeline companies and expanding their business activities once they were under his control, Standard Oil established a major market presence in the gas-producing states in Appalachia. Rockefeller's success at monopolization led to the passage of the Hepburn Act in 1906, which was intended to give the Interstate Commerce Commission (ICC) regulatory authority over interstate natural gas pipelines, even though very few existed at the time. In the end, the Hepburn Act exempted natural gas and water pipelines from regulatory oversight, but growing concern over Rockefeller's hold on the oil industry led to the U.S. Department of Justice filing suit under the Sherman Antitrust Act against Standard Oil. Curiously, in the Standard Oil breakup in 1911, the company's natural gas properties and activities remained intact within Standard Oil of New Jersey, enabling the company to maintain its standing as a major natural gas player in the Midwest and Northeast and, eventually, the Southwest.

The Battle Over Lighting

Manufactured gas commanded the market for lighting in urban areas while kerosene continued to be used in rural areas and towns not hooked up to manufactured gas. Though vulnerable to penetration by natural gas, coal gas was given a new lease on life by the discovery of a technique for making "water gas" by injecting steam into anthracite coal or coke heated to incandescence. This produced a flammable mixture of hydrogen and carbon monoxide that was sprayed with atomized oil (a new market for oil) to increase its heat content to match that of coal gas. Less costly to make than coal gas, water gas had 75 percent of the manufactured gas market by 1900.

While water gas could temporarily hold natural gas at bay, a new competitive threat entered the lighting business, affecting both manufactured and natural gas: electricity. In 1880, Edison rigged Broadway for illumination by electricity and lost no time attacking gas lighting for its odors, leaks, fires, explosions, and transport in "sewer pipes," ignoring, of course, the risk of electric shock, electrocution, and fires from exposed wires. In 1882, the Pearl Street generating station provided electricity to 1,284 lamps within one mile of the plant. Edison used existing gas statutes for permission to install electric wiring under streets and set up a system to supply electricity that mirrored gas as closely as possible to make it easier for customers to switch. The gas distribution companies knew that electricity would replace gas for lighting and responded with a two-pronged program to meet the new competitive threat. The first was to shift the emphasis of gas from lighting to cooking and heating and the second was to pursue corporate consolidation to strengthen their position.

As the availability of electricity spread throughout the nation, it did not take long for managers of consolidated gas companies to see the virtue of expanding their merger activities to include electricity-generating firms ("if you can't beat the enemy, embrace him"). The coke by-product from coal gas production could be burned to make electricity and mergers would result in major savings in corporate overhead. The first merger occurred in Boston in 1887, setting the example for the creation of innumerable gas and electric or electric and gas utility companies across the nation. Consolidating gas companies and merging with electricity-generating companies into independent gas and electric utilities further evolved into the public holding company, which owned controlling interests in independent electric and gas companies.

The first public holding company was formed by Henry L. Doherty, who started out as an office boy and rose to chief engineer of a natural gas company. Noticing that poorly designed gas stoves were a drag on natural gas sales, Doherty increased gas sales by working with stove manufacturers to improve their product. He switched to marketing, where he was an instant success because of his ability to motivate and lead salespeople, initiating all sorts of promotional activities, and setting high standards of customer service. Doherty then established his own company to provide advice on the reorganization, management, and financing of public utility companies. He began to attract investor interest and in 1910 formed Cities Service Company, the first public holding company. As the name suggested, the company was to serve cities across the nation with gas and electricity and, by 1913, Cities Service controlled fifty utilities in fourteen states.

Cities Service was a model for a much larger public utility empire created by Samuel Insull, who started out as the English representative of a U.S. bank representing Thomas Edison's interests in London. He ended up working directly for Edison as his private secretary by day and learned the electricity-generating business at the Pearl Street plant by night. He eventually rose to third place in the newly formed General Electric, a merger involving Edison Electric, then to chief executive of Chicago Edison, and finally to chairman of Peoples Gas in Chicago, where he managed a corporate turnaround. This string of successes led to the 1912 founding of Middle West Utilities and later to Insull Utility Investments, both holding companies for electric and gas utilities. By 1926 Insull's utility empire encompassed 6,000 communities across thirty-two states, and by 1930 it had grown to 4 million customers and 12 percent of the nation's electricity-generating and gas-distribution capacity.

The War Industries Board encouraged the formation of nationwide industrial organizations to carry out its mandate to coordinate the nation's industrial activities during World War I. Natural gas suppliers responded by combining several predecessor organizations into the American Gas Association (AGA) in 1918 to centralize the exchange of information, set industry-wide standards, and encourage cooperation and coordination among its members. The AGA also represented the industry viewpoint to the public, at Congressional hearings, and before natural gas regulatory bodies. The complete conversion of natural gas from lighting to cooking and heating took place at this time, symbolized by natural gas being sold in units of energy (British thermal units) rather than units of illumination (candlepower).

Long-Distance Transmission

The discovery of huge natural gas fields in the Southwest, the Panhandle Field in northern Texas in 1918, followed by the Hugoton gas fields in the Kansas, Oklahoma, and Texas border areas in 1922, changed the nature of the gas business. Both fields covered over 1.6 million acres and accounted for much of the nation's reserves. Exploiting oil found in these fields resulted in venting enormous quantities of associated natural gas to the atmosphere. The discovery of natural gas fields in the Southwest occurred just as natural gas fields in Appalachia were beginning to wane. The promise of commercial reward to be gained by substituting "free" Southwest gas for Appalachian gas spurred R&D efforts in long-distance pipeline transmission, which resulted in the development of thin-walled, high tensile strength, large diameter, seamless pipe segments joined together by electric arc welding. Technological improvements were also made in gas compressors for moving large volumes of natural gas at high pressures and in ditch digging and filling machinery for laying pipe. An indication of the progress in gas transmission can be seen in pipeline diameters and design working pressures. In the 1920s and 1930s large transmission pipelines were between twenty and twenty-six inches in diameter and could sustain up to 500 pounds working pressure; in the 1940s, diameters were up to thirty inches and working pressure was up to 800 pounds; by the 1960s up to thirty-six inches in diameter and 1,000 pounds; in the 1970s diameters were up to forty-two inches and 1,260 pounds, and after 1980 pipelines were fifty-six inches in diameter and working pressure up to 2,000 pounds.

Manufactured coal gas companies financed many of the first long-distance pipelines built in the 1920s for mixing cheap natural gas into water gas to raise its heat content. The first long-distance pipeline was 250 miles long and made of twenty-inch diameter pipe built by Cities Service in 1927/28 to connect the Panhandle field with Kansas City. This was quickly followed by Standard Oil of New Jersey's 350-mile, twenty-two-inch line from the Texas-New Mexico border to Denver. In 1929, natural gas was pipelined 300 miles from the San Joaquin Valley to San Francisco, the first metropolitan area to switch from manufactured to natural gas. These pipelines carried a considerable amount of gas that had to find a "home," spawning an intense marketing effort to induce consumers to buy gas-powered appliances such as space and water heaters, stoves, and clothes dryers. Individual burners for manufactured gas had to be adjusted to handle the higher heat content of natural gas, no small effort when a city had hundreds of thousands of individual burners. In 1930, the industry accepted a standard and distinctive mercaptan odorant to detect escaping gas.

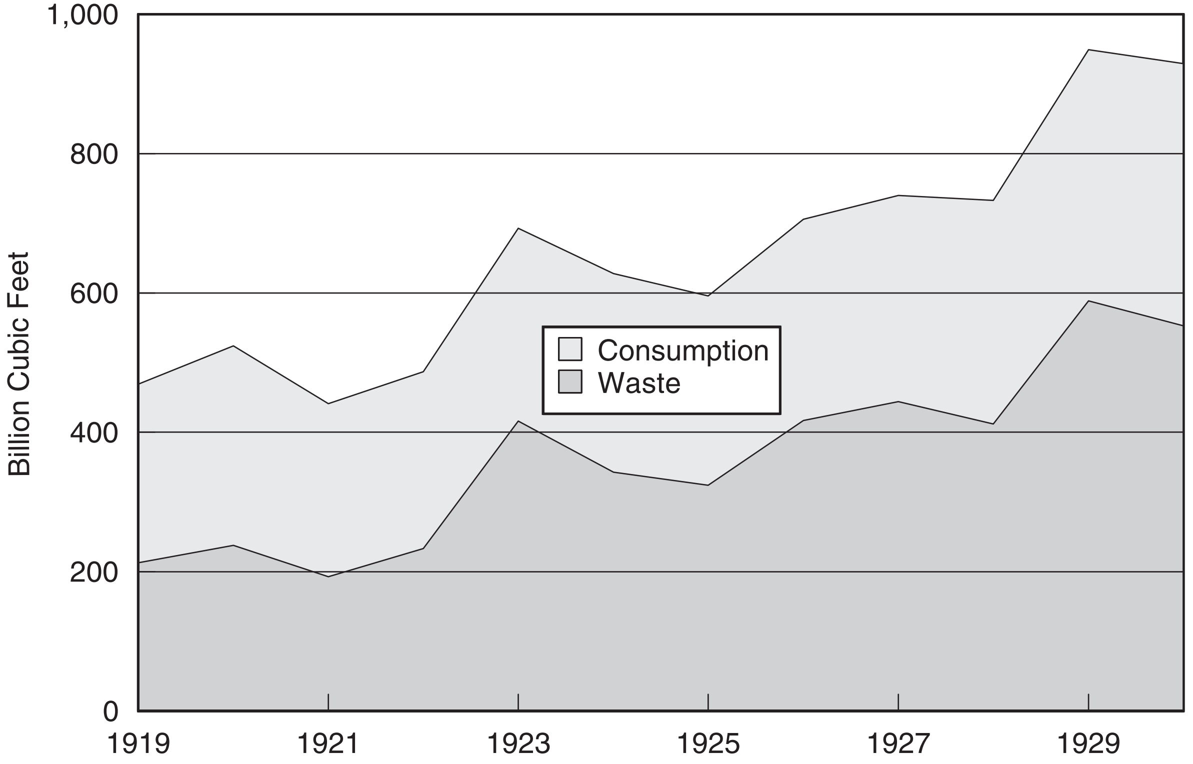
A consortium of companies controlled by Doherty, Insull, and Standard Oil of New Jersey built the first 1,000-mile pipeline (actually 980 miles) from north Texas to Chicago, called the Chicago Line. The twenty-four-inch diameter pipeline, started in 1930 and completed one year later, primarily served Insull's Chicago area utilities. The pipeline transmitted a sufficient volume of attractively priced gas to eventually convert Chicago from manufactured to natural gas. The three principal partners to varying degrees owned and controlled the natural gas fields and the pipeline transmission and distribution systems. A similar arrangement was behind the Northern Natural Gas 1,110-mile, twenty-six-inch pipeline from the Hugoton gas field to Omaha, then continuing on to Minnesota. Essentially the same power trust, but with different shareholdings and minority participants, controlled both pipelines. With no competition between the two pipeline systems in their respective "territories," and the same partners owning the natural gas fields, the pipeline transmission and distribution systems, consumers were at the mercy of a natural monopoly whose operations were far beyond the purview of state regulatory authorities.

The Panhandle Eastern pipeline was intended to introduce competition in the "territories." The power trusts employed various legal shenanigans to stop the building of the pipeline, including putting pressure on financial institutions not to fund the project. While ultimately unsuccessful, the trusts did succeed in drawing public attention to their power to thwart competition. Yet, it was in the public interest to build more long-distance gas pipelines to reduce the wasteful practice of venting natural gas to the atmosphere. As seen in Figure 7.1, venting amounted to over half of natural gas production in the Southwest.

An average of a little under 400 billion cubic feet were vented to the atmosphere per year for the eleven years shown in Figure 7.1; perhaps in no small way contributing to the buildup of methane, a greenhouse gas, in the atmosphere. To understand what this waste means in terms of oil, a cubic foot of natural gas contains 1,026 Btu, whereas a barrel of crude oil contains 5.8 million Btu. One Btu is the amount of energy to increase the temperature of one pound of water by 1°F; it is a small measure of energy equivalent to burning a blue-tip kitchen match. One barrel of oil is energy-equivalent to 5,653 cubic feet of natural gas, or one barrel of oil per day for one year is energy-equivalent to 2.063 million cubic feet of natural gas. Thus, 1 trillion cubic feet of gas per year is energy-equivalent to 484,648 barrels of oil per day or, in round numbers, 0.5 million barrels per day. The waste of, say, 0.4 trillion cubic feet of gas per year vented to the atmosphere was energy-equivalent to throwing away about 200,000 barrels of crude oil per day for a year, at a time when energy consumption was a minuscule fraction of current levels.

This waste was primarily natural gas associated with oil production. A lower oil production rate would result in less waste. Unfortunately, the common-law principal applicable to the ownership of underground hydrocarbons is the rule of capture, which simply states that whoever draws the oil out of the ground owns the oil (assuming vertical wells!). If three individuals or companies own the land and mineral rights over an oil reservoir, and only one drills wells on his or her property and, over time, drains the entire reservoir dry, the other two have no claim on the revenue or profits of the driller. The rule of capture forces everyone whose land lies over an oil reservoir to drill as many wells as possible and to pump as hard as possible to get their "rightful" share.

Figure 7.1 **Consumption versus Waste of Natural Gas**



Under these conditions, natural gas associated with oil production had no way to go but up.

Federal Regulation

The construction of long-distance natural gas pipelines came to an abrupt end with the Great Depression of the 1930s, when only three pipelines of less than 300 miles in length were built. Existing natural gas pipelines and electric utilities operated far below capacity and many could no longer produce enough revenue to support their operating and capital costs. Insull Utility Investments, the darling of the Roaring Twenties, had issued bonds galore to finance the takeover of many of the nation's utilities at stock prices inflated by investor enthusiasm and overvalued assets. The Great Depression wrung the last drop of excess out of the stocks in Insull's empire and the discovery of accounting irregularities hastened its demise in 1932, incurring the wrath of shareholders, bondholders, the Federal Power Commission (FPC), and Franklin Delano Roosevelt.

While criticism of the monopolistic corporate structure of public utilities was voiced before the 1929 stock market crash, the collapse of Insull's house of cards led to a 1935 ninety-six-volume FPC report assailing the electricity and natural gas industries. With regard to natural gas, the report focused on the waste of venting natural gas to the atmosphere, the monopolistic control by the same corporate entities over natural gas production, transmission, and distribution, and reckless financial manipulation. The report pointed out the practice common among natural gas and electricity-generating companies of inflating book values to increase the rate base, and their unrelenting pressure to influence state regulators to get higher rates and protect their "territories." The report noted only four companies, one being Standard Oil of New Jersey, which controlled 60 percent of natural gas pipelines in terms of mileage and a like amount of the nation's natural gas production. With natural gas in interstate trade passing through thirty-four states serving 7 million customers, the report concluded that it was high time for the public utility industry to come under public scrutiny.

The passage of the Public Utility Holding Act of 1935 dismantled the pyramid utility holding companies. The public utility business would henceforth be organized into single, locally managed entities providing electricity and gas and serving a specific area. Although electricity and gas operations were still allowed under one corporate umbrella, they were housed in separate organizations. The Act essentially put the utility business back to where it was before Doherty and Insull started building their pyramids.

In 1937 an explosion in a Texas high school from natural gas leaking into the basement killed nearly 300 students and teachers. Although it was industry practice to put an odorant in natural gas to detect accumulation of gas in a closed space, the particular source of natural gas for the high school did not have one. Perched on top of the wreckage was part of a blackboard with the intact inscription: "Oil and natural gas are East Texas' greatest mineral blessings. Without them this school would not be here, and none of us would be here learning our lessons." This tragedy hastened the passage of the Natural Gas Act of 1938, which empowered the FPC to regulate interstate natural gas pipelines. Viewed as an extension of its 1930 mandate to regulate the interstate transmission of electricity, the Act gave the FPC authority to approve natural gas pipeline rates based on a just and reasonable return, to require regulated pipeline companies to submit extensive documentation on operational and financial matters, and to order actions to be taken on a variety of matters if deemed necessary.

No interstate pipeline could be constructed without a Federal Certificate of Public Convenience and Necessity. To obtain this certificate, the promoting company or companies had to have a twenty or so year contract for an adequate supply of natural gas, must demonstrate a sound and proven ability to attract the necessary financing, and propose a rate based on a just and reasonable return on its investment. Moreover, the granting of the certificate had to take into consideration the impact of the proposed pipeline on other natural gas suppliers serving the intended market. The wording of this last point would prove to be a regulatory sticking point for certifying natural gas pipelines intended for areas without existing natural gas services and would serve as a way to obstruct the building of natural gas pipelines.

Before the passage of the Act, the owning corporations decided who had access to the pipeline and what rates would be charged. After the passage of the Act, interstate pipelines would be common carriers charging the same FPC-approved rates to all users without restrictions to access. The certificate, once granted, was a franchise for a natural gas pipeline company to exclusively serve a specified market. Thus, the concept of the franchise survived, but who determined the franchise had changed. Before the passage of the Act, corporations set up an exclusive "territory" where only the members of the cartel had access to the market. After the passage of the Act, the regulators would determine the "territory" where a pipeline transmission company had an exclusive franchise. The franchise was protected from competition by government fiat rather than corporate connivance. Having a chief executive with an "in" with the FPC regulators could be a natural gas pipeline's greatest asset.

Curiously, the Act did not grant the right of eminent domain to pipeline companies for building natural gas pipelines. Coal-carrying railroads, coal-mining companies, and manufactured gas producers would be able to obstruct the building of natural gas pipelines by taking advantage of the lack of a right of eminent domain and the requirement to evaluate the impact of a pipeline on other natural gas providers serving an area that had none. Standard Oil of New Jersey's response to this massive government intrusion into the natural gas business was to distribute shares in its natural gas and pipeline subsidiaries to its stockholders.

The War Years

World War II significantly boosted demand for natural gas to fuel factories and armament plants. This stimulated the building of new pipelines to tap Southwest gas to replace declining production in Appalachia. Despite wartime needs, a proposal to build a natural gas pipeline from the Southwest to New York was thwarted by coal and railroad interests and manufactured gas producers. As an alternative, a newly organized company, Tennessee Gas and Transmission, made a proposal to build a pipeline from Louisiana to Tennessee. The application failed because the company could not demonstrate how the pipeline would affect other natural gas suppliers in an area that had none! Attempts to correct this legislative imbroglio were bitterly opposed by railroads, the coal industry, and manufactured coal gas producers.

When reason finally prevailed and appropriate legislative changes were made to modify this obstructive requirement, the railroads, coal mining, and manufactured gas companies were granted the right to intervene directly in natural gas pipeline certification hearings. Having lost one means to obstruct the building of natural gas pipelines, they were awarded another. The saga for granting the Certificate of Public Convenience and Necessity to build the Tennessee Gas pipeline involved political intrigue right on up to the office of the president along with major changes in the source of natural gas, changes in the financing, changes in the corporate structure of the company, and changes in its ownership. Appeals were made to the War Production Board to support the pipeline application in order to alleviate the growing shortage of natural gas in Appalachia, which was critical to fueling wartime industries. Perhaps the exigency of war industrial enterprises running out of natural gas or perhaps the rare triumph of reason over vested interests finally prevailed, and the FPC granted the required certification for the building this twenty-four-inch pipeline.

Opening Up the Northeast

Manufactured gas companies, along with the coal-mining industry and the coal-carrying railroads, were intent on maintaining their last market bastion in the Northeast, the population center of the United States. All they had to do was obstruct the building of any natural gas pipeline. The Tennessee Gas pipeline extended the market reach of Southwest gas to Appalachia, but not to the Northeast.

During World War II tankers transported oil from the U.S. Gulf to the Northeast as the first step in shipping oil to Europe in support of the war effort. The eastern seaboard cities refused to turn off their lights at night, and the silhouetted tankers passing against their skylines became easy targets for U-Boats. As a countermeasure, the U.S. government built two oil pipelines, the Big Inch (twenty-four-inch) and the Little Inch (twenty-inch), from Texas to Pennsylvania and New Jersey. After the war, the oil companies reverted to tanker shipments and the pipelines, now under the Surplus Property Administration, were put up for sale. Congressmen representing the coal-producing states joined the railroads, coal companies, and manufactured gas producers to prevent the sale of these pipelines to a natural gas pipeline company. Proponents of natural gas pointed to the prospect of the continuing waste of venting natural gas to the atmosphere in the Southwest if these oil pipelines were not converted to natural gas. Opponents pointed to the millions of tons of coal that would be displaced to the detriment of the coal-mining industry and the transporting railroads if these oil pipelines were converted.

The winning bid at the first auction was Tennessee Gas, which obtained a one-year lease to pipeline natural gas to Ohio, with the promise that it would not to try to move gas into the Northeast. This converted the oil pipelines to gas transmission. At the expiration of the lease in 1947, a second auction was held and another fledging company, Texas Eastern, won with a bid of $143 million, about equal to the government's cost of construction. Later that year, a bill granting eminent domain to the natural gas pipeline industry quietly slipped through Congress. Philadelphia was the first northeastern city to convert to natural gas in 1948, and within a decade its manufactured gas industry was gone. The FPC then approved the building of the Transcontinental pipeline from Texas to New York City. Begun in 1949, the thirty-inch, 1,000-mile pipeline, including the "Costliest Inch" connection under Manhattan to five receiving gas utilities, was completed in 1951.

The conversion of New York City and Long Island from manufactured to natural gas made New England the last bastion of manufactured gas in the United States. Progress, once underway, was hard to stop. A two-year dispute between Texas Eastern and Tennessee Gas before the FPC ended up with Texas Eastern being the supplier of Algonquin pipeline to provide service to Connecticut, Rhode Island, and eastern Massachusetts, including Boston. A subsidiary of Tennessee Gas was given permission to pipeline gas from a connection in Buffalo through upstate New York to western Massachusetts. Natural gas flowing into New England in 1953 marked the end of a century of domination by manufactured gas producers.

Last Stop Before Total Regulation

The Natural Gas Act of 1938 split the natural gas business into three entities: the unregulated natural gas producers, interstate pipelines regulated by the FPC, and local distribution companies (LDCs) regulated by state or municipal authorities. In the 1940s Panhandle Eastern attempted to sell gas directly to a Detroit utility, bypassing the LDC. The irate LDC decided not to object, but to build a competitive pipeline to cut out Panhandle Eastern. Now it was Panhandle Eastern's turn to object to the granting of a certificate for building a competing pipeline. Despite legal maneuvering, Panhandle Eastern lost the case. The alternative pipeline had contracted with Phillips Petroleum for its entire gas supply. The pipeline fell behind schedule in 1950, forcing a renegotiation of the contract. Phillips exercised its monopolistic hold over the pipeline's gas supply by hiking the price by 60 percent, raising the shackles of those in Congress representing the pipeline's consumers. The upshot was the "Phillips Decision" by the U.S. Supreme Court that brought natural gas producers selling gas to interstate pipelines under regulatory control.

Now the FPC had jurisdiction not only over 157 interstate gas pipeline companies, but also over thousands of independent producers. While arriving at a just and reasonable tariff for pipelines was a regulatory possibility (because pipeline projects were relatively few in number with well-documented costs) now the FPC had to deal with thousands of independent producers who owned hundreds of thousands of individual wells, each with its unique cost structure. What is a fair and reasonable price for natural gas coming from a deep well, drilled at great expense with high operating costs, and drawing gas from a reservoir with a short life versus the fair and reasonable price for gas coming from a shallow well with low operating costs, and tapping a field that will last for decades? Pricing gas on a well-by-well basis proved impossible and the FPC resorted to the "fair field" method to regulate natural gas prices in interstate trade, which was challenged and overturned in a 1955 court decision. In 1960 the FPC, facing a monumental backlog of applications for price increases by independent natural gas producers selling gas to interstate pipelines, decided to establish common prices for five geographic areas. In 1965, five years later, the FPC published its first area price; another five years were to pass before the second area price was published.

The never-ending regulatory wrangling prevented price increases for natural gas sold to interstate pipelines. The availability of cheap surplus gas in the Southwest encouraged the building of interstate pipelines. California began to suffer from a decline in local supplies of natural gas and interstate pipelines were built to its border for transfer to intrastate pipelines (California prohibits the building of interstate pipelines within its borders). Another large natural gas market was tapped by building an interstate pipeline to Florida. The new gas pipelines absorbed excess production in the Southwest, eliminating the wasteful practice of venting, but the continued building of interstate pipelines began to outstrip supply. Ignoring the lessons of Economics 101, the FPC set a regulated price for natural gas in interstate commerce at a level that encouraged consumption but discouraged investment in developing new gas fields. With a price set too low, proven reserves started to decline after 1970 and an impending natural gas shortage was looming when the oil crisis struck in 1973.

Unraveling Natural Gas Regulation

The 1973 oil crisis sent oil prices spiraling, which, in turn, caused the prices of all forms of energy to rise sharply, with the exception of regulated interstate natural gas. While prices for regulated interstate gas remained unchanged, prices for unregulated intrastate natural gas jumped. Natural gas producers sold all they could to intrastate pipelines and reduced the volume sold to interstate pipelines to the absolute minimum contractual amounts. Natural gas shortages amounting to 1 trillion cubic feet in the Midwest and Northeast started in 1973 and continued to worsen. They became particularly severe during a long and cold winter in 1976–1977, reaching a peak shortfall of 3.8 trillion cubic feet with New York State declaring a state of emergency.

The FPC was at its wit's end trying to cope with the deteriorating situation, with its primary focus more on how to prevent an owner of an existing regulated well from enjoying a windfall profit than solving the natural gas shortage. Cracks in the regulatory facade started to occur in 1975 when the FPC allowed LDCs to make direct contract with producers to buy gas at higher than regulated prices. Interstate pipeline owners opposed this scheme because it threatened their merchant status, which allowed them to be the sole buyers and sellers of natural gas carried by their pipelines. The interstate pipeline owners initiated an advance payment program to natural gas producers that was essentially a five-year interest-free loan to encourage exploration, but this was abrogated by the FPC. Finally, the FPC abandoned area pricing and set a nationwide price for regulated gas 50 percent below intrastate gas prices. When this failed to stimulate exploration, the FPC announced a price menu that authorized higher prices for "new" gas and lower prices for "old" gas.

The Iranian Revolution of 1979–1980 worsened the natural gas situation with another spike in oil prices. The Synthetic Fuels Corporation, formed under Carter's Energy Security Act of 1978, was to make America energy independent. With vast deposits of coal, the most promising solution was a return to manufactured, or synthetic, gas made from coal. The Fischer-Tropsch process employed in Europe and South Africa produced nearly pure methane from coal without the adverse environmental consequences of the bygone era of manufactured gas. All but one of the proposed synfuel projects failed for a variety of environmental, commercial, and legal reasons. The one that succeeded came onstream in 1984, with four interstate pipelines companies forced to buy its output of synthetic gas made from coal at a resounding $6.75 per million Btu, nearly fifteen times the price of wellhead "old" gas when the project was conceived.

The second response to increasing natural gas supplies without increasing the price of old gas was to exploit the huge natural gas reserves found along with oil at Prudhoe Bay in Alaska. Three proposals were offered to tap this gas. One was a 2,500-mile, forty-eight-inch pipeline from Prudhoe Bay, along the coast of the Beaufort Sea, into Canada to the Mackenzie River Delta, also a source of natural gas. The pipeline would then proceed south along the Mackenzie River to pipeline connections in Alberta for delivery of the natural gas to customers in the Midwest and California. Environmentalists successfully opposed this project on the basis of potential damage to the Arctic National Wildlife Refuge. Although this project failed, two pipelines built in anticipation of its approval now carry Alberta gas to the Midwest and California.

A second proposal was to build a parallel natural gas pipeline to the existing crude oil pipeline that carried Prudhoe Bay oil to Valdez for shipment in tankers to U.S. markets. The natural gas would be liquefied at Valdez and shipped in LNG carriers to California. President Carter did not favor this project and preferred a third proposal: building a 1,600-mile pipeline from Prudhoe Bay parallel to the crude oil pipeline and then south parallel to the Alcan Highway through Canada, reentering the United States in Montana. Financing problems plagued the project until it was shelved in 1982.

The third response to increase natural gas supply was LNG imports. LNG production already existed in the United States since 1969, with Phillips Petroleum and Marathon Oil exporting LNG from Cook Inlet in Alaska to Japan. Anticipating natural gas shortages, Distrigas, a subsidiary of Cabot Corporation at the time and now part of the Suez LNG built a terminal near Boston to receive LNG from Algeria to satisfy peak winter demand. An early proposal to import LNG from the Soviet Union failed for political reasons. In 1969, before the oil crisis and also anticipating natural gas shortages, El Paso Natural Gas organized a major LNG project with Sonatrach, the Algerian state-owned oil and gas company. Sonatrach would receive an export price for LNG that reflected the then low value for natural gas in the United States. El Paso built a fleet of nine LNG carriers and contracted with three east coast gas companies to receive the gas at specially built LNG terminals along the eastern seaboard. Panhandle Eastern entered into a similar agreement for Algerian LNG deliveries at a terminal built at Lake Charles, Louisiana, to feed Trunkline, a pipeline subsidiary that was running low on natural gas. (If a pipeline is running out of gas supplies and LNG can be delivered to either end of the pipeline, why have LNG fill the pipeline?)

The El Paso project survived only the first shipments when Sonatrach unilaterally repriced LNG to fit market realities, giving El Paso the opportunity to walk away from the deal. The Trunkline failure took a bit longer because its deal with Sonatrach had tied the price of LNG to fuel oil. Whereas El Paso walked away from the project when Sonatrach repriced LNG, here Trunkline customers did the walking. The price they had to pay was a blended price of relatively cheap interstate gas and expensive LNG. Faced with opportunities to buy lower-priced gas from other sources, one Trunkline customer after another took their business elsewhere. As its customer base eroded, the rising portion of expensive Algerian gas in Trunkline's pipeline increased the price of the blended gas, encouraging others to walk. Between 1982–1984 customer defection halved pipeline volume, forcing Trunkline to walk away from the Sonatrach contract. Sonatrach successfully sued Panhandle Eastern for partial restitution.

All in all, natural gas regulation was a bitter lesson for the regulators, who seemed to forget the lessons of Economics 101. By pricing gas too low starting in the late 1960s, the regulators encouraged growth in demand while simultaneously discouraging growth in supply. Their reaction to the energy crisis—proposing esoteric solutions such as synfuels, developing isolated reserves, and using LNG rather than raising the price of old gas—simply exacerbated the situation.

The Road to Deregulation

The 1976 Carter campaign for the presidency pledged the "moral equivalent of war" on the ongoing energy crisis. Following through on his campaign pledge, Carter created the Department of Energy as part of the National Energy Act of 1978, which was preceded by Congress reorganizing the harassed Federal Power Commission as the Federal Energy Regulatory Commission (FERC). The Natural Gas Policy Act of 1978 permitted FERC to organize natural gas that came into production in and after 1978 into nine pricing categories with subcategories depending on well depth, source, and other factors, until 1985 when all post-1978 gas would be deregulated. Old gas in production prior to 1978 would be indefinitely regulated with three price tiers. The significantly higher-priced new gas flowing into the system had "unexpected" Economics 101 consequences: it provided an incentive for consumers to reduce demand by switching to other sources of energy and taking steps to conserve energy, at the same time providing an incentive for producers to expand supply. New gas prices dropped when controls were lifted in 1985 because the gas shortage of too little supply chasing too much demand had been transformed to a gas "bubble" of too much supply chasing too little demand. The word *bubble* was intentionally used to describe what FERC thought would be a transient state of oversupply, but “transient” was a situation that lasted for nearly two decades. Natural gas prices fell to the point where synthetic gas, Alaskan gas, and LNG were far from being economically viable. The fall in natural gas prices as a result of letting the market do its magic (higher prices spurring exploration to expand supply and conservation and energy-switching to dampen demand) made it easy for Congress to pass the Natural Gas Wellhead Decontrol Act of 1989, thereby abrogating price controls on all wellhead gas.

Unfortunately, the pipeline companies had arranged for twenty-year, back-to-back, take-or-pay, fixed-price contracts between electric utilities and natural gas suppliers at the prevailing high natural gas prices of the late 1970s. The growing presence of lower-priced gas during the first half of the 1980s placed a great deal of pressure on utilities to break their high-priced gas contracts. FERC bent to their demands and, through various rulings, allowed utilities to walk away from these contracts and buy natural gas directly from producers at lower prices and pay a fee to the pipeline companies for transmission services. This started the transformation of natural gas transmission companies from merchants to transporters, the first step in breaking a monopoly in which a consumer had no choice but to buy from the transmission company.

Once the utilities were allowed to break their contracts with pipeline companies, the pipeline companies were stuck with the other side of the take-or-pay contracts to buy natural gas at high prices. Faced with multi-billion dollar liabilities, in 1987 FERC issued Order 500, the first step toward breaking take-or-pay contracts with natural gas producers, allowing pipeline companies to set up a system of pipeline transmission credits against producers’ take-or-pay claims. While not an entirely satisfactory resolution of the matter, Order 500 turned out to be the precursor to a series of orders that ended up with Order 636 in 1992, which mandated the final solution to the national gas regulatory problem: deregulation.

To its credit, regulation either fostered or, at least, did not prevent the building of hundreds of thousands of miles of interstate gas pipelines linking natural gas producers to consumers throughout the nation. Under the sanction of the FPC, natural gas pipeline merchants acquired natural gas, transmitted the gas through their pipelines, and sold the gas in their respective franchised territories. The pipeline merchants generally made more money buying and selling gas than transmitting it. This system ended with Order 636. The whole natural gas system was unbundled into three distinct activities: natural gas producers, pipeline transmitters, and the end-use buyers, either LDCs or major consumers such as utilities or industrial plants.

Order 636 restricted the service offered by interstate pipeline companies to transmission of natural gas at a regulated rate that provided a just and reasonable return on their investments. LDCs and major consumers would be responsible for arranging their own supplies by contracting with natural gas producers to cover their needs. This created a marketing opportunity for companies to acquire natural gas production or represent the interest of natural gas producers and sell to end users. The pipeline transmission carriers were reduced to contract carriers and were obligated to set up Electronic Bulletin Boards in order for buyers and sellers to keep track of gas flows and pipeline allocations.

In 1983, 95 percent of all interstate commerce gas was purchased by the pipeline companies from natural gas producers and sold to LDCs and major consumers. By 1987 the pipeline companies arranged for the buying and selling of less than half of the gas going through their pipelines, and by

1994, they were fully converted to common carriers. Thus ended the world of the pipeline companies in which they dealt with producers for a supply of gas and then marketed it to LDCs and large end users. Now the pipeline company was no longer a gas merchant but a common contract carrier like a railroad. Marketing firms were the intermediaries between natural gas producers and LDCs and large end users. However, pipeline carriers were free to set up marketing organizations to compete with independent marketers in arranging and brokering deals between natural gas buyers and sellers. By 1993 the marketing arms of pipeline companies had more than a 40 percent share of the market, but they had to operate independently of the pipeline transmission organization; any collusion between the two, or less than arm's-length transactions, were subject to heavy fines.

This process of unbundling of services is not complete. LDCs buy natural gas from producers and pay a toll to an interstate natural gas pipeline as a common carrier at a rate determined to provide a reasonable return and supply the gas to their customers. This principal can be expanded to include customers of the LDCs, in theory down to individual households, who can arrange for their natural gas needs directly with producers, pay the interstate pipeline company one toll for the use of its transmission system, and then pay another toll to the LDC for the use of its distribution system at a rate that represents a fair return.

From Source to Consumer

Natural gas comes from the well as a mixture of hydrocarbons. For instance, Southwest natural gas has average proportions of 88 percent methane, 5 percent ethane, 2 percent propane, 1 percent butane, plus heavier hydrocarbons and impurities. A methane molecule is one carbon atom and four hydrogen atoms; ethane is two carbon atoms and six hydrogen atoms; propane has three carbon atoms and eight hydrogen; butane, four carbon and ten hydrogen. Heavier hydrocarbons of pentane (with five carbon atoms and twelve hydrogen atoms), hexane (with five carbon atoms and fourteen hydrogen), and heptane (seven carbon atoms and sixteen hydrogen) are in a gaseous state when in a natural gas reservoir, but "fall out," or condense, to a liquid when brought to the surface. Condensates are separated from natural gas and sold separately to refiners. Ethane is fairly expensive to separate with its low liquefying temperature, and normally remains in the natural gas stream. Propane and butane are more easily separated by fractionation, or a cooling of the natural gas, and are sold separately as liquefied petroleum gases (LPG). Impurities such as hydrogen sulfide, carbon dioxide, nitrogen, and water have to be removed. "Pipeline-quality" natural gas is primarily methane and some ethane, cleaned of impurities and stripped of condensates and petroleum gas liquids, with a heat content 1,000–1,050 Btus per cubic foot at standard atmospheric conditions.[2]

The cleaning and stripping functions are performed in the pipeline-gathering system connecting the producing wells with transmission pipelines; the last step is raising the pressure of natural gas to transmission system pressure. Transmission pipelines typically are twenty-four- to thirty-six-inch diameter operating between 600 and 1,200 psi pressure, although there are wider diameter pipelines operating at higher pressures. Compressor stations are located about every seventy miles and the speed of the gas varies between fifteen and thirty miles per hour, depending on the gas pressure, compressor capacity, and pipeline diameter. Monitoring devices and shutoff values are strategically placed about every five to twenty miles to deal with potential pipeline ruptures and routine pipe maintenance. Both the inner and outer pipe surfaces are coated to protect against corrosion. The inner surface is kept clean by running a "pig" through the pipeline for routine maintenance; a "smart pig" can also transmit data on the internal condition of the pipeline. There are also routine maintenance inspections of the external condition of pipelines to detect leaks and other potential problems.

Storage facilities are available along a pipeline. In the Gulf region, natural gas is stored under pressure in salt caverns where the salt has been leached out. In other areas of the nation, abandoned or played-out natural gas reservoirs are used for storage. Natural gas is reinjected into these reservoirs under pressure during times of weak demand and lower prices to be withdrawn during times of strong demand and higher prices. Before deregulation, natural gas was stored during the summer and drawn out during the winter. Since deregulation, natural gas in storage is recycled several times a year in response to changing prices, not necessarily related to times of peak demand. The 300,000 miles of transmission pipelines have an inherent storage capacity that can be increased by increasing the gas pressure. Metering is a vital operation as gas enters and leaves the gas transmission system and associated storage facilities for both accurate paying of suppliers, charging of customers, and system control.

Gas planning for a transmission company depends on long- and short-term forecasting models. Long-term forecasts determine investments to ensure that the pipeline system can meet future demand. Short-term forecasts ensure that the volume of gas can accommodate current demand. Nominations are made one to two days in advance to ensure that enough gas enters the system upstream to match demand downstream without exceeding pipeline capacity. Scheduling acceptances of gas from thousands of suppliers and deliveries of gas to thousands of consumers, each with their specific needs and different contractual arrangements, is both complex and crucial to the smooth operation of the system. A system of allocation cuts based on a previously arranged and agreed-on priority ranking is activated when nominations exceed the limits of pipeline capacity.

In a deregulated climate, natural gas can be drawn from storage or obtained directly from natural gas producers or from other interstate transmission companies via hubs. Hubs connect various interstate pipelines in a common system whereby natural gas consumers connected into one interstate pipeline can obtain supplies from natural gas suppliers hooked into other interstate pipelines. There are about a dozen major hubs in the United States, the most important being Henry Hub in Louisiana, where a multitude of pipelines connect natural gas suppliers and interstate pipelines into what amounts to one huge common system. Hubs not only allow for common distribution, but also common pricing. The Henry Hub is the most important distribution and pricing hub in the United States; its price is the base price for the nation. Other hubs are generally priced at the Henry Hub price plus transmission costs with local market-related variations. As an example, three interstate pipelines serve New York City. Through local interconnections, major gas purchasers can bargain for natural gas from suppliers in three different Southwest natural gas regions. These continual negotiations for the best price create a pricing hub in New York City, where price does not vary much regardless of which transmission company is actually supplying the gas. If a gap in price does appear among the transmission companies, then consumers' continual quest for the cheapest source of natural gas tends to close the gap. Another pricing hub is in Boston, where gas from two interstate pipelines from the Southwest and two pipelines from the western and eastern provinces of Canada are interconnected, providing major customers with the opportunity to buy gas for the best price from suppliers in four different natural gas-producing regions. The search by large consumers for the best price ends up with a more or less common price in Boston, regardless of the source. Natural gas prices in New York and Boston are not the same. The New York basis price reflects the price in Henry Hub plus transmission costs. The presence of Canadian producers in the Boston market affects the basis price of Boston gas and its relationship with the New York basis price and the price in Henry Hub. It is possible for Canadian gas to penetrate the New York market by reversing the Algonquin pipeline between New York and Boston. Deregulation has introduced a dynamic in the natural gas business that was lacking under regulation.

The simple days when gas producers sold to interstate pipeline companies, which then sold to LDCs selling to residential, commercial, industrial, and electricity-generating customers at a regulated price based on costs are gone. Natural gas suppliers are no longer regulated. Pipeline transmission companies have lost their status as natural monopolies. No longer merchants, they have become common carriers with regulated long-term transmission rates. Life is now more complex for customers because they must examine many options such as buying direct from gas producers via independent marketing firms, the marketing arms of transmission companies, from storage providers, or from other gas transmission pipelines and their natural gas suppliers via hubs.

While natural gas consumers are doing their best to buy at the lowest cost, natural gas producers are doing their best to sell at the highest price. Producers look at the basis price at every hub that they have access to and net the basis price of the transmission cost to obtain the netback value for their gas and then sell to the pricing hub with the highest netback value. The enormous number of individual transactions among buyers seeking the lowest delivered cost and sellers seeking the highest netback price, in a market where no individual dominates and where buy-and-sell transactions are simplified, transferable, and transparent, leads to commoditization. Natural gas prices are set by market conditions that reflect supply and demand, not cost plus pricing, which is determined by regulators. The regulators' role is to ensure that no one tries to manipulate the price, reduce the transparency of transactions, or in any way attempts to control the market. In a commodity market, where the providers are all selling at the same approximate price, differentiation among natural gas providers becomes one of value-added services. A buyer selects a provider based not only on price, but also on reliability, dependability, and behind-the-meter services. Behind-the-meter services can include maintenance and repair of natural gas equipment owned by the buyer or advice on how to utilize natural gas more efficiently. In the future, companies may provide a bundled service in which gas is combined with other utility services such as electricity and water, or even communication, to woo customers away from competitors.

Deregulation has not made life easy for transmission companies because increasing revenue means attracting more customers by, perhaps, cutting rates on spare capacity, reducing operating expenses and capital commitments, and providing value-added services behind the meter. Nor is life easier for the consumer. With natural gas prices fluctuating widely and long-term contracts becoming less available, there is a greater exposure to adverse price fluctuations in the spot and short-term markets. Independent gas marketers can also be at great risk if their buy-and-sell commitments do not match up either in volume or time duration, which exposes them to the potential of huge financial losses from an adverse change in natural gas prices.

The risk of adverse price changes generates a need for risk mitigation among natural gas suppliers, consumers, and marketers. Banks and other financial institutions provide an active over-the-counter market for swaps tailored to meet the particular needs of suppliers and consumers. A swap protects a supplier from a low price and consumers from a high price that can threaten their financial well-being. NYMEX is the nation's leading center for buying and selling natural gas futures contracts. In addition to traditional hedging, the public trading of natural gas futures contracts opens up the opportunity for individuals (including hedgers turned gamblers) to speculate on the future price of natural gas, which is why the futures volume far exceeds the physical volume. Speculators add depth to the market by accepting the risk that hedgers are trying to shed. However, not all risks can be mitigated by financial instruments such as volume risk (a customer uses more or less than the nominated amount), counterparty risk (one of the parties to a transaction does not honor its commitment), execution risk (a transaction is not properly concluded), regulatory risk, (the possibility of a change in the rules after a transaction has been signed), operational risk (system failure), and basis risk (the possibility of the price at a hub such as Boston moving differently than expected from the price used to hedge a risk using the price at Henry Hub).

The demarcation point between transmission and distribution is the citygate where regulators reduce the gas pressure, scrubbers and filters remove any remaining traces of contaminants and water vapor, and mercaptan is added as an odorant to detect gas leaks. The nation's million-mile distribution system is made up of two- to twenty-four-inch pipe with pressures from 60 psi down to one-quarter psi (above atmospheric) for natural gas entering a home or business, and higher for industrial and electricity-generating customers. While distribution pipe has traditionally been made of steel, plastic or PVC is now used for its flexibility, resistance to corrosion, and lower cost.

To ensure an adequate supply of natural gas, LDCs contract with the transmission companies for pipeline volume capacity that meets peak demand. The rate charged to LDCs is the regulated rate that ensures a fair and reasonable return on the interstate pipeline investment. However, during times of less than peak demand, LDCs are free to sell their spare capacity to third parties. While these rates are generally less than what the LDCs are paying, the revenue so earned reduces LDC transmission costs. This creates a market for interstate pipeline capacity that generally disappears during times of peak demand. But if a LDC finds itself in a position with spare capacity during times of peak demand, it can sell this spare capacity at either the market rate or a maximum rate imposed by government regulations.

In addition to the marketing of spare transmission capacity, another opportunity has opened up to address nomination imbalances. Major customers of transmission pipelines (LDCs, utilities, industrial plants) must address an imbalance between their nominated and actual usage of more than 10 percent either way or face a monetary penalty. Rather than face a monetary penalty, customers can contact a marketing outfit that specializes in finding other users with the opposite imbalance. Swapping imbalances avoids having to pay a penalty.

The unbundling of services has created a plethora of marketing opportunities and commercial dealings that natural gas has commodized, natural gas transmission and storage capacity, nomination imbalances, and risk mitigation. Yet, there is still a great deal of regulation in the natural gas industry. FERC establishes services to be provided by interstate pipelines, determines rates based on a fair and reasonable return, and approves construction of new interstate pipelines. LDCs are regulated by the states and, in some cases, municipalities, where no two state or municipal regulators issue the same set of regulations for conducting business, determining rate structures, approving construction of new facilities, and addressing complaints by users.

Rates charged by LDCs are based on covering operating and capital costs including the purchase of natural gas. Consistent with interstate transmission companies, a fair and reasonable return for LDCs is based on determining the appropriate rate of return that induces shareholders to invest in plant and equipment consistent with the inherent risk of the business and the opportunity to invest elsewhere. A balancing account keeps track of the difference between required and actual revenue, which eventually leads to upward or downward rate adjustments. Rates also reflect customer categories to take into account the peculiarities associated with each. The capital and service requirements for hundreds of thousands of residential and commercial customers in a local distribution system are quite different than those for a few industrial and utility customers.

Various states are experimenting with unbundling LDC services, but progress is slow. It is possible that regulation may not be deregulated but transformed to incentive regulation, which gives the regulated LDC an opportunity to profit from exceptional performance. One form of incentive regulation is performance-based regulation in which the LDC's cost of procuring gas for residential customers is compared to an index value for other LDCs. If the cost is lower than the index value, then the LDC shareholders and ratepayers benefit by splitting the savings. If higher, the incremental cost is again split and the shareholders suffer along with ratepayers. Incentive regulation can take the form of benchmarking with adjustments for both inflation and productivity gains. If productivity gains exceed the impact of inflation, then both shareholders and ratepayers share the benefit; if not, both suffer. Another alternative is rate caps, a method whereby shareholders either suffer or benefit from actual rates being above or below the cap. Incentive regulation provides an opportunity for LDCs to enhance their profitability by becoming more astute in gas purchases and more eager to pursue productivity gains.

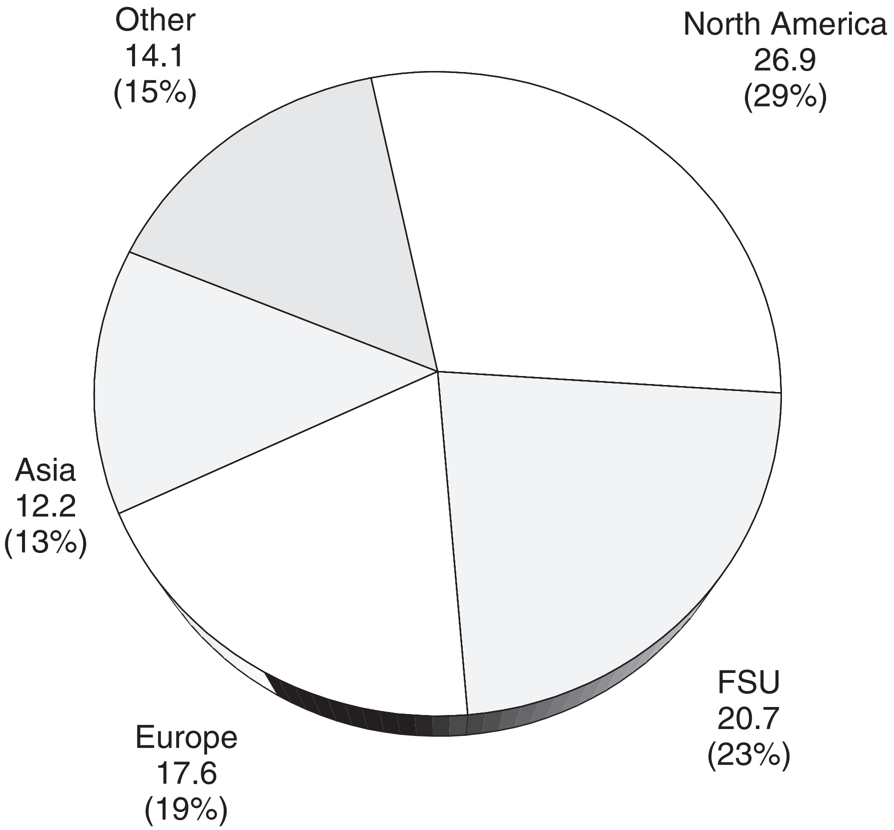
Natural Gas as a Fuel

The principal consumers of natural gas in Figure 7.2 are ranked in terms of trillion cubic feet (tcf) per year.[3] Consumption of 1 tcf per year is roughly equivalent to 0.5 million barrels of oil per day. North America and Russia account for half of world consumption. Whereas the United States is becoming an importer of natural gas, Russia has long been a major exporter and is planning to further expand its exports.

Under communism, natural gas in the Soviet Union was free, obviating the need for meters and thermostats. Torridly hot rooms were cooled by a blast of frigid Arctic air through an opened window. Since the fall of communism, meters and thermostats have been installed and natural gas usage is now charged to stop this wasteful practice, freeing up supplies for hard-currency exports.

The categories of use for any energy source are transportation, residential, commercial, industrial, and electricity generation. Natural gas as a vehicle fuel is extremely clean-burning emitting a small fraction of the precursors to ozone formation

Figure 7.2 **World Natural Gas Consumption in Tcf**



(organic gases and nitrous oxides) and carbon monoxide compared to oil-based motor fuels. Moreover, there are no particulates (soot), virtually no sulfur oxides, and less carbon dioxide emissions. The most difficult hurdle for natural gas to overcome to be an acceptable motor vehicle fuel is the logistics of refueling. For this reason, very little natural gas is actually consumed in transportation with most compressed natural gas-fueled vehicles owned by natural gas pipeline and distribution companies.

Using a rounded figure of 20 tcf of annual consumption in the United States net of natural gas consumed by pipeline compressors, 5 tcf, or 25 percent of total demand, is consumed by 61 million residential customers for space and water heating, cooking, clothes drying, pool heating, and gas fireplaces. Residential demand peaks during winter months, accounting for as much as 70 percent of annual gas consumption, which of itself can vary greatly between cold and mild winters. Residential customers pay the highest rates for natural gas because they support a distribution system connected to each individual home capable of handling peak winter needs. Moreover, residential customers consume the greatest volume of natural gas when natural gas prices are high, which, of course, are high because of increased residential demand (the chicken-and-egg syndrome). Five million commercial customers use about 3 tcf, or 15 percent of total demand, paying the second highest price for natural gas. The commercial sector consists of restaurants, motels and hotels, retail establishments, hospitals and healthcare facilities, and office and government buildings. While their natural gas usage is similar to that of residential customers, there is less of a swing in demand between summer and winter from space cooling. In addition to the weather, consumption in the commercial sector is sensitive to general business activity.

The industrial sector is the largest consumer at 7 tcf, or 35 percent of total demand, with 210,000 customers. Natural gas supplies the energy needs of a host of manufacturing industries plus waste treatment, incineration, drying and dehumidification processes, and is a feedstock for the fertilizer, chemical, and petrochemical industries. Natural gas prices directly affect food prices because natural gas is a raw material for making ammonia-based nitrogen fertilizers and is used for drying crops, pumping irrigation water, and processing food. While seasonal variations are less than those of residential and commercial users, the industrial sector experiences significant fluctuations in demand from changes in economic activity. Moreover, one-third of industrial users can switch to propane and fuel oil if natural gas prices get out of line. The industrial sector ranks third in what is paid for natural gas, while plants that generate electricity pay the lowest price. This sector consists of only 5,700 customers that consume 5 tcf of natural gas, or 25 percent of total demand. Electricity generation was the fastest growing market segment, with approximately 5,000 natural gas-fueled electricity-generating plants, not counting the cogeneration units run by commercial and industrial customers, until cheap natural gas went away with the natural gas bubble in 2000. Natural gas usage in electricity generation is affected by seasonal factors, changes in economic activity, and the relative cost of electricity generated from natural gas compared to coal, oil, and nuclear and hydropower. Some dual-fueled electricity-generating plants can easily switch between natural gas and fuel oil, depending on their relative costs.

Natural gas in electricity generation employs a variety of technologies including conventional steam generators, combustion turbines, and combined-cycle plants. Natural gas burned in a steam generator is similar in context to burning coal or oil to produce steam that passes through a turbine that drives an electricity generator. Efficiency is about 35–40 percent for new plants and about 25–35 percent for older plants, regardless of the fuel. Most of the remaining energy passes to the environment as the latent heat of vaporization (the heat consumed for water to change to steam). Combustion turbines are basically modified jet engines attached to turbines to generate electricity for peak shaving. Peak shaving occurs for relatively short periods of time such as air conditioning demand during a heat wave. Combustion turbines are best for peak shaving because their low capital cost minimizes a utility's investment in equipment that is run only for short periods of time. However, combustion turbines have a high operating cost since a large portion of the energy input is released to the environment as turbine exhaust.

A combined-cycle plant directs the escaping exhaust gases from a combustion turbine through a steam generator to drive an electricity-generating turbine. A combined-cycle plant can increase thermal efficiency up to 50 percent, higher than that of an oil- or coal-fired steam-generating plant, by the inclusion of a combustion turbine.

A combined-cycle plant has lower capital costs than coal or nuclear plants, shorter construction times, greater operational flexibility, and is the preferred choice for smaller capacity plants. With a higher fuel cost, natural gas plants of various generating capacities are built to meet the variable needs of electricity-generating utilities, leaving base load generation to large coal, nuclear, and hydro plants.

Many companies need large quantities of hot water and have historically purchased electricity for both power and for heating water. These companies are increasingly installing cogeneration plants that do not necessarily have to be fueled by natural gas. When fueled by natural gas, a cogeneration plant is a combined-cycle plant with both a combustion turbine and a steam turbine to produce electricity, with none of the line losses associated with off-site electricity generation. Moreover, water containing the latent heat of vaporization from the steam turbine condenser can be substituted for the hot water that had previously been heated with electricity. Being able to utilize the "free" hot water raises the overall efficiency of cogenerating plants to 60 percent or higher.

On a global scale, the role of natural gas in transportation is negligible, with 40 percent apportioned for residential and commercial consumers, the same as in the United States, 25 percent for industrial (versus 35 percent in the United States), and 35 percent for electricity generation (versus 25 percent in the United States). For global electricity generation, the natural gas share as an energy source grew from 13 percent in 1971 to 19 percent in 2002 and is expected to continue growing to 27 percent in 2020.[4] In the United States, nearly all electricity-generating plants built in the 1990s and the early 2000s ran on natural gas. With the passing of the natural gas bubble, utilities are beginning to look at other fuels for electricity generation, particularly coal.

The European Version of Deregulation

Europe never had a regulatory regime similar to that of the United States. European nations carried out their respective energy policies through "championed" energy companies. Championed energy companies had the support of their respective governments to dominate a nation's electricity or natural gas business. Governments exercised control over these companies by either having seats on the board of directors and/or approving the appointment of top executives. This comfortable relationship resulted in European governments being assured of a secure supply delivered in a dependable and reliable manner, in championed companies that operated profitably in a secure business environment, and in consumers who paid a high price for energy.

