World Oil: Market or Mayhem?

James L. Smith

Energy analysts sometimes speak of oil’s “golden era,” that 100-year stretch between 1874 and 1974 when the real price was relatively stable within a range from $10 to $20 per barrel (BP, 2008) in 2007 dollars. Figure 1 shows that in recent decades, that stability has ended. In October 1973, several Arabic members of the Organization of the Petroleum Exporting Countries (OPEC) announced that in response to U.S. support for Israel during the 1973 Arab–Israeli war, they would place an embargo on oil exports to the United States. That action caused real oil prices to soar from $12 to $53 per barrel within four months. Later in the 1970s, political turmoil in Iran and the Iran–Iraq war again rattled the market and by January 1981 pushed the real price up to $95. Eventually, oil prices fell back to earth with a thud, bottoming out at $21 per barrel in July 1986. The roller coaster ride of prices has continued more recently. After oil prices skidded to a low of $12 per barrel in December 1998 in the wake of the Asian financial crisis, oil stabilized again around $30 during 2000–2004 before a breathtaking ascent that reached $145 per barrel by July 2008—only to dip below $40 per barrel again before the end of 2008.

A unique combination of economic circumstances surrounds oil markets. A short list would include extremely high price volatility; the prominent role and unusual longevity of a major cartel (OPEC); the absolute size and scope of the oil industry and its important links to industrialization, economic growth, and the global distribution of wealth; nagging doubts about the sustainability of the resource base; substantial volumes of petroleum-related CO₂ emissions that pull oil

---

James L. Smith is the Cary M. Maguire Chair in Oil and Gas Management, Cox School of Business, Southern Methodist University, Dallas, Texas. His e-mail address is jsmith@cox.smu.edu.
into the center of the climate change debate; plus a host of tricky geopolitical issues that reflect the uneven distribution of oil deposits around the globe. Because excellent and detailed analyses are available elsewhere,1 we begin here with some background on the oil industry and then focus on a few key questions that have sparked recent controversy. Taking a perspective over the last few decades, what supply and demand forces can help to explain movements in oil prices? Taking a more recent perspective, why did oil prices spike in 2008, and what role (if any) did speculators play? Finally, what is the long-term outlook for the price of oil—and how concerned should we be about whether the world is passing its “peak oil” level of production?

Some Background

For many years, oil exports (crude oil plus refined products) have been the leading commodity in world trade—comprising 13 percent of total commodity

---

1 Adelman (1972) produced the first comprehensive study of the modern oil market, just before OPEC took center stage. More recent studies include Kaufmann, Dees, Karadaloglu, and Sanchez (2004), Smith (2005), and Gately (2007).
trade by value in 2006, some $4 billion per day (United Nations, 2008). Automobile exports amount to only about one-third as much ($1.5 billion/day), and iron and steel about one-quarter as much ($1 billion/day). Nearly all nations are significantly affected by developments in oil markets, either as producers or consumers—or both. At least 50 countries produce substantial volumes of oil, and two-thirds of total production is exported (BP, 2008). Countries of the Middle East, the former Soviet Union, and Africa account for the bulk of exports, whereas the United States, Europe, China, and Japan account for nearly all of the imports. There are many grades of crude oil, but they all compete in a highly integrated world marketplace with price differentials that reflect the relative desirability of grades (Bentzen, 2007).

The world oil market has changed profoundly in the last few decades, but a few outdated conceptions persist. For example, a common misconception is that world oil is dominated by a handful of multinational corporations. Forty years ago in 1969, before a wave of nationalizations reshaped the industry, the eight largest oil companies did produce 89 percent of the world’s oil (Adelman, 1972); today, those same companies account for just 12 percent of production and only 3 percent of the world’s remaining proved oil reserves (Petroleum Intelligence Weekly, 2008).2 “Big Oil” now consists of the state-owned companies of the major exporting nations, who account for about 50 percent of global output, control 70 percent of proved reserves, and operate under sovereign power beyond the reach of antitrust or regulatory authorities. Some of these national oil companies are affiliated with OPEC, some are not. Operations of the 20 largest oil producers are summarized in Table 1; the companies affiliated with OPEC are in bold type. The “proved reserves” shown in the final column of the table refer to that portion of known oil deposits that can be economically extracted at prevailing prices using available technology. Most of the oil in any given deposit will never be produced, and therefore does not count as proved reserves, because it would be too costly to effect complete recovery.

Very little oil leaves the merchant-controlled supply chain that extends from producer’s well to consumer’s fuel tank. Oil producers sell and convey their output to refiners, who in turn sell their refined products to wholesale and retail marketers. Although vertical integration has declined somewhat during the past two decades, the volume of oil refined by the 20 largest oil producers amounts to 77 percent of their crude oil production. Companies do not necessarily refine their own crude oil, however. Most oil producers have exchange agreements with other producers by which crude oil streams are swapped to minimize transportation and processing costs. In other words, I may process your oil if my refinery happens to be close to the site of your well, and vice versa.

