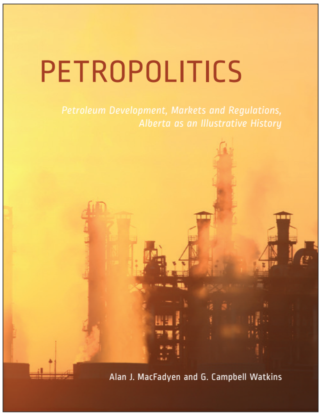


PETROPOLITICS: PETROLEUM DEVELOPMENT, MARKETS AND REGULATIONS, ALBERTA AS AN ILLUSTRATIVE HISTORY

Alan J. MacFadyen and G. Campbell Watkins

ISBN 978-1-55238-769-6

Acknowledgement: *We acknowledge the wording around open access used by Australian publisher,* **re.press**, *and thank them for giving us permission to adapt their wording to our policy* http://www.re-press.org

Part One: Overview

In January, darkness covers the Canadian prairies by five o'clock in the afternoon. The landscape appears as it has for the past century. Lights of vehicles or farmhouses, and occasional small communities, dot the landscape. Otherwise, a wide expanse of snow-covered fields stretches out, marked only by the riverbeds and windbreaks of trees. Suddenly a blazing mass appears, myriad lights from a sprawling city: glowing skyscrapers; never-ending streams of headlights spreading from the city core; hectic shopping centres; endless banks of apartments and houses in which the evening's activities begin. A modern megalopolis in the once quiet prairie provides a tangible symbol for this book. Its light and warmth in the depth of winter's cold illustrate the possibilities opened by the low-cost energy of the Fossil Fuel Age. The very existence of active, wealthy cities in the Canadian prairies, so far removed from the longer-lived centres of world power, reflects the potential offered to a region by nature's bounty of energy resources.

We want to provide an 'economic history' of the petroleum industry in Alberta, from its beginning to the present. There are many ways in which the story could be told. A scientist might emphasize the physical history – how primitive life forms in shallow seas hundreds of millions of years ago can lead to an array of wells, pipelines, and refineries that provide the energy to fuel our industrialized world. The political scientist might highlight the management role of governments and detail the complex web of petroleum legislation and regulation that results from the political interplay of local farmers, regional entrepreneurs, multinational corporations, environmentalists, government departments, and just plain taxpayers. A lawyer would trace the course of cases in civil and common law, and the judicial judgments, that give precise meaning to laws and regulations and stimulate new legislation. For the psychologist or biographer, the story might lie in a succession of determined and eccentric personalities relentlessly pursuing new ideas and opportunities. We do not dismiss the vibrancy of these approaches. However, our view is one of petroleum as an economic commodity, produced in competition with other energy products throughout the world. The history of the development of oil and gas in Alberta is in large measure an economic story. We hope that this perspective will prove valuable to readers of this book, even those who begin by thinking that energy is too important to be left to economists.

As economic commodities, oil and gas can be viewed from a purely 'private' perspective, as seen by companies and consumers. They can also be viewed from a 'social' perspective, as commodities to be utilized in the broad public interest. The importance of energy to the functioning of any economy has meant that energy is amongst the most regulated of commodities. What might appear to be purely private decisions are made within a complex and evolving web of government regulations. The title "Petropolitics" was chosen to acknowledge the importance of the legal and regulatory setting to the economics of the petroleum industry.

Our study deals with oil and natural gas in Alberta. It might, therefore, be viewed as a restricted study of narrow interest. However, the physical conditions that generated petroleum, the assorted tasks performed by the petroleum industry, the operation of economic markets, and the regulatory issues that arise are common to oil- and gas-producing regions throughout the world. It is our hope that the analysis in this book will serve as a valuable exemplar for analysts and policy-makers studying the petroleum industry in other parts of the world.

Part One provides an overview, in four parts. Chapter One deals with oil and natural gas as physical products: what they are, what the petroleum industry does to/with them, and what these physical realities imply for an economic depiction of the industry. Chapter Two provides an initial perspective on the petroleum industry in Alberta, contrasting the operations of the industry in the late 1940s with the year 2010 and briefly reviewing some critical policy issues that arose over this period. Chapter Three sets the Alberta petroleum industry in a global context. Finally, Chapter Four reviews the formal concepts and constructs that economists utilize to help understand how economies and markets function.

Part Two looks at the Alberta crude oil industry from what we call a 'private' perspective': it is largely concerned with the evolution of the industry from the viewpoint of oil producers and consumers. In this part we look at the crude oil resource base, the evolution of crude oil markets (prices and production), the development of Alberta's non-conventional oil sands resources, and various 'models' that economists have built to help us understand this complicated industry.

Part Three examines the Alberta crude oil industry from a 'social' perspective: it deals with the government regulatory environment. Three important policy areas are covered: pricing and trade regulations; production conservation regulations; and tax and royalty regulations.

Part Four includes three sections. First, it covers the history of Alberta natural gas, focusing on aspects that differ from crude oil. We discuss both natural gas markets and government regulations, with a particular emphasis on the restriction of gas exports to sales in excess of 'domestic requirements.' Next, the perspective is broadened from an emphasis on crude oil and natural gas markets to the role of the petroleum industry in the Alberta provincial economy. Finally, we conclude with lessons from the Alberta experience that may be of value to decision-makers elsewhere in the world.

While we have incorporated most of the economic issues of significance for the Alberta crude petroleum industry, we should alert readers to several exceptions. Our primary interest is with Alberta's abundant natural petroleum wealth. Hence we focus on the production and marketing of crude oil and natural gas. However, we do not examine in any detail natural gas liquids such as ethane, butane, and propane, or the sulphur often produced in conjunction with natural gas; nor do we delve deeply into the many issues associated with the subsequent shipment and processing of petroleum such as pipeline operation, oil refining, natural gas processing, or petrochemicals.

Finally, we do not engage in any detailed analysis of the environmental impacts of the industry. This is not because we think they are unimportant. However, this book is already very long, and the environmental questions involve scientific issues with which we have no expertise.

CHAPTER ONE

Petroleum and the Petroleum Industry: What Are They?

Readers' Guide: Chapter One is aimed at readers who have little familiarity with the petroleum industry. It describes the activities of the industry in terms of a number of different stages required to transform petroleum from a resource in nature to a product that consumers willingly purchase. Readers familiar with the industry may wish to move on to Section 3 of this chapter.

1. What Is Petroleum?

A dictionary will note that the word 'petroleum' is derived from Latin, meaning 'rock oil,' and is almost always used to refer to those mineral oils provided from below the earth's surface that consist mainly of mixtures of hydrogen and carbon molecules (i.e., hydrocarbons). Petroleum is, therefore, a natural resource. Sometimes the term has been broadened to include 'manufactured' hydrocarbons that are identical to the natural resource; this would include, for instance, liquid oil or natural gas derived from coal or biomass. However, such 'synthetic' products have not as yet been produced in large volumes. The term 'petroleum' is also applied to refined petroleum products like motor gasoline and fuel oil, which are derived from processing the natural resource.

Naturally occurring hydrocarbon deposits vary greatly in physical composition but are generally grouped into two broad classes, depending upon whether the main output is liquid (crude oil) or gaseous (natural gas). The greater is the proportion of carbon to hydrogen in the deposit, the heavier and more viscous the petroleum. In extreme cases, such as the bitumen in oil sands deposits around Fort McMurray in Northern Alberta and kerogen in oil shale deposits in Colorado, the hydrocarbon is so viscous that it will not flow of its own accord beneath the surface. To date relatively little petroleum of this very heavy type – frequently labelled 'non-conventional oil' – has been produced, with production concentrated in Alberta and Venezuela. 'Conventional' crude oil refers to liquid hydrocarbons derived from natural underground deposits ('pools' or 'reservoirs') in which the liquid is fluid enough beneath the surface that some of it can be lifted readily through wells.

Conventional crude oil is a liquid mixture of paraffinic and other hydrocarbons spanning a wide range of molecular weights and containing varying amounts of sulphur, nitrogen, and other elements. It varies in specific gravity (relative to water) from about 0.8 to 1 (API gravity from 50° to nearly 10°). API stands for the American Petroleum Institute, which instituted the API degree scale in the late 1800s. If (sg) is the specific gravity of the oil, the API degree is given by the following formula:

$$\text{API} = (141.5/\text{sg}) - 131.5.$$

The lower the specific gravity (the lighter the oil), the higher the API degree number. Since water has a specific gravity of 1, oil as heavy as water would have an API degree number of 10.The bitumen found in the Alberta oil sands is heavier than water, with API values around 5. Natural gas is a gaseous mixture of normal paraffinic hydrocarbons, mainly methane

(CH_4), which is often contaminated with water vapour, nitrogen, carbon dioxide, and hydrogen sulphide.