The first European leader to react against high-priced energy was Prime Minister Thatcher, who cut subsidies to coal companies, privatized national energy companies, and started the process of major consumers having direct access to energy providers to negotiate their supplies. The ability to choose suppliers introduced competition to what had previously been a natural monopoly. Thus, the two paths taken by the United States and the United Kingdom were basically parallel and arrived at the same destination. For the United States, the path was deregulation of a regulated industry; for the United Kingdom, the path was cutting the umbilical cord to subsidized energy companies, privatizing previously nationalized companies, and giving major consumers third-party access to natural gas and electricity-transmission systems in order to be able to select suppliers. Both paths led to the introduction of a competitive marketplace where buyers can negotiate with suppliers to lower electricity and natural gas costs.

While the unbundling of the UK natural gas and electricity markets in the late 1990s is very similar to what happened in the United States, it stands apart from Europe. European nations have been reluctant to liberalize their energy markets and give buyers the ability to negotiate with suppliers. While high-cost energy was widely recognized as a deterrent to European economic growth in the mid-1980s, the first EU directives for liberalizing electricity and natural gas did not appear until the latter half of the 1990s. These directives established a time frame for specified percentages of natural gas and electricity that had to be satisfied in a competitive marketplace. In 2003, another EU directive was issued to accelerate the process of liberalization. Independent transmission and distribution system operators were to be created to separate services formerly provided by integrated transmission and distribution companies, with a target year of 2007 for the unbundling of the gas and electricity markets.

With regard to natural gas, the objectives were to give major consumers access to gas-transmission networks with the ability to negotiate firm or interruptible, short- or long-term service contracts not only with gas suppliers, but also with operators of gas storage facilities and LNG terminals. Natural gas charges were to reflect actual costs, thereby avoiding cross-subsidies, and capacity allocation was to be transparent and nondiscriminatory. Interconnecting pipeline hubs and pricing hubs were to be established to give buyers access to various suppliers so they could negotiate price and terms. Yet in the early 2000s, despite some liberalization, competition in Europe is limited in scope with gas prices still linked primarily to oil prices, which thwarts price competition, and transactions being more opaque than transparent. Existing long-term take-or-pay contracts between customers and integrated energy companies impede the pace of unbundling, as do high transmission costs. Liberalization is not heartily endorsed by national governments; in particular, Germany, the powerhouse of Europe, sees no need to unbundle the services of its championed integrated energy companies. These companies have worked well in the past, providing security of supply, which to the German government is more important than cost.

Yet for competition to be effective, the number of natural gas supply sources must be increased, a market for physical and financial trading of natural gas has to be developed, the link between gas and oil pricing has to be severed, new entrants must be permitted in the energy business, and governments have to become more supportive of liberalization and less willing to shield their championed companies from competition. This foot-dragging by European nations is at variance with the EU energy bureaucrats in Brussels. In 2005 the EU warned a number of nations within the European Union that they will be brought before the European Court of Justice and face stiff fines unless they open up their energy markets.

Spain and Italy have taken steps to adopt an infrastructure of freedom of choice for consumers, with less progress being made in Austria, Ireland, Sweden, Belgium, and the Netherlands, and still less in Denmark, France, and Germany. A lack of uniformity of approach to the problem of dealing with supply, transmission, storage, and LNG terminals in a common natural gas pipeline grid among the various nations of Europe is the focus of the Gas Transmission Europe, an association of forty-five European companies in thirty countries. This organization not only deals with proposals concerning the hardware of interconnections between pipelines, storage facilities, and LNG terminals for improved network access, but also in the software of internal controls, gas flow and transaction information systems, nomination procedures, and other operational matters.[5]

Even with unbundling, natural gas prices may remain high if the primary sources of natural gas (the North Sea, the Netherlands, Russia, and Algeria) form a common front against lower prices. Natural gas buyers with access to LNG terminals may be able to break this common front were it to occur. LNG suppliers in Latin America, West Africa, and the Middle East are less controllable than traditional European gas suppliers. For instance, they may be tempted to get rid of excess production by selling low-priced cargoes into Europe. With third-party access to LNG terminals, a surfeit of LNG cargoes may exert sufficient commercial pressure on the traditional suppliers to break their common front, if one were to exist and create a truly competitive market.

LPG: Prelude to LNG

Liquefied petroleum gases (LPG) are primarily propane and butane. LPG is formed as a by-product of oil refining and is stripped out of a natural gas stream by a fractionating unit. In the 1910s light-end gases from a barrel of crude were kept in a liquid state under pressure and fueled early automobiles, then blowtorches for metal cutting, and, in 1927, gas stoves. Butane was one of the propulsion fuels for dirigibles until the market for butane-fueled dirigibles crashed in 1937 with the *Hindenburg*. An entrepreneur began selling unwanted bottles of butane as fuel for gas stoves in Brazil. All went well until he ran out of dirigible fuel. To replace the butane, he began to import cylinders of pressurized butane on the decks of cargo liners, thus marking the humble beginning of the international trade of LPG.[6]

A 1927 court decision made the fractionating process available to industry, which opened the door for the development of the LPG business. With an increased availability of supply came the opportunity to develop a new market. The industry grew slowly, with bobtail trucks delivering pressurized propane in a liquid state to refill cylinders that fueled stoves, water heaters, clothes dryers, and space heaters in rural areas and towns not served by natural gas pipelines. LPG was also used for crop drying. Propane was a preferred fuel for bakeries and glass-making facilities and other commercial and industrial enterprises that required a greater degree of control over temperature and flame characteristics. Cleaner-burning propane became a motor fuel of choice for forklift trucks and other vehicles that operated in semiclosed environments such as terminals and warehouses. Butane eventually found a home as fuel for cigarette lighters and taxicabs. LPG became a gasoline blending stock and a feedstock for steam crackers to make ethylene, the precursor to plastics, and other petrochemicals. Railroad tank cars were the primary means of moving LPG over long distances until they were replaced by pipelines in the late 1960s.

Wells drilled into a salt dome at Mont Belvieu in Texas leached out the salt to form a cavern to store LPG. This would eventually become the nation's principal storage hub and central marketplace for LPG, with extensive gas liquids pipeline connections to major suppliers and buyers in the Gulf Coast, Midwest, and Northeast. Another storage hub in Kansas served the upper Midwest via pipeline. LPG was carried in pressurized tanks mounted on vessels and barges along the eastern seaboard and the Mississippi River. In 1971, President Nixon put price controls on oil, which happened to include LPG. Not surprisingly, this encouraged the consumption of LPG because of its lower cost compared to other fuels.

In the United States, refinery-produced LPG is generally consumed internally for gasoline blending or pipelined as feedstock to an associated petrochemical plant. The commercial market for LPG was primarily supplied by stripping gas liquids from natural gas. In contrast, LPG development in Europe was based on refinery operations and rail-car imports from the Soviet Union because there were no indigenous supplies of natural gas until the late 1970s, when the North Sea natural gas fields came onstream. LPG consumption was primarily propane in the north and butane in the south because propane vaporized more easily in warm climates. In the 1970s, Italy promoted the use of butane as a motor vehicle fuel. Small shipments in pressurized tanks installed on vessels carried LPG from refineries in Rotterdam to destinations in northern Europe and from refineries in Italy, Libya, and Algeria to southern Europe and the eastern Mediterranean.

In Japan the LPG market grew out of the desires of Japanese housewives and restaurant owners to cook with propane rather than kerosene. Switching from kerosene to propane was a sign of a rising standard of living. Japan also encouraged the use of butane-fueled taxicabs. Unlike the United States and Europe, Japan had to import much of its LPG aside from that produced as a by-product in domestic refineries. Shipping LPG at sea was costly because LPG was carried in cargo tanks built to withstand the pressure necessary to keep LPG liquid at ambient temperatures. The weight of the steel in the cargo tanks was about the same as the weight of the cargo, restricting the cargo-carrying capacity of the vessel.

To counter high shipping costs, Japanese shipyards developed and began building fully refrigerated LPG carriers in the 1960s. The temperature of the cargo was reduced to keep LPG liquid at atmospheric pressure (−43°C for propane and −1°C for butane). The cargo tanks had to withstand a lower-temperature cargo and had to be insulated to minimize heat transfer from the outside environment. An onboard cargo refrigeration unit kept the cargo at the requisite temperature to prevent pressure buildup in the cargo tanks. Fully refrigerated cargoes made it possible to use a simpler design for the cargo tanks, which could be built to conform to the shape of the hull, because they did not have to satisfy the structural requirements of a pressurized cargo. This allowed an order of magnitude increase in the carrying capacity from several thousand cubic meters to 30,000–50,000 cubic meters. Parenthetically, with a specific gravity of about 0.6, a cubic meter of LPG weighs about 0.6 metric tons, compared to a cubic meter of crude oil, which weighs close to 1 metric ton. A fully loaded LPG carrier transports less cargo weight-wise than a crude carrier of an equivalent cargo volume. The first large-sized LPG carriers were employed shipping LPG between Kuwait and Japan. By 1970 the Japanese LPG carrier fleet numbered a dozen vessels, with the largest being 72,000 cubic meters.

The United States was primed for large-scale imports with adequate storage at Mont Belvieu, inland pipeline connections (the Little Inch was converted from natural gas to gas liquids), and LPG terminals, originally built for export, in the U.S. Gulf, with importing terminals in the Northeast. All that was missing was an LPG shortage, the appearance of a major new export source, and a means of transport. All the missing elements fell into place following the oil crisis of 1973. Shortages in natural gas stemming from the consequences of government price regulations of natural gas in interstate commerce reduced the domestic supply of LPG. Fractionating plants built in the Middle East and Europe to strip out gas liquids from natural gas greatly increased the overseas supply. Transport was available as more shipyards began building large-sized, fully refrigerated LPG carriers.

While all the elements fell in place for the United States to become a major LPG importer, large-scale imports never quite got off the ground. The appearance of new supplies of natural gas, along with deregulation of "new" gas, increased the domestic availability of LPG. LPG demand slackened when oil (and LPG) price controls were partially lifted, and finally dismantled, by President Reagan in 1981. Increased availability, coupled with a decline in demand, reduced the need for large-scale imports. Without the U.S. import market developing to any significant degree, the enormous capacity of new LPG export plants in Saudi Arabia and elsewhere in the Middle East and in Europe created a glut. There is nothing like a glut to present an opportunity for entrepreneurs to develop a market, as had already happened with the glut created by discoveries of huge reserves of natural gas in the U.S. Southwest.

The international price for LPG swung between a premium or a discount from the price of crude oil on an energy-equivalent basis and thus was more volatile than oil. These price swings reflected the success or failure of entrepreneurs in finding a home for the new supplies of LPG. Those who bought a cargo of Middle East LPG and loaded it on a ship without a firm commitment from someone to buy the cargo when delivered in the United States or Europe were at the mercy of a fickle market while the vessel was at sea. Millions of dollars could be made or lost during a single voyage, depending on whether the buyer was on the right or wrong side of a price swing. Some of the founding firms instrumental in developing the international LPG market were merged or liquidated when they eventually found themselves on the wrong side. The same thing happened to independent LPG carrier owners when more vessels were delivered from shipyards than there were cargoes to fill the vessels. Lining up long-term deals between suppliers and buyers, along with the ships to carry the cargoes, would drastically reduce the degree of commercial risk; but long-term deals were not always available, and, when they were available, they were not always to the liking of either the buyer or the seller.

While Western firms were enjoying a financial bonanza or going bust, the Japanese LPG players just kept rolling along in a secure business environment, the result of how business is conducted in Japan. Of course, steady growth in propane consumption as a substitute for kerosene and as a substitute for naphtha for steam crackers to produce petrochemicals would provide an element of stability anywhere. But the Japanese are particularly adept at calming the financial waters. In the case of LPG, they developed a fully integrated logistics supply chain consisting of long-term contracts arranged with Middle East exporters, vessels to move the cargoes, terminals to unload the vessels, storage facilities to store the cargoes, and cylinder bottles to distribute LPG to consumers. The Ministry of International Trade and Industry (MITI) coordinated activities with the cooperation of an industry made up of a relatively few participants who respected each other's "territories." MITI was also in a position to dictate the amount of LPG to be consumed by the petrochemical industry, which smoothed out any bumps and wrinkles in the supply chain. LPG carriers received a "regulated" rate to cover costs and ensure an adequate return on vested funds over the life of the vessel. The modest return on investment reflected little risk of unemployment for vessels built to serve a single trade for the duration of their physical lives. This investment philosophy was shared by the other elements of the supply chain. The price of LPG sold in Japan took into consideration the cost of acquiring the LPG, and the capital and operating costs of transmission (by ship, not pipeline), storage, and local distribution. The Japanese people, accustomed to paying a high price for energy, did not object to this arrangement. There was no political advantage for a Japanese government body that guided an industry to curry the favor of the electorate by underpricing a fuel. Government guidance of energy policy in Japan proved to be superior to the regulatory experience in the United States, where energy policies seem to be a series of "fits and starts" that eventually have to be scrapped. Managing LPG imports on a systems, or supply chain, basis would turn out to be the prelude to LNG imports.

The history of LPG consumption is a series of developing markets that started at a point in time and reached maturity at another. The U.S. market began in earnest in 1950 and reached maturity around 1975; for Europe the growth stage spanned 1960 to around 1980, for Japan from 1965–1985, and for Korea from 1980 to the late 1990s. In the early 2000s, China entered the growth phase of a new LPG market, with India slated to be next. While the rate of growth in aggregate LPG consumption is somewhat constant, its center of activity travels around the world as one market begins to be developed and another matures. Thus, what appears to be a stable business growing at a modest rate to outsiders is, in reality, a continual opening up of new opportunities for insiders including entrepreneurs, marketers, traders, and suppliers. The LPG business, like so much of the energy business, is a challenge for those who like to be on the cusp of change where money can be made by correctly assessing its twists and turns.

In 2000 the United States was the world's leading LPG consumer at 51 million tons annually and was essentially self-sufficient, importing and exporting only about 1 million tons annually. The second-largest consumer was Europe, at 31 million tons, exporting 7 million tons from the North Sea to the Mediterranean and Brazil and importing 15 million tons from Algeria, Venezuela, and the Middle East. While simultaneous importing and exporting may not make immediate sense, LPG is made of two distinct products, propane and butane, each of which can be long or short on a regional basis. Furthermore, it may pay to export from one location and import into another, rather than ship directly between the two, to take advantage of price disparities. The third-largest consumer was Japan, at 19 million tons, of which 15 million tons were imported primarily from the Middle East and the remainder was produced in domestic refineries. Saudi Arabia was by far the world's largest LPG-exporting nation, followed by Algeria, Abu Dhabi, Kuwait, and Norway.

In the early 2000s the fastest growing importers were China, followed by Korea, both having negligible consumption in 1980. Korea consumed about 7 million tons in 2000 and China over 13 million tons, each importing about 5 million tons mainly from the Middle East. While Korea is reaching maturity, China is far from maturity and India is just beginning to move into the growth stage. In a way, Korea, China, and India mimic Japan. Burning propane for cooking is a status symbol and indicates a rising standard of living. In Japan, propane displaced kerosene, while in Korea propane displaced charcoal briquettes, and in China propane is displacing coal and biofuels (charcoal, wood, and agricultural waste such as straw and animal dung). The next market to be developed is India, where coal and biofuels also dominate home cooking. Substituting LPG for coal and biofuels is a big step toward a cleaner environment because it does not produce air particulate (smoke), carbon monoxide, nitrous oxides, and carbon dioxide emissions, and, in the case of coal, sulfur oxides and metal (arsenic and mercury) emissions. However, a high price for oil becomes a high price for LPG. For millions of the world's poor, high-priced LPG means continued cooking with coal and biofuels, ingesting the pollution along with the food.

Natural gas liquids in international trade in the early 2000s required an LPG carrier fleet of a hundred large carriers over 60,000 cubic meters (most between 70,000–80,000 cubic meters) and another seventy vessels between 40,000–60,000 cubic meters. There are also more than sixty semi-refrigerated LPG carriers between 10,000–20,000 cubic meters in size, used more for local distribution than long-haul trading, in which the cargoes are cooled, but not enough for the gas to remain liquid at atmospheric pressure. LPG carriers can also carry liquid cargoes of ammonia, butadiene, and vinyl chloride monomer (VCM). Another twenty ethylene carriers carry liquefied ethylene at a much lower temperature than LPG, but not low enough to carry LNG. While ethylene carriers can carry LPG as backup when no ethylene cargoes are available, they make their real money carrying ethylene cargoes that cannot be carried by LPG carriers.

Liquefied Natural Gas

As mentioned repeatedly in this text, natural gas is constrained by logistics. The development of long-distance pipeline transmission was crucial to natural gas becoming a commercial energy resource. Pipelines are fixed installations connecting a specific set of suppliers with a specific set of consumers and are an inflexible mode of transmission. Most pipelines are within a single nation because political considerations enter the picture when pipelines cross national borders. A proposed pipeline from Iran to India that would cross Pakistan was, for many years, considered impossible because of the rivalry and bitter feelings between Pakistan and India. However, in 2005 progress was made in advancing this proposal in response to Pakistan's own need for energy and the potential earnings from transit fees. A pipeline connection between Turkey and Greece, long-time bitter foes, is under construction and scheduled to be completed in 2006. The pipeline link is part of a larger system to supply Caspian natural gas to the southern Europe natural gas pipeline grid. Another example of long-time foes cementing better relationships through energy is a pipeline project to supply Egyptian gas to Israel. Among allies, a healthy international trade already exists in the form of natural gas pipelines crossing national borders. Table 7.1 lists the largest international pipeline movements of natural gas.[7]

Figure 7.3 shows the locations of the largest reserves of national gas for total world reserves of 176 trillion cubic meters (tcm). With world consumption at 2.6 trillion cubic meters in 2003, the reserve-to-consumption ratio is nearly sixty-eight years, double that of oil.

Reserves can be misleading. For instance, proven reserves in Alaska incorporated in Figure 7.2 exclude potentially vast gas fields in the North Slope that have not been sufficiently assessed to classify them as proven. Nevertheless, with the exception of Russia and the United States, much of the natural gas reserves would be stranded if pipelines were the only means of transmission. Construction of long-distance undersea pipelines to connect remote fields in Iran, Qatar, Nigeria, Venezuela, Indonesia, and Malaysia with industrially developed nations, with pumping stations every 50–100 miles, is prohibitively expensive. Even Australia's natural gas fields in the northwest part of the nation are too remote from the principal cities in the southeast for pipeline transport. The natural gas reserves for these nations

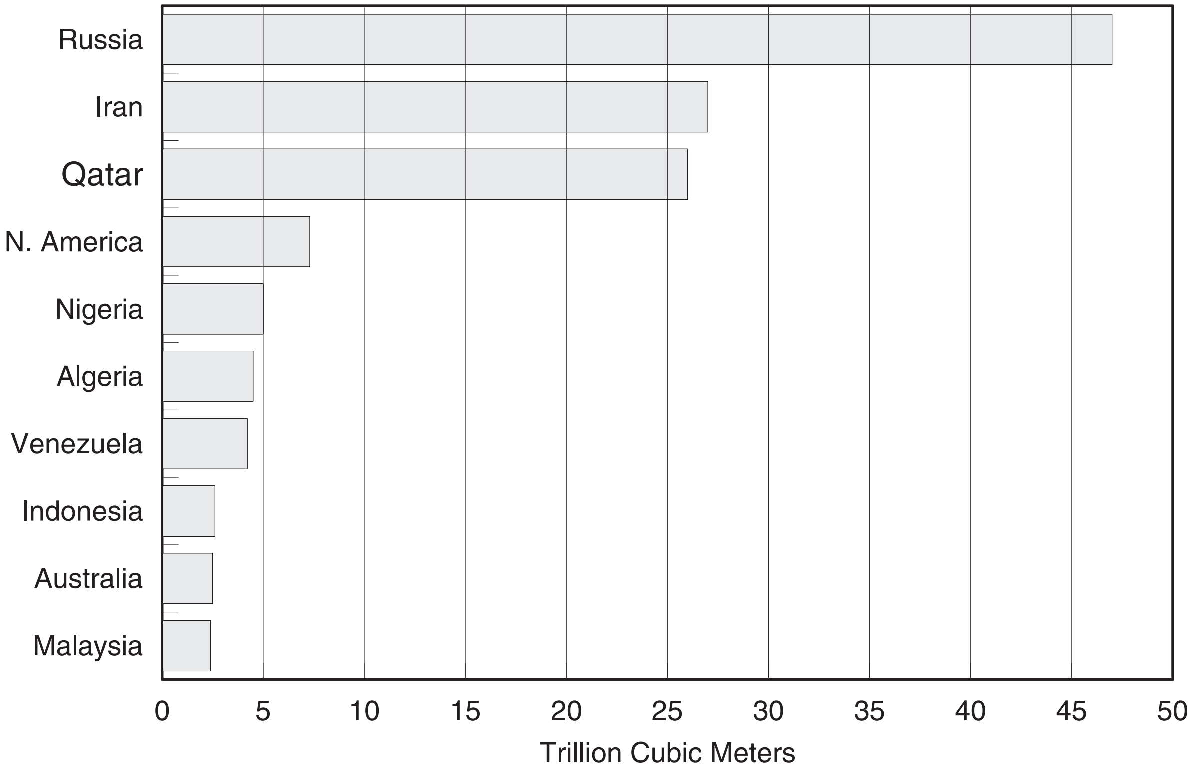
Table 7.1

International Natural Gas Pipeline Movements (bcm)

Importer	Domestic Consumption	Pipeline Imports	Supplier	Percent of Importer's Consumption
Europe*	498	132	Russia	27
United States**	630	91	Canada	14
Europe	498	31	Algeria***	6

* Europe, excluding the Russian Federation
** Net of U.S. exports to Canada
*** Via trans-Mediterranean pipelines to Italy and Spain

Figure 7.3 **World's Largest Natural Gas Reserves**



remained stranded with no commercial value until a new means of transmission was devised.

Compressed natural gas (2,000–4,000psi) can be transported in specially built tanks. The problem is the cost of building large-capacity cargo tanks that can withstand this magnitude of pressure with a cargo still four times greater in volume than in a liquefied state. However, there are special circumstances in which compressed natural gas carriers are feasible such as small natural gas fields in remote areas of the Amazon River, where reserves are not sufficient to justify building a long-distance pipeline or a liquefaction plant.

Just as liquefied gas liquids (propane and butane) are refrigerated for transport as a liquid at ambient pressure, so too can natural gas. As a liquid, natural gas takes up 600 times less volume than at ambient conditions with a specific gravity a little less than LPG. The problem is that natural gas is a liquid at atmospheric pressure at a much colder temperature of –161°C (–258°F). This imposes severe constraints on tank design and insulation to prevent the cargo from coming in contact with the hull. The conventional steel in ship hulls, if exposed to the cold temperature of LNG, is subject to instantaneous cracking, known as brittle fracture.

A much greater technological challenge in tank design and insulation than LPG carriers had to be faced before natural gas could be transported as a liquid. The success of independent research efforts in the 1950s led to the first LNG delivery in 1964 from a liquefaction plant in Algeria to a regasification terminal on an island in the Thames River. From this time forward, Algeria would remain a major force in the LNG business, expanding its export capacity in 1973, 1978, 1980, and 1981. Small-scale LNG plants were built to export LNG from Alaska (Cook Inlet) to Japan in 1969 and from Libya to Europe in 1970. Brunei was the first large-scale LNG export project to serve Japan, starting in 1972 and followed by other large-scale LNG export projects in Indonesia and Abu Dhabi in 1977, Malaysia in 1983 (a second in 1994), Australia in 1989, Qatar in 1997 (a second in 1999), Trinidad and Nigeria in 1999, Oman in 2000, and Egypt in 2005. The relative importance of LNG-importing nations can be ranked by the number of receiving terminals, with Japan in first place with twenty-six, Europe with thirteen, Korea with six, the United States with four, Taiwan with two, and Puerto Rico with one.[8]

In the wake of the energy crisis in the 1970s, Japan adopted an energy policy of diversifying its energy sources to reduce its dependence on Middle East crude oil. The first generation of large-scale LNG projects were long-term contractual arrangements of twenty or more years for the entire output of a liquefaction plant dedicated to a small group of Japanese utilities. LNG carriers were assigned to a project for their entire serviceable lives. As such, the first LNG export projects were as inflexible as long-distance pipelines and were organized similarly to LPG projects, as totally integrated supply chains.

The price of LNG sold in Japan was based on the delivered cost of crude oil. Low crude oil prices during the latter part of the 1980s and much of the 1990s kept a lid on LNG prices and, consequently, on the value of stranded gas. New LNG projects were few and far between until the passing of the natural gas bubble in the United States. This was the dawn of a new day for LNG projects because it became clear that the United States would become a major LNG importer, spurring new LNG projects in Egypt, Qatar, Nigeria, Oman, and Trinidad. But this time building new or expanding existing LNG export plants was not in response to an energy policy to diversify energy sources, as in Japan and later Korea and Taiwan, but to the commercial opportunities associated with the prospect of large-scale LNG imports into the United States.

The LNG business is unique in several aspects. One is the sheer size of the investment. Unlike oil, coal, and other commodity businesses that start small and become large through accretion, an LNG project starts out as a large multi-billion-dollar project. LNG projects are rivaled in size, complexity, and capital requirements only by nuclear power plants. But unlike a nuclear power plant, which feeds into a local electricity grid, an LNG project involves two sovereign powers—the nation with stranded gas reserves and the nation in need of natural gas. While an LNG project is like a long-distance pipeline, which requires that suppliers and consumers be lined up before the pipeline can be built, an LNG Sales and Purchase Agreement (SPA) is more akin to a commercial agreement between two sovereign powers. One is the nation with the gas supply, whose interests are pursued by its national energy company, and the other is the nation with an energy policy that calls for greater consumption of natural gas, whose interests are pursued by its receiving utilities. In both cases, a sovereign power has made a policy decision to either export or import LNG and has delegated oversight to a national energy company or the receiving utilities.

The SPA establishes the commercial link between the buyer (the receiving utilities) and the seller (the national energy company), laying the foundation for the financial structure of the project. Laying the foundation for the physical structure is the engineering, procurement, and construction (EPC) contract. The EPC contract selects a consortium of companies with the requisite skill sets in project management and technical expertise to design the plant, procure the necessary equipment, build a liquefaction plant in a rather remote part of the world, and place it in operation. Shipping contracts have to be arranged, with the delivery of ships timed to the startup and the step-ups in liquefaction plant output. It can take as long as four years for a multi-train liquefaction plant to reach its full capacity. For Japan, and later Korea and Taiwan, it was not a simple matter of building a receiving terminal with sufficient storage capacity and berths to offload the LNG carriers, along with a regasification plant to convert LNG back to a gas. These nations had to create a market for natural gas. The first customers were electricity-generating plants located near the receiving facilities. Eventually an entire natural gas pipeline distribution infrastructure, replete with customers, had to be organized, designed, and built for natural gas to become an important contributor to a nation's energy supply. Getting approvals for the requisite permits to build a natural gas pipeline grid would have been impossible if the government had not endorsed the LNG project.

The LNG supply chain consists of three major segments. The first segment is the upstream end of natural gas fields with their wells and gathering system. A gas-treatment facility removes undesirable elements (nitrogen, carbon dioxide, hydrogen sulfide, sulfur, and water) and separates gas liquids and condensates from the natural gas stream. These, along with any sulfur, are sold to third parties to provide additional revenue. A pipeline delivers the treated natural gas from the gas fields to the second segment, the downstream end, which consists of the liquefaction plant and the LNG carriers. After the last remaining contaminants are removed, a mixed refrigeration process cools methane to its liquefaction temperature using various refrigerants, starting with propane and switching to butane, pentane, ethane, methane, and finally nitrogen. Terminal storage capacity at the liquefaction plant is about two shiploads of cargo plus berthing facilities and a sufficiently sized fleet of LNG carriers to ship the desired throughput from the loading to the receiving terminals. The third segment is the market end of the supply chain, which is made up of the receiving and storage facilities and the regasification plant at the importing nation to warm and feed LNG into a natural gas pipeline distribution system. The receiving facilities must have sufficient storage for unloading a vessel plus extra storage to ensure sufficient quantities during transient and seasonal fluctuations in demand and delays in vessel arrivals. The regasification plant must be connected to a natural gas distribution pipeline system with sufficient customers to consume the LNG.

Organizing and Financing the LNG Supply Chain

The three segments of the LNG supply chain can have different organizational structures. The simplest is to have the same participants throughout the supply chain. This “seamless” structure avoids the need for negotiating transfer price and sales conditions as natural gas or LNG passes through each segment of the chain. But this form of organization can lead to management by committee in which representatives of each segment of the supply chain vote on critical matters for a particular segment. This can have undesirable repercussions if the representatives are not well versed in the technical aspects of each segment. Moreover, funding of the project may be in jeopardy if a participant in one segment does not desire or does not have sufficiently deep pockets to fund its share of the entire project.

The second alternative with regard to ownership is the upstream and downstream segments of the project (natural gas fields and the liquefaction plant) being a separate profit center that sells LNG either free on board (FOB) at the loading terminal or delivered at the receiving terminal, where the price of the LNG includes cargo, insurance, and freight (CIF). The profit for the upstream and downstream portions of an LNG supply chain is the revenue from selling LNG less all operating and capital costs, the acquisition cost of the natural gas, taxes, and royalties. A floor price for the LNG may be incorporated into the SPA to ensure a positive cash flow and a minimum value for the natural gas. The third alternative is using the liquefaction plant as a cost center that simply receives a toll for services rendered that covers its operating and capital costs. These last two alternatives can have different participants with different shares within each segment of the LNG supply chain. Segmented ownership arrangements among the participants can create interface problems in transfer pricing and risk sharing. A conflict resolution mechanism should be established to resolve potentially contentious issues as, or preferably before, they arise. Splitting the ownership of the various supply chain segments has the advantage of having participants who are interested in dedicating their financial and technical resources to a particular segment.

An LNG project can be entirely funded by equity. The return on equity is determined by the cash flow (revenue less operating costs, taxes, royalties, and acquisition costs). The advantage of equity funding is that the participants are not beholden to outside financial institutions. The disadvantage is that the participants must have deep pockets. The dedication of funds to a single multi-billion-dollar project may preclude becoming involved in other LNG projects. At the other extreme, an LNG project can be financed entirely by debt. The debt may be supplied by the sovereign nation that borrows on the basis of its creditworthiness or provides a sovereign guarantee. Revenue is funneled into a special account from which debt service charges are drawn off first and what remains pays for operating costs, royalties, and taxes; whatever is left determines the value of the natural gas.

The proceeds of an LNG project from a government's perspective are its receipts of royalties and taxes and what the national oil company, normally wholly owned by the government, earns on its natural gas sales to the LNG project plus the return on its investment in the project. The split in government payments in the form of royalties and taxes on profits is critical for LNG project participants in the event of a drop in LNG prices. Royalty payments remain fixed and independent of the price of LNG. Taxes on profits, on the other hand, decline as the price of LNG falls. The risk of a negative cash flow can be better dealt with by favoring taxes over royalties. The cost of mitigating this risk is that more money will be paid to the government when LNG prices are high.

A popular form of raising capital is project financing. Here the LNG supply chain is set up to be self-financing, with debt holders looking only to the financial wherewithal of the project itself, not the project sponsors, for interest and debt repayment. Debt issued by an LNG supply chain project is initially limited recourse debt: The sponsors assume full liability for funds advanced only during construction. Once the plant operates at its defined specifications, the debt becomes nonrecourse and the project sponsors assume no liability for debt service obligations; debt repayment relies exclusively on the financial performance of the project.

Project financing is a mix of equity and debt, determined by a cash flow analysis that takes into account the value of crude oil and other determinants on the price for LNG, the operating costs of the liquefaction plant and the upstream natural gas field, the acquisition cost of the natural gas, the LNG carriers (if part of the project), royalties and taxes, and debt-servicing requirements. Project financing exposes the LNG supply chain to the scrutiny of third parties when they exercise due diligence prior to making a commitment. Sponsors and host governments are more likely to agree to an organizational and legal structure imposed by a third party because the benefit is external sourcing of capital. In this sense, project financing has been a healthy influence because it discourages a sponsor from insisting on conditions that would not only be detrimental to others, but would also jeopardize the external funding of the project.

Project financing removes the necessity for the sponsors to have sufficient internal funds to finance the entire project by equity alone. By reducing funding requirements, the sponsors are free to participate in other LNG projects, spreading their risk and expanding their presence in the LNG business. Project financing also allows the importing nation to participate directly in an LNG supply chain by purchasing a meaningful portion of the debt. These benefits of project financing have to be balanced by the costs of satisfying third-party due diligence requirements, managing lender-project relationships, and arranging creditor agreements with various financial institutions.

Underwriters for project financing face the challenge of making an internationally oriented LNG project attractive to prospective buyers of the underlying debt. In packaging the securities, underwriters must deal with the sovereign risk of the host country (e.g., a Middle East nation), a variety of contractual arrangements with several receiving utilities from one or more nations (e.g., Europe), vessel chartering agreements involving a number of legal jurisdictions (e.g., Korea as shipbuilder, London as center of operations, Bermuda as shipowner, Liberia as ship registry), and multiple equity participants incorporated in different nations with unequal shares in various segments of an LNG project. Financial institutions funding LNG projects include pension funds seeking long-term maturities, and commercial banks and private investors interested in short- and medium-term maturities. Another source of debt funding is low-interest credits issued by governments to finance exports.

Depending on the distance between the liquefaction plant and the receiving facility, LNG carriers may account for 25–40 percent of the total investment in an LNG project, the same general magnitude of investment as the liquefaction plant. The remaining investment is primarily the development of the natural gas fields. The regasification system is usually the responsibility of the receiving utility, but for Japan, Korea, and Taiwan, a natural gas pipeline distribution system also had to be built. The development of the natural gas fields, the construction of the liquefaction plant, the building of the ships, and the receiving terminal, including the regasification plant and a natural gas distribution system with a sufficient number of consumers to absorb the LNG imports, have to be coordinated on a fairly tight schedule for all the elements of an LNG project to fall into place in a synchronous fashion. As large and as complex as LNG projects are, a number have been completed and the LNG trade has blossomed. As a point of reference, the volume of the international trade by pipeline of natural gas listed in Table 7.1, which is not all-inclusive, is 254 bcm compared to 169 bcm of LNG in 2003. Figure 7.4 shows the principal sources and destinations of the international trade in LNG.

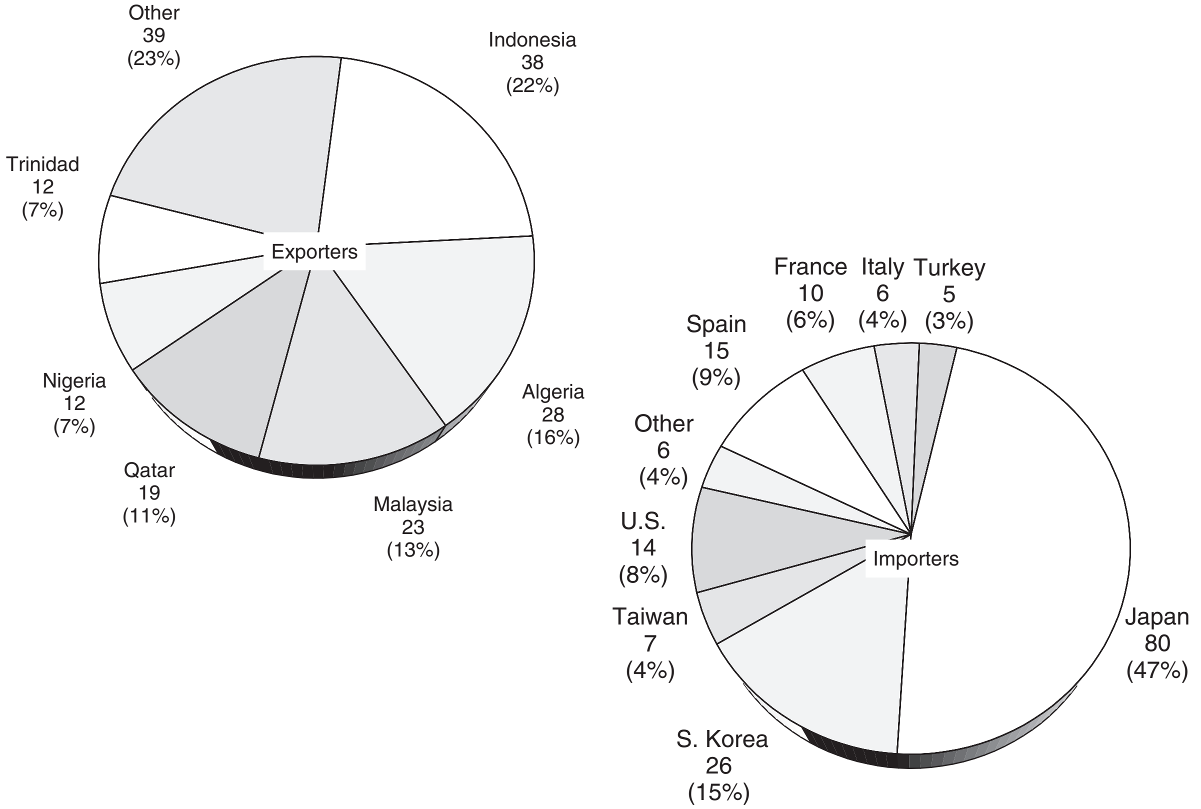
Japan, South Korea, and Taiwan consume two-thirds of the world's LNG and absorb the entire output of LNG export plants in Brunei, Indonesia, Malaysia, and Australia, plus much of the output of LNG export plants in Qatar, Oman, and the UAE (United Arab Emirates). Europe receives 25 percent of the world movement of LNG, most of which comes from Algeria and Nigeria. Small quantities of LNG are exported to the Dominican Republic and Puerto Rico, mainly from Trinidad. The United States receives 8 percent of the world trade, mainly from Trinidad, with additional supplies from Algeria and Nigeria, and a little from the Middle East. The United States is also a minor exporter of LNG from Alaska (Cook Inlet) to Japan.

Despite best intentions, LNG projects, if not organized properly, can fail. The El Paso and Trunkline projects are prime examples, but another occurred in India. The largely completed $3 billion Dabhol project foundered with a change in the local government in 2001. The winners of an election campaigned on a platform that the Enron-sponsored project was rife with political corruption and lacked competitive bidding and transparency. They opposed a local state government contract to purchase nearly all the electricity produced from the imported LNG at a price that consumers could not afford. The project was abandoned as a result of the election. In 2005, international energy companies were negotiating for the possible purchase of the largely completed plant, with the intent to bring it into operation in a manner that would have government and popular support.[9]

LNG Carriers

LNG carriers can be owned by the project for delivered sales where the price at the receiving terminal includes insurance and freight or they can be owned by the buyer for purchase at the loading terminal for free on board sales. Alternatively, the vessels do not have to be owned by either the buyer or seller, but can be chartered from third parties (independent shipowners, energy companies, or financial institutions) under a variety of arrangements. Charters shift the responsibility for raising capital to finance the vessels from the LNG project to the vessel owners.

Figure 7.4 **Sources and Destinations of LNG** (Billions Cubic Feet)



LNG carriers are classified by their containment systems: spherical or membrane. In the spherical containment system, a thick aluminum spherical shell covered by insulation and an outer steel shell is supported by a freestanding skirt that accommodates expansion or contraction of the cargo tank. Propagation of a crack, should any occur, is very slow, with little chance of leakage. While there is no need for a full secondary barrier, a partial barrier prevents any LNG leakage from coming in contact with the hull. The spherical tanks limit sloshing of the cargo when at sea for improved ship stability, but their protrusion above the main deck affects visibility from the bridge. The principal disadvantage of spherical tanks is the inefficient utilization of space within a ship's hull. Spherical tanks are also used for storage at loading and receiving terminals.

The alternative containment system is the membrane design in which the cargo tanks conform with the shape of the ship's hull, increasing a vessel's cargo-carrying capacity. Rather than thick aluminum, the membrane is a thin primary barrier covering insulation installed on the inner hull surface of the ship. This considerably reduces the weight of the metal in an LNG tank. Membrane tanks are not self-supporting, but an integral part of the ship's hull, which directly bears the weight of the cargo. The structure holding the insulation material must be strong enough to transfer the weight of the cargo to the inner hull, be an effective insulator in its own right, and prevent any liquid gas from coming in contact with the ship's hull.

The membrane for the Gaz Transport system is made of a special stainless steel alloy called invar of 36 percent nickel with a very low coefficient of thermal expansion, eliminating the need for expansion joints. Both the primary and secondary insulation consists of a layer of thin (0.7 mm) invar membrane covering plywood boxes filled with perlite, a naturally occurring insulating material made from volcanic glass. The primary and secondary insulation provides 100 percent redundancy. The membrane for Technigaz system is thin (1.2 mm), low-carbon corrugated stainless steel with a relatively high coefficient of thermal expansion. The corrugation is designed to accommodate expansion and contraction of the metal caused by temperature changes. Earlier LNG carriers of this design used balsa wood as insulation material. Now two layers of reinforced polyurethane panels, separated by a secondary membrane made of a thin sheet of aluminium between two layers of glass cloth, form the primary and secondary insulation. The latest membrane system (CS1) combines the Gaz Transport and Techigas technologies with a membrane of invar and insulation of reinforced polyurethane panels.

Of the LNG fleet of 175 vessels in early 2005, about half were spherical tank design and half membrane. However about 80 percent of the 110 vessels on order were of the membrane design. The membrane design requires less material, but construction is more labor-intensive. Spherical tanks require more material, but their construction is more automated. Thus, the comparative cost of LNG carriers of the spherical or membrane design depends on shipyard labor costs. An LNG carrier of spherical tank design costs less in Japan than membrane design because labor costs are relatively high; the opposite prevails in Korea, where labor costs are relatively low.

The first generation of LNG carriers built in the 1970s was 75,000 cubic meters, but these were quickly followed by what turned out to be a standard size of 125,000 cubic meters. The 1980s was not an active time for new LNG projects and there was little demand for new LNG carriers, but vessels built in the 1990s were typically 135,000 cubic meters. In early 2005 the order book of 110 LNG carriers was dominated with vessels of 145,000 cubic meters, with eight vessels over 200,000 and the largest being 216,000 cubic meters. These vessels are restricted to trades where the storage capacity of loading and receiving terminals and berth limitations can accommodate their larger size. Shipyards capable of building LNG carriers are in Korea, Japan, Spain, France, and China (a newcomer), with an aggregate capacity of delivering over forty LNG carriers a year, of which Korea accounts for half.

Heat passing through the insulation can warm the cargo and increase the pressure within the cargo tank. Unlike LPG cargoes in which a refrigeration plant keeps the cargo cool enough to remain liquid at atmospheric pressure, an LNG cargo is kept liquefied by boil-off, which removes the heat transmitted through the insulation into the cargo. The better the insulation, the less the boil-off; typical boil-off rates for modern vessels are about 0.15 percent of the cargo volume per day. While nearly all merchant vessels have diesel engine propulsion plants that burn heavy fuel oil, LNG carriers have dual fuel steam turbine propulsion systems that burn either heavy fuel oil or LNG boil-off, which typically provides 60 percent of the fuel requirements. This avoids wasting boil-off by flaring or venting to the atmosphere. Not all the LNG cargo on a vessel is discharged at the receiving terminal. A heel or small amount of LNG is left in the cargo tanks to keep the tanks cold via boil-off on the ballast voyage to the loading port. Keeping the tanks cold eliminates the necessity of cooling the cargo tanks before loading the next cargo and minimizes stress from repeated thermal cycling. The ship is charged for the boil-off burned for ship propulsion on an energy-equivalent basis with heavy fuel oil.

Some thought has been given to having diesel plant propulsion on LNG carriers because diesel engines are inherently more fuel-efficient than steam turbines. This necessitates the installation of a reliquefaction plant to handle boil-off or development of a diesel engine that can burn either diesel fuel or natural gas. The economic decision to switch to diesel propulsion depends on the relative fuel efficiency of diesel versus steam propulsion plants, the cost of heavy fuel oil versus the value of delivering a larger quantity of LNG, and the capital and operating costs of an onboard liquefaction unit.

The LNG cargo is pumped into LNG storage tanks at the receiving terminal. LNG has to be warmed to a gaseous state before entering natural gas pipelines for distribution to consumers. The most common way to heat LNG is to pass it through a seawater heat exchanger where the seawater is cooled and the LNG warmed to about 5°C (41°F). A gas-fired vaporizer is available if needed. A few Japanese import terminals tap the “waste cold” of LNG to cool brine or Freon for freezing food and for chemical and industrial processes that require cooling.

LNG Pricing

Pricing LNG in Japan is formula-based on the blended cost of crude oil imports (Japan Customs Cleared Crude) on an energy-equivalent basis, later adopted by South Korea and Taiwan. Thus the natural-gas exporting nation received for its natural gas a price that reflected the blended cost of crude delivered in Japan less the operating and capital costs associated with the natural-gas gathering system, the liquefaction plant, and the LNG carriers (regasification facilities are owned by the receiving nation’s gas utility). A minimum floor on the LNG price, if incorporated in the SAP contract, assured the exporting nation of a minimum price for its gas exports and the debt providers of a positive cash flow. The price relationship between LNG and crude oil is tempered to partially protect LNG importers from oil price shocks. Pricing of LNG imported into Europe is based on a formula reflecting the prices of European pipeline gas from natural gas fields, Brent crude, high- and low-sulfur fuel oil, and coal. The pricing of LNG imported into the United States is based on the price of natural gas at Henry Hub as indicated by near-term futures trading of natural gas contracts on the New York Mercantile Exchange (NYMEX).

The early LNG projects were based on fixed-quantity, twenty-or-more-year contracts between importers and exporters. Beginning in the latter part of the 1990s, spot LNG cargoes began to appear. These cargoes were the result of liquefaction plants producing more LNG than their nominal nameplate capacity (from conservative design features), improved productivity, and debottlenecking (the removal of constraints that restrict production). The first liquefaction plants were eventually able to produce 25 percent more than their indicated design capacity at a time of retrenchment in Asian economic activity, particularly in Korea, which could not absorb its contractual volumes. Now there were cargoes without a home. The third element was the availability of laid-up LNG carriers from the failed El Paso and

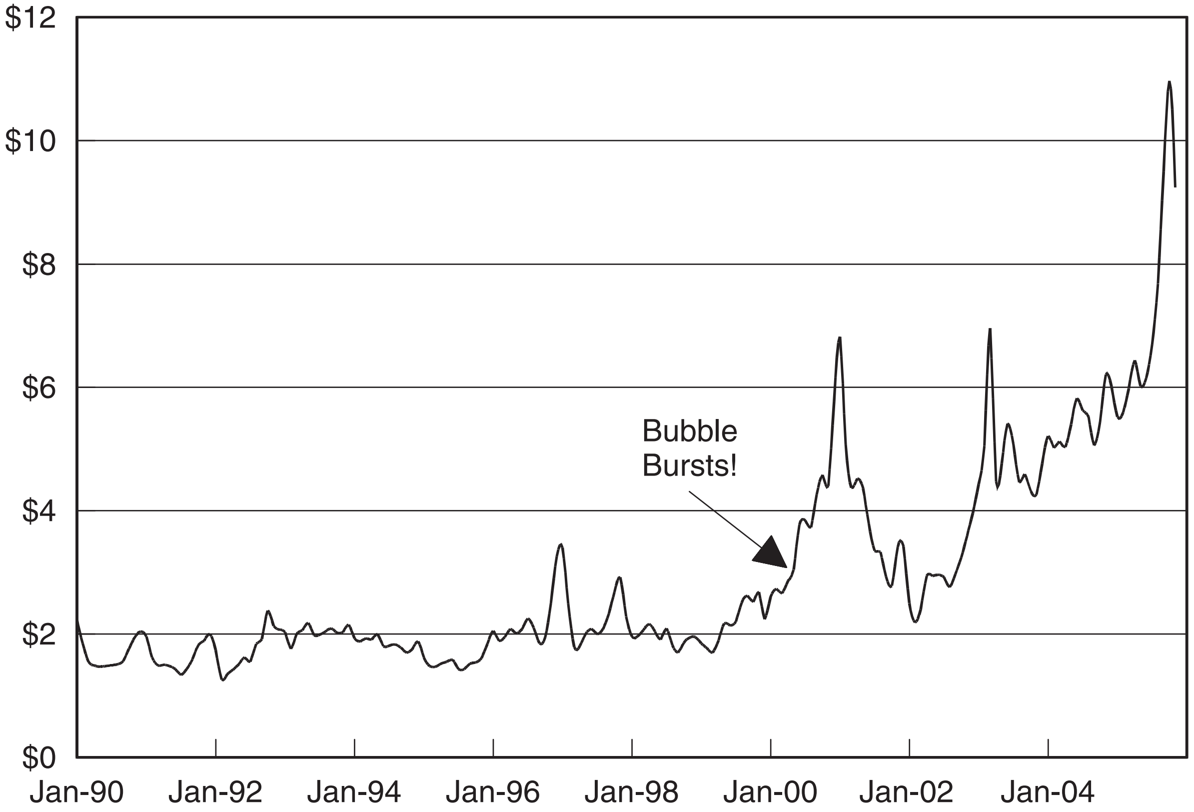
Trunkline projects plus a few built on speculation. The fourth element that caused the emergence of a spot trade in LNG was market demand in Europe and the United States to absorb these cargoes.

Since then other factors have come into play, transforming the LNG business from fixed long-term contracts based on the price of oil to a more commercially oriented business like oil and coal where the market realities of supply and demand play an important role in determining price. One was the continual decline in the cost of LNG carriers. In 1995 the cost of building a large LNG carrier was three times that of a large crude oil carrier ($280 million). Since an LNG carrier can only transport about one-third as much energy as a large crude carrier of the same cargo volume, LNG shipping costs were nine times higher than crude oil on a Btu basis. By the early 2000s, the shipbuilding cost had fallen to $150–$160 million, or a 50–60 percent premium over large crude carriers, knocking down the premium on an energy basis to about five times. Lower shipyard costs were partially caused by moving down the learning curve where repetition tends to iron out or eliminate problems that were previously experienced. Gains in shipyard productivity from automation also contributed to lower shipbuilding costs. Whereas only a few shipyards were capable of building LNG carriers in the 1970s, in the early 2000s there were a dozen. Increased competition to keep building berths busy narrows shipyard profit margins and provides an incentive to further improve shipbuilding technology. Lower shipyard prices, combined with the economies of scale of larger-sized LNG carriers, have reduced shipping costs.

Greater output and improved system design of liquefaction trains built in the early 2000s have resulted in a one-third reduction in capital and operating costs for liquefying natural gas. Two or three engineering, procurement, and construction (EPC) consortia capable of undertaking a massive and complex LNG project in the 1970s increased to about five in the early 2000s. Energy companies with the requisite project management skills, technical expertise, and the financial wherewithal to organize LNG supply-chain projects have grown from three or four in the 1970s to about ten in the early 2000s. Project funding has become more sophisticated and innovative with project managers and financial underwriters well versed on how to structure LNG projects in order to make their underlying debt attractive to potential investors. Greater reliance on debt in the financial structure of LNG projects reduces capital costs. While all these factors have lowered the delivered cost of LNG, there has also been a concomitant rise in natural gas prices, illustrated in Figure 7.5.

The average wellhead price of natural gas was about $2 per million Btu in the 1990s and tripled in the early 2000s. The significant rise in natural gas prices, coupled with a significant fall in the cost of producing and shipping LNG, has made LNG projects commercially attractive without the need for fixed long-term contractual arrangements to cover the entire output. Contemporary LNG sponsors are not so much interested in protecting against the commercial risk of an LNG project as in taking advantage of commercial opportunities. LNG project sponsors look to a mix of long-term commitments to cover the minimum financial requirements with a portion of total capacity dedicated to short-term deals to enhance profitability. The willingness of LNG sponsors to accept commercial risk and the desire of buyers not to have to commit to twenty-year, take-or-pay-contracts have encouraged the emergence of spot and short-term markets.

Figure 7.5 **U.S. Average Natural Gas Wellhead Price** ($/MMBtu)

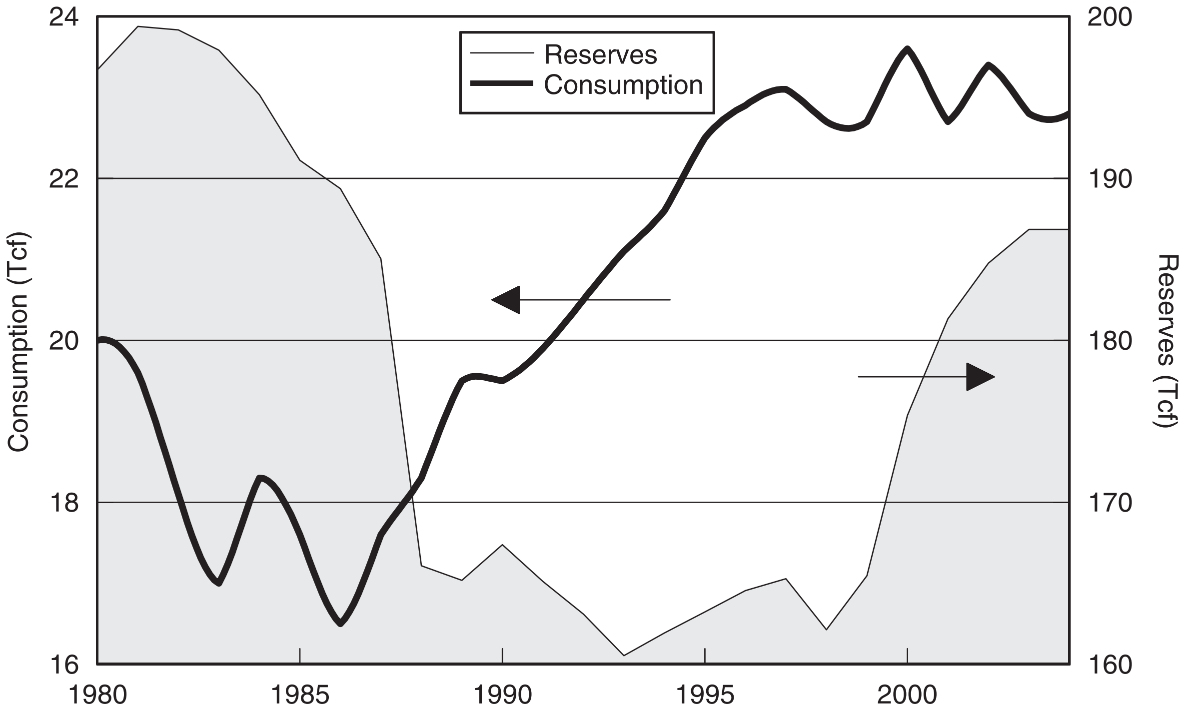


LNG buyers, no longer contractually chained to a single supplier for twenty years at a fixed formula-based price, can now select their LNG providers on a variety of short- and medium-term deals, creating a diversified portfolio of LNG purchases at prices that reflect market realities. The possibilities of commercial opportunities have led to the construction or expansion of liquefaction facilities in Nigeria, Australia, Malaysia, Oman, Qatar, and Trinidad, with new facilities under construction in Egypt, Norway, and Russia (Sakhalin Island). Nonproducing LNG nations seriously contemplating monetizing their stranded gas by building liquefaction plants include Iran, Yemen, Venezuela, Bolivia, Peru, Angola, and Equatorial Guinea.

The Future of LNG

The LNG market with the greatest potential of growth is the United States. The U.S. government recognizes the need to build LNG terminals to avert a potential shortage of natural gas. Various projections call for the United States to be importing from

Figure 7.6 **U.S. Natural Gas Reserves versus Consumption** (Trillion Cubic Feet)



12–20 percent of its natural gas needs as LNG by 2025. Figure 7.6 shows that there has been an overall decline in U.S. natural gas reserves since 1980, although there has been a marked increase in reserves since 2000. If the decline resumes in the face of growing annual demand, then this will be the prelude to large-scale imports of LNG.[10]

In addition to four receiving terminals (Lake Charles in Louisiana, Elba Island in Georgia, Cove Point in Maryland, and Everett, Massachusetts), about forty active applications for building receiving terminals are in the works for the Bahamas, the Gulf Coast, the Northeast, California, and Mexico. The Bahamas terminal, if built, will serve Florida by pipeline. The proposed projects in Mexico either supplant U.S. natural gas exports to Mexico, making natural gas more available in the United States, or are a backdoor way for LNG to enter the United States. This avoids the regulatory hassle of siting an LNG receiving facility in the United States. LNG can enter the United States via Mexico as natural gas or electricity generated from natural gas. Natural gas from a proposed LNG terminal in eastern Canada can be moved via an underutilized pipeline carrying eastern Canadian gas into New England.