Oil producers and refiners also trade physical cargos of crude oil in a marketplace. Spot trades involve cargos scheduled for immediate delivery, whereas for-

2 The original eight included Esso, British Petroleum, Shell, Gulf, Texaco, Standard Oil (California), Mobil, and Compagnie Française des Pétrole. Through various mergers and consolidation, the eight have been reduced to five: ExxonMobil, BP, Shell, Chevron, and Total.
ward trades specify a fixed delivery date within the next few months. Sometimes these trades proceed directly between oil companies, but often the crude oil is routed through intermediaries (so-called “trading companies”) who provide the managerial acumen and industry contacts to charter tankers, negotiate port fees, and generally arrange for the physical transfer of oil. In the market for Brent (UK) crude oil, the world’s largest oil market, these trading companies serve as middlemen in roughly five out of every six transactions (Weiner, 2006). In contrast, financial traders like hedge funds, banks, commodity index funds, and others play no role in the physical market for crude oil. Their trading activities are confined to the futures market, where “paper” barrels are traded. The links between physical trading, spot prices, and the futures market are discussed later in this paper. Commercial inventories of crude oil are large and mostly owned by the companies that produce, refine, or market oil (U.S. Energy Information Administration, 2008c). Large strategic stockpiles of crude oil are held by governments, but these are seldom tapped.

### Table 1

**Twenty Largest Oil Companies, Ranked by Production, 2007**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Company</th>
<th>Country</th>
<th>State ownership</th>
<th>Production (thousand barrels/day)</th>
<th>Proved reserves (million barrels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Saudi Aramco</td>
<td>Saudi Arabia</td>
<td>100</td>
<td>10,413</td>
<td>264,200</td>
</tr>
<tr>
<td>2</td>
<td>NIOC</td>
<td>Iran</td>
<td>100</td>
<td>4,401</td>
<td>138,400</td>
</tr>
<tr>
<td>3</td>
<td>Pemex</td>
<td>Mexico</td>
<td>100</td>
<td>3,474</td>
<td>12,187</td>
</tr>
<tr>
<td>4</td>
<td>CNPC</td>
<td>China</td>
<td>100</td>
<td>2,764</td>
<td>22,447</td>
</tr>
<tr>
<td>5</td>
<td>Exxon Mobil</td>
<td>US</td>
<td></td>
<td>2,616</td>
<td>11,074</td>
</tr>
<tr>
<td>6</td>
<td>KPC</td>
<td>Kuwait</td>
<td>100</td>
<td>2,600</td>
<td>101,500</td>
</tr>
<tr>
<td>7</td>
<td>PDV</td>
<td>Venezuela</td>
<td>100</td>
<td>2,570</td>
<td>99,377</td>
</tr>
<tr>
<td>8</td>
<td>BP</td>
<td>UK</td>
<td></td>
<td>2,414</td>
<td>10,073</td>
</tr>
<tr>
<td>9</td>
<td>INOC</td>
<td>Iraq</td>
<td>100</td>
<td>2,145</td>
<td>115,000</td>
</tr>
<tr>
<td>10</td>
<td>Rosneft</td>
<td>Russia</td>
<td>75.16</td>
<td>2,027</td>
<td>17,513</td>
</tr>
<tr>
<td>11</td>
<td>Petrobras</td>
<td>Brazil</td>
<td>32.2</td>
<td>1,918</td>
<td>9,581</td>
</tr>
<tr>
<td>12</td>
<td>Shell</td>
<td>UK/Netherlands</td>
<td></td>
<td>1,899</td>
<td>4,887</td>
</tr>
<tr>
<td>13</td>
<td>Sonatrach</td>
<td>Algeria</td>
<td>100</td>
<td>1,860</td>
<td>11,400</td>
</tr>
<tr>
<td>14</td>
<td>Chevron</td>
<td>US</td>
<td></td>
<td>1,783</td>
<td>7,523</td>
</tr>
<tr>
<td>15</td>
<td>ConocoPhillips</td>
<td>US</td>
<td></td>
<td>1,644</td>
<td>6,541</td>
</tr>
<tr>
<td>16</td>
<td>Adnoc</td>
<td>UAE</td>
<td>100</td>
<td>1,574</td>
<td>52,800</td>
</tr>
<tr>
<td>17</td>
<td>Lukoil</td>
<td>Russia</td>
<td></td>
<td>1,552</td>
<td>12,572</td>
</tr>
<tr>
<td>18</td>
<td>Total</td>
<td>France</td>
<td></td>
<td>1,509</td>
<td>5,778</td>
</tr>
<tr>
<td>19</td>
<td>NNPC</td>
<td>Nigeria</td>
<td>100</td>
<td>1,414</td>
<td>21,700</td>
</tr>
<tr>
<td>20</td>
<td>Libya NOC</td>
<td>Libya</td>
<td>100</td>
<td>1,368</td>
<td>30,700</td>
</tr>
</tbody>
</table>


Note: Affiliates of OPEC members appear in bold type. The “proved reserves” shown in the final column of the table refers to that portion of known oil deposits that can be economically extracted at prevailing prices using available technology. Most of the oil in any given deposit will never be produced and therefore does not count as proved reserves because it would be too costly to effect complete recovery.
Demand and Supply Shifts in the Oil Market in Recent Decades

The erratic price trajectory in oil markets in recent decades can largely be explained by demand and supply curves. Although price movements are of course determined by changes in both supply and demand, the separate impact of these two factors can be disentangled. In Figure 2, for example, equilibrium moves from \((p_1, q_1)\) to \((p_2, q_2)\) when demand and supply are perturbed. The shift in demand, which is related to the combined effects of income, population growth, and other factors, is measured by the increase (holding price constant) from \(q_1\) to \(q_2\), or in percentage terms by \(\frac{q_2}{q_1}\). Although \(q_2\) is not observable, its value can be deduced based on an estimate of the elasticity of demand. If the elasticity of demand is a constant, then one can imagine a family of demand curves, with only one demand curve going through any particular point. Similarly, one can imagine a family of supply curves with a single elasticity, with only one supply curve going through any given point, and shifts in supply due to the standard factors of resource depletion, technological innovation, and cost inflation. In this framework, any shift from one equilibrium combination of price and quantity to another equilibrium can be disentangled into a shift from one demand curve to another and from one supply curve to another. Given reasonable estimates of supply and demand elasticities, this method allows one to approximately identify actual shifts in supply and demand curves based on observed prices and quantities.\(^3\)