The general belief is that crude oil and natural gas were formed millions of years ago from the remains of aquatic plant and animal life. For this reason oil and gas (as well as coal) are called fossil fuels. In the prevailing view, petroleum originated in sedimentary basins – areas where thick layers of sediment were deposited at the bottoms of shallow seas. Over time dead plant and animal matter settled with the sediment; eventually these overlying layers of mud and silt created great pressure and high temperatures. Finally, when these beds had sunk thousands of metres deep, the plant and animal matter became chemically converted to oil and natural gas.

As pressures on the original sedimentary rock intensified, and as the earth's crust shifted over time, the oil and gas 'migrated' through pores, cracks, and fissures where it became trapped in porous rock in underground structures. These petroleum yielding structures are known as reservoirs and consist of several types of 'traps' – the most typical being structural (or fault) traps, stratigraphic traps, and a combination of both types.

Structural traps are caused by local deformations that 'fold' or 'fault' the reservoir rock. Anticlines, resembling elongated arches, and domes, resembling inverted bowls, are the main types of folds that serve as traps for oil and gas. A fault is a fracture in the earth's crust along which movement has taken place; these shifts can bring non-porous rocks in contact with porous ones, thus forming a trap.

A stratigraphic trap is one in which the chief trap-forming element is some variation in the nature of the reservoir rock. These traps represent the most difficult oil and gas accumulations to find since there are no structural features associated with them. A common stratigraphic trap consists of a wedge-shaped sandstone formation squeezed between impervious rocks and lying at an inclined angle. Oil or gas becomes trapped where the sandstone 'pinches out' against the impervious rock.

The trap holds the oil and gas in place so that they cannot escape until released by drilling a well. In an oil pool, the three elements that are usually present – water, oil, and gas – occur largely in layers. Water is at the base and gas, when present, tends to the top. Oil lies between since it is of intermediate density. In most cases oil deposits contain some natural gas in solution, mixed with the oil and held there by the high pressure in the reservoir; natural gas in gaseous form is also found as a gas cap above an oil trap. Such gas is called 'associated' since it is found in association with crude oil in the reservoir. Sometimes natural gas is discovered in a free state in a reservoir not in association with crude oil; this is called 'non-associated' gas. Frequently non-associated natural gas is relatively 'wet,' including hydrocarbons heavier than methane; that is, the molecules have more than a single carbon atom. These may be removed as natural gas liquids (NGLs), such as ethane (C_2), propane (C_3), butane (C_4), and pentanes plus (C_{5+}). While water tends to lie beneath the oil in a reservoir, it is also common to have some water molecules adhering to the rock pore spaces in the portion of the reservoir holding oil.

Petroleum is of interest primarily for its energy content: the electromagnetic (chemical) bonds holding together the various atoms in the hydrocarbon compounds can be released quite easily (e.g., by application of heat) with an attendant release of energy (again, in the form of heat), which can be harnessed to do work. More than 90 per cent of the world's use of hydrocarbons is for their energy content. In the remaining instances petroleum is used for its matter; that is, the particular hydrocarbon compounds are desired, by petrochemical and plastics companies, fertilizer manufacturers, road pavers or others, for their structural or other physical features.

Scientists note that petroleum, as a physical product, is subject to various laws of nature including the first two laws of thermodynamics (Foley, 1976, chap. 4). In simplistic terms, the first law of thermodynamics is a conservation law that states that the energy content of petroleum can be neither created nor destroyed; instead it changes form on use. Expressed in quite different terms, the utilization of petroleum *necessarily* creates waste energy and matter. The second law of thermodynamics is the famous Entropy Law. It states, in essence, that the utilization of energy necessarily reduces it to a less usable or available form (i.e., increases entropy) so that, while it is possible to use energy more or less efficiently, it is not possible to recycle it. The entropy law, combined with the tremendously long time span involved in the generation of petroleum deposits, make petroleum in nature a non-renewable or exhaustible natural resource. How efficiently the economic system recognizes this non-renewability is a matter of widespread debate. For example, Georgescu-Roegan (1973) and Daly (1973) emphasize the critical importance of entropy; on the other hand, Adelman (1990) and Watkins (1992) question whether the concept of finite physical resource limitations is meaningful to economic analysis of the petroleum industry.

2. What Is the Petroleum Industry?

The petroleum industry consists of six main sectors: **exploration** for reservoirs of crude oil and natural gas; reservoir **development**; production or **lifting** of oil and gas; **transportation** or transmission; **refining** of crude oil into refined petroleum products; and the **marketing** (distribution) of these refined products and of natural gas. Exploration, development, and lifting (or extraction) are generally referred to as the 'upstream' activities of the petroleum industry (or the 'crude petroleum industry') while the refining and marketing of petroleum constitute 'downstream' operations. Transportation provides the link between the 'upstream' and 'downstream' segments; this book treats it as part of the industry's downstream activities.

In the material that follows, we shall refer to the production of oil; unless otherwise specified, similar factors hold for natural gas.

A. What Constitutes 'Upstream' Activity?

A formal definition of the 'upstream' segment of the petroleum industry is (Canada Petroleum Monitoring Agency, 1986): "activities and operations related to the search for, and development, production, extraction and recovery of crude oil, natural gas, natural gas liquids and sulphur, as well as the production of synthetic oil."

1. Exploration and Development

a. Geological and Geophysical Work and Land Acquisition

The search for underground accumulations of oil and natural gas begins with looking for the type of rock formations in which petroleum deposits are likely to be found. (Gow, 2005, provides a useful overview of the geological and technical dimensions of petroleum industry activity, with specific reference to Alberta.) This typically restricts the search to regions with deep overlays of sedimentary rock (i.e., a sedimentary basin), in which it is believed that adequate source and reservoir rock (geological formations) were laid down in the distant past. The first stage in the exploration effort is to select the regions in which effort will be expended. This depends on a mix of factors including the physical prospects for finding petroleum, accessibility, the economic and political climate of the region, and proximity to markets. Once a region is selected, a preliminary geological survey is undertaken. Visible rocks are examined for any clues they may provide as to what type of formations lie beneath the area. Rock samples are taken to compare with samples from previously discovered hydrocarbon deposits. A detailed geological map is prepared, which provides information on the prospects for finding oil or gas in the area.

At this time the company will begin the process of acquiring 'land' in the region; more specifically, it must acquire 'mineral rights,' i.e., the property right that conveys the legal right to explore for and recover petroleum, if found, at particular locations. In Canada the majority of mineral rights are 'Crown'; that is owned by governments, mainly provincial governments. Petroleum exploration rights have been issued primarily through competitive bidding sales. Some mineral rights are 'freehold,' that is owned by private parties with whom the oil companies must negotiate. In addition to obtaining mineral rights for the subsurface petroleum resources, oil companies must negotiate with surface rights owners (e.g., farmers and ranchers) to obtain rights of use for the land needed for roads, drilling sites, etc. Before any oil can be produced, it is necessary to acquire production rights as well as exploration rights. Typically areas covered by exploration licences may be converted in whole or part into leases that allow petroleum extraction. In addition, primary landowners, like provincial governments on Crown land, often directly issue leases that permit exploration and production.

The next step is to investigate the underground rock structures. Geophysical surveying is the application of the principles of physics to the study of subsurface geology. Geophysical surveys measure the thickness of sediments and map the shape of structures within the sediments. The most common type of geophysical study is seismic, in which explosive charges are detonated at or near the ground's surface. The ensuing shock waves are recorded by geophones after they strike and rebound off underlying layers of rock. With this information geophysicists are able to locate structures that might contain oil or gas. Gravimetric and magnetic surveys are other methods employed by geophysicists to obtain subsurface data. Recent technological developments, like 3-D seismic mapping, and now 4-D seismic (with time as the fourth dimension) have expanded the role of seismic activities and have led to much re-evaluation of previously studied geological strata.

b. Exploratory Drilling

Geological and geophysical surveys undoubtedly improve the chances of finding oil or gas, but they can at best map the underlying geological structures, not pinpoint the presence of petroleum. Since many potential petroleum-holding traps are dry, the only way to prove the existence of an underground reservoir where large accumulations of oil or gas occur is to drill a hole. Thus the next stage in the search for petroleum is exploratory 'wildcat' drilling. This is done on a site recommended by the geologist or geophysicist, predicated on the survey work done earlier.

On occasion the first hole that is drilled in a new territory will 'prove' oil or gas in quantities large enough to be exploited commercially. The normal occurrence, though, is the drilling of a number of dry holes, or holes suggesting only small (non-commercial) amounts of petroleum. The additional information obtained from these holes will lead to a final decision on whether to proceed with more exploration of the area.