Proposals to build LNG terminals in Mexico and eastern Canada reflect the difficulty of siting LNG terminals in the United States. Licenses are required from the Department of Transportation, the Coast Guard, and the Maritime Administration, along with permissions from the Research and Special Programs Administration, which enforces deepwater-port pipeline regulations, and the Department of the Interior for pipeline right-of-way. The Fish and Wildlife Service is concerned with the ramifications of an LNG receiving terminal on endangered species; the Minerals Management Service is concerned with potential hazards and underwater artifacts of archeological interest; and the Environmental Protection Agency is concerned with carrying out the provisions of the Clean Air Act. The Department of Energy must issue an import certificate and the FERC must issue a Certificate of Public Convenience and Necessity. Other Federal agencies involved are the Department of Defense, the Department of State, the Department of Commerce, the National Oceanic and Atmospheric Administration, the Bureau of Oceans, the Army Corps of Engineers, and the Advisory Council on Historic Preservation. Besides these, various state bodies involved with coastal zone management, pollution control, wildlife and fisheries, and historical preservation present their own hurdles for building a receiving terminal. Most importantly, an LNG receiving terminal cannot be built without a permit from the municipality within which the receiving terminal would be located. For this to occur, the local population must be in support of an LNG facility built in their midst.[11]

In early 2005, a number of proposed projects have received FERC approval. Four of these, located on the Gulf Coast, have also received the requisite permissions from federal, state, and local authorities to proceed with construction, with the possibility of another eight receiving all their prerequisite permissions. Proposed terminals in California and the Northeast have to combat the "not in my backyard" or "Nimby" syndrome represented by local citizens who can exert sufficient political clout to stop a municipality from issuing a permit. Where once "Nimby" was concerned with the sight and smells and sounds of having an industrial plant in someone's backyard, now there are concerns over potential terrorist actions against an LNG facility. Resolution of these issues not only determines the future volume of LNG trade, but also the prospects of owners who have ordered LNG carriers on speculation without any commitment for employment by either an LNG supplier or buyer.

An innovative approach to bypass "Nimby" occurred in 2005 when a fifth U.S. LNG terminal of a radically different design was inaugurated for service. A normal LNG terminal is located in a port with a storage facility and a regasification unit for transforming LNG to natural gas as needed. The Gulf Gateway Deepwater Port, located 116 miles off Louisiana, is a floating buoy connected to an already existing and underutilized natural gas pipeline. Located outside municipal and state jurisdictions, approvals are primarily federal. Specially constructed LNG carriers with built-in regasification units tie up to the buoy and, over five days or so regasify their cargoes for discharge into the natural gas pipeline. The pipeline is connected to several transmission pipeline systems that can contract for the natural gas. Without associated storage facilities, the regasified natural gas must enter a pipeline transmission system directly from the ship, which can cause operational problems with other natural gas suppliers. This, of course, can be corrected if access can be gained to a storage facility or an accommodation can be worked out with the suppliers. The additional cost of $25 million per vessel for an installed regasification unit adds to shipping charges, which shifts the economics in favor of nearby sources of LNG such as Trinidad. As it happened, the first shipment of LNG was from Malaysia.

Another development to move LNG terminals out of ports into offshore waters is the proposed building of a gravity-based terminal that will be sunk about ten miles offshore from Venice. The terminal is to be built in a graving dock in Spain, and, when completed, the graving dock will be flooded to float the terminal. Then the terminal will be towed to offshore Venice, and, when onsite, its empty ballast tanks will be filled with water to sink it. Later, heavier material will be added to permanently ballast the terminal. LNG carriers will off-load their cargoes into the terminal's storage tanks. LNG will be regasified at the terminal and pipelined to onshore natural gas connections. The development of such terminals can bring LNG directly into populated areas where it is needed, bypassing local opposition to building LNG terminals and storage tanks within the confines of populated areas.

The United States is not the only nation with a growing appetite for LNG. The United Kingdom, famous for its North Sea oil and gas finds, is now facing declining production and reserves. The Interconnecter, a natural gas pipeline between the United Kingdom and Europe, was built for two-way flow, exporting gas to Europe during the summer and importing gas during the winter. One might expect that the Interconnecter will be flowing in one direction as output from UK North Sea gas fields dwindles, but this may not happen. A pipeline is being built to connect the United Kingdom with a Norwegian gas field and LNG terminals in the United Kingdom are being expanded to handle a higher throughput. How these developments play out will determine the future direction of flow through the Interconnecter.

Italy, Spain, and France are expanding their LNG terminals as Europe becomes more committed to natural gas to meet its carbon dioxide emission obligations under the Kyoto Protocol. China has an enormous thirst for energy, including LNG imports and, through its balance-of-trade surplus, has the capital to invest in LNG terminals and build a natural gas infrastructure. India is another nation with an enormous thirst for energy, but that nation is stymied by balance-of-trade deficits and relatively limited capital reserves. Nevertheless, a number of companies are investigating the possibility of LNG projects in India, with at least one company expressing a willingness to accept payment in Indian rupees rather than U.S. dollars.

LNG terminals in Japan serve regional needs with no interregional pipelines. Yet the natural gas price is essentially the same throughout Japan because the price of LNG imported into each region refers to the same pricing formula. Without price differentials, there is no economic justification to build interregional pipelines in Japan. With this in mind, LNG terminals will have a major impact on basis pricing of natural gas in the United States. For instance, suppose that an LNG terminal is built in New England. The LNG export plant in Trinidad, or possibly one built in Venezuela, has nearly the same shipping distance to the U.S. Gulf as to the Northeast. Thus, it is conceivable for LNG to enter both regions at the same price. If sufficient volumes were imported at both locations, the price differential between the two regions would shrink, raising havoc with tolls for the pipelines connecting the Southwest with the Northeast.

An LNG plant in Nigeria has about the same shipping distance to the U.S. Gulf, the U.S. Northeast, and Europe. This permits arbitrage trading that would tend to equalize natural gas prices in all three regions. If price differentials are large enough to absorb the extra shipping costs, LNG cargoes from Trinidad or Venezuela (if an LNG plant is built) could also be sold in Europe and LNG cargoes from Murmansk (if an LNG plant is built) could also be sold in the United States. Moreover, LNG terminals in the Middle East with excess capacity can sell cargoes west or east (Atlantic or Pacific basins), depending on netback values. The upshot of spot trading in LNG cargoes is that the price of natural gas in Europe and the United States and Asia will not be materially different, making natural gas a globally traded commodity similar to oil.

The price of natural gas may remain closely tied to oil on an energy-equivalent basis. The price of LNG in Japan, Korea, and Taiwan is directly tied to oil. The price of LNG imported into Europe is partly tied to the price of Brent crude and fuel oil. The price of natural gas in the United States was only weakly related to oil prices during the 1990s, when natural gas was in surplus. With the passing of the natural gas bubble, a closer relationship between natural gas and crude oil prices on an energy-equivalent basis has been established. Thus, spot trading in LNG may not only equalize the price of natural gas on a global basis, but also maintain parity between the price of natural gas and crude oil. This has significant ramifications for energy consumers if crude oil supply cannot keep up with demand.

Gas to Liquids (GTL) Technology

Reservoirs of stranded gas too remote for access by pipeline, and lacking sufficient reserves to support an LNG export project, can be made accessible to the market through gas to liquids, or GTL, technology. Combining methane with air at high temperatures produces a mixture of carbon monoxide and hydrogen, which, via the Fischer-Tropsch process, in the presence of iron or cobalt catalysts, can create longer hydrocarbon chains, resulting in a combination of diesel fuel, naphtha, and wax.[12] The Fischer-Tropsch process is very versatile because it can also create shorter hydrocarbon chains when employed to produce motor vehicle fuels from coal.

Diesel fuel produced by the GTL process is very clean-burning, with significantly less nitrous oxide and particulate emissions and virtually no sulfur oxide emissions. Shell Oil has been in the forefront of GTL development and has been producing 12,000 bpd of liquid petroleum products from its Malaysian plant for a number of years. Qatar has been actively seeking joint venture partners to build GTL plants capable of producing as much as 400,000 bpd of diesel fuel by 2010. Russia is looking into GTL production for isolated gas fields in Siberia and there is a proposal to build barge-mounted GTL plants to reach isolated gas fields located near water in Southeast Asia and elsewhere. Large-scale GTL projects are necessary for the International Energy Agency projection of 2.4 million bpd by 2030 to hold true. Selling GTL diesel fuel is a virtually unlimited market from the perspective of natural gas producers, whereas LNG is ultimately limited by the throughput capacity of LNG receiving terminals. One drawback to GTL production is the cost of the GTL plant, which could be reduced with further technological advances and by economies of scale in building larger-sized production facilities much as the cost of building LNG liquefaction plants has fallen over time. The other major drawback is that the GTL process is about twice as thermally inefficient as an LNG liquefaction plant. This means that a lot more of the original energy content of natural gas is lost when converted to a liquid petroleum product than to LNG.

Methane from Coal

Methane found with coal has been responsible for the death of many miners. Coal bed methane (CBM) is “mining” coal beds not for their coal, but for their methane. CBM works best for methane-rich coal beds that are fractured and submerged in groundwater. Water surrounding the coal beds, located 200–3,000 feet underground, absorbs and retains the methane as long as the water is under pressure. With a low solubility in water, methane readily comes out of solution when the water pressure is dropped.

A well is drilled to the coal seam to allow the water to rise to the surface, where it is kept under pressure. From time to time water is pumped out of the well to lower its pressure. The released methane is diverted to a gathering system for eventual distribution to a natural gas pipeline. After the release of methane, pumping stops and the well is capped to increase its internal pressure. Once the concentration of methane is restored, the well is pumped again. The principal region for CBM wells in the United States is the Rocky Mountains (Wyoming, Colorado, and New Mexico), with some production in Appalachia. CBM already contributes about 8 percent of the natural gas production in the United States and is expected to continue growing in the future, particularly with the rise in natural gas prices. CBM reserves are estimated at 700 tcf, of which 100 tcf is recoverable—enough to cover current U.S. consumption for five years. China has double the estimated reserves of the United States, but its CBM output is quite small at this time.

Environmentalists object to the pristine Western wilderness being crisscrossed with gathering pipelines and its quiet disturbed by the noise of equipment pumping water out of the well and compressors moving gas in pipelines. Much of the water from CBM wells is saline and can damage agricultural and natural plant life. Saline water is kept in ponds, but some could accidentally seep into the surface groundwater. Reinjecting saline water into the CBM well avoids the risk of surface water contamination.[13] There is speculative thought about using methane-exhausted coal beds for sequestering carbon dioxide emissions from coal-burning plants. Sequestering or entrapment of carbon dioxide to replace the methane in the coal bed would keep carbon dioxide out of the atmosphere. Carbon dioxide from burning CBM locally for electricity generation could also be sequestered in the coal bed. Sequestering carbon dioxide in coal beds would require separating carbon dioxide from flue gases and pipelining it from an electricity-generating plant to a CBM site, which, as one might expect, is quite costly.

Methane from Shale

The United States consumes about 23 tcf of natural gas annually, of which 19 tcf are produced domestically. The rest is mainly imported from Canada as pipeline gas and the remainder is imported as LNG. About 23 percent of natural gas production is in hard-to-access sources such as coal beds and shale rock. Total reserves of natural gas in shale rock are estimated to be over 500 tcf, but recovery is difficult. In 2005, about 35,000 natural gas wells were extracting a paltry 0.6 tcf from shale. With the sharp escalation of natural gas prices, efforts are being made to extract more natural gas from shale and other difficult sources. A new method has been devised to crack open shale using water and sand under pressures up to 3,500 psi. As the shale begins to fracture, the water-sand mixture fills the crack, causing the crack to propagate through the rock. After the crack has reached its desired length, water is pumped out and sand holds the crack open to allow natural gas to flow through the crack and up the well. Horizontal drilling is another way to tap natural gas trapped in shale. These methods were not economical when natural gas was selling at $2–3 per million btu, but at double or triple this price, these unconventional sources become more attractive and provide a means of increasing our supply and putting a lid on further escalation of prices.

Methane Hydrates

Methane hydrates are essentially natural gas molecules trapped in a lattice of ice whose structure is maintained in a low-temperature and high-pressure environment. Methane hydrates look like ice crystals. An ice ball of methane hydrates is suspiciously like those carefully sculpted by Calvin (in the "Calvin and Hobbs" comic strip) to throw at Suzie. The only difference is that the methane hydrate ice ball can be ignited. One cubic meter of methane hydrates contain 160 cubic meters of embedded natural gas. Methane hydrates are found beneath large portions of the world's permafrost as well as in deep-sea sediments. They are thought to have been formed by migrating natural gas or seep gas that came in contact with cold seawater at deep depths or by the decay of marine life in bottom sediments. Cold and pressure keep the methane entrapped in the ice lattice but it is released if warmed or the pressure is reduced. Some climatologists fear that global warming of the tundra regions could release methane now entrapped as methane hydrates in the permafrost. This would lead to runaway global warming because methane is twenty times more effective than carbon dioxide in reflecting back infrared radiation from the earth. The challenge is how to "mine" methane hydrates, considering their inherent instability.[14]

Methane hydrates are not limited to the arctic regions. Large deposits of methane hydrates have been found in coastal regions around Japan, both coasts of the United States, Central and South America, and elsewhere. The known world reserves of natural gas are about 6,300 tcf, while the worldwide estimate of methane in methane hydrates is over 100 times greater at 700,000 tcf. For the United States, the estimate is 200,000 tcf versus natural gas reserves of 187 tcf, over 1,000 times larger. There is an awful lot of methane locked up in methane hydrates and such a potential cannot be ignored. A joint funding program by the U.S. Department of Energy and major oil companies, a joint venture between Japan and Canada, and the U.S. Geological Survey are involved in various projects to map and assess methane hydrate deposits. Understanding the nature of methane hydrate deposits would be the first step toward dealing with the technological challenge of how to mine them.

As an aside, brine pools with an extreme concentration of salt have been found in certain areas of the ocean. These pools are also rich sources of methane, which only the surrounding colonies of mussels, which have formed a symbiotic relationship with methane-metabolizing bacteria that live on their gills, know how to tap. Methane-metabolizing bacteria have also been found living symbiotically with worms in methane hydrate deposits at the bottom of the Gulf of Mexico. Methane hydrate deposits and brine pools are fairly recent discoveries, as is methane in the atmosphere of Saturn's moon, Titan, which may have lakes of LNG. We live in an amazing world in an equally amazing universe.

Notes

1. Primary sources for the history of coal gas are Christopher J. *Castaneda, Invisible Fuel* (New York: Twayne Publisher, 1999) and Arlon R. Tussing and Bob *Tippee, The Natural Gas Industry* (Tulsa, OK: PennWell Books, 1995).

2. Bob Shively and John Ferrare, *Understanding Today's Natural Gas Business* (San Francisco: Enerdynamics, 2004).

3. *BP Energy Statistics* (London: British Petroleum, 2005).

4. International Energy Agency, *World Energy Outlook* (Paris: Author, 2004).

5. See the Gas Infrastructure Europe Web site at www.gte2.be.

6. Colin Shelley, *The Story of LPG* (New York: *Poten* & Partners, 2003).

7. The statistics for pipeline movement of natural gas presented in Table 7.1 are from *BP Energy Statistics* (London: British Petroleum, 2005).

8. Michael Tusiani, *The Liquefied Natural Gas Industry* (New York: Poten & Partners, 2005).

9. Minority Staff Report U.S. House of Representatives Committee of Government Reform, *Background on* Enron's Dabhol Power Project, a (February 22, 2002), available at www.democrats.reform.house.gov/ Documents/ 20040830150742–77212.pdf.

10. *BP Energy Statistics* (London: British Petroleum, 2005)

11. *LNG in North America Markets* from a Citigroup Smith Barney report of August 27, 2004.

12. An animated explanation of the GTL project is available on the Conoco Gas Conversion Technologies Web site at www.gassolutions.conoco.com.

13. National Energy Technology Laboratory, *Integrated Process of Coalbed Brine and Methane Disposal,* available at www.netl.doe.gov; Environmental Protection Agency, *Coal Mine Methane Use in Brine Water Treatment,* available at www.epa.gov.

14. National Methane Hydrate Program of the National Energy Technology Laboratory at www.netl.doe.gov, and the National Research Council's C*harting the Future of Methane Hydrate Research* (Washington, DC: Author, 2004).

8

Nuclear and Hydropower

Oftentimes nuclear and hydro are linked together in energy statistics. Of course, they are quite different: Nuclear power is generally viewed as dangerous while hydro is viewed as benign. This really is not quite true for either. Despite Three Mile Island and Chernobyl, the safety record for nuclear power speaks for itself: There have been over 11,000 reactor-years of safe commercial plant operation, coupled with an equivalent span of safe operation for nuclear-powered warships. The worst accident by far was Chernobyl, a case of an unsafe reactor design unsafely operated. Hydropower has its opponents and dam failures are not unknown phenomena. This chapter covers the principal aspects of both nuclear and hydropower as energy sources.

Background

Nuclear power is the outgrowth of the nuclear weapons program to transform the world's most destructive weapon to peaceful uses. The 1953 launching of the Atoms for Peace program foretold a world where commercial nuclear energy would be clean, abundant, safe, and too cheap to even meter! Nuclear power is clean, because it does not generate emissions that contribute to global warming, but dirty because the spent fuel must somehow be disposed. Three Mile Island buried the myth that nuclear power was inherently safe and Chernobyl showed how dangerous it could be. And cheap it is not, with cost overruns in the billions.

Yet, despite predictions of a phaseout of nuclear power plants and general pessimism over the prospects for nuclear power as an energy source, nuclear power plants continue to be built. There are some who believe that we may be at the dawn of a new age in nuclear power. One cannot cavalierly dismiss the fact that nuclear power is free of greenhouse gas emissions. The promise of standardized "cookie-cutter" plants, built the way Ford manufactured Model T's, would eliminate the enormous cost overruns associated with the one-of-a-kind nuclear plants that dominated past construction. Advancements in nuclear power technology, coupled with series production of a standard plant design, would make the cost of electricity from nuclear plants quite attractive compared to fossil fuel plants, particularly if fossil fuel prices continue to rise. What has to be accomplished before any renaissance of nuclear power becomes possible is reassuring the public that the human errors and circumstances responsible for the Three Mile Island incident and the Chernobyl catastrophe cannot happen again.

Human error played a major role in both the Three Mile Island incident and the Chernobyl catastrophe. Yet, nearly all of the radioactive release of the Three Mile Island incident was kept within its containment system, as it was designed to do. Soviet nuclear power plants do not have containment systems built to withstand the pressure generated from a ruptured reactor system. Soviet reactors are housed in buildings. Nor did the Soviet Union select a safe plant design. Whereas most reactors shut down when the water moderator in the core boils away (an example of

Table 8.1

U.S. Electricity Generating Units

Megawatt Nameplate Output	Number	Output in Gigawatts	Percent of Total Nameplate Output
Less than 1	2,298	1.4	0.1
1–10	6,135	18.1	1.8
10–50	3,781	96.6	9.4
50–100	1,861	132.0	12.8
100–250	1,791	286.9	27.8
250–500	441	152.4	14.8
500–750	257	156.7	15.2
750–1,000	125	107.2	10.4
1,000–1,500	66	80.4	7.8
Total	16,755	1,031.7	100.0

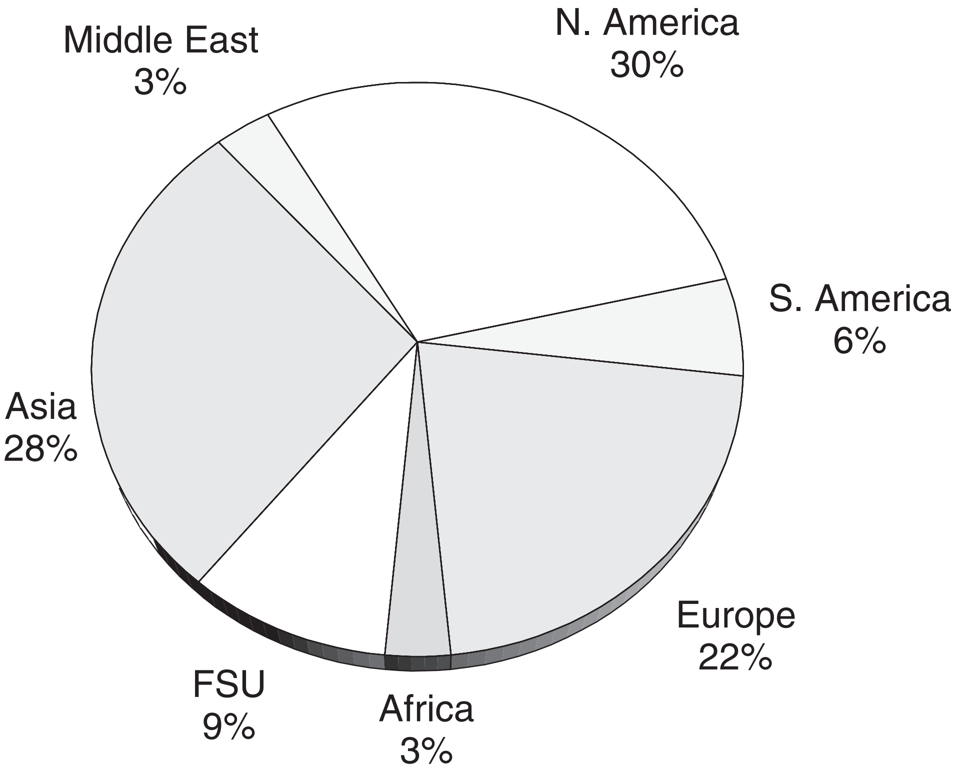
a negative feedback system), the same phenomenon with the Soviet graphite-moderated reactor led to a runaway power surge (an example of a positive feedback system). While reopening the nuclear door is going to be a hard sell in the United States and elsewhere, the fact is that reactors are currently under construction in Argentina, Canada, China, India, Iran, Japan, Korea, Romania, and Russia.

What Does a 1,000-Megawatt Plant Mean?

The typical large-sized nuclear power and coal-fired plants have an output between 1,000–1,500 megawatts, or 1–1.5 gigawatts (there are 1,000 megawatts in a gigawatt). To understand what this means, the United States, which represents 24 percent of the world's usage of electricity, has an incredible 16,755 electricity-generating units.[1] Many of these are small; for instance, the listing includes a college with four natural gas-fueled electricity-generating backup units. Other small units include natural gas cogenerating units at industrial sites, mini and micro hydrodams, and generating units powered by wind, solar, geothermal, waste products, industrial waste heat, and methane captured from landfills. Table 8.1 shows the importance of the relatively few larger-sized units in satisfying U.S. electricity demand.

The total of 1,032 gigawatts is nameplate capacity. If typical peak demand in the United States is 800 gigawatts, and if base-load demand is about 60 percent of peak demand, or say 500 gigawatts for a U.S. population nearing 300 million people, then a 1-gigawatt or a 1,000-megawatt plant can handle the base-load needs of a city of 600,000 people. This, of course, includes the proportionate share of electricity demand from industrial plants and commercial enterprises plus public lighting, traffic lights, and so forth. However, the United States is a very energy-intensive nation. Total world electricity generation capacity was 3,509 gigawatts in 2003. About 2 billion people, out of a world population of 6 billion people, are not connected to an electricity grid. Adjusting for base load, a 1-gigawatt or 1,000-megawatt plant can serve the needs of about 2 million people. Many people in the developing world who are connected to the electricity grid have minuscule electricity usage amounting to one or two electric bulbs and a small appliance. A common rule of

Figure 8.1 **Distribution of Electricity-Generating Capacity**



thumb is that a 1-gigawatt plant can serve the needs of 1 million people, a compromise between nations with high and low per capita consumption. Figure 8.1 shows the location of the world's 3,509 gigawatts of electricity-generating capacity.[2]

Base-load needs are most commonly handled by large nuclear and coal-fired plants. A 1-million kilowatt or 1,000-megawatt (1,000) or 1-gigawatt (KW) coal-fired power plant releases about 6 million tons of carbon dioxide, plus potentially a large quantity of sulfur dioxides into the environment each year, depending on the sulfur content of the coal and the effectiveness of the plant's scrubbers (if any). There is also pollution in the form of nitrous oxides, plus potential health-affecting emissions of mercury, cadmium, and arsenic. A nuclear plant of the same size consumes about 25 tons of uranium, enriched from 3.5–5 percent of the isotope U235, per year, which requires over 200 tons of uranium oxide concentrate produced by mining 25,000–100,000 tons of ore, depending on the uranium concentration. The annual waste from a nuclear power plant is less than 30 tons of spent fuel, a highly radioactive and toxic waste. If reprocessed (chopped up and dissolved in acid to recover fissionable material for recycling), spent fuel can be reduced to about 1 ton of waste. Though even more dangerous and toxic than spent fuel, such relatively small quantities should be effectively managed for transport and storage, or at least so it was thought.[3]

Uranium is as common as tin and is mined both on the surface and underground. Half of the world's uranium ore production is in Canada and Australia, followed by Kazakhstan, Niger, Russia, Namibia, Uzbekistan, and the United States. The ore is first finely ground, then leached with sulfuric acid to remove uranium in the form of uranium oxide, called "yellow cake," which is then transformed to uranium fluoride gas. Both the gaseous diffusion and high-speed centrifuge processes take advantage of the fact that U235 is slightly lighter than U238. These processes create two streams of uranium fluoride gas: one enriched with U235 and the other depleted of U235. Starting with a concentration of 0.7 percent U235, the enriched stream has a concentration of about 3.5–5 percent, depending on the type of reactor, and the depleted stream is nearly pure U238. The depleted stream is 1.7 times denser than lead and can be used for reactor shielding. While most is stockpiled, some has been drawn down in recent years to mix with highly enriched uranium released from the Russian and U.S. weapons programs for transformation to reactor fuel.

Uranium reserves for conventional reactors can last for over a century. Reserves can be extended by a factor of 100 or more by reprocessing spent fuel to reuse the plutonium generated by the fission process, by breeder reactors designed to create their own fuel, and by utilizing thorium, which becomes fissionable when transformed to U233 in a nuclear reactor. Taking into consideration uranium life extension through reprocessing, breeding, and transforming thorium to fissionable material, some view nuclear energy as a sustainable source of energy, something virtually inexhaustible.

Enriched uranium fluoride is converted to uranium oxide, pressed into small cylindrical ceramic pellets, and inserted into thin tubes of zirconium alloy or stainless steel to form fuel rods. These are then sealed and assembled into reactor fuel assemblies and placed in the core of the nuclear reactor. The core of a 1,000-megawatt or 1-gigawatt reactor contains about 75 tons of enriched uranium. The presence of a moderator such as water or graphite slows down the neutrons sufficiently for the U235 isotope to fission (or split) in a chain reaction that produces heat to transform water to steam. From that point on, the generation of electricity is the same as in a fossil-fueled plant.

Physics of a Nuclear Reactor

U235 can fission or split into fission by-products such as barium and krypton, releasing about 2.5 prompt or fast neutrons and other products. The fission by-products also decay, releasing delayed (or slow) neutrons. Both prompt and delayed neutrons are necessary to maintain criticality (a constant rate of fission). Slowing down neutrons in a moderator such as graphite or water is necessary for the neutrons to be absorbed by fissionable material in a conventional reactor. The exception is fast breeder reactors that depend only on prompt or fast neutrons to maintain criticality. The total mass of fission by-products is less than the original U235 atom and the heat released is equivalent to the loss of mass multiplied by the square of the speed of light (Einstein's famous $E = mc^2$).

From the perspective of converting matter to energy, a nuclear bomb and a nuclear reactor are similar. But a nuclear bomb is designed to have a runaway reaction, whereas a nuclear reactor is designed to prevent a runaway reaction. A nuclear bomb concentrates over 90 percent fissionable material for a single explosive nuclear event.

A nuclear reactor disperses a low concentration (3.5–5 percent) of fissionable material within a fuel assembly, along with channels for coolant to pass through and to insert neutron-absorbing control rods. It is impossible for a nuclear reactor to sustain a nuclear explosion, but it is possible for the core to melt down and release radioactivity if the coolant is lost.

A reactor is shut down when control rods are fully inserted. To operate a reactor, control rods are pulled out until a critical mass is formed in which a self-sustaining chain reaction can occur (a constant number of fissions over time). Pulling the control rods out further increases the fission rate and the power output of the reactor. A reactor control system scrams (or shuts down) the reactor by rapid insertion of control rods if system performance does not fit a tight set of specifications. Heat is generated within a reactor by the transfer of kinetic energy from fission by-products to molecules in the fuel rod and then to molecules in the coolant and by slowing down of neutrons in the moderator. With exceptions, coolant is normally water flowing through channels within the assemblies of fuel rods and control rods. Nearly all fission products are locked in the fuel rod to ensure that the water coolant has a low degree of radioactivity. The water not only transfers heat from the reactor to the steam generators to drive the electricity generators, but, with exceptions, also serves as a moderator to slow down the neutrons.[4]

Nuclear Incidents and Accidents

A nuclear incident occurs when released radioactivity is contained; that is prevented from escaping to the outside environment, with no resulting loss of life and with minimal impact on the health of those exposed to radiation. The history of nuclear accidents starts in 1952 with a partial meltdown of a reactor's core at Chalk River near Ottawa, Canada, when four control rods were accidentally removed. The resulting radioactive release was contained in millions of gallons of water and no injuries resulted. In 1957, Windscale Pile No. 1 north of Liverpool, England, sustained a fire in a graphite-moderated reactor and spewed radiation over a 200-square-mile area. In the same year, an explosion of radioactive wastes at a Soviet nuclear weapons factory in the South Ural Mountains forced the evacuation of over 10,000 people from the contaminated area. In 1976, a failure of safety systems during a fire nearly caused a reactor meltdown near Greifswald in former East Germany. Of all nuclear accidents, two stand out: Three Mile Island and Chernobyl.

Three Mile Island Incident

The Three Mile Island incident in March 28, 1979, was preceded by the release of the movie *China Syndrome* on March 16, 1979, a case of Hollywood prescience or fiction preceding fact. *China Syndrome* was about a nuclear plant with internal problems that, if unattended, could have led to a core meltdown, which would then burrow its way toward China. The film dealt with management's decision to ignore and cover up the plant's problems.

The Three Mile Island incident proved that nuclear power plants were not immune to accidents, despite claims to the contrary. In this case a malfunction of the secondary cooling circuit caused the temperature in the primary coolant to rise, shutting down the reactor as expected. What was not expected was the failure of a relief valve to close and stop the primary coolant from draining away. The open relief valve did not show up on the instrumentation panel, making it difficult for the operators to diagnose the true cause of the problem. As a result, the coolant continued to drain away until the core was uncovered. Without coolant, the residual decay heat in the reactor core raised the temperature within the core and lead to a partial core meltdown.

Although the instrumentation panel failed to show that the relief valve was still open, the blame for the accident was eventually assigned to inadequate emergency-response training on the part of the operators. In other words, despite faulty indication of the relief valve, the operators should have identified the true cause of the problem and taken proper action before it was too late. The containment system performed as it was designed to, preventing nearly all the released radioactivity from escaping to the outside environment. Contrary to the *China Syndrome* plot, management did not hide the plant's problems from the public and the core did not melt through the earth.

There were minor health impacts and no injuries from the Three Mile Island incident. Even though the nuclear power industry took remedial steps to improve training and operations to make reactors even more safe and reliable, the Three Mile Island incident dealt a deathblow to the U.S. nuclear power industry. The incident halted all further orders of nuclear power plants in the United States and the cancellation of over forty orders for plants not yet started. Most plants under construction were completed, although a few were converted to fossil fuel plants. The public concern over nuclear safety generated by this incident was sufficient to prevent the Shoreham plant on Long Island from becoming operational when it was completed in 1984. If a more serious incident than that of Three Mile Island were to occur at the Shoreham plant, the few bridges and tunnels connecting Long Island with the mainland would preclude any large-scale evacuation. The plant was dismantled in 1992.

Chernobyl

In one respect the two events were similar: both involved human error. At Chernobyl, a runaway reactor occurred during a test, ironically one associated with reactor safety: How long could turbines supply power when cut off from reactor power? What made Chernobyl so much worse than Three Mile Island was the nature of its reactor design, actions operators used to defeat safety features, and the absence of a containment system (the reactor housing was not built to contain a pressure buildup from a reactor rupture). In conducting the test, the automatic reactor trip mechanisms were disabled and the emergency core cooling system was shut off. With the valves locked, none of the operators knew who had the keys! Having disabled the reactor's safety features, the two principal operators started "doing their own thing" without communicating to each other what they were doing.

The reactor design made a bad situation worse. The Soviet reactor used graphite as a moderator and water as a coolant. Graphite has several undesirable features as a moderator. At too high a temperature, graphite can burn or react violently with steam to generate hydrogen and carbon monoxide, both combustible gases. In a U.S. reactor water, as both moderator and coolant, shuts down the reactor when water boils in the core. Void spaces in boiling water reduce the number of neutrons being slowed down to keep the reactor critical (negative feedback). In the Soviet reactor, the creation of void spaces in boiling water allowed a larger number of neutrons to reach the graphite moderator, increasing the fission rate (positive feedback). From a low power condition, the operators retracted more control rods than recommended and the reactor went supercritical, generating enough heat to turn the coolant to steam, which further increased the number of fissions. The resulting power surge ruptured the fuel elements and blew off the reactor cover plate. The graphite moderator burst into flames when air gained access to the core, and the resulting blast, along with the escaping steam, ruptured the roof of the building housing the reactor. Large chunks of the reactor core and graphite moderator were scattered outside the building, releasing far more radioactivity than the nuclear bombs dropped on Hiroshima and Nagasaki.

Death quickly followed for those in contact with the radioactive debris or caught in the radioactive cloud close by the plant. About 200,000 people living within a thirty-kilometer radius of the plant had to be resettled, and increasing the exclusion zone a few years later required resettling another 200,000. Those caught in the radioactive cloud that reached to eastern Europe and Scandinavia now suffer from a higher incidence of cancer and birth defects. While Russian inspectors monitor food for radioactivity from farms, they miss large quantities of contaminated berries and mushrooms gathered by individuals from forests that “all but glow in the dark.” Many believe that the actual death toll far exceeds the official death count of a few hundred. Even so, this does not include the shortening of life from a higher incidence of cancer and the large number of babies born with serious birth defects.

Since the Chernobyl accident, Russian reactors have been retrofitted with modifications to overcome the deficiencies in the original design. Moreover, there has been significant collaboration between Russian and Western nuclear engineers to advance safety in nuclear reactor design and operation. The hurried Chernobyl reactor entombment is showing signs of deterioration, which will have to be revisited in order to ensure that the large amounts of radioactive material still inside the building do not escape to the environment. Nevertheless, the legacy of these two events will live on. The Three Mile Island incident cast a pall over the U.S. nuclear power program and the Chernobyl nuclear catastrophe had far-ranging global implications.

The most recent nuclear incident occurred in 1999 in Tokaimura, Japan, in a uranium-reprocessing nuclear fuel plant. Workers inadvertently mixed spent uranium in solution in a container large enough to create a critical mass. Although there was no explosion, the liquid went critical, giving off large amounts of radioactivity. As the liquid solution boiled, void spaces stopped the chain reaction (lack of a moderator to slow down the neutrons). When cooled, the solution became critical again. This lasted for twenty hours before a neutron absorber could be added to the tank to keep its contents subcritical. Twenty-seven people were exposed to very high levels of radioactivity and two died, and more than 600 others were also exposed to less dangerous levels of radiation.

Japan does relatively little fuel reprocessing; much of its spent fuel is shipped to the United Kingdom and France. Fissionable material from reprocessing is returned to Japan as a mixed oxide fuel for fabricating new fuel elements. Three years after this incident, in 2002, a scandal broke out when it was learned that Japanese utility management hid the fact that there were cracks in nuclear power plant piping (shades of *China Syndrome*). All nuclear power plants in Japan were shut down for inspection and repair, if necessary. No reactor incident came of this, but there was a justifiable loss of confidence in management, raising doubts about Japan’s future reliance on nuclear power.

Weapons Proliferation

The chain reaction transforms some of the U238 in the reactor core to various plutonium isotopes. What is of concern is fissionable plutonium 239 that remains in the spent fuel when about 75 percent of the U235 has been consumed. A typical light-water reactor breeds about 8 kilograms of plutonium 239 per month of operation, although one-third undergoes fission, supplying more power to the reactor. A fast breeder reactor is designed to create more plutonium 239 from irradiating uranium 238 than the fissionable material consumed. A fast breeder reactor depends only on prompt or fast neutrons, not delayed or slow neutrons, to maintain a chain reaction, requiring a greater degree of technological sophistication for reactor control. Fast breeder reactors can extend uranium reserves forever, at least from the perspective of human existence. Three fast breeder reactors exist and two more are being built in India and Russia.

The possibility of nuclear weapons made from plutonium 239 extracted from spent fuel has been of concern to the world community for many years. With regard to weapons proliferation, only fifteen kilograms of plutonium 239 can make a crude nuclear weapon and more sophisticated varieties require less, which represents about two or three months of reactor operation. Plutonium 239 can be separated chemically from spent fuel after it is ground up and dissolved in acid. The International Atomic Energy Agency (IAEA) was set up to ensure that nuclear materials at reactor sites and at enrichment and reprocessing facilities are not diverted to nuclear weapons manufacture. The potential, real or otherwise, for diversion of plutonium 239 from spent fuel from a reactor in Iran and another in North Korea for nuclear weapons is unsettling the world community.

In recent years a new weapon of mass destruction has arisen for the terrorists' arsenal. It consists of a metal container filled with highly radioactive spent fuel ground to fine particles surrounded by conventional explosive. When detonated, the explosion vaporizes and disperses the particles as an aerosol, spreading lethal amounts of radioactivity over a wide area. Only a few micrograms of ingested or inhaled plutonium 239 are fatal. The knowledge that terrorists have seriously considered flying an airliner into a nuclear power plant is another disincentive for building nuclear power plants. Containment systems, with walls typically four feet thick made of steel-reinforced concrete, are designed to sustain the accidental crash of a jet liner. However, intentionally ramming a jet liner at full speed into a reactor may be another matter.

Disposal of Spent Fuel

About one-third of the fuel assemblies are removed from nuclear reactors each year as spent fuel and replaced with fresh fuel. Spent fuel still contains about 96 percent of its original uranium, although its fissionable U235 content has been reduced to less than 1 percent. The highly radioactive spent fuel gives off heat and is normally stored in a spent fuel pool at the reactor site; the water shields the environment from radiation and absorbs the heat. This has to be considered temporary storage, however, because the radioactivity will persist for thousands of years, far beyond the life of the plant.

Spent fuel can either be sent to permanent storage or reprocessed. Reprocessing plants, located in Europe, Russia, and Japan, separate the uranium and plutonium. Recovered uranium is converted back to uranium fluoride and reenriched with U235. Plutonium can be blended with enriched uranium to produce a mixed oxide fuel. About thirty European reactors can be loaded with 20–50 percent mixed oxide fuel, and Japan plans to have one-third of its reactors capable of using mixed oxide fuel. This recycling of spent fuel greatly reduces the demand for uranium and the volume of spent fuel. After recycling, the remaining 3 percent of highly radioactive wastes is mixed in liquefied Pyrex glass, which contains neutron-absorbing boron, and poured into steel canisters. One ton of reprocessed waste is embedded in 5 tons of glass.

The problem is now where to store the canisters. Final disposition sites for these canisters have not been built, but geological formations made of granite, volcanic tuff, salt, or shale are being examined. One proposal is to drop the boron impregnated glass canisters into ocean trenches for "natural" disposal. The glass prevents any escape of radioactive material into the environment. The canisters are adequately shielded with five to seven miles of ocean water. If the ocean trench is also a subduction zone, over millions of years the canisters will be dragged into the earth's mantle, melted, and dispersed. It is possible that the waste could return to the earth's surface in volcanic lava in some tens or hundreds of millions of years; but by that time its radioactivity will be gone.

Public objections to dumping nuclear toxic waste in ocean trenches have ruled out what may be a very practical solution to nuclear waste. Yet, there is a precedent. About 2 billion years ago, at a place called Oklo in Gabon, West Africa, six "nuclear reactors" operated naturally within a rich vein of uranium ore that went critical after being saturated with water. The water acted as a moderator and the "reactors" remained critical, producing heat and radioactive fission by-products before running out of fuel about a half million years ago. The radioactive residue, which totals over 5 tons of fission products and 1.5 tons of plutonium, has all decayed into harmless nonradioactive isotopes. It has also been theorized that another natural reactor exists in the earth's core, maintaining its high temperature and keeping its outer layer liquid to induce the enormous flow of electricity responsible for the earth's magnetic field.

The problem with land storage is that the radioactivity will persist for many thousands of years, far exceeding recorded history. Any water seepage into the storage area could become contaminated and affect the surrounding water table. Sweden and Finland are in the process of developing permanent storage facilities. The most publicized proposed permanent storage site is Yucca Mountain in Nevada. The U.S. Congress approved this site in 2002 after $4 billion and twenty years of study. It has yet to be licensed by the Nuclear Regulatory Commission, which will examine the suitability of Yucca Mountain's geology, hydrology, biology, and climate. The factors favoring Yucca Mountain are its remote location with regard to population centers, its dry climate, and the deep depth of the underlying water table. Needless to say, there is opposition to this plan. A fairly recent and unexpected source of opposition is the state of Nevada, which appears to be having second thoughts about becoming the nation's sole nuclear waste depository (dumpsite). The state filed a lawsuit against the U.S. Department of Energy for using public rail transport to ship spent fuel to the site.

The cost of the Yucca Mountain study of $4 billion, and the billions spent by the government in research to bring about peaceful uses of nuclear power, are not included in the cost factors that determine the price of electricity made from nuclear power. Advocates for alternative energy point out that if these developmental costs, along with the impact of pollution on health from coal-burning plants, were included as cost factors for pricing electricity, then the price would rise to a level that would make alternative energy economically feasible. They have a point. It would be better if the economic signals that affect decisions on investments in energy contained the full and true cost of energy. Government subsidies, if necessary, should be monetized and charged to the users over time in order to assure that economic signals fully reflect underlying costs. Nuclear power should reflect government expenditures in enrichment plants, research, and in nuclear waste disposal.

Nevertheless, economics is having its say on nuclear energy aside from the well-publicized accidents. In the United States a new natural gas-fired combined-cycle electricity-generating plant, in conjunction with the latest jet engine technology, costs about $600 per kilowatt-hour of installed capacity. A wind turbine can be built for about $1,000 per kilowatt-hour, a large coal plant $1,300 per kilowatt-hour, and a new nuclear power plant costs a little over $2,100 per kilowatt-hour (the last group of reactors to be completed in the United States cost about $3,000–$4,000 per kilowatt-hour). Other than free sources (hydro, wind, solar, and geothermal), nuclear power has the lowest fuel costs. The thrust of nuclear technology for the future is to reduce capital costs to make the marginal cost of nuclear energy (fixed and variable costs, including refueling) more attractive as a source of electricity.

Commercial Reactors

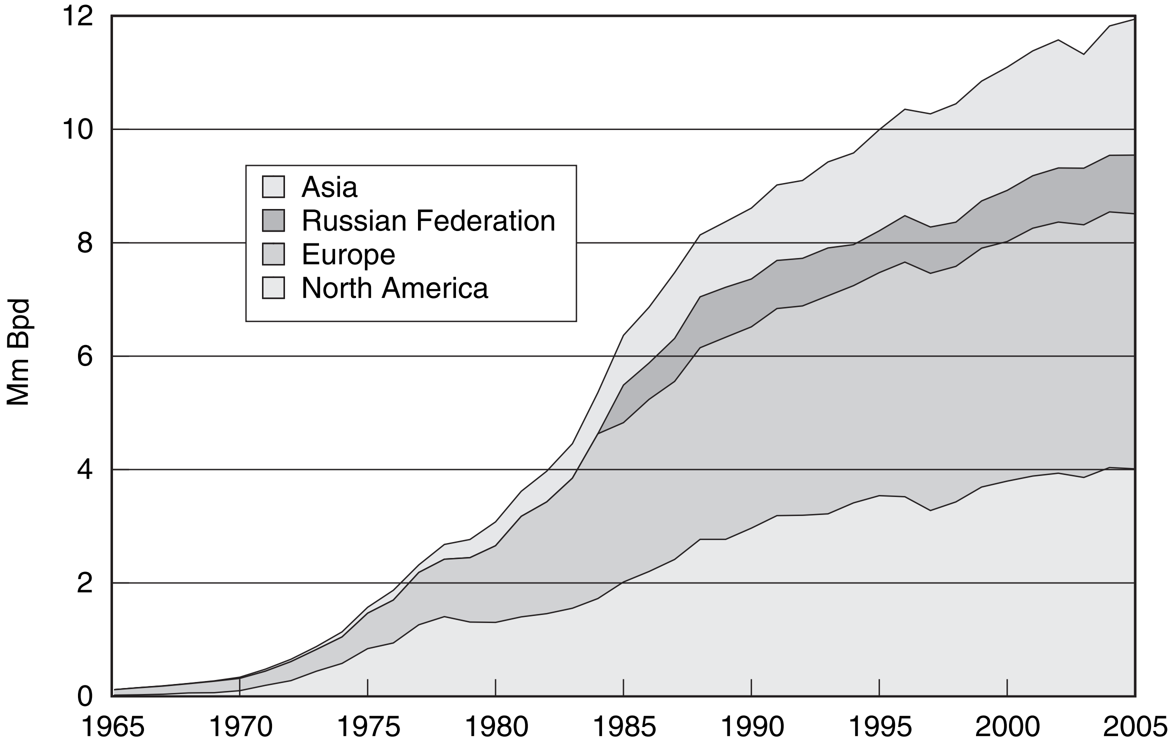
The first reactor was a small boiling water reactor (BWR) built for a nuclear submarine, a project spearheaded by Admiral Hyman Rickover in 1954. The first commercial reactor was a pressurized water reactor (PWR) built in 1957. Others were to follow, but these early reactors were really prototypes built to gain expertise to build larger plants. A BWR feeds steam directly from the reactor to the turbines that drive the generators. This introduces a low level of radioactivity to the steam turbines, condensers, and associated piping. A PWR operates under higher pressure; a heat exchanger between the reactor coolant and water in a steam generator precludes any reactor coolant from entering the steam generator, turbine, and associated equipment. The higher temperatures possible with a PWR design makes it more thermally efficient than a BWR. Most reactors in the United States are BWRs and PWRs that use light or normal water as a moderator.

Figure 8.2 shows the historical growth in nuclear power in terms of the amount of fossil fuel that would have been burned to generate the equivalent amount of electricity.[5] Generating electricity equivalent to burning 12 million barrels per day of fossil fuel is a significant reduction in carbon dioxide emissions. It has been estimated that if coal had been a substitute source of nuclear energy in the United States in 2000 there would have been an additional 2 million tons of nitrous oxides, 4 million tons of sulfur oxides, and 174 million tons of carbon emissions.

The upward sweep in nuclear power output for North America shown in Figure 8.2 did not result from building more nuclear plants. The reorganization of the electricity industry from a regulated cost-plus regime to a more liberalized competitive business environment was chiefly responsible for the higher nuclear power output. Under a cost-plus regulatory regime, there was no incentive to get more out of a nuclear power plant than what was convenient. In a liberalized competitive environment, as in the United Kingdom and the United States (and spreading elsewhere), the profit motive residing within deregulation (or liberalization) improved capacity utilization. In the case of the United States, nuclear power plant utilization increased from 65 percent in 1980 to 90 percent in 1990, from the result of better scheduling of maintenance and refueling to reduce downtime and relying more on nuclear power to take advantage of its low variable cost.

Figure 8.3 shows the amount of electricity generated in different regions of the world and

Figure 8.2 **Growth in Nuclear Power in Terms of Displaced Fossil Fuels**

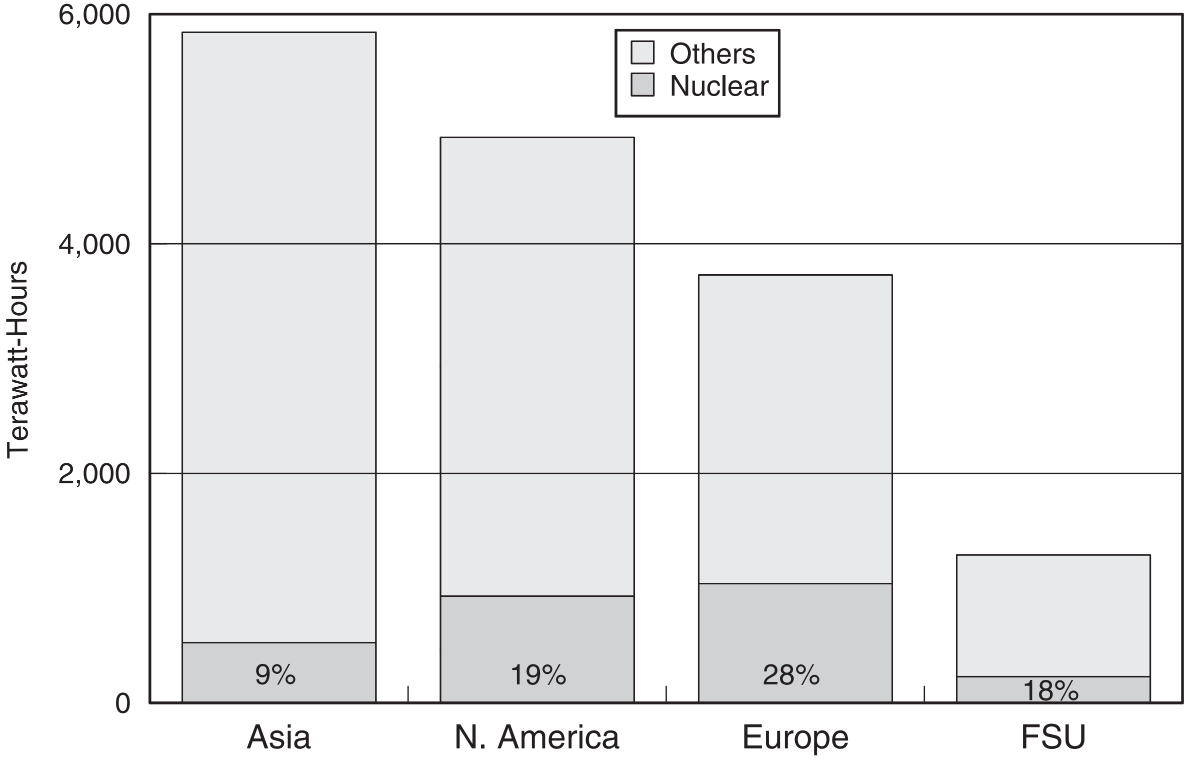


the respective role of nuclear power. Europe has the highest percentage of electricity produced by nuclear power, followed by North America, FSU (Russia and the Ukraine), and Asia.

In 2005 there were 439 operating reactors, 25 under construction, and 39 in the planning stage, with proposals for 73 more mainly in India, China, and Russia. Figure 8.4 shows the number of existing nuclear reactors plus those under construction or in the planning stage.[6]

Nuclear reactors are found in thirty nations, with half in the United States, France, and Japan. Of the 439 reactors, 263 are PWRs and 92 are BWRs. There are also twenty-six gas-cooled reactors, nineteen pressurized heavy-water reactors (popular in Canada), seventeen light-water graphite reactors (found only in FSU), and three fast breeder reactors in Japan, France, and Russia. The U.K.-designed gas-cooled reactor has a graphite moderator and carbon dioxide coolant. Carbon dioxide circulates through the core, where it is heated before passing through steam generator tubes contained within the concrete-and-steel pressure vessel. Steam from the generators passes through the pressure vessel to steam turbines that drive electricity generators. The

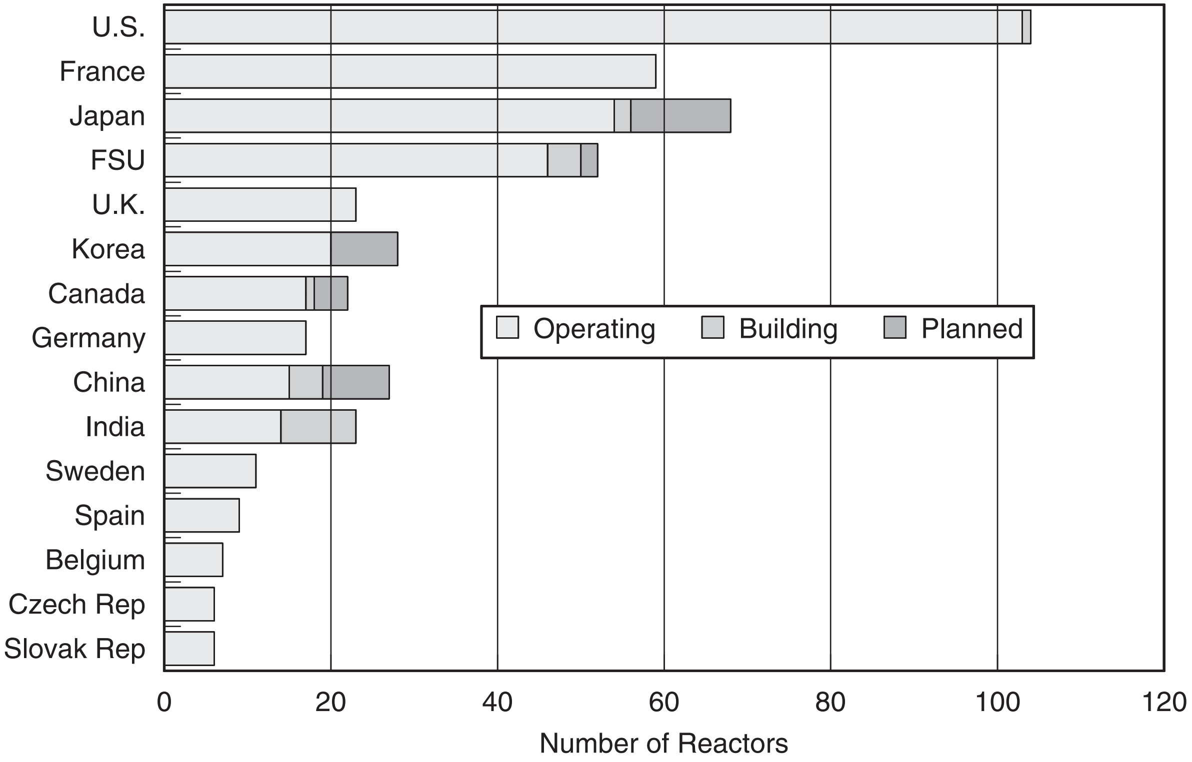
Figure 8.3 **Role of Nuclear Power in Generating Electricity**



Canadian-designed heavy-water reactors use natural, not enriched, uranium as fuel, but they require a more efficient moderator than water. Heavy water is normal water in which some water molecules have a deuterium atom (one proton and one neutron in the nucleus) rather than a hydrogen atom (one proton in the nucleus). Thus, there is a cost trade-off whether to enrich the fuel with U235 or enrich the moderator with deuterium.

Nations with the highest percentage of electricity generated by nuclear power are France and Lithuania (78 percent), the Slovak Republic (56 percent), Belgium (54 percent), Sweden (52 percent), Ukraine (48 percent), Spain (43 percent), Switzerland (41 percent), Bulgaria (37 percent), South Korea (35 percent), Czech Republic (31 percent), Finland and Germany (28 percent), Japan (26 percent), the United States and United Kingdom (20 percent), Taiwan (18 percent), and Canada (15 percent).[7] Yet, despite all the bad publicity, nuclear power is not dead. Of the twenty-five nuclear power plants under construction, nine are in India, China and Russia both with four under construction, Japan is building two, and Argentina, Canada, Iran, Korea, Romania, and the United States are each building one. Many of these plants are PWRs, while others are pressurized heavy-water reactors or advanced BWRs. Some

Figure 8.4 **World Population of Commercial Nuclear Reactors**



of these plants are improved versions of existing plants, while others represent more advanced designs.

India has six times more thorium than uranium and is aggressively advancing nuclear technology to take advantage of its ample supplies. India has inaugurated a three-stage reactor program of building a pressurized heavy-water reactor to produce plutonium. In the second stage plutonium will be the fissionable fuel for a fast breeder reactor to breed uranium 233 from thorium. In the third stage uranium 233 will be the fissionable fuel for an advanced heavy-water reactor. India, China, and Iran also have ambitious intentions to dramatically expand their nuclear power capacity.

Advances in nuclear technology have not been curtailed by Chernobyl. In 2001, the Generation IV International Forum (GIF), in which the United States, European Union, Argentina, Brazil, Canada, France, Japan, Korea, South Africa, Switzerland, and the United Kingdom participated (with Russia's collaborative interest), was chartered to examine the future of nuclear technology. India is taking an independent path to take advantage of its thorium reserves. The objective of the GIF is to obtain a standardized design for various types of nuclear reactors to expedite licensing and reduce capital costs and construction time. The intended design is to be simple and rugged, have a long life, be easy to operate, and less vulnerable to operator errors and circumstances that could lead to a nuclear accident. Technologies under consideration include a gas (helium)-cooled fast reactor in which the spent fuel is continually reprocessed to minimize disposal of long-lived radioactive wastes and a liquid lead- or lead-bismuth-cooled plant that recycles spent fuel (an adaptation of the reactors in Russian submarines). Another design involves dissolving uranium fuel in a sodium fluoride coolant that circulates through a graphite moderator, building on the experience of existing sodium-cooled fast reactors. Two other designs include supercritical water-cooled reactors capable of increasing thermal efficiency one-third higher than current light-water reactors and very high temperature gas reactors. The latter will take advantage of the experience of building and operating graphite-moderated, helium-cooled reactors in Japan, Russia, China, and South Africa.