More specifically, the difference between \(q_2\) and \(q_1\) represents a movement along the demand curve: \(\frac{q_2}{q_1} \approx \left(\frac{p_1}{p_2}\right)^{\eta_d}\). This approximation is good for small price changes, and exact for all price changes if the elasticity of demand is constant. Making this substitution allows us to identify the demand shift:
Empirical estimates of the price elasticity of demand for crude oil vary by place, time, and statistical technique (for example, Gately and Huntington, 2002; Cooper, 2003; U.S. Energy Information Administration, 2007; OECD, 2004). Estimates of $-0.05$ (short-run) and $-0.30$ (long-run) are typical, with several years required to complete the adjustment to a permanent price change. It is more difficult to produce current and reliable estimates of the elasticity of crude oil supply, due in part to confounding effects of resource depletion and technical innovation, but there is consensus that the supply of conventional oil is quite inelastic, especially in the short run. OECD (2004) reports elasticities ranging between $0.04$ (short-run) and $0.35$ (long-run). The U.S. Energy Information Administration, or U.S. EIA, (2007) and System Sciences Inc. (1985) uses even more inelastic supply curves—$0.02$ for short-run and $0.10$ for long-run—to forecast production from most regions in its international oil supply model.

In thinking about demand and supply shifts in the oil industry in recent decades, let’s assume the long-term elasticity of global demand to be $-0.3$ and the long-term elasticity of non-OPEC supply to be $0.3$, and assume further that quantities adjust to the three-year moving average price level. Figure 3 plots the sequence of implied annual shifts (relative to 1975) in global demand and non-OPEC supply of crude oil since 1973—holding real prices constant at the average 1973–75 level. The figure also shows changes in the actual quantity of oil produced by OPEC, as reported by U.S. EIA (2008a). Because OPEC acts as a price-maker, fluctuations in OPEC production can be interpreted as an attempt to influence the price, as we discuss shortly. The results shown in Figure 3 and the conclusions that follow are quite robust with respect to plausible alternative estimates of the elasticities and lag structure.

Several features stand out. First, over the long haul, demand has outrun non-OPEC supply. Although demand growth has been irregular (and sometimes persistently negative), global demand for crude oil (holding price constant) has increased by 80 percent overall since 1975, whereas actual OPEC production and non-OPEC supply have each grown by just 24 percent.

Second, during the 1980s, while global demand for oil was shrinking, the supply of non-OPEC oil was expanding robustly, which put substantial downward pressure on the market price and on OPEC producers, who cut output by nearly half between 1979 and 1985. Indeed, growth in OPEC output has been meager over the entire period since 1973. After the steep decline of the 1980s, OPEC production was not fully restored until 2004. OPEC’s production restraint represents a commercial choice, not a geological ultimatum or a reflection of high marginal costs—a point to which we will return later in the paper.

Third, until quite recently, resource depletion had no significant effect upon supply. For 30 years, until roughly 2003, non-OPEC producers were able to offset depletion of known reserves via exploration and technological innovation, and man-

\[
\Delta_D = \left(\frac{q_2}{q_1}\right) \times \left(\frac{p_1}{p_2}\right)^{a}. \]

By similar means, the underlying shift in supply can be inferred from the presumed elasticity of supply: \[
\Delta_S = \left(\frac{q_2}{q_1}\right) \times \left(\frac{p_1}{p_2}\right)^{s}. \]
-aged to increase supply in concert with demand. Moreover, a significant portion of the decrease in supply that occurred after 2003 resulted from rapidly escalating factor prices (like cost of pipe and drilling rigs) rather than resource depletion. Despite the decrease in non-OPEC supply since about 2003, the marginal cost of non-OPEC production remains lower today (and supply higher) than in 1975.

Finally, the steep ascent in the price of oil between 2004 and 2008 coincided with the first significant decrease in non-OPEC supply since 1973 and an unprecedented surge in demand. From 2004 to 2008, global demand increased by 33 percent, while non-OPEC supply decreased by 23 percent. Although OPEC members responded by increasing their production, they lacked sufficient capacity (after years of restrained oil field investments) to bridge the growing gap between global demand and non-OPEC supply.

What has determined the quantity of oil supplied from the OPEC nations, the cartel that controls 70 percent of global oil reserves and includes eleven of the 15 largest oil-exporting countries in the world?\(^4\) OPEC members meet regularly to

---

\(^4\) The eleven include (ordered by export volume): Saudi Arabia, United Arab Emirates, Iran, Kuwait, Venezuela, Nigeria, Algeria, Libya, Iraq, Angola, and Qatar (EIA, 2008e). OPEC’s twelfth member, Ecuador, produces and exports relatively little crude oil. Former OPEC member Indonesia suspended its membership at the end of 2008 because its production had declined and it no longer qualifies as a net exporter of oil.
“coordinate their oil production policies in order to help stabilize the oil market and to help oil producers achieve a reasonable rate of return on their investments” (OPEC, 2008b). In other words, OPEC’s goal is to set the price, and members synchronize production levels in pursuit of that goal (Smith, 2005).