When a successful exploratory well occurs, a series of appraisal ('stepout' or 'extension' or 'outpost') wells are typically drilled to determine the extent of the reservoir. Cylindrical samples of the formations penetrated (known as 'cores') are analyzed over the oil-bearing section of the rock so that its permeability, porosity, and oil content can be determined. In addition, samples of the oil are taken from the bottom of the well at full reservoir pressure so that the properties of the oil, as it exists in the reservoir, can be measured, including the unrestrained flow rate. Oil pools are heterogeneous: they vary tremendously in areal extent, depth, rock porosity, permeability, fluid content (oil, gas and water), quality of the hydrocarbons (light or heavy, etc.), and other salient characteristics.

c. Development Drilling

The drilling of development wells begins as soon as the information derived from appraisal drilling is sufficient to suggest that the oil or gas discovery is commercial and what would be the most suitable way to develop and produce the reservoir. This stage is often reached before the limits of the field have been fully delineated, and it is therefore not unusual for more stepout or outpost wells to be drilled at the same time as development wells are being sunk. (While stepout wells are commonly classified as part of the exploration process, they could just as well be considered development, since they occur after a reservoir has been found.)

The number of development wells, their spacing, and their depth will depend on the size and character of the field, as well as the land-tenure system under which the government establishes conditions about mineral rights. For example, development wells, like apple trees in an orchard, may be spaced in a regular pattern or grid system. This type of spacing pattern may ensure that the oil off-take is evenly distributed over the whole reservoir. However, such patterns are only appropriate in the development of flattish structures with relatively homogeneous subsurface rock and reservoir conditions. On steeply dipping structures, a single line or ring of wells is more likely to be drilled. And the distance between the wells depends on the size of the area that can be effectively drained by each one. In addition, governments typically set regulations about the allowable development patterns, often in the form of a minimum required spacing for wells. Such regulations often reflect a concern by the government to protect the (subsurface) property rights of adjacent land owners. (Since oil and gas are fugacious, i.e., fluid, it is possible for a producer to capture petroleum from beneath a neighbouring property.)

Until recently, development wells were almost always entirely vertical, or, in exceptional circumstances, slanted at a constant angle for the entire well depth. A slant well would be appropriate, for example, if the land is particularly sensitive for environmental reasons directly above the part of the reservoir being drained, or if a large area of the pool is to be drained from wells that start from the same location, as an offshore production platform. Technological advances in recent years have encouraged the drilling of 'horizontal wells' in which the well bore turns markedly away from the vertical to the horizontal as the well enters the producing formation. A single horizontal well is in contact with a larger volume of reservoir rock than a single vertical well. In a reservoir that has relatively high permeability and is relatively homogeneous in character, horizontal wells allow faster recovery of oil. In a reservoir that has relatively poor permeability and/or is very heterogeneous in nature (with 'pockets' of better and worse producibility) horizontal wells may increase the 'sweep' area and allow greater total recovery of oil than would be possible with vertical wells only.

Development activities are multifaceted and highly specific to the particular characteristics of the pool to be drained. Wells may be all vertical or vertical and horizontal. Development wells may include some or all of the following: appraisal (outpost) wells that prove up new volumes of recoverable oil (new 'reserves'); infill wells, spaced among previously drilled ones, that allow faster recovery of the oil; water disposal wells, to pump connate water back into underground formations; water or gas or other injection wells and associated oil-lifting wells, as part of an enhanced oil recovery (EOR) project to augment the natural productivity of the pool. The variety of physical production procedures combined with the heterogeneity of oil pools translates into an array of economic costs of producing petroleum.

2. Production (Lifting or Operation)

The rate at which oil can be extracted once wells are drilled depends largely on the permeability of the rock – the degree to which a rock will allow oil and gas to pass through it. If this is too low, the production obtained from an individual well might be insufficient to offset its cost so that the development of the reservoir would be ruled out on economic grounds. Generally the porosity – the number of spaces and openings that separate the individual rock grains – and permeability vary from place to place within the same reservoir rock. Sometimes these variations are so diverse that wells located in different parts of the reservoir may have markedly different production rates.

The reservoir crude can range from very heavy viscous (thick) oil under very low pressure containing little or no dissolved gas to extremely light straw-coloured crude under considerable pressure containing a large amount of dissolved gas. The viscosity of the oil depends largely on its specific gravity as well as on the quantity of gas that it holds in solution. The less viscous an oil, and the more gas it contains, the more readily it will flow through the crevices of the rock to gain entry to the well.

An oil or gas reservoir also typically contains some water in its pore spaces. This 'connate' or 'interstitial' water is believed to be water that was not displaced by the petroleum at the time of its accumulation and entrapment in the originally water-saturated reservoir. The connate water content may range from 5 to 40 per cent or more of the reservoir void space and plays an important role during the productive life of the reservoir.

a. Primary Production Methods

For oil to move through the pores of the reservoir rock and out into the bottom of a well, the pressure under which the oil exists in the reservoir must be greater than the pressure at the bottom of the well. As oil is removed from the rock, the pressure of the reservoir will decrease and the rate of production will decline. The rate at which the pressure decreases will affect the total amount of oil that can be removed from the reservoir over a given period of time, if only because declining production brings the well closer to being uneconomic to operate.

The connate water found in the reservoir, associated gas, and the free gas in the gas cap are the main sources of energy that drive the crude oil to the bottom of the producing wells and thence up the pipe tubing to the surface or wellhead. The production mechanisms associated with these sources of energy are referred to as 'water drive,' 'solution gas drive' (or 'depletion drive'), and 'gas cap drive,' respectively. 'Water drive' is normally the most efficient of the three displacement processes; 'solution gas drive' is the least efficient. Both gas cap and water drive reservoirs are often subject to more than one mechanism. Consequently, the terms 'partial gas cap drive' and 'partial water drive' may apply. Also a reservoir's predominant drive mechanism may change over time, as for instance when gas from solution collects by gravity segregation to form a gas cap as reservoir pressure declines.

The oil obtained as a result of these natural production mechanisms, supplemented only by pumping and simple fracturing of reservoir rock, is referred to as 'primary recovery.' As will be discussed below, 'enhanced oil recovery' (EOR) techniques may allow recovery of even greater volumes of oil.

b. Recovery Factor

As the preceding discussion suggests, there is no known economic process by which all of the oil in porous rock may be recovered. There are six groups of factors that jointly determine the 'recovery factor'; that is the fraction of the oil-in-place within a reservoir that can be brought to the surface. These are:

- reservoir rock properties, e.g., porosity, permeability, structural position, and thickness;
- reservoir fluid properties, e.g., viscosity, pressure, gas saturation;
- drive mechanism, e.g., solution, gravity drainage, water drive;
- method of production, e.g., well completion techniques (including EOR), spacing of wells, rate of withdrawal, utilization of EOR;
- economics, e.g., drilling and completion costs, production costs, prices of oil, gas, and by-products;
- government regulations including those relating to well-spacing and assorted 'conservation' practices, royalties, taxes.

In combination these factors determine the amount of oil or natural gas that can be recovered economically, or the 'reserves' in the pool. Applying the recovery factor to the quantity of oil in place generates an estimate of 'initial recoverable reserves' that is typically considerably less than the volume of petroleum in place. ('Remaining recoverable reserves' is an estimate of the amount of petroleum still waiting to be produced, so equals initial recoverable reserves minus cumulative production to date.)

Recovery factors vary markedly among reservoirs, but on average about 25 to 35 per cent of the oil initially in place in a conventional oil reservoir is recoverable, as is about 80 per cent of the gas in place in a non-associated natural gas reservoir. The recovery factor for oil is often improved by the introduction of enhanced recovery techniques in the extraction process. In Alberta, for instance, on average about 17 per cent of the oil in place is recovered by primary means and another 7 per cent by EOR (Alberta Energy Resources Conservation Board, 2013, *Reserves Report*, ST-98, p. 4-6). However, the recovery factor cannot be estimated reliably until the behaviour of the reservoir has been observed under actual producing conditions at commercial rates of off-take. Fields are generally put on production before delineation and development drilling has been completed. As a result, data on recoverable reserves are continually being revised as additional productive areas are drilled up. Typically more reserves additions are credited to 'extensions and revisions' in discovered pools than to 'new discoveries.' Expressed in other terms, the reserves in a petroleum pool typically 'appreciate' over time as development proceeds. An analysis by the Alberta Energy Resources Conservation Board in its 1969 *Reserves Report* of all except the smallest pools found an average appreciation factor of about nine times for oil pools and four times for non-associated gas pools. That is, the average pool in the province will ultimately yield nine times the amount of oil that was credited to the pool as reserves in its year of discovery. This average appreciation factor of nine for the Province masks a wide range of differences for individual oil pools.

c. Enhanced Oil Recovery (EOR) Processes

Some EOR methods, like primary recovery, rely on direct displacement with another fluid to force the oil out of the reservoir rock and are often called 'secondary' recovery techniques. Water flooding is one of the most successful and extensively used secondary recovery methods. Water is injected under pressure into the reservoir rock via injection wells and drives the oil through the rock into adjacent producing wells. Where there is considerable variation in the permeability of the rock, the rate of injection must be carefully controlled to avoid trapping and leaving behind large quantities of oil. In the gas drive process, gas is injected into the reservoir via injection wells, which are usually located on or near the crest of the structure. The injected gas is driven downwards and sideways through the reservoir and displaces the oil into producing wells; over time the ratio of oil to gas recovered falls.