Among the variants in new nuclear technology are pebble-bed reactors developed in South Africa. The pebble-bed modular reactor (PBMR) is a small reactor of only 110 megawatts, but a single large nuclear power plant can be made out of ten small ones built in a modular fashion with a common control center. The fuel is held in fuel spheres, not fuel rods, or pebbles about the size of tennis balls. A pebble is coated with graphite and contains 8 percent enriched uranium particles, each coated with a silicon-carbon barrier dense enough to prevent the escape of gaseous or metallic radioactive products. The reactor is loaded with thousands of pebbles, three-quarters of which are fuel with the remainder graphite, which serves as additional moderator to slow neutrons.

Helium gas is heated as it passes over the fuel pebbles, then flows through a turbine. Helium can be heated to higher temperatures than steam at the same pressure, thereby achieving higher levels of electricity output. Helium cannot burn or become radioactive or combine with other elements. Because of less friction, helium-driven gas turbines with magnetic bearings require less maintenance than steam turbines. Whereas conventional light-water reactors are shut down every eighteen to twenty-four months to refuel (Canadian heavy-water reactors can be refueled while in operation), fuel pebbles can be continually added to the reactor core from the top and removed from the bottom. Each fuel pebble will pass through the reactor about ten times over a three-year period before its fuel is depleted. When depleted, the sealed spent pebbles can be safely stored in lead containers. Between ten and fifteen total fuel loads will be required over a reactor's lifetime.

PBMRs are inherently safe because high temperatures will stop the chain reaction. The PBMR is designed with a low fuel density and small size, allowing heated atoms to spread apart, making it more difficult for an incoming neutron to strike a nucleus. The peak attainable temperature of 1,600°C within a pebble is below the 2,000°C melting point of the silicon-carbide coatings. The reactor is inherently safe in that there is no way for someone, inadvertently or otherwise, to pull out the rods and bring the reactor to a supercritical state because there are no rods. No expensive containment system, such as that required in a water-cooled reactor, is needed to contain a piping rupture as the coolant is a gas. In short, the accidents at Three Mile Island and Chernobyl cannot occur with a PBMR. Its simplicity of design and operation, the low cost of construction (including the possibility of using standardized, mass-produced components), and its inherent safety (because a core meltdown is physically possible) are beginning to attract attention. China is considering the possibility of building large numbers of PBMRs, each of which would serve a community.[8]

Building small nuclear power plants in a distributive fashion reduces the need for long-distance transmission systems. Small nuclear power plants already serve local communities in remote areas of Siberia and at permanent research stations in Antarctica, supplying electricity and hot water for heating. Small reactors of less than 300 megawatts are being developed in the United States, Russia, France, Japan, China, Korea, and Argentina. If successful, this means that a community may have its neighborhood nuclear power plant right next door to the local high school.

Fusion Power

Whereas fission is the splitting of heavy atoms, fusion is the uniting of light atoms. The sun and other stars produce heat when hydrogen atoms fuse to form helium, transforming matter into energy. Thus, for fusion to work on Earth, an environment equivalent to being in the center of the sun has to be created, requiring temperatures over 15 million degrees Celsius and pressures over 340 billion times greater than atmospheric pressure. Hydrogen fusion on Earth is obviously quite a technological challenge, but fusion of deuterium and tritium, isotopes of hydrogen, is less demanding than hydrogen. Deuterium can be extracted from seawater and tritium is a by-product of fission. The challenge is to design a magnetic field strong enough to contain plasma, a heated mix of electrons and ions, under conditions conducive to fusion (100 million degrees Celsius, much hotter than the center of the sun, to compensate for the sun's much higher pressure).

Neutrons are produced when fusion takes place and become a source of heat when trapped in a stainless steel containment vessel wall. This heat is transferred to water to produce steam to run an electricity generator. Once fusion is triggered, it has to be controlled and kept self-sustaining by adding more fuel from a surrounding blanket of lithium in which neutrons react with lithium to produce tritium and helium. Leakage of plasma from the magnetic field is a major problem because this can stop the fusion process. So far more energy (electricity) is consumed to maintain the plasma than is extracted from fusions. Even when these technical challenges are overcome, it is estimated that half of the electricity produced by fusion will be consumed to contain the plasma within the magnetic field.

The fusion process is inherently safe. A hydrogen bomb environment cannot be created because any "runaway" condition stops the fusion process by removing the plasma. The trick is how to keep the plasma together long enough for fusion to occur. Alternative approaches to magnetic confinement as a means of trapping the hot plasma are lasers or particle beams. The energy source for fusion is virtually inexhaustible. Radioactivity is limited to high-energy neutron bombardment of the containment system. This radioactivity is short-lived (100 years) compared to the radioactivity of a fission reactor (thousands of years). An additional health hazard is the possibility of tritium leaking into the environment. Tritium, with a half-life of 12.4 years, is easily absorbed by the human body and, once ingested, remains a serious threat to human health for a long time. The advantage of the deuterium-deuterium fusion process is that no tritium is involved.

In 1989 there was great excitement over the possibility of cold fusion, creating energy in a test tube (so to speak), which turned out to be either a case of vain hope or scientific sleight of hand. Research in nuclear fusion is being conducted in the United States, Russia, various European nations, Japan, Korea, China, Brazil, and Canada.[9] In 2005, France was selected as the host nation for a \$10–\$13 billion experimental nuclear fusion reactor, the

ITER (International Thermonuclear Experimental Reactor), to be funded by the European Union, the United States, Japan, South Korea, Russia, and China. Its goal is to produce 500 megawatts of power for hundreds or thousands of seconds at a time. Construction is to begin in 2006, with completion scheduled for 2013. It may take as long as twenty-five years before an acceptable design for a commercial fusion plant can be developed.

Future of Nuclear Power

In spite of new plants under construction, most expect nuclear power output to level off as older plants are phased out, something already in progress. It is possible that retirements could be accelerated over safety or cost concerns. Even if one accepts the more optimistic projection of some growth in nuclear power from plants under construction and others in the planning stage, earlier visions of the anticipated contribution of nuclear power to worldwide electricity generation have missed the mark by a very wide margin.

Yet, at the same time, demand for electricity continues to rise, although not as rapidly as in the past. In the 1950s annual growth averaged 8.7 percent, 7.3 percent in the 1960s, 4.1 percent in the 1970s, 2.6 percent in the 1980s, and 2.1 percent in the 1990s. While electricity consumption and economic activity are still closely linked, the energy required to produce a unit of gross domestic product (GDP) has fallen as the result of gains in efficiency and from the shift from manufacturing to service-based economies in developed nations. Nevertheless, electricity consumption is still expected to climb. Even a modest 1.8 percent annual growth would require the addition of about 355 gigawatts of new electricity-generating capacity net of plant retirements in the coming decades. Roughly 300 new electricity-generating plants of about 1.2 gigawatts each will have to be constructed to meet this demand.[10] Large plants of this order of output are invariably coal-fired or nuclear powered. There are plentiful domestic reserves of coal to meet the challenge, but coal has its environmental problems, unless clean-coal technology takes hold. Although wind will play some role, it is unrealistic to expect wind power to fulfill a significant portion of this shortfall in base-load demand.

It is difficult to imagine building this much electricity-generating capacity with no contribution from nuclear power. Moreover, the hydrogen economy will require large numbers of nuclear power plants to produce hydrogen through electrolysis of water. Producing the requisite electricity by burning fossil fuels defeats the whole purpose of the hydrogen economy, which is to do away with carbon dioxide emissions (unless they can be sequestered). The potential generation of electricity from carbon dioxide-free hydro, wind, solar, and geothermal energy sources is not even close to meeting the demands of the hydrogen economy.

The public should become aware that something has to be done before the lights go out. Those opposed to large electricity-generating plants, be they nuclear or coal, do not have a viable alternative other than wind and solar power. Certainly these should be encouraged, but even here environmentalists stopped the building of wind farms off the coasts of Massachusetts and Long Island. Yet no one, including the environmentalists, is advocating letting the lights go out. If nuclear power plants are to play a role in satisfying the demand for electricity, a technology should be selected that makes it possible to reduce capital costs by building a large number of essentially identical plants. The learning curve of building standard designed nuclear plants can generate further cost savings by eliminating the mistakes and inefficiencies associated with the construction of the first plants in a series. Siting and licensing have to be streamlined and a cadre of well-trained operators has to be created. Having the same basic nuclear plant design would ease training requirements and allow operators to be easily transferred from one plant to another. Moreover, plants of a standard design are not only cheaper to build but also less expensive and safer to run because equipment, skills, and experience can be shared from one plant to the next.[11] However, the problem of disposing of spent fuel has to be resolved and serious consideration should be given to reprocessing to reduce the quantity of radioactive waste and extend the effective life of uranium reserves. What is needed is public support for nuclear power, which can only come about if doubts over safe operation can be resolved.

The U.S. Energy Policy Act of 2005 provides significant incentives for ending the thirty-year moratorium on licensing new nuclear power plants. The most important is a 1.8 cents per kilowatt hour production tax credit for the first eight years of plant operation. This is a substantial tax credit, considering that the average U.S. retail price of electricity was 7.5 cents per kilowatt hour for the first quarter of 2005. A tax credit is not a cash subsidy, but a reduction in tax payments that the government would otherwise receive. The tax credit is for a maximum of 6 gigawatts of new plant capacity (six new plants of 1-gigawatt or 1,000-megawatt capacity) of not more than three separate designs. Financial support for certain specified delays caused by litigation or delayed Nuclear Regulatory Commission (NRC) approvals is also available.

Hydropower

Dams have a long history of supplying water to meet human needs. Ancient dams in Jordan, Egypt, Yemen, Greece, and Turkey were built to supply water for human and animal consumption, irrigate crops on land too dry to sustain agriculture, and flood control; the same purposes for building dams now. What is new is using hydropower to generate electricity. A few of these ancient dams have been in more or less continual operation for two or more millennia. Waterwheels turned by running water have lifted water for irrigation and ground grain since Roman times; a definite improvement over tread wheels operated by humans or animals. The first waterwheels were horizontal and drove a vertical shaft to rotate millstones that ground grain on a floor above the waterwheel. Vertical waterwheels were vastly superior to horizontal waterwheels because they could more efficiently translate the momentum of moving or falling water into power. Gearing was now necessary to change the direction of a rotating shaft from horizontal to vertical in order to operate millstones, something that was not always technically feasible. Over the centuries waterwheels were applied to a variety of tasks such as sawing wood, crushing ore, stamping, cutting, grinding, polishing, and powering bellows to force air into a furnace to refine metals. In the 1680s a large installation of waterwheels pumped water to supply the fountains at the palace at Versailles. Factories in England and New England, the first centers of industrialization, continued to be powered by waterwheels long after the invention of the steam engine. Waterpower had the virtue of being free, but steam from burning coal eventually overtook waterpower in the nineteenth century because steam could deliver a lot more power with greater reliability.[12]

There are 45,000 dams in the world with a vertical distance of fifty feet or more. These dams catch 14 percent of precipitation runoff, provide 40 percent of water for irrigated land and more than half of the electricity for sixty-five nations.[13] Of these, 150 dams are considered major dams in terms of generating electric power, reservoir capacity, and height. As a group they generate 40 percent of the energy produced by hydropower, but not all dams generate electricity. Some are built to provide water for some combination of human consumption and recreation, irrigation, and flood control. For those that do generate electricity, dams raise the level of water to create a hydraulic head to power electricity-generating turbines. Reservoirs compensate for fluctuations in the inflow and outflow of water. Inflow is determined by the amount of rainfall in a dam's watershed. Spillways and gates control the discharge of excess water from the reservoir while intake valves control the flow of water through a tunnel (penstock) to the hydraulic turbines that drive the electricity generators. Long-distance transmission lines are generally necessary as many dams are located far from population centers. A few dams have locks that allow ships to pass around them and others have steps or ladders to allow fish to get to and from their spawning grounds.

The principal advantage of hydropower is that it utilizes a renewable source of energy without pollution and is free of cost. While hydropower has no fuel cost and a low operating cost, it has a high capital cost and is site-specific. Unlike fossil-fueled plants, hydropower dams are not built where they are needed. Prospective dam sites require ample supplies of water plus favorable geological conditions suitable for building a dam whose reservoir is sufficiently large with a bottom that limits water absorption. The capital cost of a dam includes the preparation of a site, the construction of the dam, and the installation of an electricity-generating plant and long-distance transmission lines. From a fuel standpoint, hydropower is environmentally friendly, but other environmental concerns still have to be addressed such as the impact of dams on fish and wildlife, resettlement of people living upstream of the dam, and the potential of catastrophic structural failure for those living downstream.

There are 45,000 dams, and some fail every year, mostly without catastrophic results other than local flooding. The 1889 Johnstown flood was caused by the failure of the South Fork Dam, with a loss of over 2,200 lives. The rich folk living along the reservoir, which served purely as a recreational lake, did not bother to spend the money necessary to fix a deteriorating dam. Nor were they held financially responsible for the consequences of their neglect. In 1928, the two-year-old St. Francis Dam in California failed, leaving more than 450 dead. This occurred twelve hours after the builder (always a good source for an unbiased opinion) declared the dam safe, even though water was passing through the dam in spots where it was not supposed to. The cause of the dam failure turned out to be the unsuitable geology of the site. In 1975, unprecedented rainfall caused the Shimantan Dam in China to fail, and its floodwaters destroyed the downstream Banqiao Dam. The combined deluge of water and dam debris carried away other downstream dams and dikes, drowning over 85,000 people. In 1976, after seven months of filling the newly constructed Teton Dam in Idaho, with the reservoir only three feet below the spillway, three leaks were found: one at the bottom of the gravel-filled cement dam, another alongside one of its abutments, and still another about 100 feet below the top of the dam. Less than two hours later, the dam was breached and water poured through the dam. In a matter of hours the breach had widened, carrying away a large portion of the dam and emptying a seventeen-mile-long reservoir over a wide area of Idaho. Fortunately, only fourteen lives were lost.

Heavy rainfall can cause dams to fail, but dams are also affected by a lack of rainfall or a drought that fails to replenish the waters they hold. The California energy crisis in 2001 was started by a drought in Oregon and Washington that curtailed the export of hydroelectricity to California. That same year Brazil suffered power disruptions from a drought that significantly cut hydroelectricity generation, the primary source of that country's electricity.

Unlike other forms of energy, electricity cannot be stored. Electricity capacity must be able to meet peak demand without consumers experiencing brownouts or blackouts. Batteries cannot store sufficient quantities of electricity to smooth out the operations of an electric utility by supplementing supply when demand is high and being recharged when demand is low. Hydropower provides a way to "store" electricity through pumped storage plants. These plants have reversible pump-turbines that pump water up to a storage reservoir during periods of low demand. During periods of high demand, water flows from the storage reservoir through reversible pump-turbines to generate electricity. Motors that pump water to the storage reservoir become generators to produce electricity. Pumped storage plants reduce variability in electricity demand by pumping water to the reservoir during periods of low demand and by generating electricity during periods of high demand. This increases the base-load demand and reduces the need to invest in costly peaking plants.

The first commercial site for generating electricity was New York City's Pearl Street station, built by Thomas Edison in 1882. The plant produced direct current electricity from generators driven by coal-burning steam engines and was the progenitor of other plants to electrify the city. The second commercial site for generating electricity was Niagara Falls, where a hydropower plant built by George Westinghouse produced alternating current electricity. Construction of a tunnel to divert water upstream of the falls to a downstream power plant began in 1890. Commercial sales started in 1895 and the plant's generating capacity was continually expanded until the 1920s. With the increasing availability of electricity generated from hydropower, industry rapidly developed along the Niagara River.

The first recorded public outcry over the environmental consequences of energy was the ban on burning coal in London during the thirteenth century. At the turn of the twentieth century, New Yorkers demonstrated against black smoke emissions from the early electricity-generating plants. While the environmental movement can be traced back in time to a number of such public outcries over polluted air and water, the major thrust that propelled environmentalism to the forefront of public awareness was a dam powered by one of the cleanest sources of energy.

The Saga of the Hoover and Glen Canyon Dams

The Hoover and Glen Canyon dams mark the beginning and the end of a dam-building spree in the United States. When built, the Hoover Dam ranked first in the world in size and power generation. Although the Glen Canyon has the same electricity-generating capacity as the Hoover Dam, and is similar in size and structure, a few far larger dams were built in the thirty-year interim separating the two.[14] The Hoover dam was built during the Great Depression in the 1930s to jump-start the U.S. economy as were dams built in Appalachia under the Tennessee Valley Authority (TVA). Other major dam projects were the Shasta Dam across the Sacramento River and the Grand Coulee Dam across the Columbia River. The Shasta and Grand

Coulee dams supply water for irrigation and flood control, but of the two only the Grand Coulee Dam generates electricity; more than the combined output of the Hoover and Glen Canyon dams.

The Hoover and Glen Canyon dams straddle the Colorado River, discovered by Coronado in 1540 in his quest for the fabled seven cities of gold (actually Cardenes, a member of Coronado's party, was the first to discover the Colorado River from the rim of the Grand Canyon). Coronado named the river after the Spanish word for "red," the color of the silt-laden river. Coronado did not explore the Colorado River; in fact, the Colorado River presented an insurmountable barrier to further exploration. Exploring the river would not take place for another 300 years, when a daring individual led the first recorded expedition down the river.

The Colorado River falls 14,000 feet from the Rocky Mountains to sea level in the Gulf of California and carries more silt than any other river in the world, including the "muddy" Mississippi. The original time estimate for Lake Powell, the reservoir in back of Glen Canyon Dam, to fill up with silt was 400 years, but this was subsequently revised to 1,000 years by later estimates of the silt-capturing capacity of other dams upstream of where the Colorado River enters Lake Powell. The primary advantages of the Colorado River from the point of view of dam building are that the river flows through a canyon whose geology is ideal for damming and through a region desperate for water. The disadvantage of the Colorado is its relatively low average water flow, which varies from a summer trickle to a springtide flood that carries away the snowmelt of a large area of the Rocky Mountains.

In the early part of the twentieth century, the original idea was to build a dam at Glen Canyon first, followed by three more downstream dams whose construction would be made easier by building the upstream dam first. The problem was that the Glen Canyon reservoir would serve Arizona, which had a small population at the time. Population growth was centered in California, and by the 1920s it was clear that further development hinged on having an adequate and dependable supply of water to support agriculture and urbanization. California politicians prevailed at deliberations as to where to build the first dam: It would be built at Boulder Canyon, whose reservoir water could be easily diverted to California. It was understood at the time that another dam would eventually be built to serve Arizona.

The name Boulder Dam stuck after the original site was changed to a better location in nearby Black Canyon, about thirty miles southeast of Las Vegas. Boulder Dam was renamed Hoover Dam in 1930 after the president who authorized its construction. In 1933, New Deal bureaucrats decided that the world's most monumental dam project should not be named after the president who presided over the onset of the Great Depression and changed the name back to Boulder. The dam was completed in 1936 and another six years were to pass before its reservoir, Lake Mead, was filled. In 1947, a Republican-controlled Congress under President Truman passed a law to reinstate the name Hoover.

Dams and other capital-intensive projects cannot be funded from private sources; too much money is at risk. The risk private investors shun is accepted by the government because the risk of loss can be spread among the taxpaying public. Moreover, government cooperation is needed for land condemnation to clear the way for the reservoir, particularly when much of the land is already in the public domain. The responsibility for dam building fell under the auspices of the Bureau of Reclamation of the Department of the Interior. "Reclamation" was interpreted to mean "reclaiming" unproductive land for agricultural use by building dams to provide water for irrigation. Earlier reclamation projects were financial failures because the revenue from growing crops on irrigated land fell far short of justifying the cost of building a dam. It was the discovery that dams could also generate electricity that changed the financial equation in favor of dam building. The Department of Interior was also the administrative home for the Bureau of National Parks Service, charged with preserving and protecting wilderness areas, and the Bureau of Indian Affairs, which establishes and administers American Indian reservations. One bureau built dams whose reservoirs, at times, submerged lands set aside by a sister bureau to preserve wilderness areas or by another to establish American Indian reservations. Talk about dichotomy of purpose!

There are marked similarities between the Hoover and Glen Canyon dams. Both generate 1.3 million kilowatts or 1,300 megawatts or 1.3 gigawatts of output, enough electricity to supply a U.S. city of 1 million people. Both rise 587 feet above the riverbed, although the Hoover dam is taller by sixteen feet when measured from bedrock. Like most dams, both had huge tunnels built around the dam site to divert the waters of the Colorado River at full flood during dam construction. These were eventually plugged when the dams were completed to start filling the reservoirs, although both have diversion tunnels to reduce excessively high reservoir levels. Each required the building of a new town for the construction workers, one that started out as a disorganized tent city at the Hoover Dam site and the other an equally disorganized trailer park at the Glen Canyon Dam site. Tents and trailers were eventually replaced by carefully laid-out company towns for the construction workers and both survived the completion of the dams as Boulder City, Nevada, and Page, Arizona.

The reservoir behind Hoover Dam (Lake Mead) holds more water, but the reservoir behind Glen Canyon Dam (Lake Powell) has double the miles of shoreline and its length extends seventy-six miles farther upstream than Lake Mead. Lake Mead was named after Elwood Mead, a commissioner in the Bureau of Reclamation. Lake Powell was named after John Wesley Powell, the one-armed Civil War veteran who in 1869 successfully led the first recorded expedition of ten men in four boats down the Colorado River. Although Powell did mention developing the area, along with the need for preserving its natural beauty, what he had in mind in terms of development was far different than the development posed by the lake that bears his name. Mead is a fitting name for a dam's reservoir; Powell is not.

Both dams were built in a similar fashion—in blocks, the smallest being the size of a house. One-inch copper pipes for pumping refrigerated water through the wet cement were incorporated in the construction of the dam to speed up curing from an estimated 150 years to nineteen or so months. Most dams built before the Hoover dam were gravity dams; pyramidal in shape (thick at the bottom and narrow at the top), so that the weight of the dam held back the water. They were commonly cement or masonry on the outside and filled with rock or gravel. Arch dams of pure concrete or masonry, first built in the late nineteenth century, were thin in comparison with gravity dams. A gravity dam depends on its massiveness to hold back the pressure of the water in a reservoir whereas the arch dam transfers the pressure on the dam to thrust on the canyon wall abutments. The Hoover and Glen Canyon dams were an innovative combination of both the gravity and arch designs. While curved, they are still pyramidal, thick at the base and narrow at the top. While similar, there are differences between the two. The intake towers at

Hoover Dam were built on the canyon walls and tunnels (penstocks) were cut through the canyon walls for water to flow to the turbines. The intake towers and penstocks were incorporated within the Glen Canyon Dam. One can drive across Hoover Dam, but there is a bridge for vehicle traffic alongside Glen Canyon Dam, whose construction was a feat in itself.

Parenthetically, Las Vegas was built on the electricity generated by Hoover Dam. The gangster Bugsy Siegel saw "easy-going" Nevada, with its legalized gambling, as a land of opportunity, and built the first gambling palace, the Flamingo. Bugsy saw before others that Hoover Dam could supply cheap and plentiful electricity for air conditioning and lights and water for casino fountains built in the middle of a hot, dry, inhospitable desert. The Flamingo was the first step in transforming a backwoods desert town into the gambling Mecca of the world and one of the fastest-growing cities in the United States.

Glen Canyon Dam, started in 1958, was completed four years later when the gates to the lower tunnel to begin filling Lake Powell were closed. While the reservoir was filling, much work remained: The generators and transmission lines had to be installed and the tunnels that diverted the flow of the Colorado River had to be permanently sealed. The fill rate was affected because a minimum quantity of water must flow through Glen Canyon Dam to ensure an adequate supply of water to Lake Mead, which in turn supplies water to California and powers Hoover Dam's generators. With light snowfall in the Rockies in 1963 and 1964, Lake Mead was rapidly dropping while Lake Powell was hardly filling. The return of normal snowfalls speeded up the fill, and in 1973 a court injunction temporarily stopped the filling of Lake Powell when reservoir water was about to invade the Rainbow Bridge National Monument. A congressional law was subsequently passed that allowed water to flood land previously set aside as part of a National Monument, violating a prior agreement with environmentalists that allowed Glen Canyon Dam to be built. In 1980, seventeen years after the completion of the dam, Lake Powell was finally filled. It covers 252 square miles, is 186 miles long, and has 1,960 miles of shoreline. Considering the area and the length of Lake Powell, its average width can only be slightly over a mile of flooded canyons.

Hoover Dam was planned in the late 1920s in response to California developers who saw a lack of water as an impediment to further development of agriculture and urban areas. By harnessing a mighty river for the common good—making deserts bloom, lighting cities, and providing power to industry and commerce—Hoover Dam was "concrete" proof of America's engineering skill and industrial might. No one opposed the building of the Hoover dam. Supporters included the federal government, via the Bureau of Reclamation, private construction companies, and California politicians and developers. Thirty years later the Glen Canyon dam was also viewed favorably by the same coterie of supporters, except the politicians and developers were from Arizona. But a new entity was involved: the first environmentalist group to capture the nation's attention, the Sierra Club.

The Sierra Club was formed in the late nineteenth century by John Muir, a naturalist, to preserve the Sierra Nevada mountains in their original pristine condition. Ever interested in preserving nature, Muir persuaded Theodore Roosevelt to declare a portion of the Grand Canyon as a national monument, at the same time chiding Roosevelt for his habitual trophy-hunting of game animals. Sierra Club members were mainly conservative businessmen and academics dedicated to preserving the wilderness areas of the high Sierras. After

Muir lost a fight to prevent building a dam on a national preserve in the Sierras, the Sierra Club vowed that they would never allow this to happen again—in the Sierras. The transformation of the Sierra Club to openly fighting for conservation and preservation of wilderness beyond the Sierras started in 1949 when the Bureau of Reclamation publicized its intention to build a dam across the Colorado in Dinosaur National Monument. This marked the beginning of a dramatic change in the makeup of the membership of the Sierra Club, from one of conservative businessmen and academics to a more politically active constituency that advocated the preservation of the wilderness and conservation of natural resources far beyond the high Sierras.

To its everlasting regret, the Sierra Club acquiesced to the building of the Glen Canyon dam on condition that no more dams would be built in national parks and that something would be done to prevent flooding the Rainbow Bridge National Monument. The ban on dams in national parks also included two more intended for the Grand Canyon between the Glen Canyon and the Hoover dams. These dams (Bridge Canyon and Marble Canyon) were to be smaller in size and less intrusive than their larger counterparts. They were intended to generate electricity to pump water over intervening mountains to direct the flow of water from Lake Powell to Tucson and Phoenix. With the agreement not to build these dams, a substitute source of electricity was needed. It was first proposed that a nuclear power plant be built (this was in the 1960s, when nuclear power was considered safe and cheap). In the end the Navajo Generating Station was built near Glen Canyon with a 2.5 gigawatt output, about equal to the combined output of the Hoover and Glen Canyon dams. The plant, started in 1970 and completed in 1976, is fueled by Black Mesa coal strip-mined on Navajo reservation land and shipped in by rail. It is ironic that the environmentalists' success in preventing the building of two clean and sustainable hydropower dams led to the building of one of the world's largest coal-burning plants that spews carbon dioxide and other emissions into the atmosphere. It is also ironic that current attempts by environmentalists to dismantle Glen Canyon Dam ignore that Lake Powell draws far more tourists than Yellowstone and Yosemite Parks. The debate that continues to this day over Glen Canyon Dam has prevented any other large hydropower projects in the United States from moving ahead.

The building of the Glen Canyon dam marks a watershed in the change of attitudes toward large-scale industrial development. Once viewed as signs of the improvement of humanity's material condition, dams became viewed as an irretrievable loss of wilderness. The Sierra Club gave birth to innumerable environmental groups dedicated to stopping not only dams but just about anything that can be stopped: from oil refineries in Texas to wind farms off the coasts of Massachusetts and Long Island. Environmentalists maintain that building one dam leads to the building of another because the industrial and agricultural development allowed by the construction of the first dam creates the demand for electricity and water to justify building a second, then a third, and a fourth, and so on until the entire wilderness is submerged in reservoirs.

This phenomenon of progress creating its own demand was first observed when Robert Moses built a parkway on Long Island to give New Yorkers easy access to the "country." Once built, so many New Yorkers moved to suburbia that the subsequent highway congestion created a demand for a second parkway. This opened access to other parts of Long Island, creating more urban sprawl, more road congestion, and the need for building yet another parkway until, presumably, all of Long

Island would eventually be paved over. The same is true for power plants: building one allows a community to expand in population, commerce, and industry until there is a need for another. As one community experiences the economic benefits of a power plant, others copy it and the process continues until the nation is covered with power plants and the horizon cluttered with transmission lines. This is one of the chief complaints of environmentalists: progress continues until the last vestige of natural life is irretrievably lost. What the alternative vision of life under the rule of environmentalists would be like is left largely unanswered.

Aswan High Dam

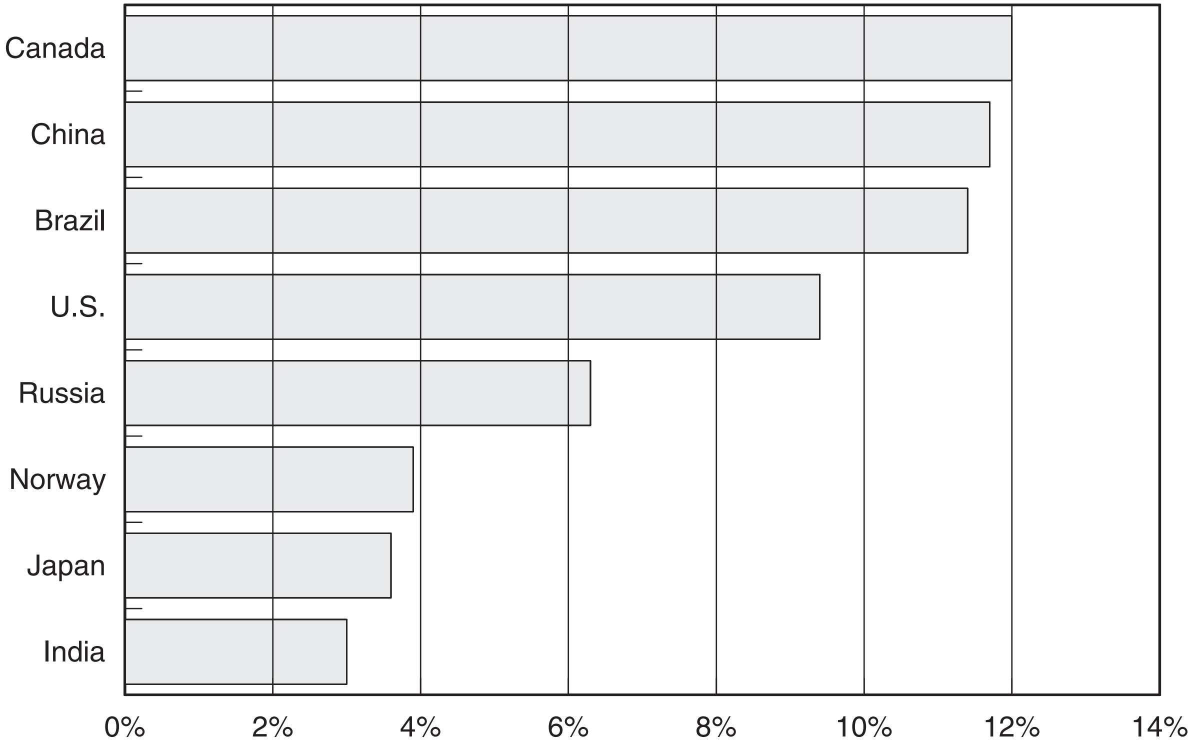
The environmental consequences of the Aswan High dam best exemplify what environmentalists fear most: The consequences are largely unknown before something is built; once built, little can be done to counter them. The first Aswan dam, built in 1889 when Egypt was under British control, was to irrigate cash crops such as cotton. The height of the dam was increased in 1912 and 1933 to enhance its water storage capacity. The sluice gates of the original Aswan dam were opened during the flood season to let the floodwaters proceed unimpeded downstream. As the flood season neared its end, the sluice gates were closed, trapping water behind the dam for crop irrigation.

The Nile flood originates in the Ethiopian highlands, the source of the Blue Nile, during the monsoon season. Silt deposited by the floodwaters formed a thick, fertile layer of alluvium that made the Nile valley and delta one of the most productive agricultural regions on Earth. After the Egyptian revolt in 1952 brought Nasser to power, the Soviet Union sponsored the building of the Aswan High dam, five kilometers long, one kilometer wide at its base, and rising 107 meters in height. This dam, called the Pyramid for the Living by President Nasser, permanently stopped the annual flooding of the Nile valley and delta.

The dam was supposed to be a major source of hydroelectric power for Egypt, but unfortunately this potential was never fully realized. Lake Nasser did not rise to its anticipated level because of its high rate of evaporation, the water diverted to irrigate cropland, and possibly leakage through the reservoir bottom. Electricity was necessary, not only to supply the needs of the people, but also to provide energy for the production of fertilizer as a substitute for the alluvial deposits formerly left behind by the annual floods. The alluvial deposits were free, but fertilizer is not. In addition to affecting the productivity of the Nile valley and delta, agricultural land has been lost by erosion of the Nile delta by the Mediterranean Sea, which had previously been replenished by the annual inundations. Penetration of saline waters from the Mediterranean into the Nile delta further decreased productivity and reduced the local fish population. Agricultural land upstream of the dam, now part of Lake Nasser, was lost, along with the livelihoods for 120,000 Nubians, who had to be resettled, but this was more than made up by bringing into production other lands bordering on Lake Nasser.

There also appears to be a correlation between Lake Nasser's water level and earthquake activity. Some geologists feel that the weight of Lake Nasser is affecting underlying faults; a phenomenon that has been observed at other dam sites. The sediments that once fertilized the Nile delta now accumulate in the bottom of Lake Nasser, over time reducing the volume of irrigation water stored in the lake. The presence of large bodies of water behind dams can affect the local climate, although this can be benign. The Aswan High dam has also been blamed for the spread of schistosomiasis, a parasitic disease that leads to chronic

Figure 8.5 **World's Largest Hydropower Producers** (Percentage of Total Output)

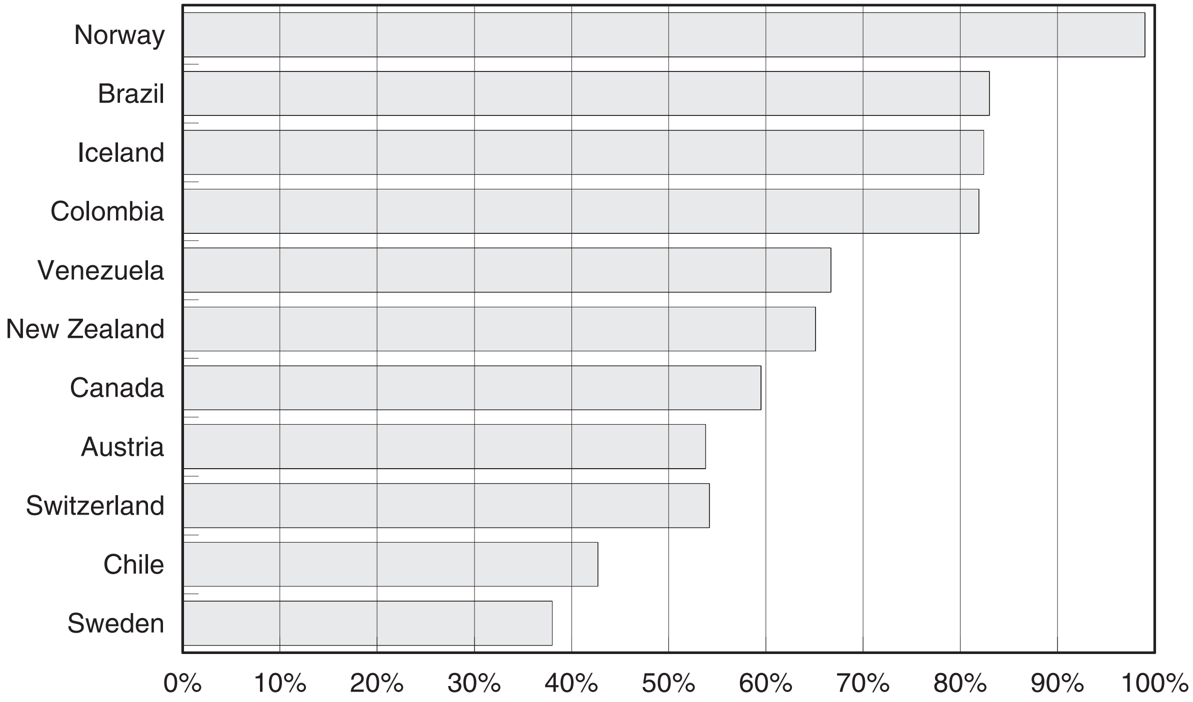


ill health that has also been associated with other large-scale water development projects. Where once the annual inundation of the Nile flushed the delta and river of snails that carry the parasite, now the snails are moving further upstream and affecting larger numbers of people.[15]

This avalanche of environmental objections over the building of the Aswan High dam has to be counterbalanced by what the proponents say. They point out that the water in Lake Nasser saved Egypt from famine in 1972 and 1973 and maintained its agricultural output during nine successive years of drought between 1979 and 1987. Lake Nasser has provided irrigation for enough new land to be brought into cultivation and to partially support a doubling of the population; but not quite enough to prevent Egypt, once a net food exporter, from becoming a net food importer. Moreover, the dam protected the Nile valley from major floods in 1964, 1975, and 1988.[16] Figure 8.5 shows the world's most important hydropower producers in 2004.[17]

Hydropower provided a significant portion of pre–World War II electricity-generation capacity in the United States. Since World War II, fossil-fueled and nuclear generating plants were built in large numbers, pushing hydropower to the background. North American hydropower development is now centered in western and eastern Canada. Hydropower plants in eastern Canada are built and operated under the auspices of Hydro-Québec,

Figure 8.6 **Nations with the Greatest Dependence on Hydropower** (Percentage of Total Electricity Generation)



with 53 hydropower generating stations encompassing 560 dams and 25 large reservoirs with an installed capacity of 31.6 gigawatts.[18] The company has access to another 5.4 gigawatts of output at Churchill Falls. Altogether 37 gigawatts of hydropower can serve a population of about 30 million people. Hydro-Québec also has another 2.3 gigawatts of output from a nuclear and several conventional plants, plus plans to build a 2-gigawatt wind farm.

The company owns and operates a 32,500-kilometer transmission system with interconnections with other transmission systems in Ontario and Québec Provinces, the Midwest, Middle Atlantic, and New England states. This arrangement helps to even out the base load where the winter peak to heat homes and office buildings in Canada is balanced by a summer peak to cool homes and office buildings in the United States. The company has spearheaded technological advances in long-distance transmission to reduce transmission losses; an imperative considering the remote location of its hydropower plants, and shares its expertise by getting involved with hydropower projects in other lands. Figure 8.6 shows those nations with the greatest dependence on hydropower for electricity generation.

Norway is almost entirely dependent on hydropower for electricity generation. Brazil, with a noteworthy 83 percent dependency on hydropower, had a national energy policy to become entirely dependent. In pursuit of this

Table 8.2

World's Largest Dams in Terms of Electricity Generation

Name	Location	Rated Capacity (GW)	Start of Operations
Itaipu	Brazil/Paraguay	12.6	1983
Guri	Venezuela	10.0	1986
Grand Coulee	U.S.	6.5	1942
Sayano-Shushensk	Russia	6.4	1989
Krasnoyarsk	Russia	6.0	1968
Churchill Falls	Canada	5.4	1971
La Grande 2	Canada	5.3	1979
Bratsk	Russia	4.5	1961
Moxoto	Brazil	4.3	1974
Ust-Ilim	Russia	4.3	1977
Tucurui	Brazil	4.2	1984

objective, the Itaipu hydroelectric power project was built between 1975 and 1991 with eighteen generating units for a total output of 12.6 gigawatts of electricity, enough to supply the needs of 13 million people. This single dam has an output equivalent to about thirteen nuclear or large coal-power plants of 1,000 megawatts or 1 gigawatt each. Itaipu is a bi-national development project on the Paraná River between Brazil and Paraguay, not far from the border with Argentina, and provides 25 percent of the electricity supply in Brazil and 78 percent in Paraguay. The height of the dam is 643 feet (196 meters) with a length of nearly 5 miles (7.8 kilometers), and a reservoir 106 miles (170 kilometers) long. The dam has become a major tourist attraction as a construction marvel, much like Hoover and Glen Canyon dams. While being built, wet cement was refrigerated before pouring to decrease the setting time rather than installing refrigerated water pipes as was done at the Hoover and Glen Canyon dams. Whereas it took seventeen years to fill the Glen Canyon reservoir, the Itaipu reservoir was filled in a matter of weeks: The water rose so fast that an intensive effort had to be made to save animals from drowning.

Brazil's dream of achieving full reliance on hydropower for generating electricity was shattered in 2001 when a severe drought lowered reservoir levels throughout the nation to the point of cutting electricity generation by 20 percent, causing widespread power disruptions and economic dislocations. Brazil is now pursuing a policy of energy diversification for electricity generation rather than total reliance on hydropower.

Dams are ranked all sorts of ways such as by height, reservoir size, and the material consumed in their construction. In terms of electricity-generating capacity, the Itaipu dam is the world's largest dam as Table 8.2 shows. There are plans to expand the Itaipu to 14 gigawatts and the Tucurui to 8.4 gigawatts, indicative of Brazil's continuing attraction to hydropower.[19]

On a global scale, hydropower supplied 18.5 percent of electricity between 1990 and 1997. Then a slow decline set in that reduced its contribution to 16 percent in 2004, which meant that other means of electricity generation were favored over hydropower. This trend may be affected by hydropower projects in India and China. The potential for electricity generation by hydropower is enormous in both nations because their major rivers start 15,000 feet above sea level on the Tibetan plateau. India intends to add between 20–30 GW of new hydropower capacity. The Indian government is a strong advocate of hydroelectricity as an alternative to coal-burning plants to reduce air pollution by utilizing a free source of clean energy. The Indian government is less enthusiastic about natural gas-generating plants because of the need to import the fuel. However, the government faces strong environmental opposition to its plans for hydropower and it is not clear how these projects will fare in the future.[20] In 2002, coal supplied 71 percent of India's electricity with hydropower second at 11 percent. The projected tripling of hydropower output between 2002 and 2030, while impressive on the surface, just keeps up with the projected growth rate in electricity generation of over 4.5 percent per year.[21] On top of this the availability of funds to support India's hydropower and other major energy expansion plans is questionable. India's chronic balance-of-trade deficit limits its capacity to raise funds to undertake capital-intensive energy-related projects.

China, on the other hand, with a significant balance-of-trade surplus, accumulates sufficient funds to support financing capital-intensive projects such as the world's largest hydroelectricity project: Three Gorges Dam on the Yangtze River. This hydroelectric dam will stretch nearly a mile (1.5 kilometers) across and tower close to 2,000 feet (600 meters) above the world's third longest river. Its reservoir will cover land 350 miles upstream of the dam, forcing the resettlement of close to 2 million people. Construction began in 1994 and is scheduled to take twenty years, at a cost of over $25 billion. Its installed capacity of 18.2 gigawatts will be the largest in the world, equivalent to that of eighteen nuclear power plants, and will supply 9 percent of China's electricity needs. A system of locks will allow ships to pass around the dam. The Three Gorges dam also figures largely as a flood-control measure for a river notorious for disastrous floods.

The benefit of flood control, along with the substitution of clean hydropower for dirty burning coal in electricity generation, has made little impact on the growing opposition to the dam. Human rights organizations criticize the resettlement plans, archaeologists are concerned about the submergence of over 1,000 historical sites, and others mourn the loss of some of the world's finest scenery. Moreover, millions of Chinese downstream of the dam would be at risk if there were a catastrophic structural failure (memories of the Shimantan dam disaster still persist). The Three Gorges dam is but one project underway in China. China plans to double its hydropower potential from 75–150 gigawatts by the year 2015. Yet, despite these Herculean efforts, hydropower is projected to just maintain its 18 percent share of electricity generation.

Large-scale dam projects are underway in Turkey at the headwaters of the Tigris and Euphrates rivers to irrigate agricultural land, supply water to towns and cities, and generate electricity. These projects have strained relations with Iraq because the Tigris and Euphrates are also the principal sources for irrigation and drinking water in Iraq. Turkish dam projects have also spurred opposition from Kurds and other indigenous people who would be displaced by the reservoirs. Once fully operational in 2010, these dams have the potential to severely reduce the water flow to Syria and Iraq, depending on the amount of water diverted for irrigation. Some believe that access to water will be the next source of conflict between nations after oil. A water conflict has already erupted between Lebanon and Israel because some

Israelis felt that a plan to divert the headwaters of a stream in Lebanon, used for irrigation and drinking water in Israel, was tantamount to an act of war.

Prospective sites for large hydropower projects are nearly exhausted in the United States and Europe. South America still has a great deal of potential that can be tapped as does Asia, the present world center of dam building. In contrast, thought is being given to mini and micro hydro plants, which are small, nonthreatening plants, not disruptive to people or the environment. Small may be back in style, but small dams lack the inherent economies of scale of megadams.

One potentially huge hydropower project under consideration is associated with the Dead Sea, 1,370 feet below sea level, the lowest spot on Earth. It is bordered by Israel and Jordan and the West Bank under the control of the Palestinian Authority. For thousands of years, the flow of the Jordan River was sufficient to replenish water lost to evaporation. Being "at the end of the road," the Dead Sea accumulates salts carried by the Jordan River. Whereas the world's oceans have a salinity content (salts of sodium, magnesium, calcium, potassium, and others) of 3.5 percent, the Great Salt Lake in Utah has a salinity of 27 percent, and the Dead Sea 33 percent, almost ten times that of ocean waters. Dead Sea waters are thought to have therapeutic properties and have an oily sensation. A person floating in the Dead Sea finds it hard to stand up and leaves the water caked with salt.

The problem is that the Dead Sea is no longer being replenished with water. The Jordan River and its tributaries have been thoroughly tapped by Israel, the Palestinian-controlled West Bank, Jordan, and Syria for the region's scarcest resource, which has cut the flow of the Jordan River by nearly 90 percent. What now flows into the Dead Sea is mainly sewage and other waste waters dumped into the Jordan River after all its clean waters have been drawn off. Depending on where it is done, baptism by submergence in the Jordan River can be hazardous to one's health. The idea to build sewage-treatment plants to remove wastes being dumped into the Jordan actually worsens the problem. Once treated, the water would probably be diverted for irrigation, reducing the flow in the Jordan from a trickle to nothing.

With this massive diversion of water for irrigation, the Dead Sea is falling about 1 meter per year and its shoreline has retreated 500 meters over the last few decades, resulting in the loss of one-third of its area.[22] To counter this, the possibility of building a 108-mile (174-kilometer) system of canals and pipelines to bring seawater via gravity flow from the Gulf of Aqaba on the Red Sea to the Dead Sea has long been considered. Pipelines would siphon water over intervening highlands. Siphoning occurs when water leaves the pipeline at a lower elevation than water entering the pipeline, eliminating the need for pumping. The end point of the Red-to-Dead project would be a hydropower plant with a hydraulic head of 500 or more meters, higher than most dams. The potential output of electricity, presently envisioned at 0.55 gigawatts, can be far larger depending on the flow of Red Sea water. Electricity generation can be partly dedicated to desalinizing water for human and agricultural use. The project requires the cooperation of Israel, Jordan, and the Palestinian Authority to arrive at a way of fairly sharing the electricity and desalinized water and its estimated cost of $5 billion. The project has to deal with environmental objections over potential damage to coral reefs in the Gulf of Aqaba caused by diverting waters to feed the Red-to-Dead Project. There are also environmental concerns over potential chemical and biological consequences of pouring vast quantities of Red Sea water into the Dead Sea basin. On the other hand, doing nothing means another environmental calamity when, in about 150 years, continued evaporation transforms the Dead Sea into a supersaturated solution of salt incapable of further evaporation.

Notes

1. This figure comes from the U.S. Department of Energy listing of generating units for the United States as of January, 2004, available at www.eia.doe.gov/cneaf/electricity/epm/epm_sum.html.

2. Figures for global electricity-generating capacity come from the *International Energy Annual* (2003), published by the U.S. Department of Energy. It is also available online at http://www.eia.doe.gov/pub/international/iea2003/table64.xls

3. A comprehensive overview of the nuclear fuel cycle and the entire nuclear power industry can be found on the Web site of the Uranium Information Centre of Melbourne, Australia, at www.uic.com.au.

4. Edward S. Cassedy and Peter Z. Grossman, *Introduction to Energy* (Cambridge, UK: Cambridge University Press, 1998).

5. *BP Energy Statistics* (London: British Petroleum, 2005). Figure 8.2 assumes a thermal efficiency of 38 percent for converting fossil fuel to electricity.

6. The numbers shown in Figure 8.4 can be found online at the Uranium Information Centre of Melbourne, Australia, at www.uic.com.au/reactors.

7. *BP Energy Statistics* (London: British Petroleum, 2005).

8. Nuclear Energy Institute, www.nei.org.

9. Uranium Information Centre of Melbourne, Australia, www.uic.com.au.

10. Nuclear Energy Institute, *Vision 2020: Nuclear Energy and the Nation's Future Prosperity,* available online at www.nei.org.

11. As a former engineer officer onboard a nuclear submarine, I could easily be transferred from one submarine to another because the reactor and steam-propulsion systems were essentially identical. There was only one plant layout and one set of manuals for operating instructions and emergency procedures to learn. Frankly, I always felt comfortable and secure with the power produced by a nuclear plant. It was safe, reliable, and gave me and other crew members confidence in a safe return (nuclear submarine casualties have been related to seawater pipe failures, accidental torpedo detonations, and navigational errors, not problems with the nuclear plant). The only time I felt nervous was when we secured the nuclear plant and ran on diesel power to ensure that the diesel engine would function if the nuclear plant became inoperative, which never (in my experience or knowledge) ever happened.

12. Vaclav Smil, *Energy in World History* (Boulder, CO: Westview Press, 1994).

13. Fen Montaigne, "Water Pressure," *National Geographic* (September 2002), vol. 202, no. 3, p. 29.

14. Russell Martin, *A Story That Stands Like a Dam* (Salt Lake City, UT: University of Utah Press, 1999).

15. This information is available online at the World Health Organization Web site at www.who.int/ctd/schisto/index.html.

16. Visit the Web site of the International Commission on Large Dams at www.icold-cigb.org.

17. *BP Energy Statistics* (London: British Petroleum, 2005).

18. This information comes from the Hydro-Québec Annual Report for 2004, available online at www. hydro.qc.ca.

19. See www.infoplease.com/ipa/A0001336.html.

20. An example of the opposition to hydropower projects is available online at the Friends of the River Narmada Web site at www.narmada.org. For a rebuttal to such opposition visit the Web site of the International Commission on Large Dams at www.icold-cigb.org.

21. International Energy Agency, *World Energy Outlook 2004* (Paris: Author, 2004).

22. Joshua Hammer, "The Dying of the Dead Sea," *Smithsonian Magazine* (October 2005), vol. 36, no. 7, p. 58.

9

Sustainable Energy

This chapter discusses the meaning of sustainability and the principal sustainable sources of energy: geothermal, wind, solar, and ocean (tidal, wave, and thermal). Their contribution to meeting world energy needs, along with their relative reliability compared to fossil fuels, will also be covered. Two sources of sustainable energy have already been dealt with: biomass in Chapter 3 and hydropower in Chapter 8. With the ability to reprocess spent fuel, some feel that nuclear power also qualifies as a sustainable energy source. The chapter ends with a discussion of the role of sustainable energy in economic development and a sustainable energy action plan.

The Meaning of Sustainability

Sustainability has various meanings for different people. For some, sustainability is associated with society. A sustainable society involves an assortment of issues to sustain life on this planet; some mix of population control, adequateness of nutrition, preservation of ecosystems, conservation of natural resources, urban livability, redistribution of wealth, controls over energy usage, industrial activity and pollution emissions, and so forth. A sustainable society is not impossible. The Indians of the Americas had a sustainable society before Europeans arrived. Tribal warfare and disease were effective means of population control. Indian culture and religion were centered on preserving the ecosystem. Hunting animals, burning wood for warmth and cooking, and raising crops did not reduce the animal and plant population or the fertility of the soil. The Indians showed the Pilgrims how to fertilize naturally by burying a fish with each kernel of corn, and helped them in other ways that ensured their survival.

The Aztec and Inca cities of Mexico and Peru were certainly livable and had no surrounding garbage dumps (the conquistadors considered them engineering marvels of unimaginable beauty before putting them to the torch). Sharing material possessions without a sense of legal ownership or property rights obviated the need for redistributing wealth. This paradise, however, did not include all Indian tribes. The Aztecs taxed neighboring Indian tribes so heavily, in addition to using them as ready sources of victims for human sacrifices, that the local tribes eagerly sided with Cortés. Without these tens of thousands of willing allies, he would have been unable to bring down the Aztec empire. Nevertheless, American Indian civilizations were sustainable and could have gone on forever. The Indians view European civilization as a temporary structure, intrinsically opposed to nature and ultimately unsustainable. As a Mohawk Indian leader once said, “Not until the last tree has fallen, the last river has been poisoned, the last fish has been caught, will man realize that money isn’t edible!”

Here the concept of sustainability is limited to energy. A sustainable source of energy is renewable and environmentally benign. While sustainable energy sources such as hydro, geothermal, wind, solar, oceanic (wave, tidal, current, and temperature differential) are seemingly inexhaustible, this is not true for biomass. Biomass is a sustainable source of energy as long as a crop is grown, burned for its energy content, and replaced by another. Under these conditions, there is no net addition of carbon dioxide to the atmosphere because the carbon dioxide released from burning is absorbed in growing the replacement crop. Deforestation adds carbon dioxide to the atmosphere as more biomass is burned than is replenished and is not sustainable; at some point the forest is gone. *Inexhaustible* means that the source of energy is always present and never diminished, but that does not infer an infinite supply. Inexhaustibility must be tempered with capacity limits. Biomass is limited by the availability of arable land for non-food crops; solar power by whether the sun is shining and the number of solar arrays; wind power by whether the wind is blowing and the number of wind turbines; and hydropower by rainfall and the number of dams.

A major difference between conventional and sustainable sources of energy is reliability. Electricity can be generated at the dispatcher's whim up to a plant's rated capacity for a generator fueled by fossil and biomass fuels, nuclear power, and geothermal energy. This is not true for other sources. Hydropower depends on rainfall. Wind and solar and wave power depend on the weather. Tidal energy is predictable, but there is no guarantee that peaks in electricity generation from changing tides coincide with peaks in electricity demand. Wind, solar, tidal, and wave sources can certainly be tied into an electricity distribution grid and contribute to the electricity pool "weather permitting," but they can only displace, not replace, conventional sources of energy.

Geothermal Energy

Geothermal energy, from the Greek words *geo* (earth) and *therme* (heat), takes advantage of hot water or steam escaping from hot spots in the earth. Geothermal sources are located where magma is relatively close to the surface of the earth and where the rock above the magma is porous and filled with subsurface water with access to the surface. Geothermal sources are found near tectonic plate boundaries that may be separating (the rift valley in Africa, or Iceland) or colliding, creating subduction zones (Japan), or sliding by one another (California). Geothermal energy sources are also found near volcanoes (Mount Vesuvius in Italy, the island of Hawaii, and the caldera at Yellowstone), where magma protrusions lie relatively close to the earth's surface. The Maoris in New Zealand and Native Americans used water from hot springs for cooking and medicinal purposes for thousands of years. The people of Pompeii, who lived too close to Mount Vesuvius, tapped hot water from the earth to heat buildings. Ancient Greeks and Romans had geothermally heated spas. Romans used geothermal waters for treating eye and skin disease. The Japanese have enjoyed geothermal spas for centuries.[1]

The earth's crust insulates us from the hot interior of the mantle. The normal temperature gradient is about 50°F–87°F per mile or 17°C–30°C per kilometer of depth and higher where the crust is relatively thin or near plate boundaries and volcanoes. Magma trapped beneath the crust heats up the lower rock layers. If the hot rock is porous and filled with continually replenished subsurface water with access to the surface, then the result is fumaroles of escaping steam or hot gases, geysers of hot water, or pools of boiling mud. As a geothermal source, the earth becomes a boiler and escaping hot water and steam, called hydrothermal fluids, are tapped for hot spring baths, heating greenhouses (agriculture), heating water to raise marine life (aquaculture), space heating for homes, schools, and commercial establishments, heating streets and sidewalks to prevent ice formation, and as a source of hot water for industrial use or steam for generating electricity. Some cities have district heating, using geothermal hot water to heat an entire area, best exemplified in Reykjavik, Iceland, where 95 percent of the city receives hot water from geothermal sources.