Here, we focus on the two major pieces of OPEC’s strategy for “stabilizing” prices: 1) “shutting in” existing production capacity, which means extracting less oil than existing wells can produce; and 2) restricting the growth of new capacity by limiting the effort to find and develop new resources. OPEC has mostly failed at the former, but succeeded at the latter. Unfortunately, consumers have suffered from OPEC’s failure as well as its success: its failure to manage installed capacity has increased price volatility, while its success in restricting capacity growth has driven up the average price level.

At first glance, OPEC’s track record for shutting in production may appear successful in raising prices. After gaining control of production by nationalizing oil reserves, OPEC recorded a quick success in raising prices in 1973 by threatening an embargo and cutting output. Shut-in capacity nearly tripled between 1973 and 1975, and the real price of oil tripled as well (U.S. EIA, 2008b). That action was not sustained, however, and oil prices soon began to fall back. Since then, there has been no comparable demonstration of OPEC’s ability to hold production off the market. The period from 1979 to 1983 (the only other time that prices and shut-in capacity both rose sharply) might seem to qualify, but those events were not the purposeful result of OPEC’s strategy. Rather, they were caused by the Iranian revolution and the outbreak of war between Iran and Iraq, which disrupted oil operations and kept nearly 6 million barrels per day off the market. Indeed, much of OPEC’s shut-in capacity has been less a calculated business strategy than the involuntary result of extraneous developments, including war, international trade sanctions, labor strife, and sabotage. OPEC’s inability to have its members shut in production—and thus to hold excess production capacity—is reflected in the members’ lack of compliance with assigned production quotas, which is of course consistent with economists’ understanding of the difficulties of maintaining collusion when each individual party has an incentive to free-ride. Since the quota system was adopted in 1983, total OPEC production has exceeded the agreed ceiling by 4 percent on average, but on numerous occasions the excess has run to 15 percent or more. In Smith (2008), I describe the origin and operation of OPEC’s system of production allocations along with the record of compliance.

Price-fixing, according to Adelman (2002), is like singing and mountain climbing: easier going up than coming down. OPEC learned this lesson the hard way in the 1980s and 1990s, and it now knows that once capacity is built it is likely to be used, whether or not to the cartel’s advantage. From OPEC’s view, better that demand outrun supply than supply outrun demand because the latter exposes OPEC’s weakness—managing excess capacity.

OPEC’s decision to limit oil production by avoiding new capacity, rather than by holding existing capacity off the market, is shown by the fact that its crude oil production capacity (34 million barrels per day) is virtually unchanged from 1973
Although the volume of its proved reserves—that is, known deposits that could have been tapped to expand capacity—doubled over that span (BP, 2008). Non-OPEC producers, working mostly in more-expensive and less-prolific petroleum areas, have expanded their production capacity by 69 percent since 1973. OPEC’s installed production facilities are sufficient to extract just 1.5 percent of its proved reserves per year, which is another way of measuring the low intensity of development. Non-OPEC producers have installed facilities sufficient to extract 5.6 percent of their proved reserves each year (BP, 2008). OPEC accounted for only 10 percent of the petroleum industry’s upstream capital investment during the past decade (Sandrea, 2006), although it produced nearly half of global output. By holding back, OPEC has effectively allowed secular growth in demand to absorb and eliminate its excess capacity, ceding market share to non-OPEC producers in the process. OPEC apparently reckoned that the risk of expanding low-cost capacity within the cartel exceeded the potential harm from expansion of high-cost capacity outside the cartel.

OPEC recently initiated numerous projects to tap its underdeveloped reserves and finally expand capacity, investments that would amount to some $40 billion per year between 2008 and 2012 (OPEC, 2008a). That effort pales in comparison to the five largest international oil companies (the “super-majors”), who collectively own just 3 percent of global oil reserves but according to SEC filings spent about $75 billion during 2007 to develop new production capacity. OPEC, with 20 times the reserves, spends only about half as much. OPEC’s restraint is also reflected in the upstream plowback rate: in 2007, the super-majors reinvested 25 percent of their gross production revenues to expand capacity, whereas OPEC members are investing only about 6 percent of their net export revenues on such projects (U.S. EIA, 2008d).

One may ask whether OPEC’s restraint in developing new production capacity has gone too far for its own good. Some insight is gained by comparing the marginal cost of new capacity with the marginal revenue of additional sales. The marginal costs of oil production for OPEC members are very low, probably no higher than $5 per barrel for Saudi Arabia and $10 elsewhere in the Middle East (OECD, 2004; International Energy Agency, 2005; Adelman and Watkins, 2008). The marginal revenue of OPEC producers is much higher. Hamilton (2009) estimates the marginal revenue of an extra barrel produced by Saudi Arabia to be about 50 percent of the sales price. Unless the market price is expected to fall below $10 per barrel (which would push marginal revenue below $5), increased Saudi production would generate increased profit—for the Saudis, that is. But an additional barrel of Saudi oil would depress the price for all producers, and so the marginal revenue for the cartel as a whole (after accounting for the reduced revenues of other members) would probably be closer to 20 percent of the price. From the cartel’s perspective, therefore, increased production would not be justi-
fied unless the price were expected to remain above $50 per barrel (MR = $50 × .20 = $10 = MC).

This comparison illustrates the internal conflict that plagues all cartels. The individual incentives of OPEC members are not aligned with the common interest. The fact that OPEC (including Saudi Arabia) refrained from building capacity during 1973–2004, when oil prices stayed mostly below $50, is an important indicator of cartel discipline: individual incentives were seemingly set aside to support the common good. Not until the price surpassed $50 per barrel around 2004 did OPEC finally begin to create incremental capacity—a long delayed change, but one that appears consistent with the common good of the cartel members. After the price of oil fell below $50 in late 2008, most OPEC members announced their intention to postpone these investments, an action that—if carried out—again appears consistent with the common good of the cartel members.