These techniques were originally developed to extract more oil out of reservoirs from which no more oil could be recovered by primary production. However, the current practice is to apply these processes much earlier in the producing life of the reservoir to forestall the decline in reservoir pressure. The application of secondary recovery techniques during the early stages of the primary production phase is referred to as 'pressure maintenance.'

In recent years, EOR processes have become more effective by the adoption of methods that improve the performance of the displacing fluids. These enhanced oil recovery methods can be broadly divided into solvent techniques and thermal techniques and are often called 'tertiary recovery techniques.'

In solvent techniques, the displacing fluid is treated with an additive to make it miscible with the reservoir oil and thus improve the efficiency with which it sweeps the oil out of the pores of the rock. For example, a 'miscible gas drive' is a process where the natural gas in a gas drive is rendered miscible with oil by the addition of a sufficient amount of liquefied petroleum gas (LPG). This process is most effective in light oil reservoirs; however, large amounts of LPG are required, which makes the process relatively expensive. Carbon dioxide is another injection fuel, the popularity of which may increase as concerns over global warming lead to policies that reward the capture and storage of CO_2.

The difficulty of recovering oil from reservoirs containing heavy viscous crude has led to the development of thermal techniques that increase the flow rate of the oil primarily by the addition of heat, as in steam injection.

Since natural gas moves so readily through pore spaces in the reservoir, EOR techniques have not been common in non-associated gas pools.

d. 'In Situ' Bitumen Recovery

Some oil deposits contain crude that is so viscous that primary production is not generally feasible. The crude in these reservoirs is generally called 'bitumen.'

In some cases it can be produced through wells, some amounts by primary pressure or water and solvent injection, but most by thermal processes, in which heat is applied to the reservoir to reduce the viscosity of the oil and make it easier to displace. The heating effect might be achieved by injecting steam or hot water or by burning, underground, some of the oil in the reservoir. In this latter process, which is referred to as 'in situ combustion,' air is injected to support combustion and to act as the displacing agent. Such processes are used to recover the crude bitumen at various sites in east-central Alberta, including Cold Lake where steam is injected into the reservoir in cycles of injection and production. More recent projects have usually used steam-assisted gravity drainage (SAGD), a technique developed with the encouragement of the Alberta Oil Sands Technology Research Authority (AOSTRA), a research facility created in 1974 by the Alberta government. (In 1994 AOSTRA was melded into the Ministry and Energy. Then, in 2000, it was terminated, with its projects transferred to the Alberta Energy Research Agency. At the start of 2010, a new Agency, Alberta Innovates, took over.) With SAGD, steam (most often created by burning natural gas) is injected through horizontal wells above the bitumen-bearing rock; bitumen is then gathered by horizontal wells within the oil-bearing rock (Deutsch and McLennan, 2005; Engelhardt and Todirescu, 2005). Such output is called 'in situ' bitumen production, since the oil is treated in the reservoir to allow production.

e. Surface Mining of Oil Sands

The petroleum found in the oil sands in the vicinity of Fort McMurray, Alberta, is thick bitumen that is mixed with sand, clay, and water. Much of the bitumen-laden sands are sufficiently shallow that the sands can be strip-mined. Huge bucket-wheel excavators and dragline shovels are used to remove the overburden and mine the exposed sands. However, this method can only be used to a depth of about fifty metres. Centrifugal water flotation techniques are used to separate the sand and the bitumen, and the bitumen is then upgraded to light crude, often called 'syncrude.' As of 2012, there were only four commercial mining operations. The Suncor and Syncrude plants had been in operation for decades. The Albian sands project co-ordinated by Shell began operation in 2005, and Consolidated Natural Resources Limited (CNRL) commenced the Horizon project in 2008. At the end of 2012, the ERCB reported two new projects under construction, twelve others that had been approved and seven that had made application; these numbers include expansions by the four current producers (ERCB, *Reserves Report,* ST-98, 2013, Table S3.2). The Imperial Oil and Exxon Mobil Kearn mining project began production in the spring of 2013, and is the first mining project to produce bitumen without upgrading it to synthetic crude oil.

Needless to say, these bitumen recovery methods are expensive and require prospective oil prices robust enough to warrant the high investment. Recent technological improvements have markedly reduced operating costs.

f. Lifting of Crude Oil

After wells have been completed, conventional oil has to be brought up from the bottom of the holes to the surface. This can be accomplished in several ways, depending on the nature of the oil, the potential energy available in the reservoir, and the specific form that prior development has taken. The reservoir may have sufficient energy to cause the well to flow naturally for many years – as is the case in almost all Middle East fields – or it may continue for only a short period of time without further investment such as 'gas lift' or pumping. And, of course, these methods are also applied to wells that have never had sufficient energy to push oil to the surface.

g. Surface Treatment of Crude Oil and Gathering

When the crude has been brought to surface, the next step is to reduce it to the form in which it will be delivered to the refinery for processing. Oil as produced at the wellhead varies considerably from field to field not only because of its physical characteristics but also as to the amount of gas and water that it contains. These have to be separated from the oil.

The oil from each producing well is conveyed from the wellhead to a gathering plant, occasionally by truck but usually through a flow line. The gathering plant, which is located at some central point to handle the production from several wells, is equipped to separate any gas and water from the oil. The size and nature of the plant required for the separation of the gas from the crude will depend on the volume of gas dissolved in the oil as it exists in the reservoir and on the pressure at which it issues at the wellhead. Crude with very little gas in it, and possessing little or no pressure when it comes out of the wellhead, can be separated from the gas in a one-stage operation. High pressure crude with considerable gas content is subject to a multistage separation process.

The crude oil is then typically gathered up in pipelines and moved from the field to meet with petroleum from other fields (generating a 'blended' crude) and is transported by pipeline to a main shipment and storage point in the region (e.g., in Alberta, for crude oil, the Edmonton terminals of the Enbridge and the Kinder Morgan pipelines. Enbridge began as the Interprovincial pipeline, moving oil eastward from Alberta; Kinder Morgan began as Trans Mountain, subsequently Terasen, moving oil westward).

As mentioned above, bitumen mined from the oil sands is separated from the sand, clay, and water; up to now it has then been substantially upgraded in Alberta to yield a synthetic crude oil (SCO), as has a portion (7% in 2012, ERCB, *Reserves Report,* ST-98, 2013, p. 3–14) of the bitumen from in situ projects. Indeed the physical characteristics of the bitumen extracted from the tar sands have made further processing mandatory. As one of the executives involved in Canada's first oil sands project wrote (McClements, Jr., 1968):

> Below 50 degrees Fahrenheit it [bitumen] is almost solid and at ambient temperatures above that it is a sticky asphaltic material. It cannot be burned in any but special equipment and it cannot be pumped through a pipeline, even during summer months. Additionally, the market for it is very limited. Consequently, it must be upgraded to pumpable, saleable material competitive with conventionally-produced crude oils.

Recently, a number of refineries capable of handling bitumen have become accessible to Alberta producers. Before shipment, the bitumen is diluted with light hydrocarbons so it will flow readily through the pipeline.

Natural gas with high hydrogen sulphide content is called 'sour' gas. When the hydrogen sulphide is removed from the gas, it is converted to elemental sulphur. Natural gas that is produced from a gas reservoir not associated with oil may require little treatment (other than sulphur removal) before it is delivered to market. However, some non-associated natural gas, and natural gas produced in association with oil, is generally 'wet' and thus is processed at plants near the field to extract useful by-products – propane, butane, and pentanes plus – and remove unwanted ingredients like hydrogen sulphide and carbon dioxide before it is shipped by pipeline to consumers. Typically, the natural gas is then combined with other natural gas from the region and shipped on a main large-diameter ('trunk') pipeline through a natural gas plant ('straddle plant'), where the natural gas liquids are removed or gas that has already been processed is reprocessed. Some 'deep cut' gas processing facilities have been located near to the field to remove NGLs prior to treatment at a straddle plant. This has generated dispute between local deep cut and straddle plant operators about who has primary claim to the NGLs.

B. What Constitutes 'Downstream' Activity?

As mentioned earlier, the term 'downstream' is used to refer to movements of petroleum beyond the main shipment point in the producing region. It includes transportation, refinery, and marketing activities for liquid oil and transportation and distribution for natural gas.