There are three types of geothermal power plants for generating electricity. The first is a dry-steam geothermal reservoir in which emitted steam directly spins a turbine. These are relatively rare and were the first dedicated to generating electricity. One in Tuscany has been in operation since 1904 and The Geysers, 90 miles north of San Francisco, has been in operation since 1960. The Geysers represents over 90 percent of the geothermal energy in the United States and generates enough electricity to supply a city the size of San Francisco. A falloff in steam pressure in recent years has been successfully countered by water injection to replenish the geothermal reservoir. Injected water is waste treatment water from neighboring communities, an innovative and environmentally safe method of disposal. Some thought has been given to tapping the world's most productive source of geothermal energy, Yellowstone, the caldera of a supervolcano that last erupted 600,000 years ago. (Another eruption of that magnitude would wipe out half of the United States and emit an ash cloud large enough to send the planet into a "volcanic" winter.) Yellowstone, as a national park, cannot be commercially developed.

The second and most common form of geothermal power plant is driven by geothermal reservoirs that emit hot, pressurized water between 300°F–700°F. The drop in pressure inside a separator allows the liquid to flash to steam, which is then directed into a turbine. Any gases in the geothermal water such as carbon dioxide, hydrogen sulfur, and nitrous oxides pass to the atmosphere, but these are a tiny fraction of the emissions from a coal-burning plant with an equivalent power output. Water and steam remaining after flashing are usually reinjected to replenish the water in order to maintain the reservoir's pressure. If reservoir pressure can be maintained, then geothermal becomes a sustainable source of nearly nonpolluting clean energy.

Shallower sources of geothermal energy in which the water temperature is between 250°F–350°F require a binary power plant where a heat exchanger transfers the heat of the geothermal water to a second or binary fluid such as isopentane or isobutane. The binary fluid boils at a lower temperature than water and its vapors pass through a turbine and are then condensed to a liquid for recycling. Binary plants are closed systems in which the hydrothermal fluid, along with any entrapped gases, is reinjected into the reservoir. A binary system may be necessary for water with a high mineral content to prevent forming a harmful scale on the turbine blades. Hybrid plants, part flash and part binary, are also available such as the one that supplies 25 percent of the electricity for the island of Hawaii.

As of 1999, there were 250 geothermal power plants supplying 8.2 gigawatts (GW) of power, equivalent to the production of eight large-sized nuclear or coal-fired plants, enough to supply the electricity needs of 8 million people. The United States leads with 2.9 GW, mostly in California and Nevada; second are the Philippines with 1.8 GW, then Italy with 0.8 GW (one center of activity is near Mount Vesuvius on the outskirts of Naples), Mexico with 0.7 GW, Indonesia with

0.6 GW, Japan with 0.5 GW, and Costa Rica and El Salvador, each with 0.1 GW. Iceland generates 17 percent of its electricity from only 0.1 GW of geothermal energy. Thermal uses, distinct from generation of electricity, total another 10 GW for hot springs, agriculture and aquaculture, and district heating. Geothermal heat pumps use the constant underground temperature either for cooling or heating residences. Heat pumps are effective over a limited range of temperatures and require supplemental cooling and heating to handle more extreme variations in temperature.

Geothermal sources are limited primarily to porous hot rock permeated with subsurface water that can escape to the surface. If water is trapped by a cover of impermeable caprock, then the geothermal reservoir must first be discovered before it can be exploited. The same technology for discovering oil fields and drilling wells to tap oil and natural gas reservoirs is used for discovering and developing geothermal reservoirs. The future of geothermal energy is limited as long as current efforts are, for the most part, restricted to developing known sources of geothermal energy, not in discovering new ones.

Hot rock underlies the entire crust. Its usefulness is a matter of depth and whether it is porous rock filled with subsurface water. The presence of subsurface water, however, may no longer be necessary. The feasibility of drilling two wells deep into the earth's crust to reach hot rock, and then fracturing the rock separating the two employing methods practiced by the oil industry, has long been considered. Water would then be pumped down one well under pressure and forced through the fractured hot rock, where it would be heated and rise to the surface via the other well as pressurized hydrothermal fluid. This would then be flashed to produce steam and drive electricity generators or heat another liquid medium in a hybrid plant.

The idea is now becoming a reality, thanks to the high cost of fossil fuels. Hot rock from a magma protrusion of up to 570°F has been discovered two and three-quarter miles below the surface in southern Australia. Two wells have been drilled into the hot granite, which has been fractured to allow the flow of water from one well to the other. It is hoped that the wells will establish the feasibility of generating commercial volumes of geothermal electricity. If successful, electricity generators will be built and connected into the electricity grid, at the same capital cost as a coal-fired plant, without the associated carbon dioxide and other emissions and with no cost for the fuel. It is possible that this hot rock formation may one day supply all of Australia's electricity needs.[2]

A more esoteric idea is to mine heat from magma as a geothermal ore in places where it is accessible by current drilling technology. In this case, a hole is drilled through the crust and a sealed pipe with a concentric inner pipe is thrust into the magma. Water is pumped down the inner pipe where it flows into the outer pipe at the bottom. There it is transformed into high-pressure steam by heat from the magma and flows up the outer pipe to power electricity generators. The major obstacle to overcome is finding materials that can withstand the extremely high temperatures and corrosive nature of being inserted into magma.[3] If the inner heat of the planet could be tapped, geothermal energy could conceivably satisfy the world's electricity demand.

Wind Energy

Wind results from differences in air temperature, density, and pressure from the uneven solar heating of the earth's surface. Like ocean currents, wind currents act as giant heat exchangers, cooling tropical and warming polar regions. The history of wind energy dates back about 2,500 years ago to China, when wind was used for pumping water. About the same time, woven reed sails mounted on a horizontal wheel ground grain in Persia. Windmills of various types continued to pump water and grind grain in Persia until 900 CE. Windmills entered Europe in the eleventh century via the Crusades and were used extensively in Holland and England. Holland had 8,000 windmills in 1650, England had 10,000 windmills in the early 1800s, and Germany had more than 18,000 mills in the late 1800s.[4] The most memorable use of windmills occurred in Holland to drain lakes and marshes in the Rhine River delta, grind grain and saw wood, and later pump out encroaching seawater that would otherwise flood the low country. Windmills in Europe were eventually replaced with the convenience and reliability of coal-fired steam engines. Windmills were vital in the development of the American West and supplied water for farmers, cattle, and steam-driven locomotives. And, of course, wind was the preferred source of propulsion for sailboats, first recorded on the Nile River as early as 5,000 BCE. It took millennia for wind power to replace those who performed what must have been the most tiring and tedious of jobs: oarsmen.

Windmills generating electricity are called wind turbines. Most wind turbines are associated with wind farms or parks that are tied into electricity distribution grids. The wind turbine is the opposite of a fan. A fan consumes electricity to power a motor that turns a rotor with attached blades to move air. In a wind turbine, moving air turns the blades attached to a rotor that drives a generator to produce electricity. The beauty of a wind turbine is that the energy source is free. The disadvantage is that the wind turbine cannot produce energy if the wind speed is too low or too high. Even when wind conditions are right, the output of a wind turbine varies with wind speed. Since wind turbines do not respond to the desires of dispatchers, but to wind conditions, wind energy can only displace, not replace, fossil fuels. A backup source of power must be available when the wind is too calm or blows too hard. Widely dispersed wind farms and placing them in areas where acceptable wind conditions are prevalent much of the time increase the reliability of wind power. As with dams, wind farms are site-specific and must be built in locales where wind conditions are favorable, which may not be near existing transmission lines or population centers.

Like water flowing in a river, wind energy can be converted to electricity. The amount of electricity is determined by the energy contained in the wind passing through the area swept by the wind turbine blades over time, called the wind-power density. The blades are designed to rotate with a frequency that ensures optimum efficiency and a maximum yield of wind-power density that can be converted to electricity with minimum tower oscillation. The rate of rotation changes for different wind speeds, producing a variable-frequency electrical current, which has to be converted to a fixed frequency required by a utility. Wind-power density depends on the cube of the wind speed (when wind speed doubles, the wind-power density goes up by a factor of eight) and also on air density and temperature (lower altitudes and cooler temperatures increase wind-power density).

Although wind-power density escalates rapidly with wind speed, this does not mean that it can be entirely transformed to electricity. A turbine has four output phases depending on wind speed. No power is generated when wind is below a minimum speed. Above the minimum speed, electricity output rapidly rises with increasing wind speed until the wind speed attains a threshold level. Above this, electricity output is constant at the turbine's rated capacity even with increasing wind speed to avoid overstressing the wind turbine. Turbine blades are purposely designed to become less efficient at wind speeds that can damage the tower supporting the turbine or the blades themselves. When wind speed is too high, the wind turbine stops producing electricity and assumes a mode of operation that protects the blades and tower against physical damage.

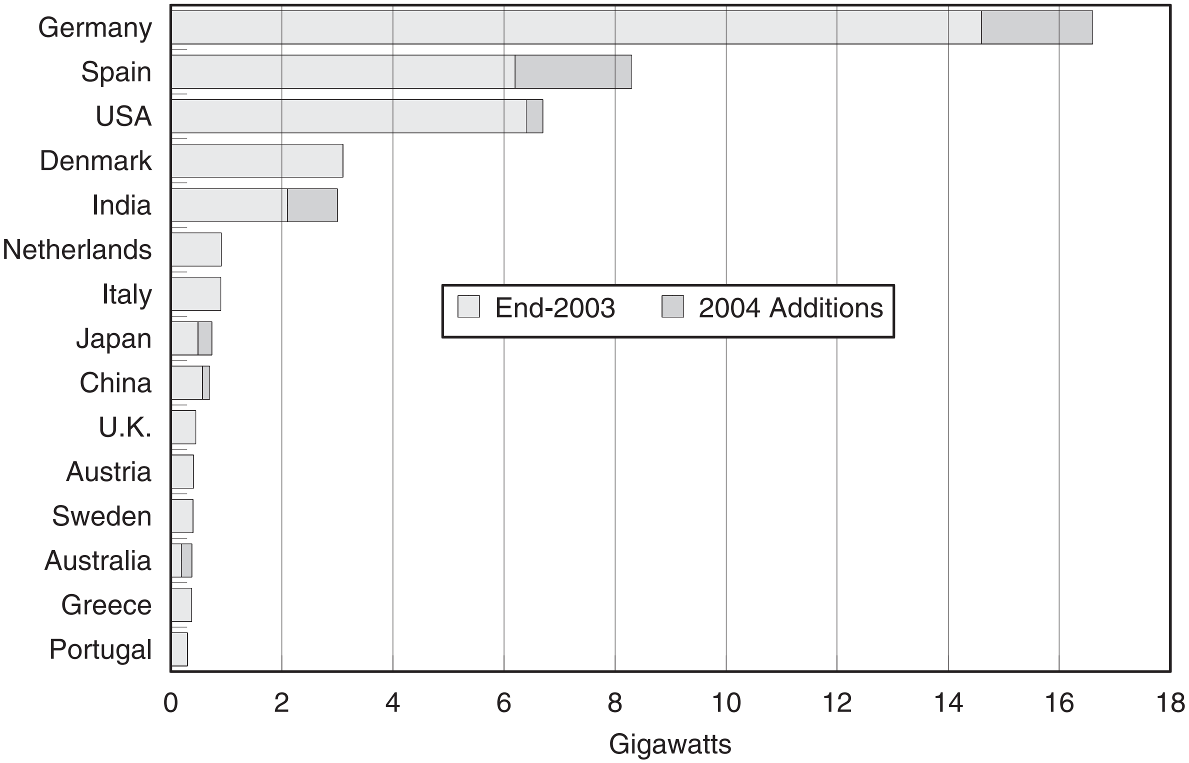
The wind energy industry started in the United States as a result of the Public Utility Regulatory Policies Act (PURPA) of 1978. This Act required state regulatory commissions to establish procedures for non-utility companies to sell electricity to utilities generated from renewable energy sources, waste, and cogenerating plants run on natural gas. California state regulators, fearing that oil-fueled electricity-generating plants would be threatened with falling oil production in California and Alaska, were particularly aggressive in carrying out PURPA provisions. They required California utilities to buy electricity generated from wind farms at a premium over conventional sources to induce the development of wind energy. As a result, California became the home of over 14,000 wind turbines, each of which can produce between 50–300 kilowatts of electricity. A single wind turbine producing 300 kilowatts or 300,000 watts can generate enough electricity for 3,000 100-watt light bulbs. Major California wind farms are located at the Altamont Pass east of San Francisco, Gorgonio Pass near Palm Springs, and at Tehachapi, south of Bakersfield; areas that experience persistent winds much of the time.

Electricity from wind is experiencing explosive growth. From the tiny acorns planted in the 1980s in California, mighty oaks are beginning to grow. Global wind-energy capacity totaled 47 gigawatts in 2004. A jump of 8 gigawatts in 2004 alone is indicative of the nature of what is happening with wind energy. Current estimates are that wind energy will reach 160 gigawatts by 2012, over a tripling of 2004 levels. As Figure 9.1 shows, the center for wind power development has shifted from the United States to Europe.[5]

Denmark led Europe's shift toward wind energy and generates 20 percent of its electricity from wind. Denmark is adding more wind energy projects at a slower pace because 20–25 percent may be the maximum contribution wind energy can make, considering its reliability. As a consequence of being the first to seriously pursue the wind option, Denmark has become a global center of wind energy technology in the design and manufacture of wind turbines and is a major exporter of hardware (wind turbines) and software (knowledge and expertise) to other nations.

Germany leads by far in wind energy capacity, generating 6 percent of its electricity (25 percent of the electricity in Schleswig-Holstein). Spain overtook the United States for second place in 2004 and is aggressively pursuing the development of wind energy, which supplies 4–5 percent of its electricity needs. Spain's fast growth is spurred by a national requirement for utilities to pay a premium price for electricity produced by wind over the first five years of operation. This program is similar to the one that launched the German project and, before that, the California wind energy industries. The United States, now in third place, is expected to add 2 GW of wind energy capacity in 2005, up considerably from 2004, when most wind projects were put on hold when production tax credit incentives expired. The United Kingdom has in the works 0.7 GW of wind development projects scheduled for completion over the next two years, slightly more than doubling its present capacity. Permit applications in 2004 would add 7 GW of capacity if all were approved. Portugal, with 0.3 GW of wind-generated electricity, intends to

Figure 9.1 **Principal Wind Energy Generating Nations**



expand its present capacity more than tenfold, to 3.75 GW, by 2010. India, Australia, China, and Japan have initiated plans to greatly expand their use of wind energy for generating of electricity.

For wind turbines to be efficiently employed, dispatchers should have a forecast of wind conditions in order to plan the next day's schedule of operation for the utility's generating resources. Even so, dispatchers cannot schedule future production with confidence considering the reliability of weather forecasts. Whereas dispatchers can depend on a fossil-fueled or nuclear power plant to produce, on average, 70–90 percent of its rated capacity, the average output of a wind turbine is only 25–40 percent of its rated output because of variance in wind speeds plus an allowance for downtime for maintenance. Thus 160 GW of anticipated wind capacity represent at most 80 GW of conventional electricity-generating plants.

To put wind power into perspective, the estimated world electricity capacity in 2005 was about 3,800 GW, made up of nearly all conventional plants. If 160 GW of wind capacity are equivalent to 80 GW of conventional plant output, then the projected wind capacity will represent a bit more than 2 percent of current capacity in terms of average output. While wind energy cannot be relied upon to deliver as much of its rated output capacity as a fossil fuel or nuclear power plant, this does not mean that wind energy should be ignored. Every watt contributed by wind energy is one less watt that has to be contributed by burning fossil fuels and that much less carbon dioxide and other emissions added to the atmosphere.

Ideal places for siting wind farms are geological formations that funnel wind through narrow passes, as in California, or persistent winds found along ridges of the Rocky and Appalachian mountain chains or across the central and northern portions of the Great Plains. The northern Great Plains (Montana, the Dakotas, and Wyoming), self-dubbed the Saudi Arabias of wind energy, have the potential to supply the United States with all its electricity needs, weather permitting. Other favored places are along coastal areas, such as the northern portions of the east and west coasts of North America. A far larger potential area than coastlines is offshore waters. There are close to 58,000 square miles (150,000 square kilometers) of water less than 115 feet (35 meters) deep available for development along Europe's coasts, with enough potential to generate all of the continent's electricity needs, weather permitting. An acceptable location is one with an annual average wind speed of at least fourteen miles per hour; an ideal location is one with a persistent wind speed between twenty-five and thirty-five miles per hour.

It is anticipated that 10 gigawatts of electricity capacity will be generated in European offshore wind parks by 2010. Whereas environmentally sensitive Europeans see offshore wind parks as a way to eliminate the consequences of burning fossil fuels, environmentalists in the United States have succeeded in delaying, if not preventing, offshore wind parks from being constructed in Massachusetts and Long Island. The proposed offshore Massachusetts wind park, located five miles off Cape Cod in the shallow waters of a shoal, would consist of 130 wind turbines delivering a total of 420 megawatts; enough to supply three-quarters of the electricity demand of Cape Cod and offshore islands.[6] The primary reasons why environmentalists object to wind parks are that the tips of the wind turbine blades protrude one-half inch above the horizon when viewed from land and the wind turbines mar the open ocean vistas of boating enthusiasts. These are the same reasons why environmentalists and yacht owners object to oil drilling in the waters off Florida. Moreover, the wind park sponsors cannot guarantee that a single bird will not be killed.

In the 1980s wind turbines were built with lattice-style structures of 50–300-kilowatt capacity and a blade 49–98 feet (15–30 meters) in diameter. In the 1990s, the tubular structure was adopted with turbine capacity of 300–750 kilowatts and a blade diameter of 98–164 feet (30–50 meters). In the 2000s, power output is up to 1,500 kilowatts (1.5 megawatts), and General Electric (GE) has built 1,200 in worldwide operation. General Electric has a 3.6-megawatt wind turbine with a rotor blade diameter of 341 feet (104 meters) in operation in Spain.[7] Thirty of these, if the wind conditions are right, could provide the power equivalent of a large coal or nuclear power plant of 1 gigawatt, enough to supply a city of 1 million people. These units would have been used for the 420-megawatt offshore Cape Cod project.

Objections to Wind Power

Not surprisingly, there has been little opposition to locating wind turbines on farmland, given the long history of windmills on American farms that lasted until the mid-twentieth century. Farming operations are minimally disturbed by the presence of wind turbines, which are normally sited in pastures or along edges of fields, and are viewed favorably by farmers as a source of incremental income. Much of the opposition to developing wind farms comes from suburbanites, real estate developers, and environmentalists, for a variety of reasons.

Wind turbines are highly visible. A typical wind turbine may stand 390 feet above the ground with blades 130 feet in diameter affixed to a rotor 260 feet above the ground. The largest GE wind turbine has a blade diameter of 341 feet (104 meters). As a point of reference, a Boeing 747–400, with a full load of 436 passengers, has a wingspan just over 211 feet (64.4 meters) and a length just under 232 feet (70.7 meters). The diameter of the GE wind turbine blade is 50 percent longer than the wingspan or the length of a large commercial passenger aircraft. No matter what the size, suburbanites do not want to look at them and environmentalists object to locating wind farms in scenic areas or along coastlines.

Some of this opposition can be overcome by designing less obtrusive and/or more pleasing designs such as foregoing lattice for tubular-style towers and blending a line of wind turbines with the contours of the land. Unsightly transmission lines can be removed by burying them. Some object to the swishing noise of the blade passing through the air, which can be heard within a few miles of a large wind farm and within several hundred feet for an individual wind turbine. While noise from a wind farm is less obtrusive than normal motor vehicle traffic, the fact that a quiet night is no longer quiet bothers people. To counter this, increasing the separation between a wind farm and residential areas reduces noise levels. Larger wind turbines can be quieter than smaller ones, depending on the speed of the blade tip and the design of a blade's airfoils and trailing edges. Local opposition can be mollified if the citizenry receives a monthly stipend or a reduction in property taxes for permitting a wind farm to be built. Opposition from outside the area of a wind farm cannot be bought off so easily.

Another objection was locating early wind farms in mountain passes that turned out to be bird migratory paths. Birds were killed when they flew into the rotating blades. Siting wind farms now takes into account bird migratory paths. Even if built in a migratory path, migration is a seasonal phenomenon. Radar can be used to stop the wind turbines when a large flock of birds is about to pass through or near a wind farm. The leading cause of bird fatalities by human interference is not birds flying into wind turbines, but into buildings, windows, high-tension lines, communication towers, and motor vehicles, plus fatal encounters with pet cats and pesticides.

Nevertheless, wind turbines pose a hazard for birds, which cannot dodge a blade whose tip speed is over 200 miles per hour. A proposed design for wind turbines with no external blade would avoid killing birds. Wind enters a vertical cylindrical structure where airfoils direct the wind against blades on a rotating vertical shaft. The vertical shaft allows the generator to be at ground level for greater ease of maintenance and reduced interference with radio, television, and communication signals. The vertical axis wind turbine is shorter in height than the traditional wind turbine, less obtrusive in appearance, creates less noise from its lower speed of rotation, and has a surrounding wire mesh to prevent birds from entering the turbine. While the vertical axis wind turbine has an optimal wind speed similar to that of traditional wind turbines of 28–33 mph, it can operate with wind speeds up to 70 mph versus 50 mph for traditional wind turbines. This higher range of wind speed increases the overall average output from 30–40 percent of rated capacity for traditional wind turbines to 40–45 percent.[8]

As with any utility project, there are a number of organizations a wind farm developer must successfully negotiate with before construction can begin. State governments have environmental boards that require an environmental impact assessment for the wind farm and its transmission lines. Permits are required from land commissions before a project can move ahead. A public utility commission must grant a certificate of need. County- and community-planning boards ensure compliance with zoning ordinances and land use requirements. As with any real estate development, clearing land for access roads and wind turbine foundations must be done in a manner that avoids or minimizes soil erosion. These boards can also address the possibility that a wind farm might interfere with radio and television reception. If the wind project is on land, then the Bureau of Land Management or the Forest Service will be involved, along with the Fish and Wildlife Service, to ensure minimal hazard to birds and other wildlife.

The views of community groups greatly influence the permit process. Such groups can challenge a site proposal through the court system if they believe that laws, regulations, and legal procedures have not been properly followed. To ensure public support, it is important for project organizers to make the public aware of the benefits of wind power, including any contribution that the wind farm is making in the form of paying local taxes in addition to steps being taken to minimize environmental objections.

Evaluating a Potential Site

While a single wind turbine can have large fluctuations in power output from abrupt changes in wind speed and direction, a wind farm covering a wide area tends to dampen the aggregate impact of shifting winds. Wind patterns can be affected for days from passing storms and weather fronts and for months from seasonal variations (winds are generally more intense in winter and spring). What counts in determining the feasibility of a potential wind farm site is not short-term wind fluctuations or seasonal swings, but the average speed throughout the year. In addition to average annual wind speed, wind patterns near the ground are critical in selecting the height of the hub (the center of the rotor). Wind shear is the change in wind speed with height, which is influenced by solar heating, atmospheric mixing, and the nature of the terrain. Forests and cities tend to increase wind shear by slowing the speed of air near the surface. A differential in air speed between the blade's lower and upper sweeps can damage the blade. Wind shear can be greatly reduced with an abrupt change in terrain height such as a sea cliff or mountain ridge. Cliffs and ridges also accelerate wind speed, as do mountain passes.

Financial Incentives

Tax shields to induce investment come in various forms, commonly as accelerated depreciation. A tax shield reduces a corporate or personal tax by the tax rate; for a corporation paying a tax rate on its profit of 35 percent, a $100 tax shield is worth $35 in reduced taxes. A tax credit is a far more powerful incentive than a tax shield because a tax credit is a dollar-for-dollar reduction in taxes; a $100 tax credit is worth $100 in reduced taxes as long as the corporation is paying taxes. In 1992, Congress enacted a production tax credit of 1.9 cents per kilowatt-hour for the first ten years of a wind turbine's life. For a tax-paying company, lower taxes can be viewed as a direct subsidy that covers the higher incremental cost of wind energy—up to 1.9 cents per kilowatt-hour more than conventional energy sources. If the installed cost of a wind turbine is no more than 1.9 cents per kilowatt-hour over that of a conventional source of electricity, then the 1.9 cents in tax credits make the wind turbine economically competitive with conventional plants. In the United States, the problem with production tax credits, both at the state and federal levels, is that they normally expire after a short period of time and require frequent legislative renewals, which are not always forthcoming.

When production tax credits expire, those units built before the expiration date keep their tax credits, but new units built after the credits expire do not receive the tax incentive. U.S. production tax credits expired in 1999, were renewed in 2000, expired in 2001, were renewed in 2002, and expired in 2003. This on-again, off-again tax credit has predictable results. Following the expiration of production tax credits in 1999, 2001, and 2003, installations of new turbines fell considerably in 2000, 2002, and 2004. The renewal of the production tax credit in 2004 (effective in 2005) was different in that eligible projects were expanded from wind to include solar geothermal, micro hydroelectric, open-loop biomass, refined coal, livestock and municipal solid waste, and landfill gas. The renewal in 2004 would have expired at the end of 2005, but the Energy Policy Act of 2005 extended the expiration to the end of 2007. While an improvement, the stop-and-go tax credit still complicates the planning and investment process.[9]

There are other ways to encourage growth in sustainable or renewable energy besides tax credits. One way is to award a grant to a company to construct a renewable energy source. Grants have not been successful because there is no mechanism for ensuring efficiency of operations. A more successful way to encourage growth in wind energy and other renewable sources is for utilities to offer customers the option to pay a premium on their electricity rates that will support generation of electricity from renewable sources. The utilities enter into a contract to buy the output of a renewable power project at commercial rates that apply to conventional plants. Since this rate cannot economically support an investment in electricity produced from renewable energy, the utility has to find enough consumers willing to make up the difference in the form of a rate premium. Care has to be exercised to ensure that the amount of electricity that carries the premium rate covers, but does not exceed, the capacity of the renewable energy source. In the United States, more than 300 utilities offer their customers a voluntary premium rate that is applied to electricity generated from renewable energy sources. Utilities in the south generally favor solar power while those in the north favor wind farms. Some utilities give consumers a choice of power supply such as solar, micro hydro, and wind, each with a different premium to cover their respective extra costs.[10] This method does not assure efficiency of operation, however, because the consumer is picking up the entire incremental cost associated with the capital and operating costs of a renewable energy electricity-generating plant.

The approach of the European Union is to set a goal to be attained, not by utilities, but by member nations. The 2001 EU Directive specified an overall EU target to increase renewables' share of electricity from 14 percent in 1997 to 21 percent in 2010. Europeans recognize that conventional power production from coal and nuclear power benefits from state aid and there is no reason why renewable energy sources should be exempt. Wind energy is subsidized to make it competitive with coal and nuclear energy. The subsidy has only a modest impact on electricity rates when spread across the entire rate base. Germany and Spain went one step beyond and established a quota that had to be filled by wind power, coupled with an associated electricity rate that would spur its development.

Policy directives, consumers voluntarily bearing the extra cost, production tax credits, and quotas with a fixed price have their pros and cons, but they are not as effective as the renewables portfolio standard (RPS) approach. If energy from renewables is desirable because of security of supply or environmental concerns over fossil fuels, then a more forceful approach is warranted than tax gimmicks, appeals to a consumer's environmental conscience, and policy goals. In the United States, as of January 2005, eighteen states have some form of renewable power requirement where a certain percentage of the power generated by a utility must come from clean, renewable energy. The RPS establishes a goal of a certain percentage of electricity being generated from renewable energy that starts low and slowly increases up to 10 or 15 or 20 percent or more of the utility's portfolio over a reasonable period of time. It is up to the utility to select the most economical form of renewable energy source. The RPS is essentially a quota system without a price. A quota with a set price does not provide an incentive to enhance efficiency. A quota without a set price creates a competitive environment among renewable energy providers to offer electricity at the lowest possible price by enhancing efficiency and improving technology.

The gap between the cost of electricity generated by conventional and renewable sources has narrowed considerably with the tripling or quadrupling of natural gas prices and the doubling of coal prices in 2004 and 2005. The gap is further narrowed by technological progress in lowering the capital and operating costs of generating electricity from renewable energy sources, particularly wind turbines. Wind is clean and free and the price of electricity generated by wind is not affected by OPEC, but only by its capital and maintenance costs. Wind farms are a hedge against rising fossil fuel prices and can be built in stages in response to growing demand, quite unlike large coal and nuclear plants that add large increments to capacity that may not be entirely usable during the early years of a plant's life.[11]

Small Can Be Beautiful

As with hydropower, wind farms do not have to be large to be effective. Isolated areas that are not connected to electricity grids via transmission lines can be served by hybrid power systems consisting of a local distribution system to supply electricity from small wind farms coupled with diesel generators as backup when the wind is calm. In addition to small wind farms, there is an increasing effort to develop small wind turbines to allow individual homeowners, farmers, businesses, and public facilities to generate their own clean power to reduce electricity bills or for use in areas not served by an electricity grid. Wind turbines for individuals have a capacity of 1–10 kilowatt hours while intermediate wind turbines of 10–100 kilowatt hours can serve villages, and can be augmented by solar power when the sun is shining. A diesel power backup covers those times when the wind is not blowing and the sun is not shining.

Solar Energy

Like wind, solar energy is also free, but like other free sources of energy, particular conditions apply. Solar power can only be generated during daylight hours, with peak output on clear days when the sun is directly overhead. Several factors affect the efficiency of solar power: cloud cover, which markedly reduces solar output; times of day when the sun is near the horizon (early morning and late afternoon); seasons during which the sun does not rise high in the sky (winter at high latitudes). As with wind, solar power can contribute to power generation, but cannot be relied upon without a backup. Solar power is more expensive and its overall contribution is smaller than that of wind. Solar power has one advantage over wind: it is produced only during daylight hours when electricity demand is highest, reducing the need for peaking generators.

There are two primary types of solar energy: as a source of hot water that can be used for heating or for making steam to generate electricity and the direct conversion of solar energy to electricity.

Thermal Solar Energy

Thermal solar energy can heat water for space heating, household appliances, swimming pools, and for various commercial and industrial processes. Solar water heaters with copper collector tubes in a glazed housing can be installed on the roofs and sides of homes most exposed to sunlight. When the sun is shining, water is pumped through the collector and the heated water is stored in a tank. A thermosiphon solar water heater has the storage tank above the collector tubes, eliminating the need for a pump. While thermal solar energy is associated with warming a house, heated water can drive an absorption or desiccant air-conditioning system for cooling a house. For colder climates, a water/glycol mixture is pumped through the collector, which then requires a heat exchanger to heat water for appliances and space heating. Backup substitute power is required for cold, cloudy, blustery days when snow and ice cover the solar panels.

Buildings can be designed for passive solar energy by having large south-facing windows complemented with building materials that absorb and slowly release the sun's heat. There are no mechanical aspects to passive solar heating and a well-designed system can significantly cut heating bills. Passive solar designs also include natural ventilation for cooling during hot weather. Hybrid lighting concentrates sunlight and feeds it through fiber optics into a building's interior. "Hybrid" means that a backup power source for interior lighting is necessary for times of little or no sunlight, and certainly for nighttime illumination. Thermal solar output for heating water, including a small portion for electricity generation, was nearly 70 GW in 2001. This dwarfed the installed output of 1.1 GW of photovoltaic output at that time (in 2005, photovoltaic output was about 3 GW).[12] China led with 22.4 GW of solar thermal capacity followed by the United States with 17.5 GW, Japan with 8.4 GW, Turkey with 5.7 GW, and Germany with 3.0 GW.

Solar thermal energy can also heat water for conversion to steam for driving turbines to generate electricity. The types of thermal solar power systems that can generate electricity are parabolic troughs, power towers, and dish/engine systems. These technologies are normally hybridized with fossil fuel (natural gas) to maintain electricity output when the sun is not shining or is covered with clouds. This gives the system the necessary reliability required by dispatchers and enhances the economic performance of the system (the generator is producing revenue whether or not the sun is shining). Natural solar power provides 2.7 megawatts per square meter per year. While this may sound impressive, it is actually a low rate of energy transfer considering that this is over a year's time. Thus, solar thermal systems require a great deal of area for mirrors to collect and concentrate the requisite solar energy. However, the land area does not compare unfavorably with coal-fired plants when mining and storage is taken into consideration.[13]

Solar thermal energy for electricity generation requires a location where there is sunlight much of the time and sufficient available space for mirrors. The ideal location is a desert. The first system to commercially convert solar thermal energy to electricity was built in the 1980s in the Mojave Desert in California. Nine solar thermal electricity-generating plants have a combined output of 354 megawatts (one-third the output of a large coal-fired or nuclear power plant), the world's largest installation of solar power. Trough-shaped parabolic mirrors automatically follow the sun and focus the sun's rays at thirty to sixty times their normal intensity on a receiver pipe filled with synthetic oil. The oil is heated to 735°F and passes through a heat exchanger to produce steam for a conventional steam turbine electricity generator. Natural gas serves as a supplemental fuel for cloudy weather and nighttime operation.[14]

The power tower, also developed in California, stores solar energy in the form of molten-salt. A circular field array of heliostats (large mirrors) individually tracks the sun. The heliostats focus sunlight on a central receiver mounted on top of a tower to heat molten salt, such as a mixture of sodium and potassium nitrate, to 1,050°F for storage in the "hot" tank. As power is needed, molten salt flows from the hot tank through a heat exchanger to produce steam for electricity generation and travels thence to the "cold" tank, where it remains molten at 550°F until it is needed for heating in the tower. Depending on the size of the hot tank and the ability of its insulation to reduce heat loss, a hot tank can supply energy to generate electricity for some hours after sunset, an advantage over trough-shaped parabolic mirrors. Moreover, the system is more reliable because dispatchers can depend on the system to produce power even when clouds temporarily cover the sun by generating electricity from energy stored in the hot tank. To further enhance reliability, the system can be hybridized with a fossil fuel, such as natural gas, to produce power when it is needed by a dispatcher at any time.

Trough-shaped parabolic mirrors and power towers require water to generate steam, a commodity in short supply in a desert, but no water is required for the dish/engine solar energy system. Parabolic dish-shaped mirrors, mounted on a single support frame, focus solar energy on a receiver to heat a transfer fluid to nearly 1,400°F. The heated fluid transfers its heat to a gas such as hydrogen or helium to power a Stirling engine, which is similar in construction to an internal combustion engine, or to a Brayton engine, which is similar to a gas turbine engine (sometimes referred to as a micro-turbine). In neither case is there combustion; the engines run off the energy of the heated gas and drive an electricity generator. Solar dish engines have the highest efficiency of thermal solar systems, converting nearly 30 percent of solar energy to electricity. Trough-shaped parabolic mirrors and power towers best serve an electricity grid whereas solar dish engines best serve isolated areas beyond a power grid. However, there is nothing that precludes connecting dish engine arrays to a grid.

The latest idea, based on one first advanced by Leonardo da Vinci, is a solar power tower shaped like a chimney that directs hot surface air up to cooler air at higher altitudes. One company plans to build such a tower that will direct heated surface air up the circular tower to cooler air almost 3,300 feet (1,000 meters) above the Australian outback. The power driver is the air temperature differential between the bottom and the top of the tower. The tower is surrounded by sunlight-absorbing material that further heats the incoming air. Air rushing in the bottom of the tower passes through wind turbines that can generate up to 200 megawatts of rated capacity, enough to supply electricity to 200,000 Australian homes.[15]

Photovoltaic Energy

The earth receives an average of 1,367 watts of energy per square meter (about 11 square feet) at the outer edge of the earth's atmosphere. The atmosphere absorbs and reflects most of the X-ray and ultraviolet radiation, reducing the energy that reaches sea level at high noon on a clear day to a maximum of about 1,000 watts per square meter. One hour's worth of solar energy striking the earth is greater than all the energy consumed by the world's population in one year. Desert land 100 miles on a side (10,000 square miles, which is equivalent to 9 percent of the area of the state of Nevada) could generate enough electricity to supply the United States, weather permitting (the Southwest could become the Saudi Arabia of solar power!). However, the intent is not to concentrate the nation's solar power in one location, but to install solar power plants on rooftops and over parking lots throughout the nation to reduce reliance on electricity from conventional sources.[16]

Photovoltaic means "electricity from light." The photovoltaic effect was first observed in 1839 when Edmund Becquerel noticed that an electrical current through two metal electrodes submerged in a conducting medium increased when exposed to light. In the 1870s the photoconductivity of selenium was discovered and this led to the 1883 invention by Charles Fritts of the first photovoltaic (PV) solar cell made of selenium. A PV solar cell is made up of two layers of semiconductor material that can convert sunlight to electricity. One layer has an abundance of electrons and the other a shortage. Sandwiching these together forms an electrical field at the interface, which acts as a battery when exposed to sunlight. The resulting direct current has to be converted to alternating current for use in a home or for feeding into an electricity grid.

In 1941 Russel Ohl invented the first solar cell made of silicon, and in the 1950s AT&T built the first solar arrays to power earth satellites. These solar arrays transformed sunlight to electricity with an efficiency of 6 percent, which was improved to 10 percent by Hoffman Electronics in 1960. In 1963 Japan built the largest ground-based PV array installation of 242 watts with a storage battery to supply electricity for a lighthouse. In 1973 the University of Delaware built "Solar One," the first roof-integrated solar array that supplied electricity to a home by day as well as feed (sell) excess power back to the utility that supplied the home with electricity at night and during times of cloud cover. Many nations are pursuing the solar option, and research is being conducted under a wide assortment of public and private programs sponsored by governments, universities, and private enterprises. The objective is to make electricity from solar power competitive with conventional sources by reducing front-end costs such as material costs for semiconductors, manufacturing and installation costs of solar arrays, and enhancing efficiency. The greater the efficiency in converting sunlight to electricity, the smaller the solar array has to be to deliver a given amount of electricity.

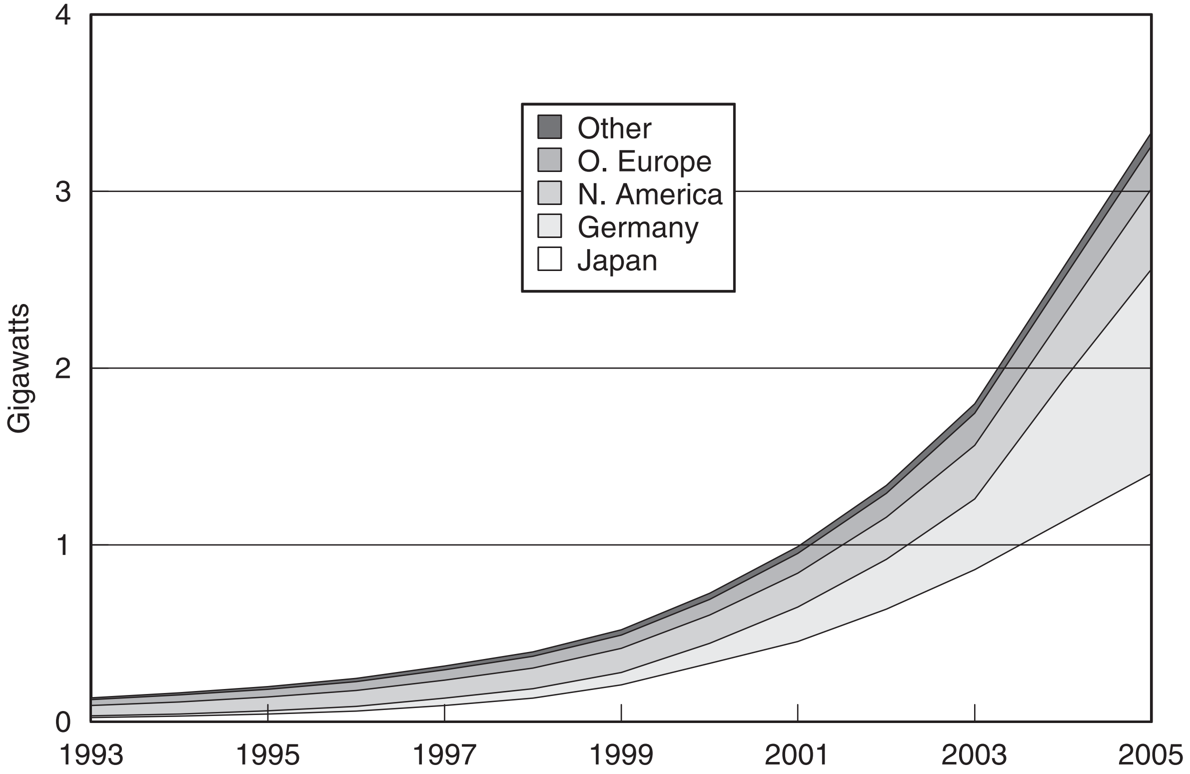
Most commercial PV solar cells are made of crystalline silicon cut in wafers as thin as 200 microns, usually between two and three square inches (12.5 to 20 square centimeters) in area. Single-crystal PV cells are grown and have a commercial efficiency that ranges between 15–18 percent in converting solar energy to electricity. Solar cells have a higher efficiency if surrounded by cool rather than warm air. The space program normally uses more expensive PV cells made of gallium arsenide, whose efficiency in transforming solar energy to electricity can exceed 30 percent.[17] Multicrystalline PV cells depend on a less expensive melting and solidification process, but have a marginally lower commercial efficiency of 14 percent. An even lower-cost solar cell is a film of extremely thin layers of PV semiconductor materials, such as amorphous silicon, copper-indium-gallium-diselenide, or cadmium telluride, deposited on a backing of glass, stainless steel, or plastic. While cheaper to make, thin-film PV arrays have a lower efficiency that ranges between 7–13 percent, so they have to cover more area to produce the same output than conventional solar panels. The advantage of thin films is avoiding the glass covering and mechanical frames of conventional solar panels. Thin films on a plastic covering can be made to look like roofing shingles and designed to fulfill the twin roles of protecting the roof from weather plus generating electricity from the sun. The savings in not having to install roofing shingles reduces the cost of the solar power system. There is also research on employing nanotechnology to produce organic solar cells of molecular polymers and other esoteric materials.

A PV, or solar, cell is the basic building block, small in size, and capable of producing 1 or 2 watts of power. These are combined into larger units, called modules or panels, which produce 50–300 watts of power, which are then joined together to form solar arrays sized to meet the desired power output. Solar arrays are particularly useful for serving isolated buildings in sunny climates such as lodges in national parks, lighthouses, and other buildings and facilities far from an electricity grid. Smaller solar modules can light signs, streets, gardens, pools, and provide power for remote telephones or automatic teller machines, or any need with similar power requirements. These applications normally have an associated battery that is charged by day in order to supply power at night or during times of inclement weather.

Solar arrays are given serious consideration by government bodies to help launch economic development in areas too remote and/or sparsely populated to justify building an electricity grid. An independent solar power system, with a battery to store electricity for times of cloud cover or at night, in isolated locations obviates the need for building a generation, transmission, and distribution system. A good example of this is a $48 million solar power project on the island of Mindanao in the Philippines. The project, funded by the Spanish government and built by BP Solar, will supply electricity to 400,000 people in 150 villages plus provide electricity for irrigation and drinking water systems and for schools and medical clinics.

Solar power can bring electricity to a remote area at less cost than building a conventional generation, transmission, and distribution system, including the purchase of normally imported fuel. Solar and wind power, either singly or together, with a diesel backup, are viable means of supplying electricity on a local or distributive basis to the 2 billion people who live in isolated communities far from electricity grids or in areas of low population density. One of the chief benefits of introducing hybrid renewable electricity to remote locations is improving the health of the people. Females in some parts of the world spend nearly every waking hour collecting and transporting dung and wood on their heads or backs for cooking and heating. Some have to walk twenty miles a day, rising early in the morning and going to bed late at night, ruining their health in the process. Moreover, cooking with biomass in closed environments is another major health hazard that shortens life expectancy. Electricity from renewable sources eliminates these time-consuming and onerous tasks and their adverse health consequences. Electricity allows women to sleep longer, improving their health; and, when they are awake, have the energy to improve their lifestyle.

Figure 9.2 **Growth in Solar Photovoltaic (PV) Power**



If an electricity grid is available, solar arrays can be connected into the grid for a power supply by night or on cloudy days, eliminating the need for a battery. Arrangements can be made for excess production of a solar array to be fed into the grid, the revenue of which can be part of the economic decision to install a solar power system. In addition to the solar panels, the capital investment also includes the cost of a mounting structure and the installation of the array, an inverter to convert the direct current output to alternating current, a storage battery for off-grid solar systems, along with a charge controller for battery operation or modifications to an existing electricity grid to allow the sale or purchase of electricity.[18] As Figure 9.2 shows, growth in solar power is exponential, similar to wind, but provides less than 10 percent of wind power's contribution.

The development of solar power, as with wind, started in the United States (California, to be precise) as a result of PURPA legislation. Since the mid-1990s Japan and Germany have become centers of solar energy development. The largest suppliers of PV cell production are the Japanese companies Sharp, with a 29 percent share of the market, Kyocera, with an 11 percent share, Mitsubishi Electric, Q-Cells of Germany, and BP Solar, each with a 7 percent share, and Shell Solar, with a 6 percent share. A large number of companies are involved with other aspects of solar power including supplying semiconductor materials, producing PV modules, and installing solar panel arrays. In addition, there is a great deal of entrepreneurial effort by companies trying to establish a niche in this emerging business.

Significant government monies are being invested in the development of solar power. In 2004, the four nations that spent the most for solar energy research and development were the United States ($86 million), Japan ($61 million), Germany ($30 million), and the Netherlands ($16 million). The two nations that spent the most for demonstrations and field trials were Japan ($103 million) and the United States ($11 million). For solar energy market stimulation, Germany ($309 million), the United States ($180 million), and Japan ($49 million) spent the most money. In terms of overall spending, Germany led at $339 million, the United States was second at $277 million, and Japan was third at $212 million, followed by a large gap to fourth place, Italy, at $35 million.[19] Japan, Germany, the United States, and other nations offer incentives for individuals and businesses to install solar energy. Like those provided for converting to wind and other alternative sources of energy, these come in various forms such as a direct grant or rebate paid to the individual or business for installing solar panels, various tax benefits, soft loans at below market interest rates and long payout periods, and the right to sell excess production back to the utility at above-market rates. In the United States, thirty-five states have some sort of program to encourage the development of solar energy, ranging from personal, corporate, sales, and property tax exemptions plus loan and grant programs as a means of inducing homeowners to install solar power. Significant rebates are offered by various states such as New York and New Jersey. In addition, the government has a production tax credit that can be applied against corporate taxes for companies that install solar power and other forms of renewable energy.

Economics of Solar Power

The economics of solar power depend on many variables. For example, an energy-efficient 2,000-square-foot home needs about 2 kilowatts of output from a solar array mounted on the roof. If the cost of installation is $16,000 and if there is a rebate available for $8,000, the net cost to the consumer is $8,000. The amount of electricity that the solar array can produce is 100,000 kilowatt hours over its twenty-five-year life, assuming that the sun is shining an average of 5.5 hours per day (2 kilowatts per hour × 5.5 hours per day × 365 days per year × 25 years). The 5.5 hours per day takes into consideration reduced output when the sun is near the horizon and during times of cloud cover. At higher latitudes, three hours per day might be a closer approximation for the equivalent of sunlight directly overhead with no cloud cover. If the average cost of electricity over the next twenty-five years were eight cents per kilowatt hour, then the avoidance of the need to purchase 100,000 kilowatts from a utility would equal the net investment of $8,000. Without the rebate, there is no economic justification for installing solar power.

Solar power works better if electricity rates are based on time-of-day metering when rates track actual demand. This would improve the economics of solar power immensely since the day rate for electricity is much higher than the night rate, reflecting marginal rates charged by base-load electricity providers. If electricity rates during daylight hours were sixteen cents per kilowatt hour, then the

Table 9.1

Economic Analysis of Actual Solar Energy System

Aggregate Savings (Costs) Over 25-Year Life of Project	
Avoided Electricity Purchases	$2,415,000
Avoided Transformer Losses	$ 50,000
Estimate of Sales of Excess Electricity Back to Utility First 4 Years Only	$ 315,000
Maintenance of Solar Energy System and Roof	($110,000)
Aggregate Savings	$2,670,000
Cost of System net of Rebate	$1,145,000
Internal Rate of Return over 25 Year Period	8.3%

investment of $8,000 (after the rebate) will generate savings of $16,000 in avoided electricity purchases over a twenty-five-year period, providing a 2.8 percent return on the investment. If there were no rebate, the savings would only compensate for the investment, assuming the money has no time value. If it is profitable to install solar power, then one can consider oversizing the array and selling the excess power back to the utility. Regardless of the economic analysis, one may still choose solar power just for the satisfaction of having a home that does not require burning a fossil fuel or relying on nuclear power.[20]

Installing a 454-kilowatt solar array for Monmouth University in New Jersey had a capital cost, including installation, of $2,860,000. A substantial rebate was available from the state of New Jersey in the amount of $1,715,000 (60 percent of capital cost). This reduced the capital investment to $1,145,000. Table 9.1 shows the economic analysis of the installation assuming a cost of purchase of ten cents per kilowatt-hour, with 3 percent escalation over the twenty-five-year life of the solar panel.

The net capital investment of $1.1 million earns a healthy return, primarily in the form of avoided electricity purchases. Any hike in electricity rates above ten cents per kilowatt-hour, which occurred as a consequence of much higher natural gas prices for the utility, increases the rate of return on the investment. Moreover, sales back to the utility are only assumed for the first four years. The project, under the management of PowerLight Corporation, will take about a year to complete, with material procurement and installation each taking about six months. From the environmental viewpoint, this solar power system reduces carbon dioxide emissions by 5,000 tons because fossil fuels will not have to be burned to generate the displaced electricity over the life of the project. This is equivalent to planting 1,500 acres of trees, removing 1,000 automobiles from the road, or driving automobiles 13 million fewer miles. But, as the analysis plainly shows, the internal rate of return is positive only because of the significant commitment on the part of the government to support the development of solar power.

One of the over 300 utilities that offer electricity from renewable energy to consumers is Arizona Power Service. For a premium rate, consumers can participate in the Solar Partners Program and a portion of their power requirements will be satisfied by solar power. The premium rate compensates for the extra cost of solar power over conventional sources of electricity. Located in the Southwest, where the company can take advantage of the 300 days of sun, and with plenty of available desert land for installing solar arrays, the utility has become a leader in promoting solar power to its customers. The company runs a 100-kilowatt High Concentration PV (HCPV) plant that consists of relatively inexpensive plastic lenses that concentrate sunlight 250 times over normal sunlight before reflecting the sunlight on a solar array. The concentrated sunlight reduces the required area of a solar panel array by a substantial amount to produce the same amount of power. This reduced capital investment results in lower-cost electricity. At another site, the company is building a 6,200-kilowatt (6.2 megawatt) solar power plant, of which 4.7 megawatts is a conventional solar panel array and 1.5 megawatts is a HCPV plant.[21]

Further government support for solar power occurred at the end of 2005 when the California Public Utilities Commission unveiled a plan to install 3 gigawatts (3,000 megawatts) of capacity over the next eleven years. This plan would double existing global solar power capacity and would supply 6 percent of California's peak electricity demand. The California Solar Initiative provides $3.2 billion in rebates over the next eleven years, with the objective of installing solar panels on 1 million homes and public buildings. Funding would also be eligible for solar water heating along with other solar power technologies. The new initiative is actually an expansion of an existing program that adds a surcharge to consumer utility bills with the proceeds dedicated to rebates for solar power installations. In addition, the Energy Policy Act of 2005 provides for tax credits for solar installations to be made available to homeowners as well as commercial businesses, although there is only a two-year life on these credits, and a maximum of $2,000 per installation. Businesses can claim a tax credit of 30 percent, up from the previous 10 percent, of the cost of an installed solar system with no dollar limit. This benefit reverts back to 10 percent in 2008.

Ocean Energy

The oceans cover over 70 percent of the earth's surface and represent an immense reservoir of energy in the form of tides, currents, waves, and temperature differentials. Tides result from the gravitational interaction of the earth and the moon, with about two high and two low tides each day. The time between high tides is twelve hours and twenty-five minutes, which means that the timing of power output changes each day in a predictable manner, but not necessarily corresponding to the time of high demand for electricity. Tides are also affected by the relative placement of the sun and moon with respect to the earth, which causes spring (maximum) and neap (minimum) tides. The elliptical path the earth traces around the sun, plus weather and other influences, affect tides, as does the topography of the shoreline. Unfortunately, coastal estuaries that can create tidal rises of up to fifty feet are located at high latitudes, far from population centers. Tidal power output must be viewed as a supplemental power source available only about eight to ten hours a day.

The concept of tidal power is fairly old; waterwheels powered by tidal currents ground grain in eleventh-century England. Tidal power can be tapped by building a dam with sluice gates across an estuary at a narrow opening to reduce construction costs. The sluice gates are opened to allow an incoming tide to increase the height of the water. When the tide turns, the sluice gates are closed, entrapping the water. As the tide goes out, the water level differential on both sides of the dam widens until there is a sufficient head for water to pass through specially designed turbines to produce electricity.

A tidal dam must be located where there is a marked difference between high and low tides. One favored area proposed for building a tidal dam is the Bay of Fundy in eastern Canada, where the difference in water level between tides is over fifty feet, the highest in the world. Other areas with pronounced tides in the northern hemisphere are Cook Inlet in Alaska, the White Sea in Russia, and the coastline along eastern Russia, northern China, and Korea. In the southern hemisphere, potential sites are in Argentina, Chile, and western Australia. With electricity production limited to between eight and ten hours a day, a tidal dam has an effective output of only 35 percent or so of rated capacity and the timing of its output may not correspond to the timing of peak demand. Moreover, a substantial investment is necessary to transmit the generated electricity from remote sites conducive to tidal dams to population centers.

There is only one major source of tidal energy in the world: the tidal dam at La Rance estuary in France, built in the 1960s, which is capable of producing 240 megawatts of electricity. It has a maximum tidal range of twenty-six feet, operates at 26 percent of rated capacity on average, requires low maintenance, and is in service 97 percent of the time. Three other such dams are far smaller; one an 18-megawatt tidal dam at Annapolis Royal, Canada (Bay of Fundy), which serves the local area, a 0.5-megawatt dam in eastern China, and a 0.4-megawatt dam in the White Sea at Kislaya, Russia. One proposal under consideration since the 1980s is to build a sixteen-kilometer (nearly ten miles) dam across the Severn estuary in the United Kingdom. It would have a maximum output of 8.6 gigawatts, employing 216 electricity-generating turbines, and would be capable of supplying about 5 percent of the United Kingdom's electricity demand. Another under consideration since 1990 is a 48-megawatt tidal dam near Derby in northwestern Australia.[22]

A more effective way to harness tidal power is channeling tidal flow through a restricted waterway so that the tidal current powers turbines during incoming and outgoing tides. This "double flow" system provides electricity generation whenever the tides are running, but not during a change in tides when the tidal current reverses. While the double flow system has a higher effective output than the tidal dam, electricity generation is still not continuous and may not be timed to accommodate demand. One proposal is to build a tidal fence two and a half miles long (four kilometers) in the San Bernardino Strait in the Philippines between the islands of Samar and Dalupiri. The tidal fence would contain 274 turbines capable of generating 2,200 megawatts (2.2 gigawatts) at peak tidal flow.[23] The Race Rocks Ecological Reserve off the coast of Vancouver Island is building a small power plant for local use that will be powered by a tidal current.