Spikes and Speculators

From about 2004 up to mid-2008, the price of crude oil rose very sharply, from $34/barrel in January 2004 to a peak at $145/barrel on July 3, 2008, an all-time high. Then the price of oil sagged abruptly, falling once again to $34/barrel by January 2009. Can this kind of dramatic price spike be explained by shifts in demand and supply curves that are highly inelastic in the short run? Or is there reason to suspect either conscious manipulation of the oil market, or a run-up in prices fueled by speculative trading in futures and spot markets?

Annual Price Volatility from Demand and Supply Shifts

Perhaps a useful starting point is to observe that, while 2008 offers an extraordinary large price swing, volatility in oil prices is ordinarily quite high. If we denote by $x$ the percentage price change from one year to the next, then annual volatility is the standard deviation of $x$ over a series of years. The annual volatility of crude oil prices is high: 31 percent when calculated over the “modern” era (1974–2007) using the BP (2008) annual price series.\(^6\) For comparison, annual volatility averaged only 20 percent during the golden age of oil from 1874–1973. Regnier (2007), who provides volatility estimates for many products, finds that oil is now more volatile than 95 percent of all products sold in the United States.

These annual volatilities are so high because the underlying demand and supply curves are so inelastic. Demand is inelastic due to long lead times for altering the stock of fuel-consuming equipment; supply is inelastic in the short-run because it takes time to augment the productive capacity of oil fields. Price volatility provides incentives to hold inventories, but since inventories are costly, they are not

---

\(^6\) If we make the strong but convenient assumption that annual price changes follow a normal distribution, then the chance of a fluctuation beyond the stated volatility (31 percent) over the course of a year is determined (from the z-table) to be roughly 1 in 3.
sufficient to fully offset the rigidity of demand and supply. This fact means that shocks to demand or to supply can help to explain the high level of volatility in oil prices.

Surely one of the biggest shocks that started the run-up in oil prices in the early 2000s was the sharp rise in demand for oil from China and other developing nations, driven by rapid economic growth in those countries. In general, income elasticities of demand for crude oil appear to vary significantly by stage of economic development. Gately and Huntington (2002) and OECD (2004) report a nearly proportional relationship between income and oil demand in the developing countries ($\epsilon_d \approx 0.70–1.00$), whereas a weaker relationship ($\epsilon_d \approx 0.40–0.50$) seems to hold in the industrialized world. Another important shock, this one with a significantly adverse impact on oil supply, was the increase in the cost of oil production due to rapidly rising prices of steel pipe, drilling rigs, engineering services, cement, and the like that occurred between 2004 and mid-2008. Thus, some combination of unexpectedly energy-hungry growth from China and elsewhere in the world together with a negative shift in oil supply caused by higher production costs can explain a substantial rise in oil prices after about 2004.

Estimating the appropriate equilibrium price of oil in response to shocks is an imprecise calling, but if oil prices had not risen past $90–$100/barrel, which is roughly where matters stood in January 2008, these explanations might well feel sufficient to many experts in the field. But the price of oil kept on rising, by an additional 50 percent between January and July 2008. By early 2008, it seems like the impact of China’s growth and higher supply prices should already have been incorporated into prices. Therefore, in the first half of 2008 even seasoned oil market analysts found themselves asking whether some additional factors were needed to explain this final spike in oil prices.

Short-term Inelasticities of Supply and Demand

One potential answer is that in the short-run, price elasticities of supply and demand are extremely small, so that even seemingly small shocks may have large effects. To put it another way, the annual volatilities of oil prices discussed previously are sure to understate volatility (and price movements) as measured on a monthly basis.

For example, consider a supply shock that takes 1 million barrels per day out of the world market, a loss equal to roughly 1.25 percent of total 2008 output. If the price were fixed, this shock would create excess demand of 1.25 percent. To restore equilibrium in the short run, the price must rise by $\delta$ percent until the reduction in quantity demanded and the increase in quantity supplied (from sources not affected by the shock) combine to eliminate the shortage. By definition of elasticity, the additional quantity supplied would, in percentage terms, amount to $\delta \times \epsilon_s$, and the decrease in consumption would be $\delta \times \epsilon_d$. Since these must sum to 1.25, the price must rise by $1.25/(\epsilon_s - \epsilon_d)$ percent. When demand and supply are both highly inelastic, the elasticities combine (inversely) to create a large multiplier effect. Since both short-run elasticities are roughly 0.05 in magnitude (and the
price elasticity of demand is negative), each physical shock should trigger a short-run price adjustment about ten times as large.

Thus, on August 6, 2006, BP unexpectedly announced that the Trans-Alaska Pipeline would be closed immediately to fix corrosion and leakage problems. That action unexpectedly removed 400,000 barrels per day (0.47 percent) of total world supply for an indefinite period. Based on the multiplier described above, we would expect the price to jump by roughly 4.7 percent. In fact, the spot price rose the next day by just 3 percent. The difference between predicted and actual results may reflect in part the roughness of the calculation, and also the fact that the U.S. government announced before the close of trading that it would monitor the situation and consider releasing oil from the Strategic Petroleum Reserve to dampen the impact (CNNMoney.com, 2006). But the pattern of what seems like a relatively small shock causing a substantial rise in world oil prices is clear.