1. Transportation

For at least the past half century within North America, it has generally been most economic to move petroleum from the producing region to the consuming region by large-diameter high pressure pipeline (Lawrey and Watkins, 1982). (Ocean-going tankers may be cheaper for crude oil on one or two specific routes, e.g., from Alaska to California or along the U.S. Gulf Coast.) These pipelines are subject to 'economies of scale,' meaning that the average cost (unit cost) of shipment is smaller the larger the volume moved. Essentially this is because the total volume that flows through the pipeline rises more than proportionately to the diameter of the pipe. It is cheaper to ship large volumes of crude long distances than it is to ship the refined petroleum products (RPPs) derived from the crude, since the RPPs must be kept separate from one another in the line. For this reason, and because they also exhibit economies of scale, the main refineries are large and usually located close to consuming centres. The crude oil moved is normally a mix of the variety of crudes produced within a region (e.g., a blend of Alberta light and medium crudes). Some specific crudes, or types of crudes such as very light or heavy crudes, may be 'batched' and moved separately, usually with a relatively non-permeable petroleum product separating this crude from the blends at either end. Such batching involves higher shipment costs. Heavier crudes and bitumen, such as those from the Cold Lake and Lloydminster areas, are typically blended with lighter hydrocarbons ('diluents,' condensate, or pentanes plus) to increase their fluidity.

For any given diameter pipeline, the flow rate can be changed by varying the number, location, and size of pumping stations that regulate the pressure in the line. In addition, parts of the pipeline may be 'looped' by adding new, parallel pipe. During operation, pipelines must move a continuous flow of petroleum and will exhibit declining unit costs up to the level of capacity for existing equipment. Selecting the optimal capital configuration (e.g., pipeline diameter and associated pumping stations) is a difficult task, particularly when the product being shipped is a depletable natural resource, whose total availability in nature is necessarily uncertain. A system that is too small would leave new discoveries with no immediate access to market, while one that is too large will involve higher than necessary unit transmission costs.

The existence of significant scale economies in pipeline transport implies that the petroleum transmission sector tends to 'natural monopoly' status, that is, a single pipeline (a monopoly) is the most efficient shipment means. However, a monopoly on shipment implies that the pipeline may be able to charge high monopoly prices to users, thereby generating higher profits for itself. There are a number of ways in which such exercise of monopoly power may be limited. First, shippers (crude oil producers and/or refiners) may build the pipelines themselves, thereby precluding the possibility of paying high prices to a third party. Second, pipelines have frequently been subject to legal restrictions giving common carrier and/or rate regulated status. Common carrier status means that all potential users have access to facilities, so that a group of oil companies who build a line cannot deny access to other users. If access to the pipeline is entirely open, markets may develop in which buyers who contract 'space' on the pipeline can trade their space allotments; such active trading can help to keep pipeline tariffs more competitive. Rate regulation has normally insisted that pipeline tariffs be based directly upon shipment costs, and be non-discriminatory (i.e., the same rate for the same service). There are many disagreements about exactly how this is best accomplished.

Both crude oil and natural gas produced in Alberta are shipped to markets outside the province by pipeline. The much lower density of natural gas, and an associated rapid fall in pressure as it moves through the line, mean that the cost of moving natural gas, per unit of energy content, is significantly higher than the cost of moving oil. Therefore, the transportation component of the delivered energy price becomes relatively more important for natural gas relative to crude oil the further the market is from the producing region. Expressed in other terms, the competitive advantage shifts toward oil the further the market is from the petroleum-producing region.

2. Refining

Crude oil is rarely consumed by final users. Instead, the crude is processed by a refinery into a myriad of refined petroleum products (RPPs) to generate an array of hydrocarbon products that best meet the needs of users.

The basic refining process is distillation in which the application of heat to the crude oil vaporizes different hydrocarbon constituents at different temperatures, so they can be separated. Subsequently, the resulting condensed components, ranging from the very light (refinery gases) to the very heavy (asphalt), are subject to a variety of other chemical manufacturing processes to generate a wide slate of separate RPPs. Any given grade of crude oil (e.g., Alberta light and medium blend) will have a particular set of distillation factors, but there are so many further 'cracking,' 'reforming' and other processing techniques available that the final RPP mix from any specific grade of crude oil is potentially quite flexible. More processing, however, involves higher refinery costs. It is technologically possible to build a refinery that could utilize almost any particular grade of crude oil to produce almost any mix of final RPPs. However, once a refinery is built, without any further capital expenditures it can normally accept only a somewhat restricted set of grades of crude oil and produce a restricted set of RPPs. It will be clear that the refinery investment decision is a very complex one, including analysis of the availability and cost now and through the future of differing grades of crude oil as well as the anticipated current and future demands for a large number of RPPs.

It is also noteworthy that refineries are subject to economies of scale; that is the average cost of refining a cubic metre of crude oil declines the larger the refinery (up to a capacity of about 25,000 m^3/d for a typical North American refinery with a relatively high yield of motor gasoline). Thus small regional markets may have few refineries, thereby possibly conveying some market power to the refineries in their sales of RPPs and their purchases of crude oil. The extent of such market power is limited by the possibility of new competitive refineries being built, and the possibility of buyers importing RPPs from other regions (or suppliers of crude oil selling their crude to refineries in other regions).

Within North America, demand and prices (even before retail taxes) have been particularly high for relatively lighter RPPs like motor gasoline and aviation fuel. This reflects our affluence, large distances between urban centres in North America, and the absence of good substitutes for these products. On the other hand, heavy RPPs are commonly used for their heat content in simple combustion processes where natural gas, coal, and even wood will substitute very easily. Since the lighter crude gives refineries a higher yield of the more valuable RPPs, and/or saves the costs of extra equipment needed to obtain a high yield of these RPPs, higher API degree crude oils (lighter crudes) command a premium over heavier oils. A further implication is that the relative values of different grades of crude oil (crude oil price differentials) will change over time in response to changes in the relative supplies of different crude oil grades, changes in demand conditions for different RPPs, and changes in refining operations and technology.

The price for any RPP depends on the price of crude oil in the producing region and the costs of crude transportation and refining, as well as the many factors affecting demand for the specific RPP. Prices vary greatly across RPPs. For example, since crude oil can be burned, it is a good substitute for a product such as heavy fuel oil (HFO). Thus, to sell HFO its price must be lower than that of crude oil. If you wonder how HFO (which has a refining cost) can have a lower price than the unrefined crude oil from which it comes, remember that a refinery is a 'joint product process' that necessarily produces more than just HFO. By analogy, think of the relative values of an ore containing gold and of the separated rock and the gold dust (Adelman, 1972). The economic requirement is that the value of all the refined products together must cover the cost of the crude oil and refining, not that each RPP must have a price higher than that of crude oil.

3. Marketing (Distribution)

RPPs are conveyed to final users in any number of ways. Large users like major manufacturing plants may be connected to the refinery by pipeline. In other instances, the product may be delivered to the final user by railcar or truck (e.g., home heating fuel deliveries), while in still others the consumer may collect the product at a local distribution centre (e.g., a service station). Natural gas is inevitably delivered to users by pipeline. Usually a local utility company performs this service, selling the natural gas to its customers, but some large users may purchase their gas directly from the trunk pipeline or even the producer, bypassing the local utility or paying a transit fee to the utility.

It is the marketing segment that finally brings the supply side of the petroleum industry into contact with society's demand for petroleum products. Businesses and individuals demand services such as warmth or transport or process heat, which must be produced by the combination of energy, capital equipment, labour time, and other materials. From this perspective, the demand for crude oil (or natural gas) is a 'derived demand' that depends upon three major types of factors:

(i) demand for the basic services that utilize energy;
(ii) production technologies and input supplies that produce those services within firms, governments, households, and other enterprises; and
(iii) supply conditions in the transportation, refining, and marketing sectors of the upstream part of the industry, which are necessary to make crude oil usable by businesses and households.

C. How Is the Industry Organized?

As just discussed, the petroleum industry involves six linked stages that effect the transfer of oil from naturally occurring deposits to final users: exploration, development, lifting or extraction, transportation, refining, and distribution or marketing. Petroleum can be sold at any point along the continuous flow, even within any one of the six stages. For example, when Petro-Canada purchased the downstream assets of BP in Canada, it bought some oil that was partway through the refining process. Most petroleum companies exhibit some degree of vertical integration, extending over more than one of the six stages. This is so customary over exploration, development, and production activities that such firms are normally labelled crude petroleum producers and are often called 'independents.' The term 'integrated' refers to companies that do more than one of crude petroleum production, transportation, refining, and marketing. Within the Alberta petroleum industry, large vertically integrated companies have been active participants from the very beginning, but so have companies that have operated at only one level of industry activity.