The "double basin" method of tidal power provides a continuous supply of electricity because the water flows continually from a higher basin to a lower basin. Water in the upper basin is replenished during high tide and water accumulating in the lower basin is drained during low tide. Numerous ongoing research development programs are experimenting with several techniques of installing a generator with a propeller in a river. Tidal or river currents with an optimal speed is 4.5 to 6.7 miles per hour (2 to 3 meters per second) turn the propeller and the generated electricity is transmitted to shore via underwater cables. In principal, this is similar to a wind turbine. The major difference is that water is 850 times denser than air, which allows a smaller propeller to generate electricity at a lower rate of rotation. Tidal turbines with thirty-four-foot-long blades are under construction in Norway, and will be capable of generating 300 kilowatts of electricity when the tide is running.[24]

Waves are caused by wind and their enormous energy potential can be tapped by using hydraulic or mechanical means to translate the up-and-down motion to rotate a generator. Calm weather and severe storms affect the operation of these devices, but when they are in operation electricity can be delivered to shore via underwater cables. While one may feel that this energy source is futuristic, thousands of navigational buoys have long relied on wave motion to power their lights. The height of a column of water in a cylinder within the buoy changes with the up-and-down motion of the buoy, creating an air-pressure differential that drives a piston that turns a generator to supply power for the lights, sound signals, and other navigational aids of the buoy. A battery is kept charged by the wave motion in case of calm weather.[25] One wave-power system has been in operation since 1989 and produces 75 kilowatts for a remote community at Islay in Scotland. A Scottish company is deploying wave-powered elongated metal semisubmerged, sausage-shaped tubes off Portugal to provide electricity to 1,500 homes. The wave motion pumps fluid to hydraulic motors to drive generators, whose output is fed to shore via an underwater cable. If the first three units with a capacity to generate 2.25 megawatts operate successfully when they are completed in 2006, another thirty will be installed, bringing the total capacity to 20 megawatts.[26] Another company has a buoy, more conventional in appearance, capable of converting wave energy to a mechanical stroking action that drives an electricity generator. These buoys are in the process of being deployed off the coast of New Jersey, Hawaii, and Spain.[27]

The last method of extracting energy from the ocean is to take advantage of temperature differentials. The warm temperature of ocean surface water can be used to vaporize a working fluid, such as ammonia, which boils at a low temperature, to drive a turbine to generate electricity. The working fluid is cooled and condensed for recycling by deeper cold water. The warmed cold water must be pumped back into the ocean's depths to prevent cooling the surface. Ocean thermal systems are located in the tropics, where warm surface waters lie over deep cold waters. This provides the greatest temperature differential for operating a turbine; even so, the efficiency of heat transfer at these relatively small temperature differentials is only 5 percent, a technical challenge that requires building and operating a heat exchanger large enough to produce a significant amount of electricity. Demonstration plants have been built, including one in Hawaii that produced up to 250 kilowatts of electricity for a number of years. However, technical problems associated with ocean thermal energy still pose a significant barrier to developing this source of energy on a commercial scale.

One idea is to have "grazing plants" located far from shore, which would preclude having an underwater cable connecting the grazing plant to the shore. These plants would tap the temperature differential in open waters and utilize the electricity to produce hydrogen by electrolyzing water. Hydrogen then becomes a stored form of electricity that can be shipped from the grazing plants in specially designed vessels to shore-based terminals.

Sustainable Energy and Economic Development

In recent years governments are becoming increasingly reliant on market forces to allocate material, human, and financial resources for the selection of energy sources for electricity generation. Many nations are privatizing formerly government-owned utilities to introduce an element of competition that can lead to greater efficiency, lower electricity rates, and attract private capital. Governments are moving toward new regulatory frameworks for making rules and monitoring the application of those rules to ensure that markets work efficiently while, at the same time, advancing social causes. Increased globalization of corporate activity, flow of information by modern means of communication, including the Internet, and increased awareness of people of what is going on around them have changed how governments approach economic development. With a growing recognition of the failure of all-encompassing bureaucracies to provide the fundamental necessities of society, the world is moving from centralized planning by entrenched bureaucratic elites to more regionalized or localized planning with centralized guidance. This fundamental change in the decision-making process is reshaping the future world of energy suppliers and providing a major impetus for adopting sustainable sources of energy.[28]

For economic development to take place, something has to be done for the one-third of humanity who survive without electricity, relying primarily on biomass for heating and cooking. Substituting kerosene or propane for dung and wood in heating and cooking and diesel generators for lighting only aggravates the deteriorating situation in oil. Diesel generators and kerosene or propane heaters and stoves represent a low capital investment with a high operating cost in terms of fuel. For the many developing nations with no domestic sources of oil and negative trade balances, the days of recycling petrodollars are at an end. In the past, bank deposits by oil producers were lent out to developing nations to buy petroleum products. This was a good deal for buyers, who used the loan proceeds to purchase petroleum products, and for suppliers, who received cash for their oil. The cash from these sales was again placed on deposit in banks for another round of petrodollar recycling. Petroloans for many developing nations have reached a point where debt servicing is consuming a large part of government revenues, leaving little for social services and economic development. The banking community has admitted that these loans will never be paid and over $50 billion of petroloans to sub-Saharan nations are in the process of being written off the books. Having taken a loss with past petroloans, banks presumably will not be overly eager to enter into new ones. Without being able to borrow petrodollars to fund their purchases of oil, what are the many developing nations without domestic oil reserves and with a negative trade balance to do?

Yet these nations are in desperate need of energy. Economic development centered on electricity is the only way to alleviate extreme levels of poverty in areas where human efforts are primarily dedicated to collecting dung, wood, and water. Furthermore, electricity, through communications, increases the awareness of people of the world around them and, through education and training, raises the knowledge level and technical skills of the people, making them better able to help themselves. About ten or more new businesses are created for every hundred households hooked up to electricity. Electricity frees up time from performing domestic chores and provides the power for efficient and reliable manufacturing of goods to meet the basic necessities of life. When electricity replaces biomass for cooking and heating, it promotes the health of the people by improving the quality of indoor air. Electricity also improves the environment by slowing or stopping deforestation, perhaps even allowing reforestation to take place, a result that would help remove some carbon dioxide from the atmosphere.

The Human Development Index (HDI) is a measure of per capita GDP in terms of purchasing power parity, life expectancy at birth, adult literacy, and the number of those enrolled in institutions of higher learning. Studies have shown that electricity introduced into a society, even on a limited per capita basis, has a dramatic impact on improving the HDI, the well-being of the people. After the initial spurt, further gains in the HDI level off, with increasing availability of electricity.

This suggests that introducing electricity over wide areas for low per capita consumption is more beneficial to society than restricting electricity to small areas for high per capita consumption.

Using fossil fuels for electricity generation and distribution with long-distance transmission lines cannot economically serve remote areas with low population densities. The only alternative is distributive or local electricity generation based on sustainable sources of power, not fossil fuels. This is a switch from investments of low capital cost with high operating costs (diesel generators) to high capital cost with low operating costs (solar, wind, mini-hydro, and geothermal, if available). In addition, the widespread introduction of electricity from sustainable sources raises the technical expertise of the local people responsible for its operation and maintenance. With this background, they can become involved with other activities that can benefit society.

While conventional utilities can best oversee centralized generation and transmission of electricity for high-density population areas, local distributive systems are better for remote areas far from a power grid. Distributive systems push planning, policy, and decision making away from politically influenced bureaucracies in urban centers to local communities where those most affected, the local citizens, can participate in the decision-making process. A case in point is Bangladesh, where the traditional electric utilities were rife with ineffectual management and corruption. Because of managerial ineptness and a bureaucratic focus on fulfilling political objectives (winning the next election), the centralized utilities suffered significant cash shortfalls from nonpaying customers, special interest groups whose subsidized rates did not come close to covering costs, and theft (unofficial tapping into the electricity grid). This resulted in poor maintenance, unreliable service, and an inability to expand service to new areas beyond the electricity grid.

Tired of the old system not working, the Bangladesh government set up nonprofit rural electricity cooperatives operating under the general guidance of a government oversight organization. Placing the operation of the electricity generation and distribution system in local hands, with no government handouts to compensate for cash shortfalls, resulted in higher revenue generation from greater efficacy in bill collection, curtailment of subsidized rates to special-interest groups, and an end to theft. The cooperatives practiced greater financial discipline and fostered greater community involvement than traditional utilities. The cooperatives were not truly independent because the central government oversight organization set annual performance targets and conducted regular management reviews and financial audits. The oversight organization had the authority to dismiss incompetent or corrupt managers and reward those who met their performance targets with bonuses. Board members of the cooperatives were limited to three-year terms, as were meter readers. Billing was in the hands of women, who seemed more capable of rooting out dishonest practices than men.

While rural cooperatives only serve a relatively small portion of Bangladesh's electricity needs, based on their track record their role is expanding. Rural cooperatives have proven that the local community can handle the responsibility of managing and operating electricity generating and distribution systems. Moreover, as nonprofit organizations that have to cover costs, they could do what the government utilities could not: eliminate subsidies to special-interest groups whose votes were considered crucial in elections. With the establishment of property rights and government institutions oriented to electricity generation as a business rather than as a means of political patronage, rural cooperatives have been able to attract direct investments from both domestic and foreign sources. By stemming the significant cash drain on the government budget, the cooperatives have freed up capital for the central government to pursue other avenues of economic development.

A Sustainable Energy Action Plan

Sustainable energy is coming into its own, partly as a result of decentralization, which allows local people to become more involved with the decision-making process, and partly from the surge in petroleum prices. Oil price hikes from production and refining capacity constraints have also dragged along the prices of natural gas and coal. This affects the entire gamut of energy costs, from diesel fuel for local electricity generators to kerosene and propane for cooking and lighting. Moreover, growing alarm over dependency on politically volatile oil-exporting nations, and the potential climatic consequences of carbon dioxide emissions, are driving nearly every nation, rich and poor, to support the development of sustainable energy in some fashion. For example, China has announced a long-term $180-billion program with the goal of providing 15 percent of its total energy consumption from a combination of solar, wind, and biomass (e.g., capturing methane from agricultural waste).

Sustainable energy is now being viewed as a means of mitigating the risk of oil supply disruptions, fossil fuel price hikes, and the environmental impact of increased carbon dioxide emissions. Thus, the higher cost of sustainable energy, mainly concentrated in the initial capital investment, can be rationalized as an insurance premium against the risk of economic turmoil stemming from interruptions in oil flows, fossil fuel price hikes, and climatic change.

Advocates of sustainable energy maintain that greater reliance on sustainable energy is not optional, but mandatory, given what appears to be the depletion of petroleum reserves and the potential destabilization of the earth's climate by greenhouse gases, primarily carbon dioxide. Sustainable energy cannot replace petroleum products for motor vehicles aside from the role it can play in the hydrogen economy. As will be discussed in the next chapter, the day of the hydrogen-fueled motor vehicle is decades away. Sustainable energy reduces carbon dioxide emissions by replacing fossil fuels for electricity generation. This principally affects coal and natural gas because relatively little petroleum, other than certain waxy grades and the bottom of the barrel left over from the refining process, is burned to generate electricity.

The proposed action plan's approach to petroleum is to reduce consumption, not substitute oil with sustainable sources. Encouraging conservation and enhancing energy efficiency also play major roles. The action plan to begin transforming an economy from fossil fuels to sustainable energy sources is based on:

- removing government subsidies for fossil fuels and nuclear energy;
- placing a gas-guzzler tax on automobiles with low mileage to discourage their purchase and reduce petroleum consumption;
- inaugurating a tax credit program for the purchase of high-mileage motor vehicles such as hybrids and motor vehicles run on alternative fuels;
- encouraging greater government subsidies and/or tax benefits to advance the development and adoption of sustainable energy;
- passing a national building code that establishes standards to enhance energy efficiency and reduce energy consumption;

- changing local building, permit, and zoning laws that require solar water heating and electricity generated from solar arrays and/or wind turbines on new buildings, when feasible;
- supporting zero-energy building designs that will, on balance, consume no energy by means of greater insulation and energy-efficiency standards combined with solar water heating and solar panels or wind turbines for electricity generation; *zero energy* means that sales and purchases of electricity from and to a utility are equal; that is, no net purchases;
- establishing a national renewable energy standard to mandate a growing portion of electricity output to be from renewable energy sources;
- requiring utilities to make the necessary adjustments to their distribution systems to permit nondiscriminatory interconnection in order that excess electricity produced by renewable sources at homes and buildings can be sold into the electricity grid;
- requiring time-of-day automatic metering in order that electricity produced during the daytime and sold to utilities obtains a rate that reflects the higher marginal cost of electricity. This will encourage the installation of solar panels and provide a greater return on investment;
- counting sales of electricity back into the electricity grid generated by renewable sources in individual homes and businesses as part of the required purchases of green power imposed on utilities;
- instituting a public awareness campaign on the benefits of running an economy with sustainable energy;
- instituting a utility awareness campaign that business as usual, a near total reliance on centralized power plants fueled by fossil fuels with long-distance transmission lines, will have to give way, at least in part, to a distributive electricity system powered by sustainable sources of power in which producers of electricity can readily sell their output to utilities;
- encouraging political leadership to promote sustainable energy sources with the will to see this process through to its conclusion.[29]

While the contribution of sustainable energy is still minuscule compared to that of fossil fuels, one must not forget that the use of coal was once minuscule compared to biomass, as was that of oil compared to coal and natural gas compared to oil. Now the contribution of sustainable energy is minuscule compared to coal, oil, and natural gas, but that does not mean that its contribution will forever remain minuscule.

Notes

1. From the Geothermal Education Office, available online at http://geothermal.marin.org.
2. A description of this endeavor is available online at the Geodynamics Ltd. Web site at www.geodynamics.com.au.
3. Wendell A. Duffield and John H. Sass, *Geothermal Energy–Clean Power from the Earth's Heat,* Circular 1249 (Washington, DC: U.S. Geological Survey, U.S. Department of the Interior, 2003).
4. Vaclav Smil, *Energy in World History* (Boulder, CO: Westview Press, 1994).
5. The data presented in Figure 9.1 pertains basically to the situation in 2002, augmented by further data drawn from the American and the European Wind Energy Associations Web sites, www.awea.org and www.ewea.org, respectively, and the Global Wind Energy Council's Web site, www.gwec.net.
6. For more information, visit the Capewind project Web site at www.capewind.org.
7. This information can be found at General Electric's wind turbine Web site at www.geopower.com.

8. For such comparisons, visit the Terra Moya Aqua (TMA) Company Web site at www.tmawind.com.

9. American Wind Energy Association, *Outlook 2005 for Wind Power*, available online at www.awea.org.

10. For example, the Australian joint venture utility company ActewAGL offers the GreenChoice Program on its Web site at www.actewagl.com.au.

11. For more information, visit the Web site of the National Wind Coordinating Committee at www.nationalwind.org.

12. From United Nations Development Program (UNDP), *World Energy Assessment: Overview 2004 Update,* New York, 2004.

13. This information can be found online at the National Renewable Energy Laboratory's Web site at www.nrel.gov.

14. This information can be found at the Energy Efficiency and Renewable Energy Solar Technologies Program Web site at www.eere.energy.gov/solar/csp.html.

15. Information available from the Web site of the Enviromission Company at www.enviromission.com.au.

16. More information on this project can be found at the Web site of the Solar Energy Technologies Program of the U.S. Department of Energy Efficiency and Renewable Energy at www.eere.energy.gov/solar/pv_cell_light.html.

17. For additional information, visit the PV Power Resource Site Web site at www.pvpower.com.

18. From the International Energy Agency Photovoltaic Power Systems Program Report IEA-PVPS T1–14:2005.

19. These statistics can be found at the Web site of the International Energy Agency's Photovoltaic Power Systems Program at www.iea-pvps.org.

20. For more information, visit the OKSolar Company Web site at www.OKSolar.com.

21. This solar program is described on the Arizona Power Service Web site at www.aps.com.

22. Kimberly K. Smith, *Powering Our Future* (Lincoln, NE: Alternative Energy Institute, iUniverse, 2005).

23. For more information about this proposal, visit the Snowy Mountains Engineering Corporation Web site at www.smec.com.au.

24. Paula Berinstein, *Alternative Energy: Facts, Statistics, and Issues* (Westport, CT: Oryx Press, 2001).

25. Edward S. Cassedy, *Prospects for Sustainable Energy: A Critical Assessment* (New York: Cambridge University Press, 1999).

26. More information on this project can be found at the Web site of Ocean Power Delivery Ltd. at www.oceanpd.com.

27. For more information on this type of buoy visit the Web site of Ocean Power Technologies at www.oceanpowertechnologies.com.

28. Thomas B. Johansson and José Goldemberg, eds., *Energy for Sustainable Development: A Policy Agenda* (New York: United Nations Development Program, 2002).

29. Joel B. Stronberg, *Making the Transition to a Sustainable Energy Economy* (Boulder, CO: American Solar Energy Society, 2005).

10

Looking Toward the Future

This chapter deals with the hydrogen economy, climatic change, the impact of fossil fuels on the environment, legislative acts to deal with air pollution, and energy efficiency and conservation.

The Hydrogen Economy

Hydrogen is the most abundant element in the universe, making up 75 percent of its mass and 90 percent of its molecules. Hydrogen, when burned as a fuel, emits only water and heat, the cleanest source of energy by far. Though plentiful in the universe, there is no free hydrogen here on Earth. While some is locked away in hydrocarbons and other chemicals, most of what there is has already been burned and its product of combustion is all around and in us: water.

Curiously, human progress in energy has been marked with de-carbonizing fuel sources. For most of history, humans have burned wood, which has the highest ratio of carbon to hydrogen atoms, about ten carbon atoms per hydrogen atom, in comparison to fossil fuels. This means that burning wood emits more carbon dioxide than burning fossil fuels for an equivalent release of energy. Coal, the fossil fuel that sparked the Industrial Revolution, has about one or two carbon atoms per hydrogen atom, which means it emits less carbon dioxide than wood. Next is oil, with one half of a carbon atom per hydrogen atom (or one carbon atom for every two hydrogen atoms), and natural gas is last, with one-quarter of a carbon atom per hydrogen atom (or one carbon atom for every four hydrogen atoms). Thus, as people have learned to use new fuels, each one was a step down in carbon dioxide emissions for an equivalent release of energy. The ultimate step is hydrogen, which has no carbon atoms and, therefore, no carbon dioxide emissions, no emissions of carbon monoxide, sulfur, nitrous oxides, and other progenitor chemicals that create smog, and no metallic emissions (mercury, arsenic); hydrogen produces only plain water and heat.

Henry Cavendish discovered hydrogen in 1776 and Antoine Lavoisier named Cavendish's "life sustaining air," oxygen, and "inflammable air," hydrogen. Hydrogen is colorless, odorless, has no taste, and burns with a pale blue flame virtually invisible in daylight. In the 1870s Jules Verne thought that water would be the fuel of the future. In 1923 John Haldane predicted that future energy would be in the form of liquid hydrogen. Rows of windmills would generate electricity to produce hydrogen by the electrolysis of water. Hydrogen gas would then be liquefied and stored in vacuum-jacketed underground reservoirs until needed to generate electricity when recombined with oxygen. Although his idea was ridiculed at the time, Haldane's prediction is essentially where we are headed today.[1]

The fuel for the engines on German-made Zeppelin dirigibles that carried passengers between European cities and across the Atlantic Ocean to the United States varied from diesel fuel to a mixture of benzene and gasoline, augmented by excess hydrogen blow-off as a booster fuel. The crash of the *Hindenburg* in 1937 ended the days of dirigibles filled with hydrogen, which was replaced with helium. The *Hindenburg* gave hydrogen a reputation for being a dangerous fuel. It was originally thought that an atmospheric electrical charge called St. Elmo's Fire, a blue glow sometimes seen around church spires, sailing masts, and airplane wings during stormy weather, ignited the hydrogen. More recent investigations into the cause of the tragedy point to other possibilities such as an electrical discharge that ignited not the hydrogen, but the highly combustible coating of aluminized cellulose acetate butyrate dopant, a component of rocket fuel, which saturated the outer cotton fabric. Another possibility was leaking fuel dripping on a hot surface that started a fire within the internal structure of the dirigible. The investigations concluded that design faults and operating deficiencies made the *Hindenburg* a bomb waiting to be detonated. Regardless of the cause of the fire, once the hydrogen ignited, the end came quickly.

Hydrogen also has gotten a bad rap by being associated with the hydrogen bomb, which, of course, has nothing to do with combustion. On the other hand, hydrogen aficionados maintain that a tank full of hydrogen in an automobile presents no more of a hazard to a passenger than a tank full of gasoline. They argue that it may be less hazardous because a ruptured gasoline tank spills its contents on the ground, which, if ignited, will almost completely combust. In contrast to a ruptured hydrogen tank, a large portion of the fuel may escape to the atmosphere before ignition takes place. Hydrogen, the lightest of all elements, has a very high diffusion rate and disperses four times faster than natural gas and ten times faster than gasoline vapors. Moreover, hydrogen radiates relatively little heat compared to burning petroleum, and personal injuries are confined to those caused by direct contact with the flame. On the downside, hydrogen can burn when its concentration in air is between 2–75 percent, whereas the flammable range of natural gas is between 5–15 percent. This means that natural gas cannot ignite if its concentration is less than 5 percent or greater than 15 percent, a much more restrictive range than hydrogen. Not only does hydrogen have a wider flammable range and "ignite" easier than gasoline or natural gas, its nearly invisible flame in daylight is another element of danger. Whatever the virtues of supposedly being less hazardous than commonly used gasoline, propane, or natural gas, this is not the public's perception of hydrogen. There are real safety concerns when one is dealing with hydrogen, as is the case with any flammable substance.[2]

Yet hydrogen has a long history of use. Hydrogen was a component of coal-derived town or manufactured gas along with methane, carbon dioxide, and carbon monoxide. This mixture of gases was burned in homes and businesses long before the discovery of natural gas fields and continued to be burned as late as the 1950s before being fully replaced by natural gas. Hydrogen has been an industrial commodity for over fifty years with about 50 million tons of hydrogen produced annually worldwide. Most hydrogen is a byproduct of reforming naphtha in oil refineries and is largely consumed within the refinery to increase the yield of more valuable light-end products from the heavy end of the barrel with a relatively small portion pipelined to nearby petrochemical plants.

Only about 5 percent of hydrogen consumed by industry is merchant hydrogen, specifically produced (mostly by steam reforming of natural gas) for commercial purposes. It is transported either in pipelines as a compressed gas or in tanks carried by rail, barge, and truck as either a compressed gas or a cryogenic liquid at a temperature of minus 253°C, only twenty degrees above absolute zero. Merchant hydrogen is used by the food industry for hydrogenation of edible organic oils and in making margarine, by the fertilizer industry for producing ammonia for nitrogen-based fertilizers, and by the semiconductor industry. The aerospace industry relies on merchant hydrogen for fuel cells aboard manned space stations to produce electricity and potable water. The handling of hydrogen by the oil industry and the producers and consumers of merchant hydrogen has an excellent safety record because of their understanding and appreciation of its inherent risks.

In the area of transportation, a number of experimental vehicles including submarines and torpedoes ran on hydrogen in the 1930s and 1940s. The 1973 oil crisis awakened the public to the possibility of hydrogen as a motor vehicle fuel. In 1988, the Soviet Union and the United States experimented with airplanes fueled by liquid hydrogen. The first hydrogen-fueled buses were in operation in Belgium in 1994, and in 1995 the city of Chicago tested hydrogen-fueled buses. A small number of hydrogen-fueled buses currently operate in several European and American cities, mainly for testing and demonstration purposes. A motor vehicle engine can be fairly easily converted from gasoline to natural gas or hydrogen. For hydrogen-fueled vehicles, the problem is fuel availability, cost, and storage, not engine design. But the role of hydrogen is not limited to its potential as a motor vehicle fuel. In 1992, the first solar home that relied on hydrogen as a means to store electricity, rather than a battery, was successfully demonstrated in Germany. In 1999, Iceland announced a long-term plan to become the world's first hydrogen economy by totally eliminating fossil fuels by 2050. Icelandic motor vehicles and fishing vessels will run on hydrogen produced by electrolysis of water and electricity will be generated from hydro and geothermal sources.[3]

The problem is how to produce hydrogen. Reforming it is a three-stage process that begins with the hydrocarbon (mainly natural gas, but coal can be used as well as various petroleum products) in an endothermic (heat-absorbing) reaction in the presence of a catalyst to form hydrogen and carbon monoxide. The second stage is combining carbon monoxide with steam in an exothermic (heat releasing) reaction to form additional hydrogen and carbon dioxide. Heat released by the exothermic reaction can be recycled to supply a portion of the heat for the endothermic reaction. The third stage is the removal of carbon dioxide and trace amounts of carbon monoxide through an adsorption process to separate the hydrogen. "Black" hydrogen results if the waste-product carbon dioxide is released to the atmosphere. An owner of a hydrogen-fueled automobile who proudly announces that he or she is not polluting the environment is suffering from a case of self-delusion if the automobile is running on black hydrogen. Hydrogen from reforming is quite expensive and efforts are underway to find a different technology such as advanced ion transport membranes to reduce the cost of separating hydrogen from hydrocarbons.

"Green" hydrogen results if the carbon dioxide emissions from steam reformers are sequestered such as in an integrated coal-gasification combined cycle (IGCC) plant. The U.S. government-sponsored FutureGen project employs ultra-low emissions technology, coupled with sequestering carbon dioxide emissions, in which coal produces electricity and hydrogen with virtually no emissions to the atmosphere. Green hydrogen can be produced from ethanol, which opens up biomass as a hydrogen fuel. While this process releases carbon dioxide, growing crops to supply ethanol removes an equivalent amount of carbon dioxide from the atmosphere. This makes hydrogen from biomass essentially carbon dioxide neutral as long as energy from biomass fuels, not oil, is consumed in the growing, harvesting, and processing of the crops. Another way to produce green hydrogen is electrolysis of water, in which the source of electricity is not a fossil fuel (unless carbon dioxide emissions are sequestered), but from nuclear, hydro, wind, solar, geothermal, tidal, or grazers (floating plants on the world's oceans generating electricity from thermal differentials).

Electrolysis is the flow of direct-current electricity between a positive and negative electrode in pure water that contains an electrolyte to enhance conductivity. Electricity splits the water molecule into its elements, oxygen, which collects at the anode, or positively charged electrode, and hydrogen, which collects at the cathode, or negatively charged electrode. The gases are drawn away from the electrodes, dried, and stored. Hydrogen can be a fuel for specially adapted conventional motor vehicle engines or for fuel cells. Oxygen can be pressurized in bottles and sold for industrial use or released into the atmosphere. One proposal is to install solar panels on hospital rooftops to generate electricity for the hospital and for electrolysis to produce oxygen and hydrogen. Oxygen can supply patients who need breathing assistance and hydrogen can supply fuel for ambulances and/or be sold to owners of hydrogen-fueled motor vehicles.

Only 4 percent of the world's output of hydrogen is by electrolysis because of the high cost of electricity compared to steam reformers that strip hydrogen from fossil fuels. A cost differential of three or four times puts hydrogen by electrolysis at a severe disadvantage. If hydrogen is to reduce carbon dioxide emissions, electricity for electrolysis cannot come from fossil-fueled plants (other than coal-burning IGCC plants that sequester carbon dioxide emissions), but from nuclear and sustainable sources of energy. From a practical viewpoint, the generation of the enormous quantities of electricity necessary for the hydrogen economy would have to depend largely on nuclear power and IGCC plants augmented by hydro, wind, and solar sources. The capacity of this combination of power sources could be expanded to the point of satisfying peak electricity demand. As electricity demand from consumers, businesses, and industry moves off its peak, excess electricity-generating capacity would be dedicated to hydrogen production. This combination of electricity-generating plants operating at full capacity would eliminate carbon dioxide emissions from electricity generation as well as the need for building plants to satisfy transient electricity demand. The uniform charge for electricity to consumers and to hydrogen producers would be the same base rate, eliminating marginal electricity rate differentials. The prospect of nuclear power and coal playing a major role in the hydrogen economy is viewed disdainfully by environmentalists, but not necessarily by hydrogen enthusiasts. These two groups should not be at odds with one another because both share a mutual desire to eliminate carbon dioxide emissions.

Hydrogen Stores Electricity

Energy (heat) generates electricity for producing hydrogen by electrolysis of water. Fuel cells reverse electrolysis by reuniting hydrogen with oxygen, producing electricity, heat, and water. Electrolysis consumes electricity when converting water to hydrogen and oxygen and a fuel cell generates electricity by converting hydrogen and oxygen to water. A fuel cell represents a twofold generation of electricity—once to produce hydrogen by electrolysis of water and again when the fuel cell converts hydrogen back to water. Hydrogen can be produced any time there is available electricity, then distributed, stored, and converted back to electricity when and as needed by fuel cells. While one might say that hydrogen fulfills the same role as a battery, there is a major difference. A battery stores chemical energy that is converted to electricity until the chemical energy is exhausted. As such, a battery is a finite source of energy that must be recharged or discarded when the chemical energy is gone. A fuel cell converts chemical energy (hydrogen) to electricity continually up to its rated capacity as long as hydrogen is fed to the fuel cell. It never runs down, never is exhausted, and never has to be recharged. Hydrogen in concert with a fuel cell is far more effective in storing electricity than a battery.

Sir William Robert Grove invented the hydrogen fuel cell in the 1830s, but, without practical use, the fuel cell faded from view. It was revived in the 1960s when General Electric developed a workable fuel cell as a power supply for the Apollo and Gemini space missions. Hydrogen, the fuel for fuel cells, must be uncontaminated and can be supplied from tanks pressurized with hydrogen gas or from associated reformers that extract hydrogen from natural gas, propane, methanol, gasoline, or other types of hydrocarbons. A fuel cell consists of a proton-exchange membrane (PEM) at its center, surrounded on both sides by a catalyst. On the outside of each catalyst is an anode or a cathode electrode connected by an electrical circuit that passes through a motor or light bulb or other electrical load. Hydrogen, passing through a flow plate at one end of the fuel cell, enters the anode catalyst, where the hydrogen is split into protons and electrons (a hydrogen molecule is two atoms of hydrogen each of one proton and one electron). The PEM only allows protons to pass through to the cathode catalyst, where they establish an electrical charge to induce the flow of the electrons from the anode electrode through the electrical load to the cathode electrode. Once through the PEM, protons pass to the cathode catalyst, where they are reunited with the returning electrons from the cathode electrode and with oxygen from the air to form water and heat. The waste products, which may contain a trace of nitrous oxides from the protons reacting with nitrogen rather than oxygen atoms in the air, pass through the flow plate on the other end of the fuel cell. A fuel cell stack of individual fuel cells is connected to others to form a fuel cell module that produces a desired output of electricity. Fuel cells have no moving parts and are about two to three times more efficient than an internal combustion engine in converting fuel into usable power. In addition to PEM fuel cells, there are also phosphoric acid fuel cells, molten carbonate fuel cells, and solid oxide fuel cells that have different operating temperatures and performance characteristics for specific applications.[5]

Water waste poses a problem when operating a fuel cell exposed to freezing weather. Waste heat prevents water from freezing during operation, but when a fuel cell is shut down, residual water can turn to ice, which damages the fuel cell. This, along with costs, presents a challenge for manufacturers of fuel cell-powered motor vehicles. Although there has been progress in reducing costs, fuel cells are still expensive to manufacture and to operate, given the cost of hydrogen. Further technical breakthroughs are necessary to make fuel cells competitive with gasoline engines. About fifty companies in North America, Europe, and Japan are dedicated to making fuel cells a viable means of supplying power to homes and motor vehicles. The current market for fuel cells is primarily as backup power supplies for critical communications and computer systems in which a power loss can have severe operational repercussions. These fuel cells, for the most part, have associated reformers that strip hydrogen from hydrocarbons. Fuel cells also supply electricity and potable water on manned space missions.

Major automobile companies are dedicating considerable resources to research on hydrogen-powered fuel cell cars, but mass production appears to be twenty or more years away. The automobile manufacturer that succeeds will have a competitive advantage over the others, particularly if there is a future shortage of gasoline. An automobile manufacturer who fails to participate will be at a competitive disadvantage, stuck with an outmoded gasoline-fueled engine while others are coming out with fuel cell-powered automobiles. Thus, the incentive for automobile companies to dedicate their resources to the development of fuel cell-powered automobiles can be viewed as either an offensive strategic move to achieve competitive advantage or as a defensive move to counter those taking the offensive. Either way, automobile companies' active involvement is a boost for fuel cell technology. In addition to funding from private sources, there is also public funding such as the U.S. government's $1.7 billion research program to develop FreedomCAR, a motor vehicle powered by a hydrogen fuel cell. California, with its tough air quality laws and calls for zero-emission vehicles, has become a national testing ground for battery-powered and hydrogen-fueled motor vehicles.

Hydrogen's low density, however, presents a serious logistics problem. When compressed to the same pressure as natural gas, hydrogen contains only about one-third the energy content. Once in an automobile as a compressed gas, the small molecules of hydrogen can more easily leak through cracks, porous material, and faulty fittings and gaskets than the larger methane molecules. The integrity of fittings and gaskets is an even greater technical challenge when hydrogen is compressed to 10,000 pounds in order to store a sufficient volume for a 300-mile driving range in a normally sized automobile fuel tank. Liquefied or cryogenic hydrogen at atmospheric pressure would require three times the volume required by gasoline, thus, three times the tank size, to deliver the same amount of energy. The reason why hydrogen-fueled buses have preceded automobiles is that large tanks can be mounted on top of buses to accommodate the storage requirements. Research is being conducted on methods of solid-state storage using metal hydrides (magnesium, lanthanum and nickel, sodium aluminum, or lithium boron) for automobiles. Hydrogen is stored within the molecular structure of metal hydrides and released by heating the storage medium.

Hydrogen-fueled motor vehicles represent the classic chicken-and-egg situation. No one is going to build hydrogen refueling stations without hydrogen-fueled cars and no one is going to build hydrogen-fueled cars without refueling stations.[6] In 2005, there were only seventy hydrogen refueling stations worldwide. California is again the national trendsetter with its hydrogen highway initiative program, which has the goal of establishing 200 hydrogen-refueling stations. California is not alone. Canada is planning to have a hydrogen highway between Vancouver and Whistler in British Columbia in time for the 2010 Olympics, when fuel cell automobiles will shuttle people between events. Another hydrogen highway is being considered between Windsor, Ontario, and Montreal, Quebec. Manitoba, the center of North American bus manufacturing, is intent on becoming the leader in hydrogen-powered buses and in creating a hydrogen fuel infrastructure that would reach down into the northern plains states of the United States.

California and Canada are attacking the chicken-and-egg situation by ensuring an adequate number of refueling stations for hydrogen-fueled vehicles. These vehicles will have to be owned by those whose driving patterns are more or less confined to a 150-mile-wide strip on either side of the highway (assuming a 300-mile range between refuelings). Once a stretch of road in California or Canada can serve hydrogen-fueled vehicles, then the population of hydrogen-fueled vehicles can expand to communities along the hydrogen highway. As the population of hydrogen-fueled vehicles grows in areas adjacent to the hydrogen highways, more refueling stations can be added, increasing the area that can serve hydrogen-fueled vehicles, allowing for another step up in the population of hydrogen-fueled vehicles, and encouraging the opening of more hydrogen refueling stations. Once started, this process feeds on itself and could mark the start of the era of hydrogen-fueled vehicles and the end of the era of gasoline-fueled vehicles.

If this process sounds vaguely familiar, it is. The first automobiles did not have gas stations for refueling. Gasoline (naphtha and other light-end products) was purchased in tins. The first gas stations were in city centers where the first automobiles were sold. Building gas stations and roads around city centers expanded the market for automobiles, which, in turn, expanded the market for gas stations and for roads. This process continued until the nation, and eventually much of the world, became blanketed with automobiles and gasoline stations and paved with roads.

While the probability of a massive switch to hydrogen-fueled vehicles may seem remote at this time, there are serious bets being made that hydrogen will eventually come out a winner. Right now there is a problem with cost. The hydrogen fuel cell, which once was a hundred times more costly than a comparable internal combustion engine, is now ten times more costly. A forty-passenger fuel cell-powered bus costs between $1 and $2 million, about ten times more than a conventional diesel-powered bus. Huge developmental efforts are still necessary to improve manufacturing processes and the expected life and reliability of fuel cells, particularly those exposed to low winter temperatures. Some believe that an entirely new fuel cell technology will have to be created for another tenfold reduction in costs. On the other hand, the automobile companies are convinced that the necessary cost reduction and performance enhancements can be achieved with present-day technology.

On top of the large capital investment is the cost of fuel. Hydrogen was about five times more expensive in energy equivalence than gasoline in 2004. Thanks to the escalation of oil prices, it was about three times more expensive in 2005. If another oil crisis occurs and gasoline prices approach $10 per gallon, hydrogen may prove cheaper. The odds of hydrogen becoming the fuel of the future depend largely on what happens in the oil patch.

Tomorrow's World of Hydrogen

While hydrogen will most likely come from stripping hydrogen from hydrocarbons, black hydrogen will eventually have to give way to green hydrogen to reduce carbon dioxide emissions. The problems in transporting hydrogen over long distances, associated with its low density and propensity to leak through fittings and gaskets, can be overcome by generating electricity and producing hydrogen locally in small plants. Each community, or group of communities, would have its neighborhood nuclear- or IGCC-power plant, augmented as much as possible by sustainable energy sources. A distributed generation system would supply electricity locally and excess electricity would be dedicated to generating hydrogen, which would be sold at nearby motor vehicle refueling stations. This would be a return to Thomas Edison's original idea for neighborhood electricity-generating plants. Electricity generators would operate close to full capacity in order to sell electricity at a low base rate for consumption by individuals, businesses, and industry or for electrolysis of hydrogen. This would negate the need to invest in generators that only operate part of the time in response to transient changes in electricity demand and also the need for marginal electricity rates. This concept, if carried to its logical conclusion, would make Westinghouse's idea of large centralized power stations with long-distance transmission lines obsolete.

The next step toward the hydrogen economy would be for every building and home to be fitted with a solar array and/or a wind turbine, an electrolysis unit, a hydrogen storage medium, and a fuel cell module. The solar array or wind turbine would provide electricity to the building. Excess electricity would be fed to the electrolysis unit to produce hydrogen, which would be stored in a tank or storage medium. Hydrogen would be fed to the fuel cell module to generate electricity when the sun was not shining and/or the wind was not blowing. Waste heat, in the form of hot water generated when hydrogen is converted to electricity, would be recycled for personal use, running appliances, and space heating.

The ultimate dream of the hydrogen aficionados is to increase the amount of electricity generated by solar energy, which would depend on vastly improving the efficacy of converting sunlight to electricity. If solar power could be significantly stepped up, then the electricity output might be great enough to produce enough hydrogen for a fuel cell-powered motor vehicle. Motor vehicles for personal use operate only about 5 percent of the time, which means that their fuel cells are idle 95 percent of the time. This is not an efficient use of any capital investment. Once parked at its destination, a hydrogen-fueled automobile would be plugged into the electricity grid to generate electricity as long as there is enough hydrogen left for the return trip. If sufficient numbers of these mobile generators are available and if a sufficient volume of hydrogen can be generated from the sun and wind, this could conceivably eliminate most of the neighborhood nuclear and IGCC plants.

Parenthetically, one might wonder why it is necessary for hydrogen-fueled automobiles to become electricity generators if solar arrays are capable of generating enough electricity to produce that much hydrogen fuel. Why not just have the solar arrays generate all the needed electricity and limit fuel cells to powering automobiles? Fuel cells would still be needed in buildings to transform hydrogen made by day to electricity needed at night. Fuel cells in automobiles could perform this function, but this would preclude driving at night.

Anyway, utilities would still be needed for power generation to cover shortfalls and serve as backup, but their primary purpose would be providing the physical connections for millions of automobiles, homes, and buildings to plug into the electricity grid where every home, building, and automobile is both an electricity-generating utility and a consumer. If the aggregate output of sustainable power sources were large enough, there would be no need for backup generators. Once this occurs, then the distributive utilities could become as obsolete as their centralized kin for electricity generation. However, they would remain as "virtual" utilities, overseeing the buying and selling of electricity among millions of users and generators and providing the technical means to dispatch and control millions of microgenerators. The transformation of an electricity grid with a few large generators into an interactive electricity network of millions of microgenerators would require advanced computer technologies, millions of sensors, and sophisticated software to allow electricity to flow exactly where and when it is needed at the lowest rate in a world where everyone is connected to everyone else. Cooperatives could be set up for buying and selling electricity among their members, possibly even taking over the utility's role of providing an interactive electricity network modeled after the worldwide communications Web. If the cooperatives also took over the responsibility of servicing and installing electrical wires and cables and connections, then the entire concept of a utility becomes obsolete.

If all this sounds esoteric, it is. It may hold true a century or two from now, but not from today's perspective. Significant technological advances have yet to be made to bring about the hydrogen economy. Other than those who are alarmed over its cost, few are against the concept. Hydrogen is virtually pollution-free with an unlimited supply. But for hydrogen to become a major energy source in the coming decades, sustainable sources of energy (hydro, wind, solar, geothermal) will not be enough. Nuclear power and coal-burning IGCC plants with sequestered carbon dioxide emissions will have to play a major role in generating the requisite amounts of electricity. For those who say that this is unacceptable, what, then, is the alternative?

Climate

First and foremost there is no such thing as a "normal" climate.[7] Climate runs in warm and cold cycles and within these cycles weather patterns change significantly. Ancient Carthage in North Africa was sustained by agricultural activities on land that is now desert. Cave paintings in the Sahara portray vibrant animal life on grasslands. Satellite photographs, which employ a technology that can see through sand, reveal a world of dried riverbeds and streams. Thus, climatic change does not have to take millions of years. Dramatic changes can occur in a few thousand years; for example, 5,000 years ago most of the greenery in North Africa gave way to an inhospitable desert. It only required a few centuries for agricultural activities around ancient Carthage to disappear. Climatic change can also occur overnight: Woolly mammoths were instantly frozen during the last ice age, and they can be found today "fit for eating," with their flesh and stomach contents intact.

With modern weather record keeping barely two centuries old, a method had to be developed to track the history of climate. The first such method, counting tree rings, was devised in the early part of the twentieth century. The width of annual tree rings is a record of the weather. Wide tree rings mark years of favorable growing conditions with plentiful rainfall, whereas narrow tree rings mark years of unfavorable growing conditions such as drought and extremes in temperatures. The first trees analyzed were ponderosa pine and giant sequoia in northern California, where overlapping sequences of cores taken from living and dead trees provided a history of climate going back 3,000 years. From this record, along with carbon-14 dating of wood in the Anasazi cliff dwellings in the U.S. Southwest, it was shown that this 500-year-old advanced American Indian civilization collapsed in the 1200s as a consequence of a twenty-six-year drought.

The next method for analyzing the history of climate examined lake and ocean sediment. Several techniques had to be devised to make this time capsule of climatic change readable. One was the discovery that certain types of plankton thrive in warm waters, others in cold, and the ratio of their calcium carbonate skeletons is a good indicator of water temperatures. Another was improving the technique to bring up deep cores from lake bottoms and ocean floors with minimum distortion to the sedimentary layers. Areas of ocean floors had to be found that were least disturbed by burrowing worms and other marine life whose activity blurs the distinction between layers. The discovery of radioactive carbon dating in 1947 was followed by the discovery of the ratio of two oxygen isotopes sensitive to changes in temperature. These two measuring sticks made it possible to obtain a record of major climatic changes in terms of ocean temperatures and the waxing and waning of ice sheets going back many thousands of years.

Just as tree rings showed that a severe twenty-six-year drought brought an end to the Anasazi civilization, the analysis of sediments from the bottom of a Yucatan saline lake showed that three periods of extreme drought within a 150-year dry spell brought an end to the 3,000-year-old Mayan civilization in the ninth century. The Maya, capable of devising a highly sophisticated calendar to keep track of time and of building massive temple complexes, had an estimated population of 15 million in the centuries before the dry spell. Droughts of unusual severity can occur even when climate is reasonably stable, for example, the 1930s Dust Bowl in the U.S. Southwest and the ongoing expansion of desert in sub-Saharan Africa.

Advances in analyzing sediment layers in lakes and oceans to detect the history of climatic change were accompanied by advances in analyzing annual snowfalls that formed distinct layers in stationary ice packs in Greenland and Antarctica. The first 1,000-foot core of ice was removed from the Greenland ice pack in 1956, then cut into segments for transport to labs in Europe and the United States for analysis. Technical advances in drilling allowed cores to be withdrawn from deeper depths in a more pristine state, and in 1966 a 4,500-foot-long core that extended down to bedrock was extracted from the Greenland ice pack. In the 1980s another core, this time 6,600 feet long, was extracted, followed in the 1990s by a core over 9,800 feet in length, both down to bedrock. Sediment from lake bottoms and ocean floors and cores from Greenland have provided a record of climatic change for the past 100,000 years, and subsequent cores drilled in Antarctica have pushed back the record to about 400,000 years.

Before the extraction of sediment and ice cores, the theory of climatic change was based on the earth's elliptical orbit and its slight wobbling about its axis (precession), which induced periods of reduced solar radiation every 22,000 years in the northern hemisphere. It was thought that this would create a 22,000-year cycle of relatively short ice ages, interspersed by long periods of a stable and warm climate. This early theory on climate was in concert with a general belief that change in the natural world was gradual and resulted from existing forces operating uniformly over eons of time. This gave animals and plants ample opportunity to shift their habitat in response to the slow pace of climatic change.

The record of climate gleaned from the lake and ocean sediments and the Greenland and Antarctica ice packs dashed the belief in gradual and uniform change as well as the implied ability to forecast general climatic conditions within reasonable bounds. The record better supports chaotic and catastrophic change, leaving little time for animals and plants to adapt to shifting climatic conditions. The story locked in sediment and ice cores is that there is no such thing as a normal climate. The only predictable behavior regarding climate is change itself, but not its direction or magnitude. Significant shifts in the ratio of oxygen isotopes testify to large changes in average temperatures over relatively short periods of time, sometimes accompanied with heavy layers of volcanic ash. The analysis of gas entrapped in the ice core showed cyclical fluctuations in methane and carbon dioxide concentrations. With the exception of the past 10,000 years, variations in temperature were much more severe, transitions between cooling and warming trends were swift (about 1,000 years), and the earth was a decidedly much colder place to live. The warmest part of the temperature cycle would be similar to today's weather, but it did not last long before the world plunged into another frigid ice age. About 14,700 years ago, the earth warmed and the climate stabilized for about 2,000 years before there was a sudden reversion to a 1,000-year ice age. Then, for inexplicable reasons, the climate suddenly reversed direction and an era of unusual warmth and stability began that has lasted 10,000 years; a phenomenon not experienced during the previous 400,000 years.

While some might consider this evidence of the transition from Genesis 1:2 to 1:3, others feel that the cause of the sudden warming was civilization. According to this hypothesis, people have been affecting the weather for at least 8,000 years, since the advent of agriculture, much longer than the 200 years, since the advent of the industrial age, as normally thought. Analysis of ice cores shows that the concentration of methane in the atmosphere rose and fell over the past 250,000 years fairly closely following the 22,000-year cycles in solar radiation. During this cycle, solar radiation in the northern hemisphere varied between 440–520 watts per square meter with methane, a greenhouse gas, tagging along and varying between 400–700 parts per billion as measured from Vostok (Antarctica) ice cores. Methane follows the 22,000-year solar radiation cycle because warm spells encourage plant life. When dead plants decay in anaerobic (without oxygen) water, copious releases of methane (swamp gas) add to natural gas seeps from underground coal, oil, and natural gas fields, increasing its concentration in the atmosphere.

This trend lasted until 5,000 years ago, when solar radiation fell as part of its normal 22,000-year cycle and methane, rather than falling to an expected 450 parts per billion, rose to nearly 700 parts per billion. This unexpected rise in the methane concentration is hypothesized to be the result of agricultural practices, especially growing rice and breeding herds of domesticated animals. The anaerobic decay of rice stalks in flooded rice paddies and the digestive processes of grazing animals are both major contributors to atmospheric methane. Thus, the earth did not cool as expected from a fall in solar radiation because the rise in methane, a greenhouse gas twenty times more effective than carbon dioxide, inhibited the escape of infrared red radiation from the earth into space.

Carbon dioxide cycles are more complex than methane. For the last 400,000 years, carbon dioxide concentrations peaked at about 280–290 parts per million every 100,000 years, with a number of minor cycles contained within each major cycle. The discovery that carbon dioxide concentrations bottomed out at about 200 parts per million in the depths of an ice age was the first evidence linking carbon dioxide with climate. Carbon dioxide peaked around 10,500 years ago, at the end of the last ice age, and began its expected retreat as it had done in the past. About 8,000 years ago, however, the retreat became an advance. By the start of the industrial era, the concentration of carbon dioxide was back to 280–290 parts per million, the "normal" peak during the previous 400,000 years. This was an estimated 40 parts per million higher than one would predict on the basis of past patterns and could be attributable to our use of biomass fuel. Methane, about 250 parts per billion, and carbon dioxide, about 40 parts per million higher at the start of the Industrial Revolution, compensated for a falloff in solar radiation. Had it not been for these additional concentrations of greenhouse gases, the earth might have experienced another period of glaciation with the reduction of solar radiation. Thus, it is conceivable that humanity's activities have helped to stabilize climate for the better![8]

Climatic Changes During Recorded History

There have been significant changes to the climate within recorded history. The period 900–1300 CE was called the medieval warm period with a warming trend similar to what is happening now. Agricultural output soared, as did the human population by an estimated 40–60 million. Vineyards sprouted in England and English wine gave French wine a run for its money. Greenland was not misnamed as some have thought by a real estate charlatan trying to induce prospective settlers to buy frozen land. Its green coastline would have allowed Vikings to establish communities where, as archaeological evidence shows, supported themselves by farming the land for food and crops to feed grazing herds of livestock, augmented by fishing and trading.

The start of the Little Ice Age in 1300 CE, known as history's Big Chill, saw average temperatures fall between 4°F–7°F. While this might not sound like much of a change, it was sufficient to bring an end to the greenery in Greenland, along with the Viking settlements and grape vineyards in England. Agricultural output plunged, as did the population from starvation and malnutrition, which weakened resistance to disease. Between 1371 and 1791 there were 111 recorded famines in Europe, with one famine in Russia claiming a half million lives in 1601. Part of the blame for the Black Death, which wiped out about one-third of the population of Europe, was rats seeking warmth in human habitations to escape the cold. The coldest period during the Little Ice Age was between 1645 and 1715, a period of minimum sunspot activity and reduced solar radiation, which lowered the average temperature by another 3°F. Alpine glaciers began to advance rapidly, swallowing up farmland and villages, and the Thames River froze over, starting a tradition of ice festivals that lasted until the early nineteenth century. Further hardships were in store for the people from sparse harvests, which contributed to social unrest, perhaps even to the French Revolution.[9]

There are three explanations for the Little Ice Age. The first is that it resulted from a .5 percent reduction in solar radiation, whose cyclicality may also be affected by periodic changes in sunspot activity and magnetic field intensity. Variation in solar radiation, coupled with variations in the earth's orbit and precession about its axis, induced severe cold spells. The second explanation is the slowing of the Gulf Stream conveyor belt. Normally, the Gulf Stream is more saline than the water in the North Atlantic. After warming the northern European atmosphere, whose latitude is the same as Newfoundland, the saline, cooled, and heavy Gulf Stream waters sink to the bottom of the Atlantic and return to the Caribbean. The sinking of the Gulf Stream waters is thought to be the driving force behind the immense conveyor of warm Caribbean waters to the North Atlantic. Thus, the preceding medieval warm period may have caused polar ice to melt, releasing vast quantities of freshwater along with a greater outpouring of freshwater from Siberian rivers. The freshwater emptying into the Arctic Ocean would eventually flow into the North Atlantic, diluting the salinity of the Gulf Stream. Less saline waters decreased the density of the Gulf Stream, reducing its capacity to sink and power the conveyor belt. As the Gulf Stream weakened, the weather in northern Europe cooled.

The third explanation is that there were several major volcanic eruptions during the Little Ice Age; the worst being the 1815 eruption of Tambora in Indonesia, which had the force of 100 Mount St. Helens. Whereas Mount St. Helens blew off 1,300 feet of its top in 1980, 4,200 feet of Tambora's top was blasted thirty miles into the stratosphere, reducing its height from 13,500 feet to 9,300. The eruption left behind a five-mile-wide caldera three-quarters of a mile deep, the largest on Earth within recorded history. An estimated 120,000 people were killed either by instant carbonizing in the 2,000°F pyroclastic flows rushing 100 miles per hour down the volcano's slopes or, more slowly, by starvation. This was just the start of the death count as Tambora, as with other large volcanic eruptions, affected global climate. The enormous plume of 200 million tons of sulfur contained in 100 cubic kilometers of ash blanketed the earth, which prevented solar radiation from penetrating the earth's atmosphere. Sulfur combined with oxygen to form sulfur oxides and then with water to form an aerosol of sulfuric acid that covered the upper surfaces of clouds, which reflected incoming sunlight. With ash shading the earth and an aerosol mist reflecting sunlight, the earth cooled by an average of 2°F. However, in the U.S. Northeast, Canada, and northern Europe, the cooling was about 10°F, making 1816 the year without a summer. It snowed in North America and northern Europe was drenched in cold rain, which prevented crops from maturing. The eventual death toll from starvation and disease brought on by malnutrition, mostly concentrated in Europe, is thought to be several times that of the death toll in Indonesia.

Volcanoes are "natural coal burners," releasing huge volumes of ash, sulfur, sulfur oxides, aerosols, and carbon dioxide that can affect global climate. On average, volcanoes emit about half of civilization's contribution of sulfur oxides and carbon dioxide annually, but individual eruptions can emit much more. The eruption of Mount Pinatubo in the Philippines in 1991 released more particulate pollution into the atmosphere in a few weeks than civilization had released since the start of the Industrial Revolution, and brought about a couple of years of cooler weather.

In 1850 the Little Ice Age abruptly ended, inaugurating a general warming trend still in progress. There are four explanations for this change: no major volcanic eruptions of the order of Tambora (notwithstanding the 1883 Krakatoa eruption); an increase in solar radiation, perhaps influenced by increased sunspot and magnetic field activity; the restoration of the Gulf Stream conveyor belt; and further addition of carbon dioxide to the atmosphere by human activities. Each of these explanations has been advanced with no consensus as to which or what combination was responsible for a warmer climate.

Where Is Climate Headed?

The evidence that the warming trend continues is ample: Glaciers are continuing their retreat, although there are counterexamples of some glaciers growing, and the Arctic ice cap is thinning and covering less area. The thawing of the permafrost is marked by "drunken" trees bending in all directions in Alaska, Canada, and Siberia, roadbeds that resemble roller coasters, and structurally damaged buildings from shifting and sinking foundations.[10] Other evidence of climatic change are the record-breaking temperatures nearly every year for the last decade, including 2005, although December 2005 ranked among the ten coldest Decembers on record for the last 100 years in the United States, with Europe and Asia suffering from unusual cold spells and China experiencing its coldest winter in twenty years. Record-breaking wet and dry spells have accompanied the record-breaking temperatures, spawning floods in some areas and droughts and wildfires in others. The sea level rose during the twentieth century, partly from melting of the Greenland and Antarctica ice packs and partly from the expansion of warmer ocean waters. Melting of the Arctic polar cap is sometimes cited as a cause of rising ocean levels, but floating polar ice has already displaced seawater and its melting does not affect ocean levels. This can be seen by filling a glass with ice cubes and then with water to its brim, letting the ice cubes protrude above the top of the glass. A paper towel under the glass will remain dry as the ice melts, showing that the water level has not changed as the ice melted. Of course, this observation does not hold true for melting snow and ice on land whose waters flow into the oceans.

Unfortunately the reduction of the surface area covered by polar ice decreases the earth's reflectivity, or albedo. Sea ice reflects up to 80 percent of the sunlight that strikes it, but reflectivity is reduced when white ice gives way to dark water. Greater absorption of solar energy by open ocean water leads to higher temperatures and more ice melting, creating larger areas of reduced albedo, an example of a positive feedback system. In tropical ocean waters, higher average temperatures from a warming Earth spawn hurricanes and cyclones. The intense hurricanes that have struck the United States in the last few years are partly caused by record water surface temperatures in the Gulf of Mexico that can transform a low-level tropical storm entering the Gulf into a Category 5 hurricane in two days' time. However, global warming may not be the only cause of the increasing frequency of hurricanes because hurricane activity is itself cyclical over a period that spans several decades. Moreover, history records devastating hurricanes and typhoons and tornadoes in the past when average temperatures were lower. For instance, a hurricane in the Caribbean in 1780 sank 1,200 ships, drowning their crews; the Galveston hurricane in 1900 killed over 6,000 people; 1955 was an extremely active year for hurricanes and tornadoes, and the 1970 typhoon killed over one-quarter of a million people in Bangladesh.[11]

A 1973 study of the carbon dioxide concentration of the atmosphere measured atop Mauna Loa, Hawaii, showed that the carbon dioxide concentration had climbed steadily from 316 parts per million in 1959 to 330 parts per million in 1972. In 1980 an analysis of Antarctic ice cores established that the concentration of carbon dioxide was around 280–290 parts per million at the start of the Industrial Revolution, already at its cyclical peak for the previous 400,000 years. For the first time in 400,000 years, carbon dioxide did not decline from its peak, but continued to rise. It reached around 300 parts per million in 1900 and a century later 370 parts per million; by 2004 the concentration of carbon dioxide had reached 377 parts per million and is still climbing. Carbon dioxide concentration increased at a faster pace during the twentieth century than the nineteenth, correlating well with our greater consumption of fossil fuels.