Can short-term shocks of this sort help to explain the spike in oil prices in the first half of 2008? It was definitely a time of significant upheavals, some with the potential for sustained disruption of supplies (U.S. EIA, 2008f). In February 2008, Venezuela cut off oil sales to ExxonMobil during a legal battle over nationalization of the company’s properties there. Production from Iraqi oil fields, of course, had still not recovered from the wartime damage, and in late March 2008 saboteurs blew up the two main oil export pipelines in the south—cutting about 300,000 barrels per day from Iraqi exports. On April 25, 2008, Nigerian union workers went out on strike, which caused ExxonMobil to shut-in oil production of 780,000 barrels per day from three fields. Two days later, on April 27, Scottish oil workers walked off the job, leading to closure of the North Forties pipeline that carries about half of the United Kingdom’s North Sea oil production. As of May 1, about 1.36 million barrels per day of Nigerian production was shut-in due to a combination of militant attacks on oil facilities, sabotage, and labor strife. At the same time, it was reported that Mexican oil exports (tenth largest in the world) had fallen sharply in April due to rapid decline in the country’s massive Cantarell oil field. On June 19, militant attacks in Nigeria caused Shell to shut-in an additional 225,000 barrels per day of production. On June 20, just days before oil reached its historic peak, Nigerian protestors blew up a pipeline that forced Chevron to shut-in 125,000 barrels per day of its production. Each of these events clearly registered in the spot market. It is not implausible to believe that, arriving in quick succession, they contributed heavily to the rapid acceleration in the spot price of oil.

Financial Speculation and Current Oil Prices

Although the rising price trend of 2004–2008 is consistent with changes in market fundamentals—surging demand and falling supply—the spectacular ascent especially in the first half of 2008 created widespread suspicion that “speculators” were responsible (Talley and Meyer, 2008; Commodity Futures Trading Commission, 2008). How might speculators in the oil market affect oil prices paid by consumers?

Most advocates of the speculative trading hypothesis focus on the rapid rise of
financial traders like hedge funds, pensions, commodity index funds, and other investors with no apparent connection to the oil industry. Büyüksahin, Haigh, Harris, Overdahl, and Robe (2008) report the market share of financial traders rose from 33 percent in 2004 to 50 percent by 2008. Conversely, the share of traditional commercial traders like oil producers, refiners, and wholesalers fell during that span from 31 to 15 percent. “Swap dealers,” who provide trades that cater to the needs of commercial entities, make up the balance.

The commercial traders referred to above are companies whose business operations are naturally exposed to the risk of oil price fluctuations. Some of these companies hedge their business risk by taking offsetting positions in the futures market; for example, an airline sells tickets and also uses futures markets to lock in the price of the fuel it will use in the future. Financial traders lack natural exposure to oil price movements, have nothing to offset, and no reason to hedge.

In practice, however, the distinction between financial and commercial traders is less clear cut. For one thing, not all financial trading represents speculation. Managers of commodity index funds, for example, play a role that is essentially passive; their goal being to assemble a portfolio of commodities that allows outside investors to participate in the average return experienced by a particular asset class. Conversely, many commercial traders do speculate, if only by selectively hedging their risks based on their view of future price movements (Stulz, 1996). For example, some oil producers elect to hedge their sales revenue only when they suspect that prices might fall, and their decision not to hedge when they suspect prices will rise amounts to speculation.

The widespread belief that the futures market, bloated by speculative trading, dwarfs the size of the underlying physical market for oil is another misunderstanding. This misconception arises when the volume of futures contracts that relate to oil deliveries that extend over many months is compared to the flow of oil production on a given day. After reconciling dates, even on the busiest trading days during 2007 and 2008, the volume of futures contracts for delivery of oil in any given month was but a fraction of the underlying physical production of oil (Ripple, 2008).

In the end, the distinction between hedging and speculative trading in the futures market is not important because neither one exerts any significant effect on current oil prices. A futures contract is a derivative; its price is derived from the price of oil sold in the underlying spot market. When I buy a crude oil futures contract today priced at $50, I commit to take delivery of one barrel of crude oil when the contract expires, say in one month, in exchange for the $50. If, at the expiration date, oil is selling in the spot market for $60, I gain $10 since I can take delivery and immediately resell my oil at the higher price. The party who sold me the contract (the futures exchange acts only as an intermediary) loses $10, since that party receives $10 less than the oil is worth. If the spot price had turned out to be $45, the tables would be turned and my $5 loss becomes the other party's gain.

By design, the futures contract references the same commodity and location as the underlying spot contract; thus, the futures price must converge to the spot price
as the expiration date draws near. This fundamental principle of convergence is clearly reflected in actual price movements. From January 1986 (when these data are first available) to date, the daily correlation between the spot price and the nearest-dated futures price (of West Texas Intermediate crude oil, as reported by U.S. EIA, 2008a) is 99.99 percent in terms of levels and 93.57 percent in terms of daily price changes. In other words, an expiring futures contract does not stray from the spot price that is required to equilibrate supply and demand in the physical market.7

Most oil futures contracts trade on either the New York Mercantile Exchange (NYMEX) or the Intercontinental Exchange (ICE). The standard NYMEX contract requires physical delivery of West Texas Intermediate crude oil (or its equivalent) to a terminal located at Cushing, Oklahoma. The ICE contract requires delivery of North Sea Brent crude oil (or its equivalent) to the Sullom Voe terminal in Scotland. In reality, however, futures traders (whether they be hedgers or speculators) are not prepared to make (or take) delivery of physical cargos of crude oil—they are in it for the cash—so they unwind their positions and extract the monetary gain or loss by selling the same futures contract they previously bought (or buying the contract they had previously sold). Hedgers use the net gains and losses to offset fluctuations in operating earnings; speculators simply count the net gains as profits or loss. Virtually all futures contracts are settled in this manner, for cash. Physical delivery almost never occurs. The NYMEX exchange reports that more than 99 percent of its WTI futures contracts are settled for cash in the manner described. The WTI futures contract traded on the ICE exchange actually requires settlement in cash; physical delivery is not an option.