There are, as a result, a number of different organizational structures used by petroleum firms. Most large integrated oil companies, for example, have a 'production' section or department whose responsibility it is to search for and establish a supply of crude petroleum that is sufficient for other aspects of the company's operations. There may even be exploration subsidiary companies. Whatever the organizational set-up, the purpose is to keep the company's reserve situation under constant review, actively bringing forward new projects to add to existing reserves, or replace depleted reserves, as well as to ensure that these reserves are scientifically and economically exploited. Which activities are delegated to the production arm of a company depends to a large extent on the individual circumstances under which a company operates. For example, the production department of a large integrated company, in addition to being responsible for the more obvious activities relating to exploration, development, and extraction, may also be in charge of processing the petroleum lifted from the wellhead to meet pipeline specifications and for arranging transportation of this crude to the company's refinery. A smaller independent production company may confine its operations to finding and bringing petroleum to the surface. And, since most of these smaller firms do not have processing facilities of their own, the petroleum is delivered to gathering systems (oil) and gas-processing plants for treatment. But these processing costs come out of the producer's pocket, and so, even though the company has no direct involvement with the processing of the petroleum, this activity can properly be included as part of the production process.

Government-owned oil companies, some vertically integrated, have been popular in many parts of the world. In the industrialized western world, these have generally operated in competition with privately owned companies but have been established for a variety of reasons. For example, for strategic reasons, the government of the UK, in 1913, provided financing and took over ownership control of the Anglo Persian Oil Company (the forefather of BP). Statoil in Norway was founded in 1972, and Petro-Canada in Canada in 1975, to increase domestic ownership in the petroleum industry and to provide a 'window' on industry activities. (Both BP and Petro-Canada have since been privatized.) Some developing nations, for example Brazil, have also established state-owned oil companies to operate in competition with privately owned companies (often with special advantages). Many developing countries (for example, most OPEC members) have established a government oil company (or nationalized private companies) to leave a single state oil company in operation.

In conclusion, the consensus would be that all activities beginning with initial geological and geophysical studies to determine whether there is petroleum underneath the surface up to the stage where the oil and gas is to be stored or transported to refiners or market should be classified as 'crude petroleum production' activity. We use this term in the broadest sense of the word, for, as we have seen, a more narrow interpretation of 'production' (also referred to as 'extraction' or 'lifting' or 'operation') represents only the final stage of the crude oil or raw natural gas production process, being preceded by exploration and development.

It is the crude petroleum production industry, so defined, which has been of such significance in Alberta. We will not emphasize the specific organizational structures and decision-making methods that various producing companies have utilized. Rather, we shall be concerned with the aggregate activities of the industry with particular emphasis on the production of crude petroleum and the operation of crude oil markets. That is, our main interest is the upstream industry. However, we do not consider in any detail natural gas liquids, natural gas processing, or gathering pipelines.

3. What Are the Economic Aspects of the Petroleum Industry?

A. What Is the Economic View?

Obviously analysis of the petroleum industry must be based on the physical realities of petroleum production and use. However, purely physical factors are an insufficient basis for private or social decision-making. Deposits of crude oil and natural gas in Canada are a natural resource available for human use. Whether humankind is broadly defined ('all people of all generations'), or more narrowly ('today's Albertans'), or more narrowly still ('shareholders in oil companies'), the fact remains that the value of the resource is not inherent in its physical characteristics but must be mediated through people. Economic analysis stresses that the production of a resource such as petroleum normally contributes both positive and negative effects to the parties concerned. There are costs involved in devoting effort to the production of oil and in depleting nature's stock, but the consumption of oil generates beneficial work or products. It is a presumption of most economic analysis that decision-makers are interested in deriving the maximum 'efficiency' from energy resources, where economic efficiency, in its broadest sense, means maximizing the excess of the benefits of resource use over the costs.

In this book we shall consider petroleum investment, production, and consumption from two points of view, the 'private' and the 'social.' The 'private' view is that of specific oil companies or consumers who make decisions about the utilization of Alberta's petroleum resources and whose activities are mediated through the operation of economic markets. The basic simplifying assumption is that these decision-makers evaluate benefits and costs from a 'private' perspective, seeking to gain the maximum net benefits for themselves. Companies producing crude petroleum, for instance, are generally understood to be trying to maximize the profits received from their activities. Economists would not claim this as an accurate description of the behaviour of every decision of every private decision-maker, but it is a useful working assumption that usually leads to relatively accurate depictions of individual transactions and aggregate behaviour in economic markets.

From a 'social' perspective, we will be largely concerned with assessing the desirability of certain public policies. We will assume that government decision-makers assess social efficiency with a view of benefits and costs that reflects the overall interests of the society they represent, not just private interests. Many differing definitions of social or economic efficiency are possible since different value systems or criteria exist and each would define social benefit and social cost within its own framework. Despite this, many economists have found one particular definition of social efficiency to be useful in policy analysis, in part because it leads to a concept of benefit and cost that is often readily measurable, but also because it derives from a particular value system that appears to enjoy wide acceptance. This view of efficiency assumes that social benefits and costs are equal to the net sum of the dollar values all individuals in society associate with the benefits (positive) and costs (negative) that they perceive. The approach is, therefore, individualistic rather than paternalistic; the measuring rod for the intensity of feeling of the individual is his or her 'willingness to pay' (that is, a particular dollar value). If, on balance, an energy policy measure generates a positive aggregate dollar sum of social benefits less social costs, it is judged by the social efficiency criterion to be a desirable policy.

It is important to realize that benefits and cost in the efficiency criterion are measured by reference to all members of society, irrespective of whether they are directly involved in the production and consumption of the product. One implication is that the social costs of petroleum may exceed the private cost to the individual oil company, for example, if there are environmental costs associated with production. Similarly private benefits of oil use to consumers may understate social benefits, for example, if there are national security benefits. This view of economic efficiency is controversial, and those who advocate its usefulness are generally careful to assert that it is only one of the objectives that social decision-makers may view as relevant. Often, for instance, the distribution of benefits and costs, as well as their net sum, matters. Moreover, there are conceptual ambiguities in measuring benefits and costs (Blackorby and Donaldson, 1990) as well as major practical difficulties in measurement and implementing policies based on economic efficiency.

There are also purely physical (thermodynamic) definitions of the efficiency of energy use. In the eyes of the economist, such measures of efficiency are an insufficient basis for social decision-making since they abstract from the values that people place on energy and its production (Berndt, 1977), whereas the economic concept of efficiency relies directly on such human-generated values.

B. How Does the Economic View Reflect Physical Reality?

If we adopt the private firms' perspective for the moment, the largely physical definition of economic activity in the earlier part of this chapter can be translated into the necessity of undertaking dollar expenditures (costs) in order to receive dollar revenues (benefits). The various stages of industry activity, then, are perceived not so much in terms of the physical activities performed as by the expenditures (on capital, labour, energy, materials, land, and taxes) and the sales receipts that result. The critical role of economics to industry behaviour cannot be denied. A strong expectation that petroleum deposits exist in nature in a region will not generate exploratory activity unless companies anticipate a positive net