Similarly, the rise of methane concentration is also accelerating. From 1000 to 1800 CE, methane was about 700 parts per billion. It rose 100 parts per billion during the nineteenth century to 800 parts per billion. In the twentieth century methane leaped by 900 parts per billion to 1,700 parts per billion. This rise is almost entirely attributable to human activities: increased rice growing and other agricultural activities, a far higher population of domesticated grazing animals, a higher methane-emitting termite population from deforestation, greater volumes of methane escaping from landfills, and from increased coal mining and oil and gas activities. And, of course, vast quantities of methane entered the atmosphere from venting of natural gas in oil fields in the U.S. Southwest before construction of the long-distance natural gas pipelines. More methane entered the atmosphere from venting natural gas in Africa and the Middle East before the 1973 oil crisis. Thereafter, natural gas was reinjected into oil fields when it was finally recognized

Table 10.1

Calculating Carbon Release

	World Consumption In Billion Toe[13]	Percent Fuel	Carbon Release Factor	Carbon Release In Billion Tons	Percent Carbon Emissions
Coal	2.78	31.0	1.14	3.2	38
Oil	3.77	42.0	0.89	3.3	41
Natural gas	2.42	27.0	0.73	1.8	21
Total	8.97			8.3	

that venting and flaring were equivalent to burning money.

Projections vary, but depending on what actions are taken to reduce carbon dioxide emissions and which computer model is selected, the most likely projected level of carbon dioxide by 2100 is 450–550 parts per million, with the outside possibility of 1,000 or more parts per million. Again, depending on the selected computer model, the projected rise in average temperatures is between 1°C–5°C, or between 2°C–9°F.[12] If the low-end estimates are correct, we will survive. If the high-end estimates are correct, we will be facing severe problems on all fronts. Shifting weather patterns may affect rainfall over the grain-growing regions, seriously affecting agricultural production. Sea levels may rise as much as ten feet (three meters), flooding coastal regions including highly populated areas (New York City, London, the Louisiana delta, and Florida) and large parts of nations (Bangladesh and the Netherlands). Increased storm activity and severity completes the climatic doomsday scenario.

Carbon Dioxide Conundrum

As previously noted, burning coal releases more carbon dioxide for a given quantity of energy than other fossil fuels. There is nothing that can be done about this, other than sequestering, as this is a chemical property of coal with its higher ratio of carbon to hydrogen atoms. Coal releases more carbon dioxide as a product of combustion on an energy-equivalent basis than oil and natural gas. Natural gas, with its relatively low ratio of carbon atoms to hydrogen atoms, is the cleanest-burning fossil fuel, releasing the least amount of carbon dioxide on an energy-equivalent basis plus water. According to the International Energy Agency, burning coal in terms of the energy release of 1 metric ton of oil equivalent (toe) produces 1.14 metric tons of carbon, burning 1 metric ton of oil releases 0.89 tons of carbon, and burning natural gas equivalent to 1 metric ton of oil releases 0.73 metric ton of carbon, all of which enters the atmosphere as carbon dioxide. With these figures as guidelines, Table 10.1 calculates the global release of carbon from the burning of fossil fuels for 2004.[13]

Table 10.1 excludes biomass. The commonly accepted figure for the annual carbon release into the atmosphere caused by human activities, including nonsustainable burning of biomass (deforestation), is something close to 10 billion tons. As noted in Table 10.1, coal adds a proportionally greater amount of carbon dioxide to the atmosphere than the other fossil fuels, representing 31 percent of fossil fuels consumed in terms of energy release and 38 percent of carbon dioxide emissions. Moreover, old coal-burning electricity-generating plants in some parts of the world, particularly India and China, are energy inefficient compared to those in developed nations. Replacing these coal-fired plants with new natural gas energy-efficient plants would reduce carbon dioxide emissions in two ways: Natural gas releases less carbon dioxide than coal for an equivalent output of energy, and less fuel would be burned for an equivalent output of electricity. Furthermore, natural gas does not emit sulfur, nitrogen oxides, or heavy metals, nor does it desecrate the landscape or affect water supplies. One can see why nations prefer to abandon coal in favor of natural gas. But the reality is that the supply of natural gas is not sufficient to replace coal, and its high price in 2004 and 2005 indicates that natural gas is itself in short supply.

Before we become too enthralled with our capacity to influence the weather, the 10 billion tons of carbon we add to the atmosphere per year is dwarfed by nature's recycling program. It is estimated that 2,000 billion tons of carbon dioxide are locked up in partially decayed plant matter in the soil, peat bogs, and permafrost. Living plant life absorbs 120 billion tons, about that released to the atmosphere by decaying plant life. As a point of comparison, agriculture and land use by humans releases and absorbs less than 2 billion tons, a rounding error in what nature recycles. The atmosphere is estimated to contain 730 billion tons of carbon in the form of carbon dioxide, of which about 90 billion tons are exchanged with the oceans, which contain 3,800 billion tons of carbon in the form of dissolved carbon dioxide. In comparison, the burning of fossil and biomass fuels adds about 10 billion tons. With a total of 200 billion tons of carbon being exchanged naturally between vegetation and the oceans with the atmosphere, scientists ponder why a paltry 5 percent addition to nature's recycling program is making such a big difference to the atmospheric concentration of carbon dioxide.[14]

A higher concentration of carbon dioxide, in itself, should spur plant growth, which would absorb greater quantities of carbon dioxide from the atmosphere. The 12,000 species of diatoms in oceans and lakes are another sink, or source of natural removal, of carbon dioxide. Diatoms are single-celled algae that convert carbon dioxide, water, and sunlight into food and release oxygen. Unlike other phytoplankton, diatoms also absorb carbon dioxide to create microscopic shells that sink to the bottom of the lake or ocean when they die. It is estimated that diatoms remove about half as much carbon dioxide as photosynthesis, a huge volume of carbon dioxide. One potential way of reducing carbon dioxide would be expanding the population of diatoms.[15]

If a higher concentration of carbon dioxide is causing global warming, which leads to greater evaporation, one would think that greater cloud cover, through reflectivity, would reduce the heating effect of sunlight on the earth's atmosphere. If this were true, then greater cloud cover would be a negative, or self-correcting, feedback system. Moreover, increased precipitation can also be a self-correcting feedback system. While the ice pack in West Antarctica is melting along its coastal regions, about 75 percent of Antarctica's total land mass is experiencing increased precipitation that is thickening the interior ice pack, keeping sea levels from rising at a faster pace! Another puzzle for one to ponder is that, with a carbon dioxide concentration far in excess of that of the medieval warm period, temperatures are still less than what they were when greenery flourished along Greenland's shores.

Projecting the Future

There are three projections for future climate: It will stay about the same; it will get a lot warmer;

or it will get a lot colder. The advocates of a much warmer Earth are worried about global warming reaching the point where the permafrost in Alaska, Canada, and Siberia will thaw, which could release enormous volumes of methane currently entrapped within an ice lattice (methane hydrates). Being twenty times more effective as a greenhouse gas than carbon dioxide, massive releases of methane could conceivably set in motion runaway global warming. Rising temperatures would release even greater volumes of methane, along with the carbon dioxide locked in the oceans and in the partially decayed plant matter in the permafrost.

Those who predict that the earth will get a lot colder look to the introduction of vast quantities of freshwater into the North Atlantic from melting Arctic and Greenland ice caps, an increased flow of Siberian rivers, and increased freshwater runoff from the thawing permafrost. This could dilute the salinity of the Gulf Stream and slow down the conveyor belt sufficiently to cause another ice age to set in. The reflectivity of the earth, its albedo, increases once ice begins to cover wider areas of the earth, further reducing temperatures, causing more snow to fall, enlarging areas covered by snow and ice, and increasing reflectivity, thereby inducing even lower temperatures. If the ice covers a sufficiently large area of the earth, then an irreversible trend may set in where the earth becomes a snowball with glaciers and frozen oceans covering the entire planet. There is evidence that the earth has been a global snowball twice, once about 2.2 billion years ago and again about 0.5 billion years ago. It is thought that the effects of extreme volcanism ended the frozen state by introducing vast quantities of carbon dioxide to the atmosphere, which initiated global warming, and by leaving behind ash deposits that decreased the earth's albedo.

Both runaway global warming and cooling are examples of positive feedback systems; once started, nothing can stop the trend from worsening other than some massive external intrusion such as extreme volcanism. For those who are wondering whether we are in danger of entering an era of runaway warming and cooling, in 2005, the National Oceanography Centre in Southampton, UK, announced that the measured quantity of warm water flowing northward in the Gulf Stream had slowed by about 30 percent, with much of the slowdown occurring between 1992 and 1998.[16] A slowdown of this magnitude would portend cooler weather in Europe, but pundits observed that Europe experienced a warming trend throughout this period with the exception of the severe winter in 2005. So who knows?

We believe in linear systems: Tugging a rope attached to a wagon causes the wagon to move, whose speed can be estimated by taking its resistance to movement into account. But climate is a nonlinear system, which is akin to tugging on a sleeping dragon's tail. Nothing happens tug after tug. Then that one tug, which may be weaker than the others, suddenly awakens the dragon. Now something happens. While nonlinear systems may lend themselves to statistical analysis, probabilities become quite meaningless if a set of circumstances (variations in the sun's radiation, variations in the earth's orbit and precession, variations in greenhouse gas concentrations) all line up to induce climate to move from one point of equilibrium to another. The shift between equilibrium points can be relatively swift, with dire repercussions for life on earth, as the frozen woolly mammoths in Siberia plainly testify.

Our climate may be warming from a naturally occurring cyclicity in solar radiation, variations in the earth's orbit and precession, Chandler's wobble (variations in atmospheric and oceanic density and snow accumulation affecting the orientation of the earth's poles), increasing greenhouse gas concentration, volcanic activity, or some combination thereof. Having a carbon dioxide concentration at a level never experienced in 400,000 years, which is continuing to rise, is a cause for worry. It is not absolutely clear how much of this increase is caused by burning fossil fuels or by something else that is affecting nature's capacity to recycle carbon dioxide. Others believe that global warming, at least in part, is caused by greater solar radiation, and point to the melting ice caps on Mars as possible confirmation that the sun may be emitting more energy. However, there may be other reasons why Mars's polar regions are retreating since weather patterns on Mars are even less understood than those on Earth. Regardless of the cause, measurements taken deep in the ocean and from space indicate that the earth is absorbing more energy than it is giving off, a surefire indication of global warming.

The risk we face is that, by adding carbon dioxide and methane to the atmosphere, in concert with other natural forces that are affecting the global climate, may be the tug that awakens the sleeping dragon initiating a major shift in climate. The new equilibrium point may make life difficult for Earth's inhabitants. Such a shift may not be triggered by heightened carbon dioxide concentrations but, perhaps, by heightened gamma ray bursts from outer space, which also affect global weather patterns. We just do not know what the trigger point for climatic change is. What makes this such a challenge is that the trigger point may not be a single combination of conditions, but a number of different combinations in which heightened levels of greenhouse gases may play an undeterminable role in triggering radical climatic change.

The Environment

Speculation concerning climatic change during the next hundred years tends to mask what is happening to the environment now, best epitomized by the Asian brown cloud.[17] The Asian brown cloud encompasses thousands of square miles and can be seen from space. It originates primarily in China and India and results from burning coal and biomass without environmental safeguards as required in developed nations. In the developed world, electrostatic precipitators capture over 99 percent of the fly ash, the solid particles in combustion emissions, which is trucked to a disposal site or consumed in the production of cement. Desulfurization units (sulfur scrubbers) remove a large percentage of the sulfur. In India and China, much of the particulate residues and sulfur emissions from burning a generally low-quality coal enter the atmosphere as ash, soot, and sulfur dioxide. Burning coal is not entirely to blame; other major contributors are burning biomass and exhaust fumes from a rapidly growing population of motor vehicles throughout Asia. Some motor vehicles made in Asia (but not in Japan) have substandard environmental safeguards compared to those produced in the West, and cannot be exported because they would fail pollution emission standards. These vehicles contribute to airborne pollution more than Western-made motor vehicles. The two-cycle gasoline engines found on small motorcycles throughout Asia are horrific polluters. However, China is embarking on a program to enhance the emission standards of its domestically manufactured motor vehicles, not only to try to clean up its environment, but also to open up export markets in the West.

The Asian brown cloud that hovers over mainland Asia and the island nations of Southeast Asia affects the health of millions of people. Air in major metropolitan areas noticeably affects the eyes, nose, and throat. At times the overhanging haze is so thick that one can look directly at the diffused light of the sun. The two-mile thick Asian brown cloud reduces the amount of sunlight reaching the ground by 10–15 percent. Sunlight warming the lower atmosphere, or troposphere, rather than the earth's surface increases the frequency and strength of thermal inversions, which trap large amounts of pollution near the earth's surface. Those immersed in a cloud of pollution suffer from an increasing incidence of acute respiratory infection, pulmonary disease, lung cancer, tuberculosis, and asthma. Combined with outdoor pollution is indoor pollution from burning biomass fuels (wood and dung) and coal inside dwellings. Indoor pollution in India is felt to cause a half million premature deaths of children below the age of five annually. With less sunlight reaching the earth's surface, the Asian brown cloud reduces evaporation and affects the amount and pattern of precipitation, reducing photosynthesis and agricultural productivity. Rice production in India is estimated to be down by about 5 percent from air pollution.

Depending on the season, prevailing winds over India spread pollution to Nepal in the Himalayas or over much of the Indian Ocean. Seasonal prevailing winds over China spread pollution to otherwise isolated, idyllic, and pristine tourist Meccas in the Pacific Islands, where it is noticed, or at times as far away as Los Angeles, where it goes unnoticed. The Los Angeles basin frequently suffers from a temperature inversion that traps pollution near the surface, where it is unable to escape over the surrounding mountains. Early explorers noted overhanging smoke from native Indian fires, but now it is caused by the hydrocarbon emissions of motor vehicles. As late as the 1930s, the probability of having a very clear summer afternoon, with a visual range in excess of thirty-five miles was 21 percent, but by the 1940s the increased population of automobiles had dropped the probability by a factor of 100 to 0.2 percent.

Various types of pollution seem to conflict with one another. Carbon dioxide and methane hinder the escape of infrared red radiation from the earth to space, heating up the atmosphere. An aerosol mist of sulfuric acid collects on cloud tops, increasing their albedo, or reflectivity, reducing the amount of sunlight that enters the atmosphere. Acid rain inhibits wetland bacteria from producing methane, as swamp gas, from decaying plant matter, a major contributor to greenhouse gases. Soot from burning coal and biomass and diesel engine emissions collect on glaciers, reducing their albedo and accelerating their melting. This same soot is responsible for "global dimming," which is no longer limited to Asia. Records over the past thirty-five years of a declining rate of water evaporation on the surface of the earth, despite higher average air temperatures, confirm that less sunlight is reaching the surface. Solar radiation reaching the surface of the earth, measured in watts per square meter, has declined on average from 191 in 1958 to 190 in 1965 to 182 in 1975 to 176 in 1985 to 171 in 1992, an accelerating trend.[18] Reduced solar radiation at the earth's surface would mean less evaporation, and perhaps, less cloud cover, which if true, means more incoming solar radiation that would increase air temperatures. Some combination of these conflicting environmental factors may be behind the heightened activity and severity of storms experienced in recent years.

Some climatologists think the connection between rising levels of carbon dioxide in the atmosphere and global warming is tangential, at best, and should be regarded with caution. Spending huge sums of money to reduce carbon dioxide emissions takes away from programs to relieve poverty and, in that sense, the fight against carbon dioxide may be deleterious, not beneficial to society. Russian climatologists nearly unanimously deny any anthropogenic linkage between human activities and global warming. They believe that variation in weather patterns reflects the normal cyclicality of climate. Western climatologists are nearly unanimously opposed to this opinion and believe that there is a direct anthropogenic link between global warming and our consumption of fossil fuels. The Russian climatologists look at drunken forests, roads that resemble roller coasters, and collapsing buildings in Siberia and blame only nature. Western climatologists look at drunken forests, roller coaster roads, and collapsing buildings in Alaska and blame only ourselves. Some cynics suggest that some Western climatologists may be guilty of purposely arousing public fear over global warming in order to win more grants to conduct research. The uneasy feeling that we may be taking the wrong path has made its way into fiction, which, on occasion, has turned out to be fact.[19]

The U.S. Clean Air Acts

The United States has been criticized for not being environmentally responsible for refusing to sign the Kyoto Protocol. These critics conveniently forget that the United States was the first nation to pass legislation specifically aimed at improving the environment. Before a series of Clean Air and Clean Water Acts beginning in the 1950s, industry was free to pollute, degrading the environment for all, with no cost to itself. This represents a market failure. Markets supposedly establish a clearing price that matches supply with demand, taking into consideration all the factors that affect costs. The market mechanism clearly failed when no cost was ascribed to the damage from environmental pollution. The Clean Air Acts, in concert with the Clean Water Act, established regulatory regimes to reduce air and water pollution. The Acts set forth standards that industry was expected to comply with or face punitive action, either in monetary terms or cessation of operation.

Unfortunately, one consequence of such actions to improve the environment in the United States has been to export pollution to areas where environmental restrictions are largely nonexistent. For example, to reduce sulfur emissions, high sulfur coke from U.S. oil refineries is prohibited for use as a boiler fuel in the United States. As a consequence, the U.S. price for high sulfur coke fell substantially in relationship to the world price, making it attractive for Far Eastern cement manufacturers to buy the coke. The consequence of the Clean Air Act, which was intended to reduce sulfur emissions in the United States, has been to increase sulfur emissions in Asia. Pollution simply changed location and, once in the atmosphere, wind currents assured its global distribution.

Some industrial enterprises, such as electricity-generating plants, cannot move, and must comply with air pollution restrictions. Even for companies that could move, but stayed and complied with the regulations, the cost of compliance could make their product prices uncompetitive in a world marketplace where competitors, mainly in developing nations, operate with virtually no restrictions on pollution, and, as well, on wages and working conditions. Pollution abatement, for all its merits, has contributed to the deindustrialization of the United States and Europe.

The 1990 Clean Air Act differed from its predecessors by internalizing an externality. The externality was environmental degradation, which affects everyone. Internalizing makes environmental degradation a cost directed at those responsible for pollution emissions. Once a cost is placed on pollution, whether in the form of acquiring allowances that permit pollution emissions or as a direct tax, industry then has an economic incentive to do something about reducing pollution.

The process for internalizing an externality—the reduction of sulfur dioxide emissions—set up by the 1990 Clean Air Act has been adopted by the Kyoto Protocol as a primary means of reducing greenhouse gas emissions.

The first legislative act was the Air Pollution Control Act of 1955, "an Act to provide research and technical assistance relating to air pollution control." While the Act recognized air pollution as a national problem, its scope of action was limited primarily to providing research grants. The Clean Air Act of 1963, "an Act to improve, strengthen, and accelerate programs for the prevention and abatement of air pollution," called for a more specific recognition of problems associated with burning high sulfur coal and oil and motor vehicle emissions. The Act recognized two general categories of pollution: stationary sources (utility and industrial plants) and mobile sources (motor vehicles, trains, and aircraft). Mobile sources are more difficult to control than stationary sources because their movement affects pollution over a wide area.

The first tailpipe emission standards for motor vehicles were established in California in 1959 to take effect for the 1966 model year. Realizing that state regulation would result in automobile manufacturers having to comply with as many as fifty different sets of pollution emission standards, the 1963 Act established the principle of a national standard for automobile emissions. One standard of pollution emissions would apply to all motor vehicle manufacturers (domestic and foreign), regardless of where a vehicle was sold in the United States. The first federal emission standards adopted in 1965, as amendments to the 1963 Act, applied to the 1968 model year. These standards were virtually identical to those adopted by California for 1966 models. The 1990 Clean Air Act, a counterexample to establishing uniform regulations, established gasoline standards to deal with automobile exhaust emissions that vary both in time and place. This has complicated the refining and distribution of gasoline in the United States, a situation made even more complex by the right of states to impose standards on gasoline sold within their jurisdiction.

California has always been a frontrunner in state-inspired initiatives to cut pollution and a model for other states and for the federal government. California, along with New Jersey and Illinois, passed state laws to force the cleanup of aircraft engine emissions over their airspace, which were subsequently adopted by the Clean Air Act.[20] The state has been particularly vigorous in setting tough standards on gasoline sold within its jurisdiction. In 2005, California again seized the initiative by setting standards for automobile exhaust greenhouse gas emissions for future model years. This legislation does not violate the Clean Air Act's uniform set of standards for automobile emissions because these standards covered specific pollutants, not greenhouse gases. If other states follow California's example, then the federal government may be forced once again to take action to set a uniform standard for automobile greenhouse gas emissions throughout the nation.

The six recognized forms of pollution in the Clean Air Act are sulfur and nitrous oxides, carbon monoxide, volatile organic compounds, particulate matter, and lead. Sulfur oxides come from burning high sulfur coal in electricity-generating plants plus emissions from papermaking and metal processing, and burning motor vehicle fuels with a high sulfur content. Nitrous oxides come from both mobile (gasoline and diesel fuel engine exhaust) and stationary sources (smokestack emissions from industrial boilers and utility plants). Sulfur and nitrous oxides damage the lungs and, when combined with water in the atmosphere, form acid rain, acid snow, fog, mist, and dust. Acid rain can have a devastating impact on marine life in lakes and streams, depending on the type of bottom. Lake bottoms of sedimentary rocks like limestone neutralize acid rain, whereas granite and similar rocks do not. Acid rain has also caused extensive damage to forests in the northeastern United States, Canada, and Europe (the Black Forest in Germany), and has eaten away limestone and marble in outdoor statues, frescoes, and facades of buildings.

To avoid local pollution, Midwest coal-burning utility plants built high smokestacks to let prevailing winds carry sulfur and nitrous oxide pollution aloft. The prevailing winds not only carried pollution away from the Midwest, but also transformed it to acid rain that eventually fell on the U.S. Northeast and Canada. This is an example of market failure because the harm to marine life and forests and damage to stone facades far from the high sulfur coal-burning utilities was in no way reflected in the emitters' costs, other than what they spent to build higher stacks.

Carbon monoxide results from the incomplete combustion of fuels. It reduces the ability of blood to deliver oxygen to body cells and can result in death in a confined space. Volatile organic compounds (VOCs) are released from burning motor vehicle fuels in the form of benzene, toluene, and other hydrocarbons and also from solvents, paints, and glues. Ground-level ozone, the principal component of smog, results from "cooking" VOCs and nitrous oxides in the presence of sunlight. Ozone reduces visibility, damages plant life, irritates lungs and eyes, and lowers resistance to colds and other infectious diseases. Particulate matter (dust, smoke, and soot) comes from burning fuels and agricultural activities. The finer the particulates lodged in the lungs, the more hazardous they are to health, aggravating or causing bronchitis, asthma, and other respiratory diseases. Lead found in leaded gasoline, which was subsequently phased out by 1985, is also emitted by lead smelters and processing plants. Lead harms the brain and nervous systems, particularly in children. Because the public was largely unaware of global warming when the Clean Air Acts were originally legislated, they were concerned with reducing air pollution, not global warming, and did not include carbon dioxide and other emissions that, as pollutants, cause global warming.

Amendments to the Clean Air Act in 1967 divided the nation into Air Quality Control Regions (AQCRs) to monitor air quality. AQCRs that meet or exceed Clean Air Act standards are designated attainment areas, whereas those that do not are designated nonattainment areas requiring special attention. Enforcement of pollution standards is primarily carried out by the various states' environmental agencies. Each state develops a state implementation plan (SIP) that outlines how it will enforce compliance with the pollution standards set forth in the Clean Air Act. While states have the right to set tougher standards than those imposed by the Act, they are obligated to meet its minimum standards. The Environmental Protection Agency (EPA) approves each SIP and has the right to take over enforcement if the SIP is unacceptable. The issuing of environmental permits under the Clean Air Act includes information on the type and quantity of pollutants being released, how they are being monitored, and what steps are being taken to reduce pollution. The EPA does not specify how to reduce pollution, but requires that the Maximum Available Control Technology (MACT), changed to the Best Available Control Technology (BACT) by the 1990 Act, be used to ensure an effective means of pollution abatement.

The Clean Air Act of 1970, "an Act to amend the Clean Air Act to provide for a more effective program to improve the quality of the Nation's air," was an essential rewrite of the previous Act. The Act established National Ambient Air Quality Standards (NAAQS) for the cited pollutants and also included New Source Performance Standards (NSPS) that strictly regulated emissions from new factories, including electricity-generating and other industrial plants. The Act set standards for hazardous emissions and gave individuals the right to take legal action against any organization, including the government, for violating emissions standards.

Rather than set guidelines that were normally not complied with, as in the previous Clean Air Acts, the 1970 Act required the EPA to perform compliance tests, enforce performance warranties from manufacturers, and impose a per vehicle fine for those that did not meet Clean Air Act standards. The 1977 amendments to the Act established a policy of Prevention of Significant Deterioration (PSD), which defined areas such as national parks where there was a general prohibition from doing anything that would result in significant deterioration of the environment. A long string of amendments, or EPA granted extensions, was necessary to give motor vehicle manufacturers time to comply with emission standards. In the meantime, lead was removed from gasoline. Oil refiners had to invest in refinery improvements to produce a lead-free gasoline whose performance standards were similar to leaded gasoline and to meet lower sulfur specifications imposed on gasoline and diesel fuel. Diesel engine manufacturers had to build engines that cut particulate emissions. These additional costs borne by industry were eventually passed on to consumers as higher prices.

Congress once again drastically amended the Clean Air Act in 1990, "an Act to amend the Clean Air Act to provide for attainment and maintenance of health protective national ambient air quality standards, and for other purposes." The Act recognized that pollution in many metropolitan areas could not be restricted to a single state. People live in one state and work in another. Trucks frequently pass over state borders. Thus, air pollution from motor vehicles cannot be handled under a single state jurisdiction when cars and trucks spread pollution throughout an entire metropolitan region. The 1990 Act set up interstate commissions responsible for developing regional strategies to clean up the air. Pollution crossing the national borders in either direction between the United States, Mexico, and Canada was addressed in a special annex to the North America Free Trade Agreement (NAFTA). The Act mandated that all power plants had to install continuous emission monitoring systems (CEMS) to keep track of sulfur and nitrous oxide (SOx and NOx) emissions. Pollution permits required every power and industrial plant to specify a schedule for meeting emission standards.

The 1990 Clean Air Act introduced reformulated and oxygenated gasoline. Reformulated gasoline combats smog by emitting a lower level of VOCs that mix with nitrous oxides to form ozone, the primary ingredient in smog. Oxygenated gasoline burns more completely, particularly during engine startup in cold weather, reducing carbon monoxide emissions. Reformulated and/or oxygenated gasolines were required for certain nonattainment areas of the nation that suffer from high levels of ozone and/or carbon monoxide pollution for particular periods of the year. As described in Chapter 3, the Energy Policy Act of 2005 substituted a minimum quantity of ethanol and other biofuels in the gasoline stream for the oxygenate requirement contained in the 1990 Clean Air Act.

The 1990 Act contained mandatory reductions in sulfur emissions on a timetable that permitted industry to adapt to the new standards in an orderly fashion. A cap was established on aggregate sulfur emissions that stepped down in time. This cap became a maximum allowance or limit for the principal sulfur dioxide emitters. Reducing sulfur emissions on an individual company basis can be done in a number of ways. One way was for a sulfur emitter to buy a particularly rundown, outmoded, inefficient, obsolete, cheap sulfur-emitting smoker in order to establish its baseline for sulfur emissions, then shut the plant down, thereby fulfilling its sulfur-reduction obligation. The more prevalent way was the method used by utilities in the Midwest and Northeast, which shifted from high-sulfur Appalachian coal to low-sulfur western coal. This caused low-sulfur coal prices to rise in relation to high-sulfur coal prices, establishing a price differential, which, if wide enough, provided an economic justification to install scrubbers. Scrubbers remove sulfur in exhaust fumes and enable the plant to burn cheap high-sulfur coal while keeping within mandated sulfur emission allowances. Retiring old coal-fired plants and replacing them with clean-burning natural gas plants, or, better yet, with sustainable solar and wind power plants, are ways for an industrial concern or utility to reduce sulfur emissions.

These actions are pedestrian compared to the innovative part of the 1990 Clean Air Act, which internalized an externality; instituting an economic benefit to encourage pollution abatement through emissions trading. (Installing scrubbers based on a coal price differential could be considered an economic incentive, but emissions trading takes a more direct approach.) Although the beginning of emissions trading can be traced back to the 1977 Act, the 1990 Act brought it to the forefront. The Act introduced a market-based system for the buying and selling of rights or allowances to release pollution emissions. A pollution-emitting plant could invest in environmental equipment to decrease its emissions. If the equipment installation lowered sulfur emissions below the cited allowance or authority of a company to emit a pollutant, the company had the right to sell emission allowances (or rights to pollute), up to its maximum allowance, to companies that exceeded their mandatory emission allowances. The value of these emission allowances was to be determined by the market forces of supply and demand just like the value of corporate shares and commodities.

Buying and selling pollution emission allowances does not in any way abrogate the reduction in total emissions required by the Act. It simply redistributes the amount of allowed pollution among emitting plants. This gives companies flexibility to reduce emissions either by taking direct action to do so by investing in pollution-reduction equipment or by buying excess allowances from other emitters. At the end of a compliance period, every emitter must have EPA issued and purchased allowances equal to its pollution emissions or face stiff fines. This ability to buy allowances is called allowance trading, or cap and trade. Cap and trade means that there is an aggregate emissions cap, or a maximum limit, on total allowable emissions. The overall cap on emissions is initially set lower than the historical level of emissions, which gives value to allowances. As the overall cap is progressively lowered with time, so is each individual pollution emitter's allowance limit. If a company's emissions are less than its allowance limit as determined by the EPA, then the company has the right to sell the difference. If a company's emissions are above its allowance limit, then it must buy the difference or reduce operations to cut pollution or pay stiff fines. Cap and trade allows companies to act in the most economical way to reduce pollution by investing in pollution-abatement equipment, by buying allowances from a company that has, by reducing operating levels, or any combination thereof that best serves their financial interests.

Rights to pollute have a positive impact in that plants that go the full mile to reduce pollution are rewarded for beating their individual allowance limit by making money selling their excess emission allowances. Plants that have not fully complied with their individual allowance limits now have a financial incentive to do so because of the cost of buying emission allowances above their allotted allowances. A plant can use the revenue from selling excess emission allowances to justify its investment in reducing pollution emissions below the legal requirement. A plant buying sulfur-emission allowances can use the cost to justify investing in emission-abatement equipment. As the aggregate cap on emissions shrinks with time, these allowances, or rights to pollute, are apt to gain value to the greater benefit of those who are below their allowance limit and to the detriment of those who are above. Bonus allowances are rewarded to power plants that install clean-coal technology to further reduce pollution emissions, use nonpolluting renewable energy sources such as wind and solar, or encourage energy conservation to reduce the amount of power that has to be generated. The proceeds from selling bonus allowances can justify taking more costly actions to reduce pollution. Allowances can also be stored for future use.

Since 1997 between 15–25 million tons of sulfur allowances have been traded annually at a generally increasing price of $100 per ton in 1997 to a little over $200 per ton in early 2004. The price escalated sharply in the second half of 2004, reaching $485 in the fall of 2004, $700 in early 2005, and broke through $1000 in October 2005. The 1990 Clean Air Act's establishment of a cap-and-trade program has succeeded in reducing sulfur emissions despite growing coal consumption at a lower cost than a command requirement. A command requirement dictates that a company must reduce sulfur emissions by a set amount regardless of cost. Cap and trade has the benefit of lowering the overall investment in pollution-control equipment by making it possible for companies to avoid high pollution abatement costs by buying excess allowances from companies with low pollution abatement costs. However, at $1,000 per ton, companies must be taking a second look at investing in pollution-control equipment rather than buying excess allowances. Nevertheless, the cap and trade emissions trading program is estimated to have reduced the overall cost of pollution abatement by about one-third over a cost-indifferent command requirement. This has resulted in smaller electricity rate hikes for consumers, who, ultimately, foot the bill for pollution abatement.

In 2005, the EPA issued the Clean Air Interstate Rule (CAIR) to reduce air pollution that moves across state boundaries. Under the ruling, CAIR caps sulfur dioxide and nitrous oxides in twenty-eight states and the District of Columbia, covering nearly every state east of the Mississippi. When fully implemented in 2015, CAIR will reduce sulfur dioxide emissions by 70 percent and nitrogen oxides by 60 percent, resulting in projected health benefits of the order of $100 billion per year. The reduction in health costs is estimated to be twenty-five times greater than the cost of instituting CAIR. The EPA will set up a state's emission budget for both pollutants letting the state decide what measures are to be taken to reduce pollution and whether to have power plants participate in an EPA-administered interstate cap-and-trade program. The reduction in sulfur dioxide emissions represents an acceleration of that contained in the Clean Air Act for the affected states.[21]

Reducing nitrous oxides is a greater technical challenge than removing sulfur oxides. Sulfur is removed as an impurity from gasoline and diesel fuel during the refining process and is barely present in motor vehicle fuels with low-sulfur specifications. Nitrogen is not an impurity that can be removed; it is part of the molecular structure of oil. Sulfur impurities are not removed from coal prior to combustion, but sulfur oxides are technologically easier to remove from smokestack fumes than nitrous oxides. The government requirement to remove nitrous oxides acts as an incentive for entrepreneurs and corporations to find a cost-effective way of doing that. The technology to remove noxious oxides does not have to be sold; the market has already been created by government mandate.

In a closely related ruling, the EPA also issued the Clean Air Mercury Rule (CAMR) to cut mercury emissions for the first time. Mercury contamination harms the central nervous and reproductive systems of adults. Pregnant mothers eating mercury-laden fish have a greater risk of bearing children with brain disorders and learning disabilities. The largest source (40 percent) of mercury is coal-burning power plants that emit about 48 tons of mercury each year. The CAMR calls for an ultimate reduction of nearly 70 percent from 1999 levels with an initial cap of 38 tons in 2010 and 15 tons in 2018. A cap-and-trade program will be set up for mercury emissions to aid utilities in attaining their allowance limits.

The Clean Air Acts have proven to be contentious, complicated, and difficult to administrate. The succession of acts and amendments to the acts manifest the challenges faced by the federal government in trying to improve the quality of the air in fifty states. While one can argue that the environment is far from pollution-free and we still have a way to go for a clean air environment, nevertheless, the Clean Air Acts have also been effective in reducing pollution. Since 1970, sulfur dioxide emissions are down by 27 percent, carbon monoxide by 31 percent, VOCs by 42 percent, particulate matter (soot, smoke) by 71 percent, and lead emissions by a whopping 98 percent, the greatest single achievement of the Clean Air Acts.[22]

The Montreal Protocol

Ozone is a molecule of three oxygen atoms, not the normal two, and is constantly being created and destroyed in the stratosphere, nine to thirty miles above the earth. In an unpolluted stratosphere, the natural cycle of production and decomposition is at the right equilibrium to maintain a protective layer of ozone to filter out harmful ultraviolet B-rays that can cause skin cancer. But manmade chemicals such as chlorofluorocarbons (CFCs) affect the ozone layer. CFCs are a refrigerant in refrigerators and air-conditioning units, a propellant in aerosol sprays, and are used in solvents and foam-blowing agents. CFCs escape into the atmosphere through leaks in refrigerators and air-conditioning units for buildings, rooms, and motor vehicles, and from the failure to remove CFCs from disposed units. As an aerosol propellant, CFCs obviously end up in the atmosphere. Once in the atmosphere, CFCs rise into the stratosphere where they decompose and release chlorine, which acts as a catalyst, speeding up the decomposition of ozone. This thins the ozone layer to the point of creating an ozone hole, first noticed over Antarctica, where it grew in size to include the southern tip of South America. Those caught in the ozone hole were exposed to ultraviolet B-rays and began to suffer from a higher incidence of malignant melanoma, a dangerous form of skin cancer. While the ozone hole is mostly restricted to Antarctica, there is evidence of thinning of the ozone layer over the Arctic that could potentially affect those living in northern Europe.

The Montreal Protocol on Substances That Deplete the Ozone Layer is a landmark

Table 10.2

Global Warming Potential and Atmospheric Lifetime of Greenhouse Gases

Greenhouse Gas	Global Warming Potential for 100 Years	Atmospheric Lifetime In Years
Carbon Dioxide	1	50–200
Methane	21	9–15
Nitrous Oxide	310	120
Hydrofluorocarbons	140–11,700	1.5–264
Perfluorocarbons	6,500–9,200	3,200–50,000
Sulfur Hexafluoride	23,900	3,200

international agreement for controlling air pollution on a global scale.[23] Originally signed in 1987 and substantially amended in 1990 and 1992, the Montreal Protocol makes it obligatory on the part of signatory nations to phase out production of CFCs and substitute other chemicals that fulfill the same function without affecting ozone. The Protocol's timetable calls for an 85 percent phaseout of CFCs by 2007, a 100 percent phaseout of CFCs and halons by 2010, and a 100 percent phaseout of methyl bromide by 2015. Even on this timetable, it may require another half century before atmospheric levels of CFCs fall sufficiently to fully restore the ozone layer. The Montreal Protocol to reduce CFCs for protecting the ozone layer was a precursor to the Kyoto Protocol to reduce greenhouse gas emissions for countering global warming.

The Kyoto Protocol

The Kyoto Protocol is an international agreement signed in 1997 to reduce carbon dioxide and other greenhouse gas emissions. The Protocol calls for a reduction in greenhouse gas emissions to an average of 5 percent below 1990 levels to be achieved between 2008–2012. Greenhouse gases are defined as carbon dioxide, methane, nitrous oxides, hydrofluorocarbons (HFCs), perfluorocarbons (PFCs), and sulfur hexafluoride. Table 10.2 shows the relative strength of the greenhouse effect of these gases compared to carbon dioxide and their lifetime once in the atmosphere.[24]

HFCs and PFCs represent two families of similar chemicals, which explains the wide spread in their global warming factors and lifetimes. Table 10.2 is a bit unsettling in that carbon dioxide has a lower impact on global warming compared to the other greenhouse gases. A given volume of methane has twenty-one times the impact as the same volume of carbon dioxide. One unit of volume of sulfur hexafluoride has the warming potential of 23,900 units of volume of carbon dioxide. Of course, there is far less methane and other greenhouse gases in the atmosphere than carbon dioxide. The lifetime of carbon dioxide varies between 50 and 200 years before it is removed from the atmosphere by natural processes (photosynthesis and absorption by the ocean and marine shells), whereas methane has a much shorter lifetime of nine to fifteen years before it is removed by decomposition. But HFCs and PFCs and sulfur hexafluoride, once in the atmosphere, remain forever from a human perspective. The largest greenhouse gas emitters in 2000 were the United States, which was responsible for 21 percent of global emissions, followed by China at nearly 15 percent, the European Union (twenty-five nations) with 14 percent, Russia and India, both with 6 percent, and Japan with 4 percent.

Voting for the ratification of the Kyoto Protocol was both a straight vote by nation and a weighted vote, depending on a country's volume of greenhouse emissions. The United States, the world's greatest contributor to greenhouse gas emissions, had a single vote as a nation and the largest weighted vote based on emissions. The ratification of the Kyoto Protocol required the approval by a minimum of fifty-five nations, representing an aggregate contribution of 55 percent of greenhouse gas emissions.[25] In 2002, over ninety individual nations had ratified the Protocol, but their aggregate contribution to greenhouse gas emissions was only 37 percent. The ratification of the Protocol was in doubt. The principal objection to the Kyoto Protocol was that, if ratified, only thirty-eight developed nations were obliged to take action to reduce their greenhouse gas emissions. The obligor nations were mainly in Europe, including those in transition to a market economy in central and Eastern Europe, plus the Ukraine and Russia. Nations outside Europe obliged to reduce their greenhouse emissions were the United States, Canada, Japan, Australia, and New Zealand. While developing nations could vote in favor to ratify the Protocol, they were exempt from taking any action to reduce greenhouse gas emissions.

The rationale for limiting the obligation to reduce greenhouse gases to thirty-eight nations (actually thirty-seven since one signatory was the European Community, whose members were separate signatories) was that these nations were most responsible for the presence of greenhouse gases already in the atmosphere. Since the developing nations were not responsible for creating the greenhouse gas problem, they should not be held responsible for decreasing their emissions. The counterargument to the obligation being limited to developed nations was that China, India, Mexico, and Brazil, representing 40 percent of the world's population, had become major greenhouse gas emitters in their own right and were expected to be the largest contributors to projected gains.

In 2003, the United States and Russia announced that they would not sign the Kyoto Protocol. Both nations feared the adverse economic impact of complying with the Kyoto Protocol on their respective economies and the accompanying economic advantage that would accrue to nations that did not have to comply. The combined voting power in terms of greenhouse gas emissions of Russia and the United States was sufficient to prevent the Kyoto Protocol from coming into force, but either one switching would swing the vote the other way. In 2004, with 123 nations having ratified the Protocol, representing 44 percent of total emissions, Russia signed the Kyoto Protocol under pressure from European nations, bringing the Protocol into force in February 2005. Excluding the United States and Australia, which also did not ratify the Protocol, the remaining thirty-five nations became responsible for putting into place policies and procedures to reduce their greenhouse gas emissions to 95 percent of their 1990 levels sometime between 2008 and 2012.

The United Kingdom and Germany are already close to compliance because of their switch from coal to natural gas for electricity generation since 1990. Other nations have a tougher row to hoe. Compliance can come about by reducing methane, nitrous oxides, HFCs, PFCs, and sulfur hexafluoride rather than focusing on carbon dioxide since these greenhouse gases have a much greater impact on global warming than carbon dioxide on a unit volume basis. Thus, a relatively small reduction in these gases counts the same as a relatively large reduction in carbon dioxide emissions at potentially less cost. Greenhouse gas emissions can be reduced by building nuclear and sustainable power plants (hydro, wind, solar, tidal), switching from coal to natural gas, and by creating carbon dioxide sinks.

A cap-and-trade program for greenhouse gases in terms of tons of carbon is being set up by the signatory nations to the Kyoto Protocol. Its structure is similar to the cap-and-trade program established by the U.S. Clean Air Act of 1990 for sulfur oxide emissions in terms of tons of sulfur. The emissions trading program for the EU is built around a National Allocation Plan, whereby each member nation is assigned an annual allotment of greenhouse gas emissions. These are broken down and reassigned to an estimated 12,000 greenhouse gas emitters targeted for emission reduction. These include power and heat-generation facilities, oil refineries, coking ovens, and production facilities for ferrous metals, cement, glass, brick, porcelain, pulp, and paper, which, in the aggregate, are responsible for 46 percent of European carbon emissions. Similar to the U.S. program, the targeted companies can participate in an offset program, which compensates for additional emissions from a new plant or an expansion of an existing plant. Targeted companies can select a downstream emissions program, which focuses on reducing greenhouse gas emissions at the point of release to the atmosphere such as during combustion, or an upstream program, which focuses on the characteristics of a fuel such as substituting natural gas for coal.

The estimated cost of greenhouse gas emissions program is of the order of 15 billion euros net of the savings achieved through a cap-and-trade program.[26] To achieve this savings, an effective emissions trading program has to be set up by first establishing an aggregate limit or cap on emissions that is stepped down over time. For trading to be successful, individual participants must have divergent compliance costs in order to create a cost-savings benefit. Companies that can reduce its emissions at a lower cost than others should do so to the maximum practicable extent. This allows the companies to profit from selling emission allowances that are below their stipulated maximum allowances to those facing a high cost for lowering emissions. By purchasing these allowances, companies can avoid making a costly investment in pollution-abatement equipment. Accurate monitoring of actual emissions and effective enforcement are also part of an emissions trading program to ensure that every greenhouse gas emitter holds enough emission entitlements and trading allowances to cover its actual emissions or face stiff fines.

Exchange markets have been established, with others in the process of being set up, in Europe and the United States for trading emission pollution allowances for SOx, NOx, and carbon dioxide. Emissions trading requires a properly organized exchange market to provide ease of trading, transparency of market information to ensure that all participants know the price and volume of every transaction, and standardization of contracts between buyers and sellers to simplify ownership transfer and settlement of disputes. A well-organized market provides access to speculators and traders, whose participation adds depth (volume) to the market. This gives hedgers an opportunity to take a position against an adverse change in the price of greenhouse gas emission rights. Hedging is important if the future value of greenhouse gas emission rights is financially supporting the purchase of pollution-abatement equipment.

An alternative to cap and trade is baseline and credit, a method in which an emissions baseline is defined. A participant takes actions to reduce emissions below the baseline. If successful, an emitter is granted credits for the difference between actual emissions and the baseline at the end of the compliance period. These credits can then be sold to those emitters requiring credits to meet their emission baselines or banked for future use or sale. Borrowing involves the use of allowances or credits on the basis that they will be credited in the future when, for instance, emission control devices have been installed. Banking and borrowing can be similar to the workings of a commercial bank where banked credits (deposits) are borrowed at some "interest" charge.[27]

In addition to emissions trading, the Kyoto Protocol established the Clean Development Mechanism (CDM) as another means of reducing greenhouse gas emissions. This allows nonparticipants, that is, nongreenhouse gas emitters, to earn carbon credits by implementing a program that reduces emissions or creates a greenhouse gas sink. Carbon credits can then be sold to emitters in need of emission allowances. The economic basis for the CDM is that the reduction of carbon dioxide at its source may cost $15 to $100 per ton of carbon, whereas developing a carbon dioxide sink may cost as little as $1 to $4 per ton.

The World Bank is a proactive investment banker for carbon finance, and manages a total portfolio of about $800 million. The World Bank has set up the Prototype Carbon Fund to demonstrate how to reduce greenhouse gas emissions through the use of CDM in a cost-effective fashion. The Community Development Carbon Fund and the Bio Carbon Fund are designed so that small rural communities can benefit from carbon finance by selling credits in sustainable energy projects.

An example of carbon financing is a project to substitute charcoal for coal in pig iron production in Brazil. A plantation of eucalyptus trees is to be established on degraded pastureland where the growing trees are treated as a carbon dioxide sink. After seven years, the trees are harvested and transformed into charcoal as a substitute for coal. The financing is based partly on the economics of the project in the form of revenue generated by selling eucalyptus trees for charcoal production and partly on the sale of carbon credits generated by the plantation as a carbon sink. Another example is an electricity-generating plant fueled by methane emissions from a landfill. The project could not be financed solely on the sale of electricity. By selling the project's carbon credits through a World Bank facility to a company obliged to cut gashouse emissions in accordance with the Kyoto Protocol, enough incremental revenue can be earned for the project to obtain the necessary financing. The financing draws on both the economic benefit of the project to generate electricity and on its environmental benefit of lowering methane emissions to the atmosphere.[28] Although the CDM program is itself in a developmental stage, in early 2005 there were about 160 CDM projects in the world in various stages of conceptualization, approval, validation, registration, and implementation. No projects have reached the point of monitoring, verification, certification, and issuance of carbon credits. Most proposed projects are in South America and Asia, and over half are associated with gas capture/destruction via electricity generation. Others have to do with increasing efficiency, fuel switching, renewable and hydro plants, and forestation projects.

Nonparticipating nations to the Kyoto Protocol are at least partially fulfilling the objectives of the Protocol indirectly. China, with much of its population submerged in a heavy haze of pollution, is paying more than lip service to cleaning up its air. China is trying to reduce its reliance on coal and is taking action to cut exhaust emissions of its domestically produced automobiles, a necessity if China wants to become a major world exporter of automobiles. China's largest automaker has teamed up with Toyota to assemble fuel-efficient hybrid automobiles in China. The EPA has teamed up with its counterpart in China to study ways to significantly reduce air pollution in Beijing in preparation for the 2008 Summer Olympics.

Several states in the United States are requiring companies to report on their carbon dioxide emissions, a possible precursor to setting up a program to place a limit on or decrease emissions. A few states require either a carbon cap or an offset requirement for new plants, a few more have set up committees to explore the possibility of carbon sequestration, and a fair number are developing climate action plans. Many states have instituted some means of keeping track of greenhouse gas inventories and/or have issued mandates for renewable energy to play a specified role in meeting electricity demand. Nine participating states in the Northeast, with five observing states including California, have a goal of setting up the Northeast Regional Greenhouse Gas Initiative to design a regional cap and trade program for carbon dioxide emissions from power plants. Mayors of U.S. cities representing nearly 30 million people have responded favorably to an idea put forth by the mayor of Seattle to have cities contact greenhouse gas emitters to take an inventory of greenhouse gas emissions. Once the amount of greenhouse gases is known, then it would be possible to establish a goal for cutting emissions by about 7 percent. Some utilities are initiating actions to cut greenhouse gases. One utility gave land to the U.S. Fish and Wildlife Service for incorporation in the national refuge system to grow trees as a carbon sink to counter its carbon emissions. Others are looking into carbon sequestration as a means to reduce their carbon emissions. Companies emitting greenhouse gases may one day face potential public and/or shareholder approbation for failing to take some sort of action.

In 2005, the California Air Resources Board adopted the first rules in the United States to reduce greenhouse gas emissions from automobiles sold within its jurisdiction. The new rules require automakers to cut greenhouse gas emissions by as much as 25 percent, beginning in the 2009 model year and increasing to 34 percent for 2016 models. Although this unilateral action appears to be in defiance of a provision in the Clean Air Act that requires one set of automobile pollution emission standards for the nation, as mentioned, the Clean Air Act does not deal with greenhouse gases. New York, New Jersey, and New England are likely to follow the California initiative. If they do, the federal government may be forced to step in to ensure that automobile manufacturers do not have to build automobiles with fifty different greenhouse gas emission standards.

Automobile manufacturers have warned California that this would increase the cost of automobiles because the engine, transmission, and air-conditioning systems would have to be redesigned. But this same warning was given to reduce pollution emissions stemming from the Clean Air Acts. Like all corporate costs, the incremental costs of pollution abatement are passed on to the consumer in the form of higher prices. If such legislation is established in the United States, then automobiles imported into the world's largest market will also have to abide by these rules. Once a company's automobiles marked for export have reduced their greenhouse gas emissions for sale in the United States, it may be cumbersome to have differently designed automobiles marked for domestic consumption. Thus, what California initiates for reducing greenhouse gas emissions from automobiles sold under its jurisdiction may impact other states, which, in turn, may force the federal government to set uniform standards for the nation, which, in turn, would end up affecting both domestic and foreign car manufacturers.

Mandatory programs for reducing carbon dioxide emissions in the United States would likely result in an active carbon emissions trading program. This would be a boon for solar and wind power providers because their sale of carbon credits would aid in financing new sustainable energy projects.

Efficiency and Conservation

Efficiency is giving up an SUV that gets ten miles to the gallon for a hybrid that gets forty miles to the gallon. Thus, driving an automobile the same distance consumes less gasoline. Conservation is driving the automobile fewer miles by combining trips, carpooling, and eliminating unnecessary travel. Efficiency is adding insulation to a house to keep it at the same temperature, thereby consuming less heating oil or natural gas. Conservation is lowering the temperature. Efficiency implies that the same function can be accomplished with less energy. Conservation implies less need for a function. Both energy efficiency and conservation have an economic benefit, because they lower energy costs by reducing demand, and an environmental benefit reducing pollution emissions.

Energy Star Program

The major impetus to energy efficiency in the United States is the Energy Star program, established by the EPA in 1992 to reduce greenhouse gas emissions by encouraging energy efficiency.[29] Computers and computer monitors were the first products to carry the Energy Star label, which was extended to other office equipment and residential heating and cooling equipment in 1995. In 1996, EPA partnered with the Department of Energy (DOE) for those products that fell under the DOE domain for enhancing efficiency (dish and clothes washers, refrigerators, room air conditioners, light bulbs, and windows). The Energy Star program now has thousands of partnerships with private and public sector organizations to deliver the technical information and tools necessary for organizations to select business solutions and management practices that enhance energy efficiency. Companies that produce energy-efficient products are provided the means to break down market barriers and alter decision-making patterns so that more consumers will buy energy-efficient products.

The Energy Star label is issued when the additional cost of enhancing energy efficiency compares favorably with the benefit of lower energy costs throughout the product's life. At times, there is no additional cost such as reducing energy demand when office equipment and home electronics (e.g., personal computers) are automatically placed in a standby mode. The Energy Star label provides consumers with a straightforward determination of whether or not to purchase an energy-efficient product. The goal of the Energy Star program is to have all manufacturers of a product offer Energy Star labeled products, making it impossible for consumer to buy cheap energy-inefficient products.

The primary focus of the Energy Star program is to enhance efficiency in order to reduce pollution from residential and commercial sources. These account for 35 percent of greenhouse gas emissions in the United States, equally divided between the two. The remaining sources are industrial (30 percent), transportation (27 percent), and agriculture (8 percent). Greenhouse gases are estimated to be 85 percent carbon dioxide, 13 percent methane and nitrous oxides, and 2 percent HFCs, PFCs, and sulfur hexafluoride. Methane emissions come primarily from agricultural activities while nitrous oxides stem mainly from electricity generation and transportation. The Energy Star program estimates that it has reduced energy costs to consumers and businesses by a total of $10 billion and, in so doing, has cut pollution emissions equivalent to 20 million motor vehicles since its inception to 2004. Its goal for the next ten years is to continue pursuing energy efficiency whose savings in energy would result in reduced pollution equivalent to 40 million motor vehicles. To accomplish this, it is necessary to increase the 60 percent of consumers who are aware of the Energy Star label and the 30 percent willing to restrict their buying to Energy Star-labeled products. In 2004 the Energy Star label was on forty different types of products (32,000 individual product models) with 1,400 participating manufacturers.

In addition to products, the Energy Star Label is beginning to appear on homes and buildings. About 360,000 homes have been constructed carrying the Energy Star label, with more homebuilders being recruited. Twelve percent of commercial buildings have been rated for energy performance including hospitals, office buildings, supermarkets, schools, and hotels. Another organization with similar aims is the U.S. Green Building Council, a coalition of industry leaders promoting the construction of environmentally sound apartment and commercial buildings.[30] According to the Council, buildings in the United States account for 65 percent of electricity demand and for 30 percent of greenhouse gas emissions, 30 percent of raw material consumption, and 30 percent of waste output plus 12 percent of potable water consumption. The Council's LEED (Leadership in Energy and Environmental Design) rating system has become a nationally accepted standard for the design, construction, and certification of environmentally friendly "green" buildings that cut energy demand by half and water demand by a third. Structure, materials, insulation, heating and cooling systems, windows and doors, water usage, shade trees and landscaping, and disposal of construction debris are some of the areas scrutinized in a LEED certification.

The Energy Star program has partnered with industrial concerns to support the purchase of highly efficient combined heat and power units in order to satisfy their hot water needs by recycling the waste heat of electricity generation, with landfill and industrial concerns to cut methane emissions, with Green Power providers to encourage building solar- and wind-generating facilities, and with a host of state and local organizations pursuing energy efficiency and emissions reduction. The Energy Star program has taken a "backdoor" approach to fulfilling the objectives of the Kyoto Protocol by working with industrial concerns to reduce their HFCs, PFCs, and sulfur hexafluoride emissions.

Light-Emitting Diodes

Light-emitting diodes (LEDs) are semiconductor materials that emit light. They have been around since the 1960s when LEDS emitted only red, green, and yellow light. In the 1990s, blue LEDs were developed, which, when combined with red and green, resulted in white light. Since then improvements have been made to reduce their manufacturing costs and improve their brightness. White light LEDs can last 50,000–100,000 hours of continuous use (nearly six–twelve years), 50–100 times longer than the conventional incandescent bulb.

An incandescent bulb, as the name suggests, heats a filament in a vacuum until it glows, releasing a large amount of waste heat. LEDs can give off as much light with far less waste heat, consuming one-tenth the electricity. If the initial upfront cost continues to decline and the brightness becomes compatible with an incandescent bulb, a crossover point may occur where conventional light bulbs can be replaced by LEDs. Complete conversion to LEDs will significantly reduce total electricity demand, saving billions of dollars for consumers in terms of smaller electricity bills. The fall in electricity demand will lower energy costs, cut pollution emissions, and reduce the need to build large generating plants. Less capital spending for electricity generators and transmission lines can be dedicated to fulfilling other needs. The fly in the ointment is that most lighting is needed at night, when electricity demand is at its lowest. LEDs reduce nighttime base-load demand, not higher daytime (daylight) demand. Thus, the fall in base demand still necessitates generators to handle daytime demand.