Because futures contracts settle for cash, futures trading by hedge funds, commodity index funds, speculators, or anyone else—even if they rush into the futures market with lots of money—does not increase the demand for oil. Because those who trade futures contracts do not take possession or make delivery of crude oil, their trades lack any conduit that could affect the physical market or the spot

7 Just as movements in the spot price of oil affect the level of futures prices, they also determine the shape (forward curve) of futures prices at different dates. The spot price of oil is ordinarily expected to rise at the net cost of carrying inventories. If this were not true, participants in the physical market would have incentives to either build up or draw down inventories, until these adjustments (which entail buying or selling oil on the spot market) bring spot prices into line with expectations. An important corollary, due to convergence of futures and spot prices, is that the prices of distant futures contracts “normally” exceed the price for “prompt” delivery. A forward curve that rises in this fashion is said to be in “contango.” The opposite pattern of “backwardation” is caused by a temporary shortage in the spot market. For example, if a supply outage pushes the spot price up, it will also pull up the near-dated futures contract due to convergence. If the outage is expected to be restored quickly, however, the price of a distant futures contract will not be affected since its value is tied (via convergence) to a spot price that is expected to return to “normal.” A permanent or long-lasting supply disruption, on the other hand, would tend to move the entire structure of futures prices up, but not cause the forward curve to flip into backwardation. Although contango may be thought of as “normal,” in reality it is normal for the spot market to be buffeted by shocks, which leads to backwardation. On 59 percent of the trading days since 1986, the forward curve of NYMEX crude oil futures has exhibited backwardation (EIA, 2008a)—no surprise, given the frequency of negative supply shocks during this period.
price. Moreover, futures trades that settle for cash should not be expected even to move the futures price itself because each trader (including speculators) must eventually sell everything the trader initially buys, and vice versa. Anyone who would buy a futures contract after “buying pressure” has lifted its price could only expect to take a loss as the inevitable “selling pressure” from those same traders arrives to push the price back down.

The only avenue by which speculative trading might raise spot prices is if it incites participants in the physical market to hold oil off the market—either by amassing large inventories or by shutting in production. If participants in the physical market are convinced by speculative trading in the futures market that spot prices will soon rise, their reaction could cause inventories to rise and/or production to fall. However, neither phenomenon was observed during the recent price spike.

With respect to inventories, commercial stocks of petroleum in the United States and the OECD nations as whole were no higher in 2007 than in 2003 (U.S. EIA, 2008b). In the United States, where data are most complete, commercial inventories of crude oil in the first half of 2008 during oil’s furious price ascent were well below the prevailing level for 2005–2007 (U.S. EIA, 2008a). This includes all crude oil held by refiners, pipelines, storage terminals, and producers. In other words, oil inventories were being drawn down; supply was not being withheld from the market.

Neither is there evidence that non-OPEC oil producers contributed to rising prices by reducing output and keeping oil in the ground. Although data on underground reserves are probably too coarse (and affected by too many other factors) to reveal what would have been a subtle shift, actual production data belie the hypothesis. The quantity of oil supplied by non-OPEC producers did not change significantly between 2004 and 2008, even though (as reported in Figure 3) their supply curve was decreasing. In other words, non-OPEC producers responded to rising prices by pushing output further up a receding supply curve—not by shutting in reserves. This pattern is also reflected in the pace of drilling to bring new oil wells on line. Figure 4 charts the historical relationship between the real price of oil and the number of new development wells drilled per month in the United States. The observations from 2007 and 2008 reveal no departure from the historical tendency for high prices to stimulate the industry’s efforts to bring new wells on stream. The pace of development drilling did not abate (due to short-run rigidities) until long after prices started declining in July 2008, which is reflected in the curly tail in the diagram.

Finally, we might ask whether price fixing, rather than speculation per se, might be responsible for the dramatic increase in price? With respect to OPEC, this has already been answered: OPEC does engage in price fixing, and oil prices would not have reached $145 per barrel if OPEC had not previously restricted investment in new capacity. Perhaps OPEC was caught “off guard” by the coincidence of dramatic growth in Asian demand and the decline of non-OPEC supplies, in which case the price spike might be blamed to some extent on OPEC’s limited foresight.
and bungled management of the market—although under this hypothesis, OPEC did not actually take any positive action in 2007 or 2008 that precipitated the price spike.

OPEC aside, there is no evidence of price fixing on the part of anyone else, which includes both speculators and the “super-major” oil companies. It would be a daunting proposition to fix the price of the single largest commodity in world trade, something a cartel might attempt if it controlled 70 percent of world reserves but not something that small fry could achieve (Borenstein, 2008). Relative to the size of the world oil market, hedge funds and even the “super-major” oil companies are small fry. To succeed, they would have had to: 1) accumulate large inventories that were diverted from the commercial supply chain; or 2) shut in a significant portion of global oil production. Neither phenomenon was observed.

**The Other Side of the Price Spike**

What combination of factors then explains the collapse in oil prices that occurred during the second half of 2008? Surely, the primary factor is that demand for oil dropped sharply around the world due to the economic decline, which in early 2008 few analysts were predicting would turn out to be so deep. In addition, on the supply side, the cost increases that reduced oil supply in the 2004–2008 period have now reversed themselves, and the supply curve of oil from non-OPEC producers has consequently expanded. Various investments to increase supply and
reduce demand, triggered by four years of elevated prices, may have also begun to bear intended fruit.