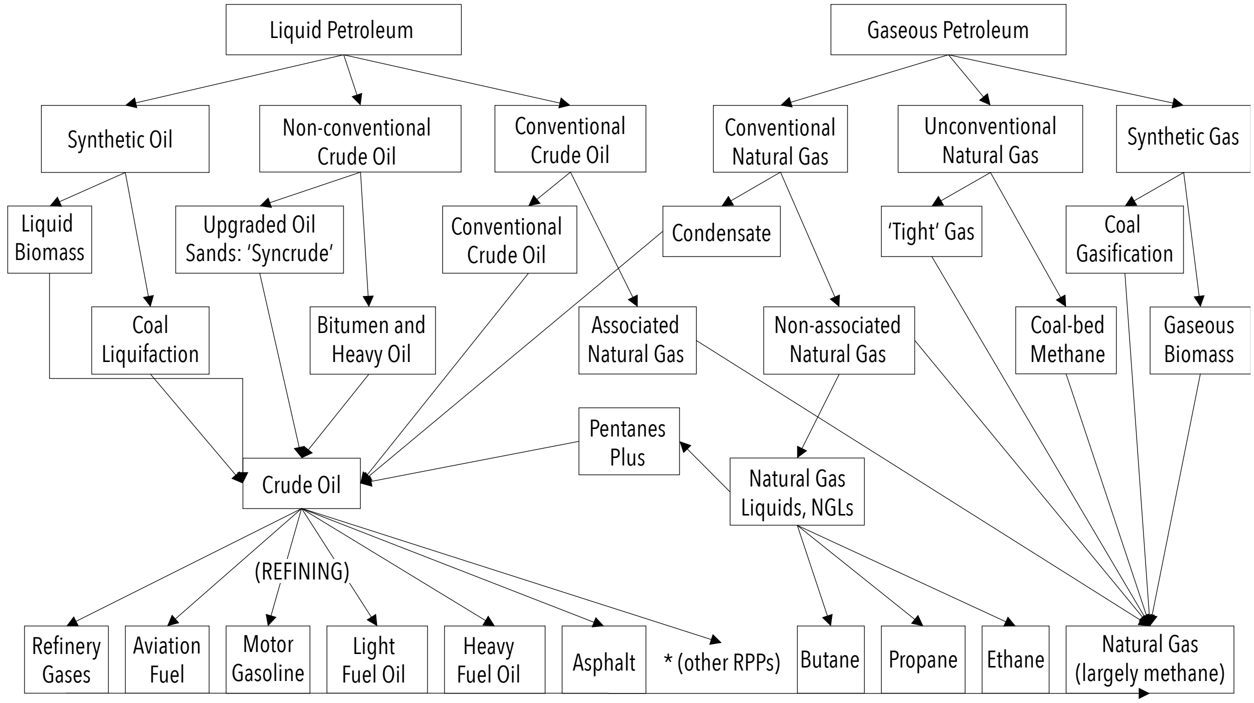


Figure 1.1 Types of Petroleum

return after tax from their actions. The clear presence of crude oil in an exploratory well core will not guarantee development unless anticipated revenues will exceed the expenses needed to develop and lift the oil. A positive petroleum flow rate from an established well will not normally be allowed by a company unless the revenues from the oil are high enough to cover the operating costs of the well (including royalties and taxes). In this manner, all the important physical attributes of the industry generate economic effects that decision-makers will consider.

While the physical realities of petroleum production will generate economic implications for those parties producing petroleum, it is difficult for the external economic analyst to examine the activity of individual decision-makers. In most cases, whether the purpose is descriptive or normative (i.e., policy generation), economic analysis provides a simplified view of petroleum industry activities, but one that captures the essence of what is occurring. Simplification is necessary in part because of data limitations, since accurate statistics on all specific transactions are not available. More fundamentally, however, there are so many complex individual actions that they could not possibly be considered in a single analysis. Instead, simplified 'models' of petroleum industry activity are necessary.

The joint product nature of the industry generates particular problems for the analyst. As mentioned above, a joint production process is one in which a particular activity necessarily generates more than a single output, so that the markets for the products are linked (at least on the supply side). For example, a petroleum refinery produces an entire slate of separate RPPs. An exploratory well generates some general geologic knowledge applicable to other drilling sites, specific knowledge about this site, some estimated petroleum reserves (if it is successful), and some equipment that is utilizable in lifting oil to the surface. The fluids lifted through a petroleum well include a mix of products (hydrocarbons such as crude oil and natural gas, and non-hydrocarbons such as sulphur).

Some analysts try to handle the joint product problem by combining the different outputs into a single product. For example, natural gas discoveries can be converted to 'volumes of oil equivalent' by assuming some equivalency factor. Relative energy content (about 6Mcf of gas per barrel of oil) or relative market values are often used, but there is no obviously correct conversion factor because oil and natural gas have separate and non-interchangeable markets, reflecting the fact that the two products are far from perfect substitutes in use. Another approach

Table 1.1: Petroleum Industry Activities: Physical and Economic Aspects

	Physical Activities		*Economic Activities (oil)*	
Stage	*Description*	*Final Product*	*Costs (Inputs)*[(b)]	*Benefits (Output)*[(c)]
1. Exploration[(a)]	Geological and Geophysical (G&G) surveys; Exploratory drilling	Knowledge of presence or absence of petroleum-bearing geologic formations	Costs of K, L, E, M; T	Price of Undeveloped reserves Price of G&G data
2. Development[(a)]	Installation of Production Facilities; Extension and Infill drilling; EOR; Pumps, Gathering lines and separation equipment; Gas plants	Established reserves; productive capacity	Linkage: undeveloped reserves Costs of K, L, E, M; T	Price of developed reserves
3. Production[(a)] (Extraction or Lifting or Operation)	Bringing petroleum to surface separation of products; gas plant processing; shipment to regional gathering location	Crude oil (and natural gas, NGLs, sulphur) at input terminal of main pipeline or transportation facility	Linkage: developed reserves Costs of K, L, E, M; T	Price of crude oil, f.o.b., in producing region
4. Transportation	Movement of petroleum from producing region to export point or city gate (gas) or refinery gate (oil)	Crude oil at border or at refinery gate; natural gas at border or delivered to distributor	Linkage: Crude oil, f.o.b. Costs of K, L, E, M; T	Price of crude oil, c.i.f., at refinery gate
5. Refining	Distillation, catalytic processing etc. of crude oil into refined petroleum products (RPPs)	RRPs (gasoline, kerosene, fuel oil, etc.) at refinery gate	Linkage: Crude oil at refinery gate Costs of K, L, E, M; T	Prices of RPPs at refinery gate
6. Marketing (Distribution)	Conveyance of RPPs and natural gas to final users	RRPs and natural gas at point of final sale	Linkage: RPPs at refinery gate Costs of K, L, E, M; T	Prices of delivered RPPs

Notes: (a) Exploration development and production together make up the crude petroleum industry, or petroleum upstream industry.
(b) K, L, E, M, are the services of capital, labour, energy, and materials used in production processes. T represents royalties, taxes and land payments. f.o.b. means 'free on board', before shipment costs.
(c) Unit benefits (prices) have been shown; total benefits or values are the price multiplied by the quantity of output; c.i.f. means 'cost, insurance, freight', that is, after shipment costs.

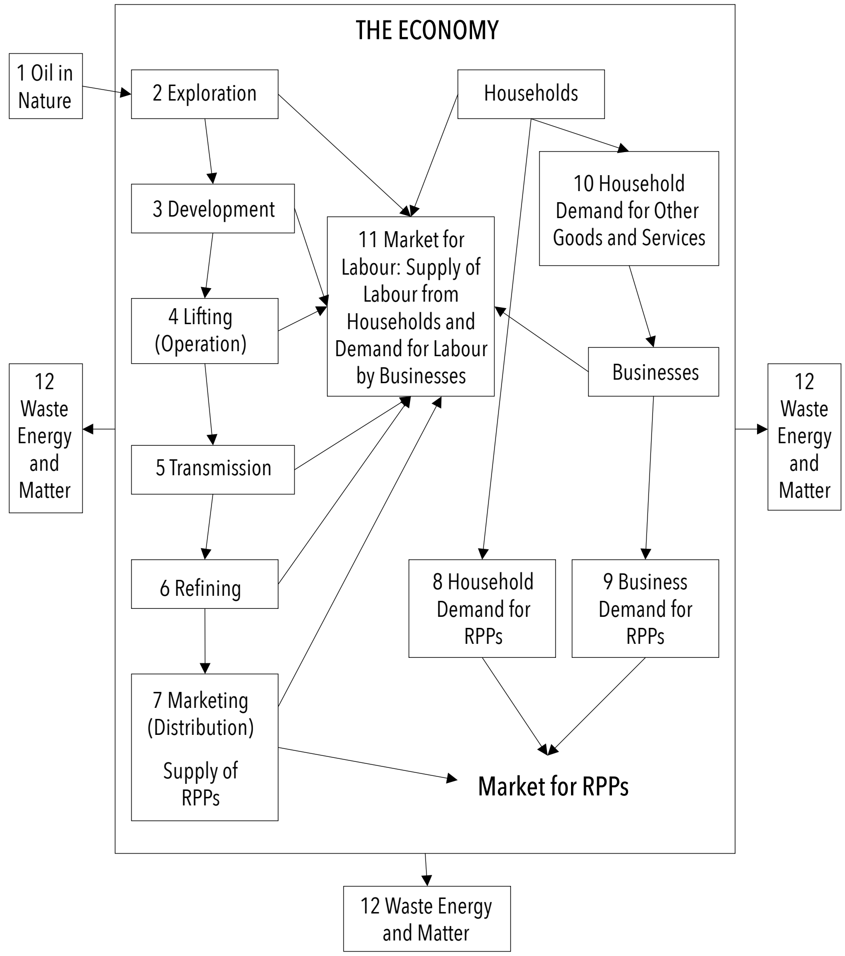
is to divide the exploratory effort between the different products; for example, the proportion of total discoveries that are oil, or the percentage of total successful well metres drilled that was in oil discoveries, might be used. The important point is that there is no correct way to combine outputs or separate inputs in a joint product process. Any attempt to do so leads to economic fictions. However, analysis is made immeasurably more complicated if one must build economic models that always include all inputs and all outputs of the joint product activity. As is invariably the case, the economist must balance the costs of a theoretically more complex but realistic model against the benefits of unrealistic simplification. Increasingly, there is acceptance of the necessity of facing up to the joint product nature of petroleum industry activities and rejecting attempts to avoid it by combining two distinct products into one, or artificially separating a single activity into two.