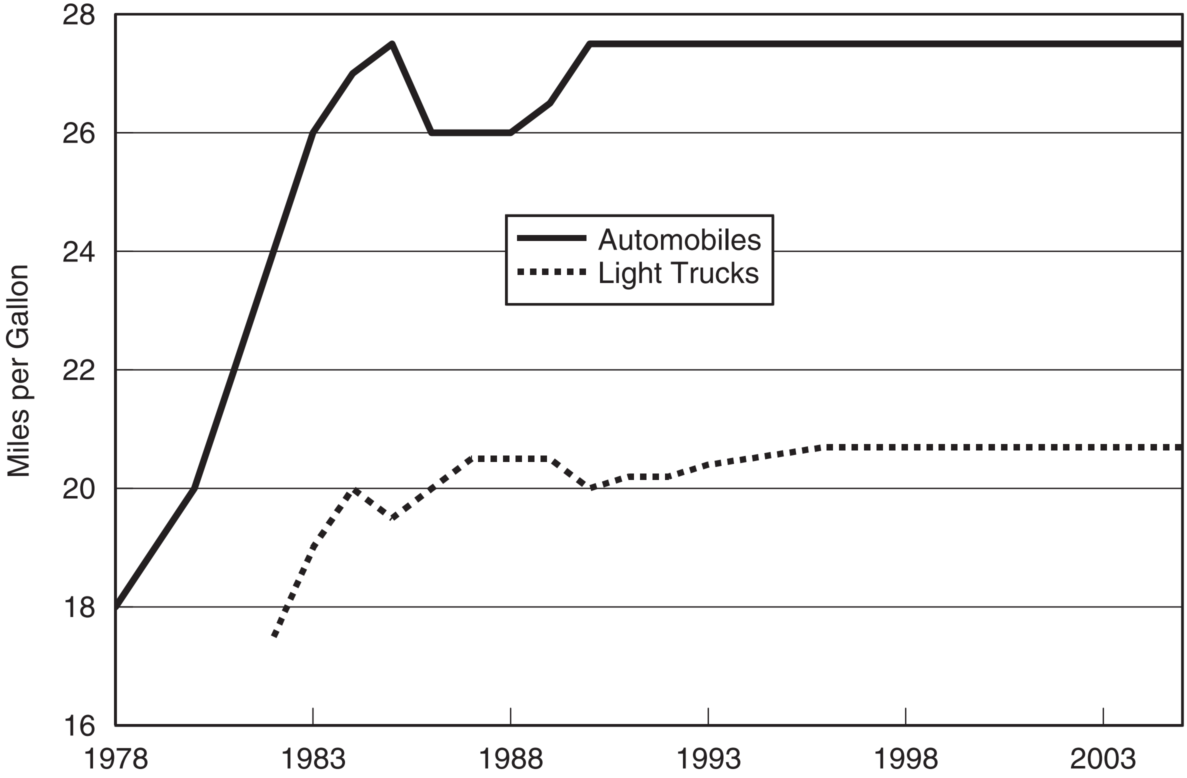
While the promise of LEDs appears positive, compact fluorescent light bulbs with energy savings of 66 percent and a life ten times longer than a conventional incandescent bulb, but with a higher acquisition cost, have barely made a dent in the light bulb market. Old habits die hard and few consumers take the time to perform a comparative cash-flow analysis between lighting alternatives. Yet this is what is required for energy conservation to take hold. Energy conservation is a particular way of thinking, or mindset, about energy, if individuals are to undertake a whole range of energy-saving actions. These can take the form of wearing sweaters rather than tropical short-sleeved shirts during the winter to lower room temperatures, cleaning furnace filters, sealing air leaks around windows and doors, and installing insulation. During the summer the largest source of energy savings is relying less on air conditioning and more on natural air circulation, perhaps enduring higher, but not uncomfortable, temperatures. Regardless of the season, turning off unnecessary lights and electronic equipment (televisions, computers) when not in use, running dishwashers and washing machines with full loads, and a host of other practical actions can save energy. While each act saves only a smidgeon of energy, the aggregate impact of many individual acts by tens or hundreds of millions of individuals can have a significant impact on energy demand, energy prices, and pollution emissions.

CAFE Standards

The Energy Policy Conservation Act of 1975 established the Corporate Average Fuel Economy (CAFE) standards for automobiles and light trucks, administered by the National Highway Traffic Safety Administration (NHTSA) under the Secretary of Transportation.[31] As originally structured, CAFE standards applied to the sales-weighted average fuel economy of a foreign and domestic manufacturer's fleets of passenger cars and light trucks of less than 8,500 pounds in gross vehicle weight sold in the United States. The original near-term goal was to double new car fuel economy by model year 1985. The EPA is responsible for calculating the average fuel economy for each manufacturer. CAFE standards, shown in Figure 10.1, have been set at the maximum feasible level consistent with technological and economic practicality, along with the need to conserve energy.

It is clear from Figure 10.1 that the objective of doubling gas mileage did not occur. The initial standard for passenger cars for the model year 1978 was an average eighteen miles per gallon, increasing to twenty-seven and one half miles per gallon for model year 1990 and thereafter. Light truck standards originally applied to vans and pickup trucks, commonly used in place of automobiles for personal travel in the United States in addition to their commercial use. Manufacturers

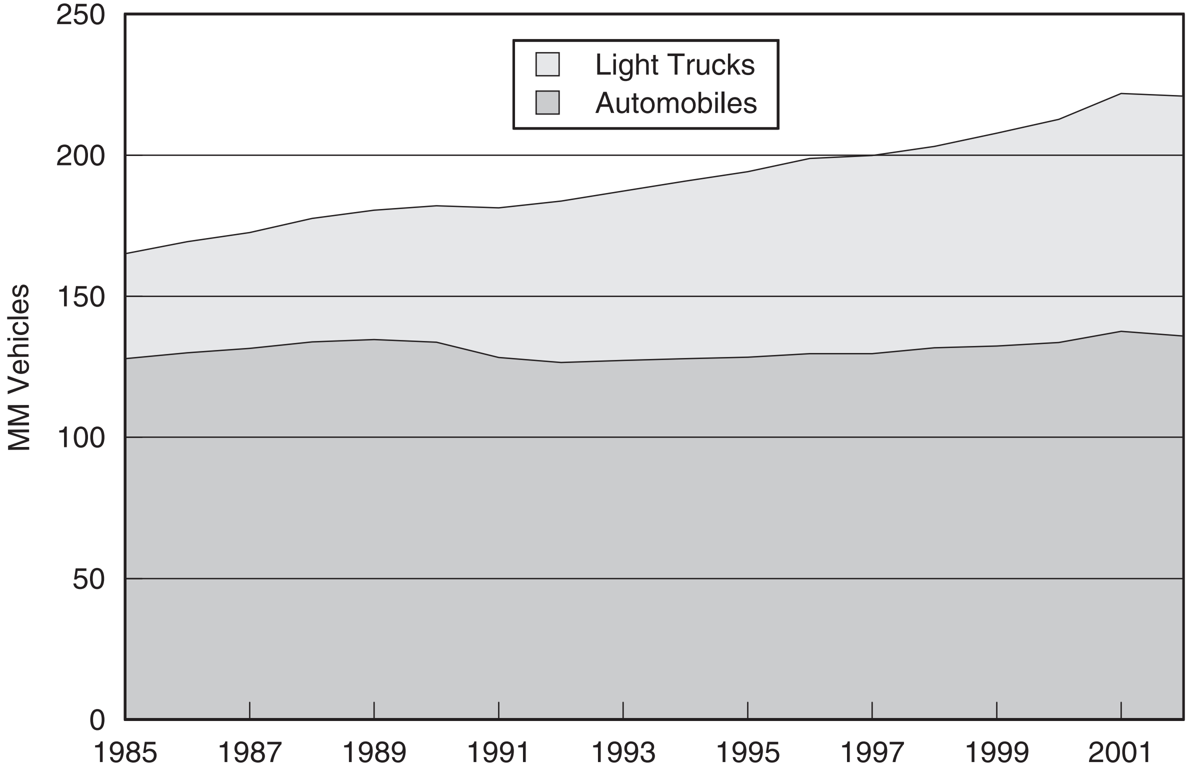
Figure 10.1 **Automobile and Light Truck CAFE Standards**



earn CAFE credits for models that beat CAFE standards, which can be applied to models that exceed standards, or face monetary penalties if average fleet performance does not meet CAFE standards. Light and other truck designations above 8,500 pounds, which includes nearly all commercial trucks, were outside the purview of CAFE standards. Figure 10.1 shows that CAFE standards have barely changed since the mid-1980s. Gasoline mileage is measured at the factory under the auspices of the EPA and does not necessarily reflect actual mileage on the road.

Critics of CAFE standards point out that the automobile manufacturers already had in the pipeline a number of technological improvements to enhance automobile gasoline mileage when the CAFE standards were instituted. These improvements responded to the public demand for fuel-efficient automobiles in the wake of high gasoline prices in the late 1970s and early 1980s, which gave a tremendous boost to fuel-efficient Japanese and European automobile imports. Thus, the move to higher-mileage automobiles was not driven by a government mandate, but by market forces. A more damning criticism can be leveled at the decision to classify sports utility vehicles (SUVs) as light trucks rather than automobiles. As light trucks over 8,500 pounds, the high-mileage, heavier SUVs

Figure 10.2 **U.S. Light Truck and Automobile Population**



avoided compliance with CAFE standards. The popularity of SUVs, which for years made up half of new car sales in the United States, increased gasoline demand at a time when gasoline prices were relatively low. Figure 10.2 shows the growth in light trucks (vans, pickups, SUVs), with much of the incremental growth being SUVs. In 2002, just under 40 percent of the U.S. population of automobiles was light trucks.[32]

CAFE standards are a subject of intense discussion whenever the U.S. energy policy is under review, when gasoline prices spike, and during elections. U.S. labor unions keep a wary eye on CAFE standards because they see small fuel-efficient cars as a potential loss of jobs compared to the jobs generated by manufacturing larger, more fuel-inefficient SUVs. U.S. labor unions also feel that pushing the market toward fuel-efficient cars opens up marketing opportunities for foreign makers of small cars, another threat to domestic jobs. Automobile companies keep a wary eye on CAFE standards because profit margins on large cars are greater than for small cars. On top of this is a continuing debate over the role of the government in determining fuel standards or of passing the responsibility from automobile manufacturers to automobile buyers through higher-priced gasoline. Some argue that CAFE standards should

Table 10.3

Proposed CAFE Standards for Light Trucks

	Miles/Gallon
Small SUVs	28.4
Midsize SUVs	27.0
Small Pickups and Larger SUVs	24.5
Minivans and Midsize Pickups	23.3
Pickups, Large SUVs and Vans	21.7
"Big Trucks"	21.2

again be increased to try to cut gasoline consumption, which continues to grow from a rising motor vehicle population that is driven more miles per year. Others argue that CAFE standards should not be increased because this would force the manufacture of smaller cars that are more dangerous in a collision than larger cars.

In September 2005, the National Highway Traffic Safety Administration (NHTSA) proposed that CAFE standards for light trucks be increased and phased in between 2008 and 2011. The standard for automobile mileage would remain unchanged at twenty-seven and a half miles per gallon. The maximum weight of light trucks covered by CAFE standards would rise from 8,500 to 10,000 pounds to bring Hummers and larger SUVs under CAFE standards. Rather than having a single category for light trucks, whereby lighter-weight and more fuel-efficient models compensated for heavier and less fuel-efficient models, the light truck category would be broken down into six categories based on size (weight), shown in Table 10.3. The Big Truck category applies up to 10,000 pounds, which excludes large commercial freight-carrying trucks.

On an overall basis, the average mileage for light trucks should increase from the current 21.2 mpg to 23.5 mpg by 2011. This nearly 11 percent increase in CAFE standards does not mean an 11 percent decrease in gasoline consumption because of the difference in fuel usage as measured by factory tests and actual road performance. An estimate of actual fuel mileage can be derived from gasoline consumption, the population of automobiles and light trucks, and an estimate of annual miles driven for both kinds of vehicles. This works out to be 22.1 mpg for automobiles versus the CAFE standard of 27.5 mpg and 17.6 mpg for light trucks versus the CAFE standard of 20.7 mpg. Nevertheless, an increase in the CAFE standard for light trucks will have some impact on reducing per-vehicle gasoline consumption.

The exemption of SUVs from compliance with CAFE standards, coupled with the sharply escalating growth of motor vehicles in India and China, played a role in the emerging oil crisis in 2004. Despite what appear to be continuing problems with oil supplies, there are those who argue against CAFE standards on the basis that more fuel-efficient lighter-weight cars lead to more injuries and deaths in collisions. Another argument against lighter-weight cars, particularly hybrids, is that fuel-efficient automobiles use less gasoline and, therefore, pay less in highway taxes. A proposal has been made to tax fuel-efficient automobiles in order to ensure that their contribution to highway trust funds remains equivalent to that of SUVs. Thus, at this point, with an incipient crisis in oil, there are advocates who want to keep U.S. highways congested with gas-guzzling SUVs to ensure that their occupants have plenty of steel around them in case of an accident and to keep highway trust funds amply endowed.

Notes

1. Jeremy Rifkin, *The Hydrogen Economy: The Creation of the Worldwide Energy Web and the Redistribution of Power on Earth* (New York: Jeremy P. Tarcher/Penguin Group, 2002).

2. This information is from an online article, “Hydrogen: The Fuel of the Future,” which can be accessed at the Plug Power, Inc. Web site at www.plugpower.com

3. For details, visit the Icelandic New Energy Web site at www.newenergy.is.

4. The workings of fuel cells and their present commercial uses are well described online at Plug Power, Inc. Web site at www.plugpower.com, and at Ballard Power Systems at www.ballard.com.

5. More information on hydrogen can be found at the U.S. Department of Energy Office of Fossil Energy Web site at www.fossil.energy.gov.

6. Steven Ashley’s “On the Road to Fuel-Cell Cars,” *Scientific American*, (March 2005), vol. 292, no. 3, p. 62 provides an excellent status report on the hurdles facing hydrogen-fueled automobiles.

7. John D. Cox, *Climate Crash* (Washington, DC: Joseph Henry Press, 2005).

8. William F. Ruddiman, “Did Humans Stop an Ice Age?” *Scientific American* (March 2005), vol. 292, no. 3, p. 46.

9. Brian M. Fagan, *The Little Ice Age: How Climate Made History (1300–1850)* (New York: Basic Books, 2001).

10. *Impacts of a Warming Arctic: Arctic Climate Impact Assessment* (Cambridge, UK: Cambridge University Press, 2004), Web site: www.acia.uaf.edu.

11. “Looking Back” and “Making Records” from the January/February 2006 issue of Weatherwise magazine, vol. 59, no. 1, pp. 10, 48.

12. *Climate Change 2001: Third Assessment Report,* published by the Intergovernmental Panel on Climate Change (IPCC), a group jointly established by the World Meteorological Organization (WMO) and the United Nations Environmental Program (UNEP), available online at www.ipcc.ch.

13. *BP Energy Statistics* (London: British Petroleum, 2005).

14. “The Case of the Missing Carbon” *National Geographic* (February 2004) vol. 205, no. 2, p. 88.

15. “Gas Guzzlers” by Deborah Franklin published in the February 2004 issue of *Smithsonian,* vol. 34, no. 11, p. 25.

16. This information can be found at the University of Southampton Web site at www.soton.ac.uk/mediacentre/news/2005/nov/05_204.shtml.

17. *The Asian Brown Cloud: Climate and Other Environmental Impacts* published by the United Nations Environment Program (UNEP), 2002.

18. G. Stanhill and S. Cohen, “Global Dimming: A Review of the Evidence,” *Agricultural and Forest Meteorology* 107 (2001), p. 255.

19. Michael Crichton’s *State of Fear* (New York: HarperCollins, 2004) is one such book of fiction.

20. Environmental Protection Agency, *The Plain English Guide to the Clean Air Act* (Washington, DC: Author, 2006); see Web site at www.epa.gov/oar/oaqps/peg_caa/pegcaain.html

21. To read the EPA Clean Air Interstate Rule; visit the EPA’s Web site at www.epa.gov/interstateairquality.

22. The “Clean Air Act Timeline,” published by Environmental Defense, is available online at www.environmentaldefense.com/documents/2695_cleanairact.htm.

23. United Nations Environmental Programme at www.uneptie.org/ozonaction.

24. Intergovernmental Panel on Climate Change (IPCC), National Greenhouse Gas Inventories Programme at www.ipcc-nggip.iges.or.jp/index.html.

25. The Convention and Kyoto Protocol is available online at the United Nations Web site at www.unfccc.int/resource/convkp.html.

26. *Emissions Trading and Utilities: What’s the Impact?* (Albuquerque, NM: UtiliPoint International, 2005).

27. *A Guide to Emissions Trading* (New York: United Nations Environmental Programme, 2002).

28. The “Carbon Finance Annual Report for 2005” and other information can be found on the World Bank Web sites at www.worldbank.org and www.carbonfinance.org.

29. Information on the Energy Star Program and the Annual Report of the Energy Star Program for 2004 are available online at www.energystar.gov.

30. More information on this organization is available at the U.S. Green Building Council Web site at www.usgbc.org.

31. See the National Highway Traffic Safety Administration Web site at www.nhtsa.dot.gov/cars/rules/cafe/overview.htm.

32. *Transportation Energy Data Book* (Oak Ridge, TN: Center for Transportation Analysis, Oak Ridge National Laboratory, 2004).

11

An Energy Strategy

Nations pursue a military strategy for winning a war. Companies pursue a business strategy for achieving competitive advantage. Business strategy involves examining a company's internal strengths and weaknesses along with its external threats and opportunities. Such an analysis tells a company how to take advantage of its strengths, deal with its weaknesses, neutralize its threats, and pursue opportunities. Strategy implies a plan for survival: for a nation, to win a war; for a company, to remain an ongoing concern. Similarly, an energy strategy has to do with maintaining a way of life: getting into a car and having a job to go to or flicking a switch and having a light go on. Once an energy strategy has been identified, then an energy policy can be devised to achieve its goals. If we address energy strategy from the point of view of survival, two issues come immediately to the fore: oil and pollution.

Oil

The International Energy Agency publishes the *World Energy Outlook,* which includes a projection for oil supply and demand. This projection is based on extrapolating past trends. This is a favorite technique for forecasters because it relies on something that is known—past data. If nation *W* has been experiencing *X* percent growth for the last *Y* years, what is wrong with the assumption that this will continue for the next *Z* years, perhaps plus or minus a bit to take into account something that may affect its future growth? This is a very comfortable approach because if continuity of trends is a valid assumption, then a projection of future oil consumption can be easily determined by mathematical means.

It is clear from viewing the history of oil prices that there is no underlying trend. The approach here is to tap the brains of experts. If a noted authority on oil projects a future oil price of $*N* per barrel and another noted oil authority projects a future oil price of $*Q* per barrel, then what is wrong with assuming that the future oil price will be 10 percent above or below the average of $*N* and $*Q*? Who else has a better sense of future oil prices than the world-recognized oil authorities? Do not they advise oil companies, producers, and government agencies and are they not quoted in the press every other day? They are by far the best source of future price assessments. If not these people, then who; and if they miss the mark, are we at fault?

The problem with oil authorities is one of credibility. If the price of oil has been in the range of $20–$25 for the last three years, the safe bet is that it will continue within the same range, or maybe +/− 10 percent above or below that range. Even if a noted authority feels that $10 or $100 per barrel oil is in the offing, his or her entire reputation is on the line if this projected estimate does not materialize. The individual can become the butt of jokes within his or her professional circle. Worse yet, consulting revenue will shrivel up. The safe bet in projecting the future is to assume that things do not change much from the current situation. Outlier projections have a low probability of occurrence.

Table 11.1

Average Annual Energy Costs for Oil and Natural Gas

	West Texas Intermediate $/Bbl	Henry Hub $/MMBtu
2002	$26.16	$3.33
2003	$31.07	$5.63
2004	$41.49	$5.85
2005	$56.59	$8.90

Too many things can happen that will make an outlier projection look silly. The real problem is that too many things happen that make a "business as usual" projection look silly. The difference is that an outlier projection can knock an individual forecaster out of the box, whereas having about the same projection keeps all forecasters in the game.

In the world of forecasting, the concept of chaos is very uncomfortable. It means that we cannot forecast with confidence and this means that nearly all, if not all, forecasts turn out wrong. Table 11.1 is the average cost of oil and natural gas between 2002 and 2005. No one in 2002, 2003, or, for that matter, 2004, projected that the average oil price would be over $50 per barrel and natural gas prices would be close to $10 per million Btu in 2005 (as with any average, spot prices exceeded the average). Energy experts looked at the cost of energy in 2005 and shook their heads in disbelief.

The 2002 issue of the *World Energy Outlook,* based on 2001 data, projected a 2020 oil price in the range of $20–$27 in constant 2000 dollars, based on projections made by a leading oil company, a leading oil economics consulting firm, and a leading government energy agency.[1] The 2004 issue of the *World Energy Outlook,* based on 2003 data, projected oil prices *declining* to $22 per barrel in 2006, remaining there until 2010, then stepping up to $29 per barrel in 2030 in terms of constant 2000 dollars. In addition to this reference price scenario, there was a high oil price scenario of $35 per barrel throughout the projection period. Natural gas prices were to follow the reference oil price scenario and coal prices were to experience a slow increase after 2010.

As seen in Table 11.1, the projected oil price was blown to bits within months of the report's publication. The same can be said for the projected prices for natural gas and coal. Even the high-price scenario, which had a lower implicit probability of occurrence than the reference price scenario, was smashed by events. When industry observers shook their heads in disbelief in 2005 over how high energy prices had jumped, what they were admitting to is an element of chaos, the occurrence of unpredictable events that confounds our predilection for continuity of trends.

The *World Energy Outlook* is unique in the forecasting business for publishing its past projections on oil prices and performing a critical review of their accuracy. Most forecasters forget about their past forecasts and hope their readers do likewise. The criticism is not aimed at *World Energy Outlook* projections, but points out the difficulty of making projections. Most forecasters employ the same methodology used in the *World Energy Outlook* and with similar results. By default, forecasters assume a rational world where trends remain essentially intact or slowly change over time. The world of forecasters is not the real world, although they would like to believe it is. They believe in the continuity of trends and in the ultimate rationality of the world around us. Unfortunately for forecasters, life is chaotic. It is like a card game that starts with no wild cards but, as the game continues, certain cards are declared wild; then, throughout the game, the cards that are wild change constantly, making a mess of the odds of winning.

Table 11.2

2030 Projected Oil Demand and Supply

	MMBpd
World Demand 2002	77.0
Projected World Demand 2030	121.3
Incremental World Oil Gain 2002–2030	44.3
Incremental Demand China	8.1
Incremental Demand India and Rest of Asia	9.4
Incremental Demand United States & Canada	6.9
Incremental Demand to be Supplied by OPEC	36.6
Incremental Demand to be Supplied by Middle East OPEC	32.8

Should projections be done? Absolutely, as long as readers recognize their inherent limitations; the true value of projections is the thought process of gaining insight into and understanding of the underlying forces and external factors affecting energy. From this corporate planners and government policymakers can formulate courses of action or strategies to deal with emerging trends and the changing nature of energy usage.

An Unlikely Projection

The projection of oil demand and supply for 2030 in the *World Energy Outlook* for 2004, based on extrapolating from past trends, is summarized in Table 11.2.

The big three producers in the Middle East with the potential for vastly expanding their production are Saudi Arabia, Iraq, and Iran. There are others, but they are small in area and the chances of discovering huge oil deposits on their soil are likewise small. In 2004 Iran produced 4.1 million bpd, about one-third less than it did under the shah. With domestic consumption of 1.6 million bpd, the nation exported 2.5 million bpd. Sixty percent of Iranian production is from steadily declining oil fields that have been in production between forty and seventy years. Iran has to have a fairly active oil development program just to stand still. There is a plan to increase production to 5.8 million bpd in five years, but Iran will have to attract $70 billion in outside investments, plus import technical expertise, no small task for the xenophobic mullahs. A possible partner is China, but even if this project materializes, Iran's exports will be up by only around 1.7 million bpd from current levels, far from the needed 33 million bpd in Table 11.2.

Iraq is regrouping as a nation, which has to be accomplished before it can resume being a major exporter. Its production of around 2 million bpd depends on the level of security of its pipelines to prevent sabotage. Iraq actually has a potential for the discovery of a megafield because large areas of the nation have not been explored. But from today's perspective, Iraq has a long way to go before it can pay a great deal of attention to exploiting its oil resources.

This leaves Saudi Arabia. Unlike Iraq, Saudi Arabia has been thoroughly explored. While it is true that Saudi Arabia can expand its output of about 11 million bpd, of which 1.7 million bpd are consumed internally, not even Saudi Arabia claims that it can triple production in order to be the chief contributor in satisfying the 33 million bpd of incremental demand projected in Table 11.2. If the Middle East nations as a group, with Saudi Arabia being the principal contributor, cannot increase their production by 33 million bpd, then the projection of oil demand based on extrapolating oil trends will run into a brick wall sometime between 2004 and 2030.

The world ran into the brick wall in 2005. The wild card, which no one foresaw, was more rapid growth in oil consumption in China, India, and the United States than had been anticipated.

A straight-line interpolation between 2002 and 2030 of 44.3 million bpd from Table 11.2 infers an annual gain of 1.6 million bpd, or projected consumption for 2005 of 81.8 million bpd versus an actual consumption of 84 million bpd. Both China and India experienced incredible economic growth on the order of 10 percent and 7 percent per year, respectively, which doubles the size of their economies every seven (China) to ten years (India). Energy demand may not match, but it certainly follows economic growth. Part of energy growth is oil as the necessary fuel to get raw materials and workers to factories and products to market. India and China are intent on joining the ranks of developed nations and having the same accoutrements, including a car in every garage. While India and China were emerging as powerful industrial nations, the United States was on a gas-guzzling SUV binge.

In this mad rush for oil, we came up against two physical constraints missed by those extrapolating demand trends in 2003. One was the capacity of the world to produce oil. In 2005 every well in the world was running at full capacity and oil prices kept rising. Repeated calls by OPEC to increase production to get prices under control went unheeded, primarily because production was already full out. The estimate for world spare capacity was about 1–2 million bpd, nearly all of which was in Saudi Arabia. This degree of surplus did not provide much comfort when world demand hovered around 85 million bpd, and was insufficient for a sustained interruption in exports from Venezuela, Nigeria, Iran, or any other exporter of 2 million bpd or more. As a point of comparison, in the late 1970s, there was enough spare capacity in Saudi Arabia to compensate for nearly a 5 million bpd loss of production during the Iranian Revolution, with some to spare. This level of security to meet unexpected jumps in consumption has long since disappeared.

OPEC is supposed to control price within an acceptable range by adjusting its volume to world demand. If prices become too low, OPEC cuts production to raise the price. If prices are too high, OPEC hikes production to decrease the price. One may wonder why OPEC would increase production to lower price. Those within OPEC with long memories worry about high prices. High oil prices in the late 1970s and early 1980s induced a worldwide economic contraction, encouraged programs to enhance efficiency and conservation, promoted switching away from oil, and financed the development of non-OPEC sources of oil, most notably the North Sea. By the mid-1980s, Saudi Arabian exports were headed toward the basement and were subsequently restored when Saudi Arabia cut the price of oil by flooding the market. The 2004 *World Energy Outlook's* high oil price scenario came to the same conclusion: High-priced oil cuts overall demand and encourages growth of non-OPEC oil such that by 2030, OPEC revenue would be below that of the reference-priced scenario because of a lower volume of exports. This was reflected again in the 2005 *World Energy Outlook* revision in which projected oil prices were raised and world consumption was lowered along with Middle East exports.

Prior to 1970 the Texas Railroad Commission was responsible for world oil prices by controlling oil well production in Texas and Oklahoma. The United States lost control over prices in 1970 when the Texas Railroad Commission ordered every well to pump at full capacity as the United States emerged as a major world oil importer. It took OPEC a few years to realize that the chips were now on their side of the table. Since the 1973 oil crisis, OPEC has influenced price by controlling production. But OPEC cannot play this role of price arbiter when running at full capacity any more than the Texas Railroad Commission could. While high oil prices fill the coffers of the oil exporters, they also provide plenty of incentives to explore and develop more costly fields, to pursue more sophisticated forms of tertiary recovery methods from existing wells, to seek alternatives to oil, and to more vigorously pursue efficiency and conservation, all of which leads to less consumption of OPEC oil.

Another physical constraint that we bumped against in 2005 was refinery capacity. World refiners were operating at nearly full capacity to meet demand. Some blame high oil prices not on production constraints, but refinery constraints. If refinery output is less than demand, then consumers start bidding up the price of oil products. As oil product prices rise, it becomes relatively easy for oil producers to hike crude oil prices. In other words, the villain for high crude prices in 2005 was not crude oil exporters, but gasoline consumers. Our economic system works well as long as there is a surplus. All hell breaks loose once demand gets too close to supply, a scenario discussed in the California electricity debacle in Chapter 2. Prices become chaotic as bidders try to outdo one another and grab what they believe to be their rightful share of a dwindling resource.

Scarcity Playing Itself Out

There are only two alternatives when dealing with scarcity: increase supply or cut demand. One or the other or both must occur. We are dealing in a world of physical realities; demand cannot exceed supply. If demand appears to be getting ahead of supply, someone is going to come up short.

Increasing Supply

While no one questions oil companies' eagerness to develop new oil fields as fast as they can, there is a question about their eagerness for wildcatting: the unabashed search for oil in areas that have not been explored. It is relatively safe to explore near existing oil fields, but it is relatively risky to sink a well where no one else has even with the benefit of three-dimensional seismic analysis.

Part of the problem of the apparent unwillingness to wildcat is the Western petroleum geologists' obsession with the organic (biotic) origin of oil. If it is believed that oil has a biotic origin, the search for oil is constrained to sedimentary rock. If much of the sedimentary rock has been explored, then where does one drill a wildcat well? Russian petroleum geologists believe that the origin of oil is not biotic, but abiotic, and comes from the earth itself. Natural gas from the earth's mantle seeps up through fissures in the bedrock and is transformed to oil as it migrates through sedimentary rock, picking up evidence of a marine origin along the way, until it becomes trapped within an impermeable rock formation. If this abiotic origin of oil is true, then this opens up a whole new avenue of exploration—deep fissures in the earth's crust. An abiotic origin does not solve the problem of running out of oil because we can be draining oil fields faster than they are being replenished. However, if oil has an abiotic origin, we should have more time before reserves are exhausted.

The origin of oil is, however, immaterial; what is material is whether or not oil lies below sedimentary rock. A conference should be held where Russian petroleum geologists can make their case; not regarding the origin of oil, but their discovering oil in bedrock beneath sedimentary rock. If oil can be found in bedrock, a whole new realm opens up for exploration. This would absorb funds now being diverted to buying back oil company stock, a form of self-liquidation.

Discovering new sources of crude oil does not mean greater availability of gasoline because crude oil has to be refined before consumers can purchase oil products. Refinery construction is strong in the eastern hemisphere (the Middle East, India, and China) and nearly non-existent in the western hemisphere. Environmentalists have successfully prevented any grassroots refinery from being built in the United States for decades, where aggregate refinery capacity expands through enhancements to existing refineries (refinery creep), but not enough to keep up with demand. Refinery construction in Europe and South America is minimal. India and China are ensuring that their incremental oil needs will be refined whereas the United States is intent on drawing down on whatever spare refinery capacity can be found overseas. We have already exhausted spare refinery capacity in the Caribbean and South America, and perhaps in Europe. As U.S. consumption increases, we will become increasingly reliant on spare refinery capacity in Russia, if it exists, and the Middle East.

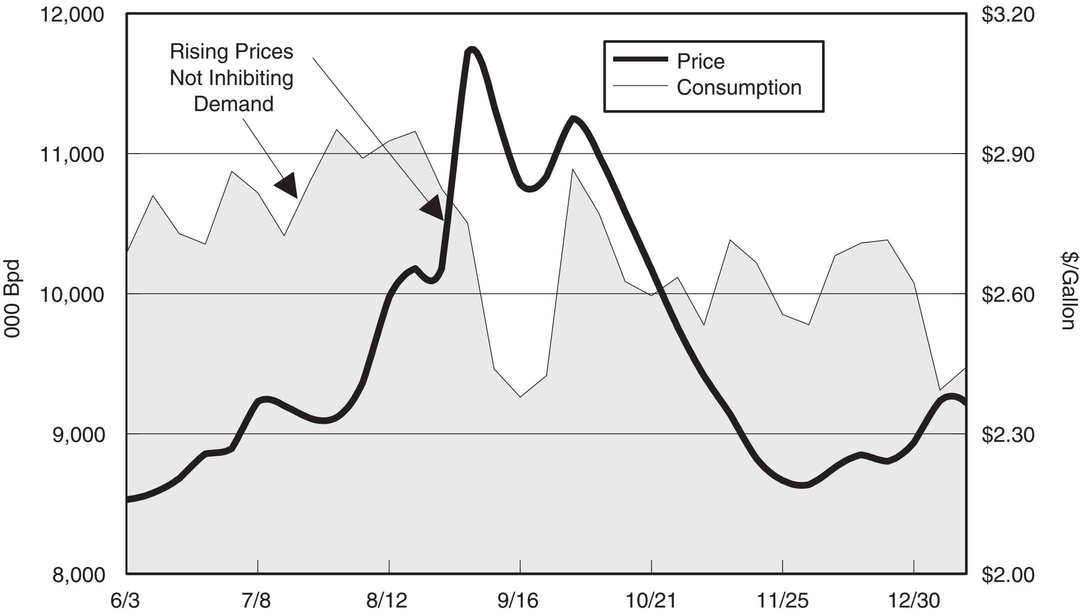
One can reasonably ask why something as obvious as oil supply and refinery constraints were not anticipated before 2005. Both have been discussed in the petroleum press for some time, so they were not really surprises. The surprise was the unexpected surge in oil consumption in 2005 that was sufficient to hit both limits. The question is not so much why this was not foreseen, but what we are going to do about it now that we have encountered real constraints. Developing new oil fields and building new refinery capacity takes time. It is an eventual solution, but eventuality is not now or in the near future.

Decreasing Demand

In the short run we will be relying on the other way of dealing with scarcity: demand destruction. Demand destruction occurs when a higher price cuts demand. The first to cut consumption from higher prices are those who can least afford to pay. Demand destruction has already occurred for the third-world nations: The poorest of the poor have to learn to consume even less oil than what little they consumed in the past. Petroloans were a way for developing nations with a negative trade balance and without oil resources to obtain oil. Western banks took deposits from oil producers and recycled the funds as petroloans to developing nations to allow them to buy oil. The system worked fine for the oil exporters, who could now sell oil for cash to nations that had no cash. The system worked fine for the oil importers who could now buy oil without having the requisite cash in hand. Petroloans let countries buy oil they could not afford just as credit cards let individuals buy things they cannot afford. As the decades rolled on, petroloans just kept being rolled over into larger and larger loans, plus accrued interest until it became apparent to all that these loans could not and would not ever be repaid. In 2005 the Western banking system admitted defeat (or accepted reality) and started to write off these debts in the tens of billions of dollars (the true beneficiaries of petroloans, the oil exporters, suffered no losses). If developing nations could not afford oil at $20 per barrel unless financed by petroloans, who is going to enter into new petroloans when oil is $60 or $70 per barrel while old petroloans are being written off? Certainly not the oil exporters!

The plight of the poor in developing nations is irrelevant to a solution. Their consuming less is not going to save the day for us. Demand destruction has to occur in the United States, the consumer of one-quarter of the world's oil production. Even though demand destruction has a terrible ring about it, we have to do it. We cannot rely on the rest of the world to cut oil demand in order for us to remain the most profligate consumer in the world. No one has any sympathy for us. India and

Figure 11.1 **U.S. Gasoline Price and Consumption, 2005**



China have teamed up to form a common front to coordinate their oil plans by not placing themselves in the position of bidding against one another for a scarce resource. In the future, they will both be bidding against us as a united front: It is 300 million of us versus 2.5 billion of them.

In 2004, 380 million Europeans consumed, on a per capita basis, an average of 450 gallons of gasoline and diesel fuel. Of course, this is not all automotive fuel, but also oil consumed by trucks, trains, and aircraft. The corresponding figure for the United States, with a reduction in gasoil consumption to reflect heating oil, was 690 gallons per person, or 50 percent higher. We have to ask ourselves, are we 50 percent happier than the Europeans? If not, then we can live with less per capita oil consumption without sacrificing our happiness. Those who reject the idea that the United States should cut its consumption of motor vehicle fuel retort that the United States is not Europe. We have longer distances between cities and more extensive rural areas. But about half of Americans live in urban areas with driving needs not much different than Europeans. If Europeans have a greater reliance on public transport, particularly fuel-efficient rail for commuting and traveling, that is their foresight and our lack thereof.

Demand destruction will only occur if the price of gasoline is raised. The question is, how much of a price rise is necessary to impact gasoline consumption? Figure 11.1 suggests that the run-up in gasoline prices in the summer of 2005 was not enough.[2] Oil consumption continued to climb throughout the summer, the peak season for gasoline consumption, despite higher gasoline prices. The interruption in consumption in late August was not caused by high gasoline prices, but the impact on domestic refining and imports from hurricanes Katrina and Rita. The falloff in the autumn with declining prices is normal after the peak summer driving season. With prices climbing along with consumption during the summer and prices falling along with consumption during the fall infers a positive relationship between the two. This confounds normal economic theory of a negative relationship whereby higher prices are supposed to depress volume. The real culprit for gasoline prices going up and down was not demand, but the changing price of crude oil.

The central message in Figure 11.1 is that gasoline consumption is not sensitive to prices below $3 per gallon. Yet there was an impact. Sales of gas-guzzling SUVs declined while hybrid car sales soared as gasoline prices increased. High gasoline prices put Ford and General Motors's strategy of near-term profit maximization by concentrating on large SUVs in jeopardy. In sharp contrast to Ford and General Motors's single-minded focus on the here and now, Honda and Toyota took a longer-term view and developed the hybrid. The business strategy of Ford and General Motors mirrors the American public's nonchalant concern over energy. The business strategy of Honda and Toyota reflects an entirely different mindset toward oil: respond to the market today, but keep an eye on the future, particularly the possibility of another oil crisis. It remains to be seen whether Ford and General Motors can survive their here-and-now approach and bring hybrid and other fuel-efficient automobiles to market fast enough to compete against Honda and Toyota. Even if Ford and General Motors succeed, they are up against two companies already firmly entrenched in this market. As a people, it is time for us to stop thinking like Ford and General Motors and start thinking like Honda and Toyota.

Internalizing an Externality

The price of gasoline covers the cost of extracting, processing, and distributing as evidenced by oil company profits. The price of gasoline covers the cost of building and maintaining highways as evidenced by highway trust funds. But the price of gasoline does not cover oil security—the hundreds of billions of dollars being spent to ensure that the Middle East remains a reliable and ready supply of oil. It is high time that this externality, now paid for by the taxpayer, be internalized and paid for by those who directly benefit. Apartment dwellers in city centers without an automobile should not be bearing this cost. Most taxpayers being also automobile owners must not sway the argument that those who benefit pay.

Europeans are willing to pay about $6–$7 per gallon for motor vehicle fuel and have adopted a way of life or, better, a way of travel conducive to high-priced fuel. Europeans want a high price on gasoline and diesel oil to discourage unnecessary travel in order to reduce pollution. While Europeans are amenable to high-priced motor vehicle fuels, Americans definitely are not. Yet, in order to bring about demand destruction, the price of gasoline and diesel fuel should be set at $4 per gallon, possibly with planned step-ups to $5 per gallon. This is still less than what Europeans pay and takes into consideration the longer distances between population areas and more extensive rural areas.

European motor vehicle fuel taxes are constant, which means that the price of gasoline and diesel fuel fluctuates with changes in the underlying price of crude oil. Rather than having a flat tax, it is possible to have a flat price for motor vehicle fuels. This can be accomplished by having the oil security tax fluctuate with the price of crude oil. With the oil security tax keyed to crude oil prices, the government earns more tax revenue when crude oil prices fall and earns less when they rise. This keeps the price of gasoline and diesel fuel more or less constant, assuring the public that there will be no further price hikes if crude oil prices should rise as long as prices remain below a prescribed cap. An oil security tax not only internalizes an externality, but would help to restore fiscal responsibility by going a long way toward balancing the national government budget that has been thrown out of whack by our military adventure in oil producing Iraq. It will also restore a bit of stability to the world oil situation by cutting demand.

The argument has been made that Congress would never permit a huge increase in the gasoline tax. Americans would be upset and would extract their revenge by ensuring that those who voted in favor of it were not reelected. This might well be true. But it might also be true that Americans would support an oil security gasoline tax if they saw it as part of an overall program to bring back the troops from the Middle East, never to return. A grassroots initiative might be necessary for Americans to let their representatives in Congress know that they would not be thrown out of office if they voted to impose an oil security tax as part of an overall program of getting our armed forces out of the Middle East.

A Plan to Get Out of the Middle East

Demand destruction is one phase of a plan for the United States to extricate itself from the political quagmire of the Middle East. We are vulnerable to regimes that are not friendly toward us, to put it mildly. Our continued presence in the Middle East only aggravates a bad situation. We are in danger of ending up not only occupying much of the Middle East, but also in coming into conflict with China and possibly India over access to Middle East crude oil. We have to get out or be willing to spill American blood for the privilege of maintaining our insouciant attitude toward oil consumption.

We cannot substitute other conventional sources of crude oil for Middle East supplies. The Atlantic basin (South America, Africa, and Europe, including Russia) does not produce enough oil to meet our needs. That is why we are dependent on Middle East supplies. We have to substitute nonconventional sources of crude oil for conventional sources to cut the umbilical cord to Middle East oil.

Dealing in round numbers, humanity has consumed 1 trillion barrels since Edwin Drake first discovered oil. The experts argue over whether another 1 or 2 trillion barrels are left, depending on assessments of undiscovered oil and the effectiveness of enhanced recovery technologies. Reserves of conventional oil are dwarfed by the estimated reserves of 7 trillion barrels of nonconventional oil. These reserves are located right here in the western hemisphere, about equally divided among the United States, Canada, and Venezuela.[3] Even if only a relatively small portion can be tapped for making synthetic crude, it will see us through the next few decades (if not longer) without importing a drop of Middle East crude. Bringing nonconventional sources of oil into production means that undiscovered oil offshore the United States and in Alaska can remain undiscovered—a potential strategic reserve for future generations.

In the United States, nonconventional oil takes the form of oil shale. Mining and processing oil shale present an unacceptable environmental impact in terms of waste disposal and water demand. However, in situ processing has been demonstrated to work and is ripe for further development. In situ mining is heating oil shale in place, which causes kerogen in the shale to flow to the surface for collecting and refining with minimal environmental impact. Further development should be pursued, but it will be another decade or more before this alternative can be fully exploited. In conjunction with developing domestic resources of oil shale, it may also be in the interests of the nation to inaugurate a TVA-like effort to develop and build coal to liquid plants similar to FuturGen, but with gasoline and diesel fuels as output rather than hydrogen. The hydroelectric plants built by the TVA have turned out to be profit-making enterprises that help to fulfill our electricity needs. In like manner, coal to liquid plants will bring the vast coal resources of the United States to help alleviate the oil import problem, which can be profitably employed at oil prices prevalent in 2005 and 2006.

Tar sand has been mined and processed in Canada for decades. Historically, tar sand was economically processed when crude oil was $15 a barrel. The subsequent rise in the price of natural gas may have pushed the cost closer to $30 per barrel, but tar sand is still economical considering the high price of crude oil. Tar sand production is being expanded, and further expansion should be encouraged. The drawback of Canadian tar sand production is that it consumes large quantities of water and natural gas. Natural gas is used to heat the tar sand for extraction when the ground is frozen. After transport to production facilities, natural gas heats water necessary to separate the tar from the sand. Finally, natural gas is a source of hydrogen, which transforms bitumen to an acceptable grade of synthetic crude that can be processed by a conventional refinery. Natural gas consumed in tar sand processing could also be collected and pipelined down to the United States, where there is a ready market. Moreover, there is concern over the adequacy of natural gas supplies for processing huge quantities of tar sand and the energy input (natural gas consumed) to output (syncrude produced) relationship.

This leaves Venezuela. Four syncrude plants in Venezuela process 600,000 bpd of Orinoco bitumen for export to be refined in conventional refineries. Whereas new syncrude plants are under construction in Canada, none are being built in Venezuela. In 2005, Venezuela unilaterally changed the tax and royalty status of existing plants. Venezuela might have had good reason for doing what it did, but the consequence was that it scared off investors.

From the point of view of attracting new investment, it is better if a nation's revenue from an oil project incorporates a tax and royalty formula that fulfills two objectives. One is that enough funds flow to a project to ensure that investors are receiving a fair and reasonable return on their investment, considering its inherent risk when oil prices are low. The other is that the host nation benefits to a far greater extent than investors when oil prices are high. A carefully structured tax and royalty regime reduces risk for the investors when oil prices are low while, at the same time, increasing the benefits to the host nation when oil prices are high. Investors can live with any tax and royalty plan—it is a matter of whether the cash flow is adequate to cover their investment. Investors cannot live with the uncertainty of changing tax and royalty rules because this makes cash-flow projections meaningless. Without a cash-flow projection, investors cannot judge what their return will be or whether there will be any return. The best thing for investors to do under these circumstances is to look elsewhere.

Venezuela has two advantages over Canada. The bitumen is not mixed with sand and frozen in the ground, making extraction easy. The bitumen is thinned with naphtha for pipeline transport to the syncrude plant where the naphtha is separated and pipelined back for recycling. Therefore, no natural gas is consumed heating bitumen for extraction or heating water to separate bitumen from the sand. Natural gas is used only for transforming bitumen to syncrude. Syncrude production monetizes reserves of stranded gas and bitumen that at present have no commercial value. Monetizing stranded natural gas and bitumen reserves can fund social services for Venezuela's 27 million people, many of whom live in poverty, and fund economic development to reduce the extent of poverty.

Developing syncrude projects in Venezuela serves the mutual benefit of both the United States and Venezuela. Venezuela gains by greatly increasing its revenues by monetizing its stranded gas and bitumen reserves. The United States gains by reducing its dependence on Middle East oil and the world gains in relieving the pressure on conventional oil supplies. If we can cut motor vehicle fuel consumption by 10 percent, or by 1,500,000 bpd, and increase synthetic crude production in Canada and Venezuela by about the same amount, we can essentially eliminate our dependence on Middle East imports (in 2004, Middle East exports to North America were close to 2,700,000 bpd). Once out of the Middle East, much of our foreign relation burdens can be left behind and our troops can be brought home without body bags. As a first step, we must improve our relations with Venezuela and change their perception of us from a hostile power to a partner in a venture that has mutual benefits. The advantage for the United States is a means of extricating itself from the Middle East. The advantage to Venezuela is a means to create a "Greater Venezuela."

Pollution

A drawback to synthetic crude being a substitute for Middle East oil imports is that gasoline made from synthetic crude emits more carbon dioxide than does natural crude. The reason has nothing to do with gasoline, but with the energy (natural gas) consumed in producing synthetic crude. However, consuming less gasoline via a high oil security tax partially compensates for higher carbon dioxide emissions of motor vehicle fuels made from syncrude. Moreover, a high security tax on diesel fuel may shift some of the freight that is currently moved by trucks to rail, which would cut congestion on interstate highways. The piggyback service (carrying truck trailers and containers on rail) offers significant energy savings as steel wheels on steel track are much more fuel-efficient than rubber wheels on concrete.

Pollution in the form of sulfur and nitrous oxides, mercury, ozone (smog), and others is clearly harmful to the health of people and should be reduced as a matter of principle. The case against carbon dioxide and other greenhouse gas emissions being responsible for global warming is less clear. The connection between the two has not been established with absolute certainty. Spending billions of dollars to fix something that may not be broken (cutting greenhouse gas emissions that ultimately are not to blame for global warming) precludes helping people in need of aid or undertaking economic development projects. On the other hand, doing nothing and letting the world bear the consequences of runaway global warming or slip into another ice age may not be the most brilliant move either.

We are left in a position of spending money for reducing greenhouse gas emissions to prevent global warming without a scientifically proven connection between the two. Statistical evidence of a link is not proof. There once was statistical evidence of a link between the whale population and stock market performance, but that is not proof that fecund whales necessarily create a bull market for stocks. We do not know with certainty the consequences of reducing greenhouse gas emissions or of doing nothing.

Blaise Pascal, a seventeenth-century mathematician with a strong religious bent, noted that individuals face uncertainty with regard to believing in the existence of God. Individuals have free will and can choose whether to accept or reject God. Pascal's solution was to treat the matter as a wager: either God does or does not exist. If we believe in God and we live a life conducive to salvation, and he does not exist, then at most we have lived a "foolish" life. Foolish in the sense that we could have done things that we chose not do in order to avoid explaining why we did what we did to a living God. If we do not believe in God and live a life of sin, and God exists, then we face eternal damnation. Pascal's point was that the consequences of being wrong were not comparable. The consequence of being wrong and assuming that God does not exist when he does, and living a life of sin, is eternal separation from God. The consequence of being wrong and assuming that God does exist when he does not, and living a life conducive to salvation, is a presumed loss of pleasure. Since there is no pleasure in this life even close to compensating for eternal damnation, Pascal concluded that individuals would be better off if they bet on the existence of God and lived their lives accordingly.

We face the same situation with carbon dioxide and other greenhouse gas emissions. The earth appears to be warming up (ignoring the harsh winter of 2005–2006 experienced by Europe, Russia, and northern Asia). It is not a foregone conclusion that the cause is a buildup in carbon dioxide in the atmosphere. Some believe it could be simply part of a natural cycle caused by wobbles in the earth's rotation around the sun or the sun radiating more energy or some other explanation unrelated to greenhouse gases. Of course, there is the possibility that burning fossil fuels, and the consequent carbon dioxide buildup in the atmosphere, has led to global warming. The consequences (changing weather patterns, more severe weather, flooding of coastal areas, runaway warming or the start of another ice age) of our doing nothing if carbon dioxide emissions are responsible for global warming exceed the waste of money of doing something if carbon dioxide emissions are not responsible for global warming.

On this variant of Pascal's wager, we should do something to reduce carbon dioxide emissions. Reducing motor vehicle fuel demand by placing a stiff tax on gasoline and diesel fuel consumption is one step toward cutting both pollution and greenhouse gas emissions. We should pursue efficiency and conservation to further reduce our energy needs, which, in turn, would cut pollution and greenhouse gas emissions. This can be done without affecting our quality of life, as Europeans and the Japanese have learned. The federal government should set up some sort of national carbon dioxide emissions program, which does not mean that we have to be a signatory to the Kyoto Protocol. A national program is preferable over having individual states set up a hodgepodge of separate programs. The fact that states are initiating carbon-emission programs shows public support for cutting greenhouse gas emissions. A carbon-emission trading program would be a financial shot in the arm for the development of wind and solar power projects because they could sell rights to emit greenhouse gases.

We should reinvigorate the nuclear power program. Nuclear power plants have no carbon dioxide emissions. In particular, we should investigate the claims of advocates of pebble-bed nuclear reactors with regard to their simplicity of design, their low cost of construction and operation, their inherent safety, and the ease of disposal of spent pebbles. Pebble-bed reactors cannot become supercritical and cause a disaster similar to the one that happened at Chernobyl because there are no rods to pull out. If these claims are true, then we should push ahead with their development, and if we are successful, begin series production of "cookie-cutter" plants of identical design. Pebble-bed reactors are relatively small in generating capacity and can be built throughout the nation to serve local communities, reducing the need for long-distance transmission lines.

Building nuclear power plants is sure to arouse the ire of environmentalists. Yet, why is it that Europeans and the Japanese, who are as environmentally conscious as anyone, permit the building of all types of electricity-generating plants, including those that are nuclear? All nations with a seacoast permit the building of wind farms in offshore water as a source of clean energy except the United States. Offshore wind farms cannot be built in the United States because environmentalists object to the tips of the blades being visible from shore, yacht owners object to having the view from their yacht spoiled, or the lack of assurance that not a single bird will ever be killed. The Europeans and Japanese have found a balance between the needs of energy providers and consumers and the need to protect the environment. We have not. We have not even recognized the possibility that the ultimate outcome of the "not in my backyard" syndrome is a real energy shortage and our becoming a second-rate economic power.

In addition to nuclear power plants, clean-burning coal plants should be built. Whereas pebble-bed nuclear power plants can be dispersed within population areas, coal-fired plants have to be built either near coal mines or on railroad tracks capable of handling high-volume coal shipments. Economies of scale dictate large coal-burning plants that require long-distance transmission lines. Coal-burning IGCC electricity plants should be favored in areas where sequestering carbon emissions is possible. Both nuclear power and coal-fired plants will reduce demand for natural gas for electricity generation. Cutting back on natural gas demand will soften prices for homeowners, businesses, industrial concerns, and utilities and will also prevent this nation from becoming overly dependent on foreign sources. Importing large volumes of LNG will only compound the problems that already exist with importing large volumes of oil. Moreover, stranded gas in the Middle East and Russia should not rely entirely on LNG as a way of monetizing an asset, but should also include gas-to-liquids (GTL) plants that produce a high-grade diesel fuel to further reduce demand on crude oil.

We should go to time-of-day electricity rates for all consumers. This gives homeowners and businesses an economic incentive to shift part of the electricity load from times of high demand to times of low demand. Homeowners can save money on their electricity bills by washing and drying clothes and dishes at night and on weekends. Businesses can shift some of their electricity demand to times of lower rates. Shifting demand from weekday to night and weekends increases the base load, which can be economically served by nuclear and coal-fired plants, and decreases the day load and the need for building more natural gas plants for handling the daytime peak load. Decreasing variability in electricity demand lowers average electricity rates.

Another advantage of time-of-day electricity rates is they enhance the economic incentive to install solar power. There are tens of millions of rooftops without solar arrays in the nation's sun belt. One cannot criticize homeowners for not installing solar arrays if the electricity savings barely recoup the cost of the installation over a thirty-year period. By coupling the falling price of solar arrays with a higher day rate for electricity, solar arrays become an investment whereby owners may earn a reasonable return on their investment in the form of reduced electricity bills. While a single array has little impact on the grand scheme of things, millions of residences with solar arrays producing electricity during times of peak demand significantly reduces the need to build new generating capacity, although backup generators will still have to be available to cover those times when the sun is not shining.

We Need Time

If we can sustain ourselves for a few decades using these various means, then we can give technology associated with the hydrogen economy time to overcome its current technical and economic challenges. Time lets us look at synthetic fuels from coal, biomass, and solid waste, at alternative fuels for automobiles such as ethanol, not made from corn, but from agricultural wastes and nonagricultural plants. Time allows further development of sustainable sources of electricity (wind and solar). Time gives us an opportunity to develop advanced storage batteries for electricity and hydrogen that would place electricity generation on a near-constant base load, eliminating variability and driving electricity rates down to their lowest possible level. Time may make it possible to find a way to mine methane hydrates and build thorium breeder reactors, superconducting transmission lines, and nuclear fusion plants. Maybe the challenges posed by these technologies can be overcome, maybe not, but we need the time. Cutting consumption by placing an oil security tax on motor vehicle fuels and developing the vast resources of nonconventional oil, pursuing the nuclear and coal options, going to time-of-day electricity rates, and pushing sustainable sources of energy (wind and solar) gives us that time.

We have another choice: Do nothing. Keep the price of gas low, keep the gas-guzzlers on the road, and watch traffic congestion and oil demand continue their relentless upward spiral. We face the real possibility of $100 per barrel oil ($5–$6 per gallon gasoline) when demand is pushed beyond the limits of supply and/or when turmoil breaks out in politically unstable oil-exporting nations. Now the proceeds of high-priced gasoline do not flow to our government, but to oil exporters. We will need to have a continuing presence in the Middle East to ensure oil security. As we hog world supplies, we risk conflict with China, and possibly India, as they stake their claims to a dwindling resource.

No matter what choice we make, we face higher gasoline prices. If we raise gasoline taxes now to cover the cost of oil security, demand destruction will occur and our government will benefit; or we can wait and let circumstances raise the price of crude oil and allow others to benefit. We cannot pass the consequences of our inaction to future generations. Those who are alive today will still be alive when the consequences of what we do not do come home to roost.

Notes

1. *World Energy Outlook* (Paris: International Energy Agency, 2002 and 2004) provides one of the best outlooks for energy available. No criticism is being leveled at this organization; the problem is with the forecasting process itself, not with those who face the challenge of attempting to forecast the future.

2. Gasoline consumption is derived from domestic refinery supply plus imports adjusted for changes in inventory. The falloff in apparent gasoline consumption during hurricanes Katrina and Rita was caused by refinery shutdowns in the United States Gulf plus interruptions to shipments of oil product imports. For more information, visit the U.S. Department of Energy Web site at www.eia.doe.gov.

3. The 2004 edition of *World Energy Outlook* takes a strong stand on the necessity to develop nonconventional sources of crude oil to meet future demand.

Index

About the Author

Roy L. Nersesian has spent most of his professional career observing the energy industry as part of his involvement in energy transportation. He is an adjunct professor at Columbia University for the Center for Energy and Marine Transportation in the School of International Public Affairs and a professor at the School of Business at Monmouth University. He brings a fresh perspective of the energy business—where it was and where it is heading.


We Trip The Light
FANTASTIC