**The Future of Oil Prices: Peak Oil?**

The sharp spike in oil prices in mid-2008 brought renewed prominence to the “Hubbert curve.” More than a half-century ago, Hubbert (1956) advanced a novel approach to forecasting the date at which oil production would enter an inexorable decline. Hubbert’s idea was that, in the case of oil production, prices and other incentives are superfluous; it is all a matter of time. His resulting model of production behaves like a ballistic missile, first rising and then falling of its own accord. Hubbert’s 1956 prediction that U.S. oil production would peak around 1970 was famously borne out. The Hubbert curve has become influential in certain quarters, and is now widely applied to predict that global oil production will soon peak, with cataclysmic effects according to some analysts (Hirsch, Bezdek, and Wendling, 2005).

The main counterargument is that the timing of the peak of production (or peaks, since there may be local maxima) and the nature of its aftermath are determined as much by economics as geology. From an economic view, a peak is an artifact of an intertemporal allocation of resources that is determined by market forces. Models that describe the market-clearing process for exhaustible resources show how equilibrium prices and production should trend over time (Hotelling, 1931; Herfindahl, 1967) and demonstrate that both the timing and the economic consequences of a peak are highly sensitive to small changes in underlying economic parameters like population growth rates and elasticities (Holland, 2008). The phenomenon of peaking may occur early or late in the life cycle of a particular resource, and may signal either scarcity or relative surplus, the onset of rapidly rising prices or the arrival of a price plateau—all depending on underlying structural relationships and parameters within the economy.

Given this sensitivity to underlying parameters, it comes as no surprise that Hubbert’s (1956) approach has often proved quite faulty. For example, while the model did accurately predict that U.S. oil output would peak in 1970, the prediction was for output to peak at three billion barrels per year, whereas actual production in 1970 reached 4.1 billion barrels. Hubbert’s predictions regarding oil production from many major oil basins around the world have been substantially off (Nehring, 2006a,b,c; Brandt, 2007). Hubbert also predicted that U.S. natural gas production would peak at 14 trillion cubic feet per year in 1973; actual production was 20 trillion cubic feet in 2007. Hubbert further predicted that global coal production would peak in 2150 at about 6.4 billion metric tons; actual production reached that level in 2007 and is still growing rapidly.

For the global oil market, the crucial insight is that while oil is constantly being “used up,” the world is not “running out” of oil (Adelman and Watkins, 2008). Despite global consumption (and consequent depletion) of almost 700 billion
barrels of crude oil during the past quarter-century, the stock of remaining proved reserves has doubled from 700 billion barrels in 1980 to 1,400 billion barrels—and now stands at an all-time high. The ratio of reserves divided by annual production has also grown from a multiple of 29 years in 1980 to a multiple of 45 years in 2008; in other words, we now extract a smaller fraction of remaining oil reserves each year than several decades ago.

In the longer term, recent assessments by the International Energy Agency (2005) and Aguilera, Eggert, Lagos, and Tilton (2009) identify at least five trillion barrels—the equivalent of 160 years of current oil consumption—of unconventional petroleum resources in forms like heavy oil, oil sands, and oil shale that could eventually augment or supplant conventional crude oil at prices well below $100 per barrel. Thus, while oil prices may experience short-term peaks (and valleys), there is no geological reason to believe that oil prices are likely to plateau in the foreseeable future at or above the sky-high levels of mid-2008.

**Conclusion**

The world oil market operates subject to the familiar laws of supply and demand, and market fundamentals are the dominant influence on price. The market is subject to shocks, and when these shocks are taken together with short-run rigidities and high costs of adjustment, the resulting price volatility is largely inherent, rather than contrived by speculators, cunning producers, or anyone else.

In the longer run, the effect of shocks will average out and the effect of structural trends is paramount. The most conspicuous trend, by far, is the rapid pace of economic development in China and other emerging nations. If that continues, oil’s high income elasticity implies a proportionate increase in demand. However, the global economic slowdown of 2008 also suggests that an appreciable decline in Asian growth has contributed to a substantial fall in the expected demand for oil and consequent increase in excess production capacity—which brings us back to OPEC. The long-run trend has been for OPEC to restrict the expansion of new production capacity. Given the cloudy economic outlook and indications of sharp reductions in oil demand in 2009, the cartel will almost certainly adhere to form, even to the extent of canceling some production investment projects that have already been announced. But many OPEC members also have a fundamental tendency to ignore the cartel’s attempts to rein in surplus production. For as long as the current economic slowdown persists, it will be difficult for OPEC to boost the price of oil of its own volition.

The sustainability of oil supplies from non-OPEC producers is also of fundamental importance. Proven oil reserves of non-OPEC producers have been rising—but resource depletion puts constant upward pressure on costs. For decades, the oil industry has been able to use technological innovation to offset the impact of depletion by finding and producing oil in ways that held the marginal cost of
output in check. Although we cannot expect further technical advances to prevent the supply of conventional oil from ever declining, in the longer-term, ample supplies of unconventional petroleum resources and other substitutes for crude oil should prevent oil prices from surpassing the mid-2008 peak on any sustained basis, although too many technological and political uncertainties exist to permit a definite prediction.

The author is indebted to M. A. Adelman, Larry DeBrock, Richard Gordon, James Hamilton, Dwight Lee, John Parsons, Douglas Reynolds, Ron Ripple, Rex Thompson, and the editors for their many helpful comments on an earlier draft.

References