4. Conclusion

Our introductory chapter concludes with three summary depictions of the petroleum industry.

Figure 1.1 sets out the main physical products the petroleum industry generates, divided into liquid and

Figure 1.2 Petroleum in the Economy



gaseous products including those drawn directly from nature and 'synthetic' petroleum and RPPs that are not immediate natural resources. Petroleum drawn directly from nature includes 'conventional' oil and natural gas, which are drawn from deposits where some of the petroleum flows to the surface through wells drilled into the deposit; this is the traditional ('conventional') way of producing petroleum. Non-conventional crude oil comes from oil sands and oil shale deposits holding such heavy, viscous oil that none will naturally flow to the surface through a well, so 'non-conventional' production techniques must be used. Synthetic oils are liquid hydrocarbons generated from some other natural resource such as coal or biomass. Historically, almost all of the world's petroleum has come from conventional crude oil, natural gas, and natural gas liquids. This book will focus on conventional crude and natural gas, which have been the mainstay of the Alberta petroleum industry. Alberta is one of the few areas in the world with significant non-conventional oil production from its oil sands, so we also incorporate this resource.

Table 1.1 provides an overview of physical and economic views of the six stages of petroleum production, including physical descriptions of the activities and output of each stage and brief summaries of the economic costs and benefits to private decision-makers. Exactly how an oil company perceives the costs and benefits will differ depending on whether the company is vertically integrated or not. A vertically integrated company absorbs all the costs involved in the various physical stages of activity but generates benefits only at the downstream stage where sales occur. In the table, we have made some allowance for this by indicating at each stage a 'linkage cost' that ties this stage to the one just upstream. For a vertically integrated firm, the linkage cost is the *costs* of the previous upstream activities, while for a non-vertically integrated firm it is the *sales price* of the output of the adjacent upstream activity. Of course, vertically integrated firms are never perfectly 'in balance' throughout all stages of petroleum industry activity; they will buy or sell some oil at most stages of activity.

Figure 1.2 is a flowchart that sets the activities of the petroleum industry within the larger economic system. The entire economic system is indicated by the large rectangular box, with the (unknown) natural endowment of petroleum deposits situated outside (1). The six stages of petroleum industry activity are shown (2 to 7), as is refined petroleum product (RPPs) demand by households (8) or business and other enterprises including government (9). The business demand for petroleum products derives from the production of goods and services, which is in turn derived from the household demand for goods and services other than petroleum products (10). Some indication of aggregate economic effects of the petroleum industry (the "macroeconomic" effects) can be seen in the box showing the labour market (11). The six stages of petroleum production serve, along with other businesses, as a source of demand for the labour services of households. Similarly, although it is not shown in Figure 1.2, the petroleum industry demands the goods and services (capital, material, and other energy products) produced by businesses. The more active is the petroleum industry, the higher these demands. Finally, all those activities from exploration through to final consumption involve the ejection of waste heat and matter (12) back into the environment.

Chapter Two provides an overview of the Alberta petroleum industry, illustrating how the stages of industry activity developed there from 1950 to 2010, and introducing our theme of 'petropolitics' in the form of questions about the role of government in regulating the industry.

CHAPTER TWO

An Overview of the Alberta Petroleum Industry

Readers' Guide: Chapter Two provides a bridge from the discussion in Chapter One of what the petroleum industry does to the substantive description of the crude petroleum industry in Alberta, which occupies the remainder of this book. In this chapter we illustrate industry activity in Alberta for select years from 1950 through 2010. We also introduce the concept of 'petropolitics,' as we set out possible reasons for, and forms of, government regulation. Finally, we touch on some aspects of Alberta industry activities downstream from the crude petroleum phase, which will not be dealt with in this volume.

1. Introduction

Contrasts are often suggested between the neighbouring prairie provinces of Alberta and Saskatchewan; the Albertan may regard Saskatchewan with a certain old-fashioned nostalgia, and the Saskatchewan resident may be drawn to Alberta's dynamic multifaceted society. These images probably speak more of stereotypical simplification than reality, but they do draw attention to differences in the economic history of the two provinces.

In 1946 Alberta's population was 800,000, Saskatchewan's 830,000. By 2013 Saskatchewan's population had grown to just over 1,000,000, while Alberta's was approaching 3,900,000. Before 1946, both provinces had been molded primarily by the wheat boom early in the century and their agricultural resources. Both economies have diversified since the end of World War II, but it is hard to avoid the conclusion that development of Alberta's oil and gas industry lies at the foundation of a coherent explanation of the very different provincial growth paths. Alberta's petroleum resources, especially crude oil, provided the base for a classic natural resource boom: strong export markets for oil and, later on, natural gas attracted new capital investment to the industry, and drew immigrants. The influx of people, plus backward and forward linkages from the crude petroleum industry into input supply and transmission and processing industries, generated further investment spending, especially in construction. The financial and service sectors expanded and government coffers swelled.

This chapter follows up on Chapter One's general review of the petroleum industry by providing a broad overview of the activities of the industry in Alberta. We illustrate how the industry's activities evolved from 1945 to the present and set out the major policy decisions faced by corporations and governments. We will begin by looking at the position of the Alberta petroleum industry in 1946, just prior to the start of the 'Oil Boom.' We then provide preliminary information on the operation of the six stages of the petroleum industry in Alberta. Subsequent chapters in this book will explore in greater detail a number of the issues introduced here with regard to the Alberta crude petroleum industry.

Table 2.1: Selected Statistics for the Alberta Crude Oil Industry

		1950	1960	1970	1980	1990	2000	2010
1	Petroleum and Natural Gas Leases outstanding (10^3 acres)	4,410	30,004	46,348	60,807	65,550	81,706	87,898*
2	Western Canada Survey Crew Months	1495	632	936	n/a	n/a	n/a	n/a
3	Net G&G Expenditures 10^6 current $	24.5	33.0	80.3	452.3	452.0	824.1	716.1
4	Net G&G Expenditures 10^6 1990$	160.1	164.2	292.0	739.1	452.0	695.4	488.8
5	Number of Exploratory Wells	224	403	963	2,623	2,286	3,114	799
6	Exploratory Drilling Expenditures 10^6 current $	11.0	44.0	83.8	1,621.6	950.2	2,364.8	2,145.2
7	Exploratory Drilling Expenditures 10^6 1990$	71.9	218.9	304.7	2,469.7	950.2	1,995.6	1,464.3
8	Initial Oil Reserves Discovered 10^3 m^3	81,809	11,910	9,439	29,357	12,990	11,842	735*
9	Initial Gas Reserves Discovered 10^6 m^3	31,616	126,337	24,004	135,545	32,177	53,884	5,079*
10	Oil Discoveries per Exploratory Well 10^3 m^3	366	30	4	11	6	4	1.8
11	Gas Discoveries per Exploratory Well 10^6 m^3	141	313	25	52	14	17	11.4
12	Exploration Success Ratio	.25	.33	.25	.61	.53	.79	.84
13	Oil as % of Exploration Successes	62	31	23	23	32	19	42
14	Initial Established Conventional Oil Reserves 10^6 m^3	174.8	711.6	1,734.3	1,913.2	2,256.1	2,534.4	2,829.7
15	Remaining Established Conventional Oil Reserves 10^6 m^3	152.6	525.0	1,207.9	719.9	510.4	291.4	236.9
16	Initial Established Gas Reserves 10^9 m^3	146.8	935.9	1,593.2	2,647.1	3,286.8	4,063.5	5,213.5
17	Remaining Established Gas Reserves 10^9 m^3	129.0	926.8	1,352.0	1,812.1	1,694.1	1,210.7	991.4
18	Number of Development Wells	788	1,363	884	4,425	1,969	8,953	7,103
19	Development Success Ratio	.94	.83	.76	.88	.81	.94	.95
20	Oil as % of Development Successes	97	84	37	31	55	35	49
21	Development Expenditures 10^6 current $	67.5	165.0	327.7	2,340.7	2,566.7	8,105.6	15,600.3
22	Development Expenditures 10^6 1990$	441.2	820.9	1,191.5	3,824.7	2,566.7	6,840.2	10,648.5
23	Conventional Oil Output 10^6 m^3	4.3	20.8	52.4	63.2	53.0	43.5	26.6
24	Oil Sands and Bitumen Output 10^6 m^3	0	0	1.9	8.0	19.9	35.3	84.7

/continued

2. Before the Boom: 1946

The embryo of an active petroleum industry was well-established in Alberta by 1946. Geological potential had long been recognized. Rocks of the Canadian Shield lie exposed in the North East corner of the province, and the Rockies run along the southwest border, but 90 per cent of Alberta's 662,000 square kilometres is overlain with sedimentary strata, part of the extensive central North American sedimentary basin that runs from the Mackenzie River delta south to the long-established petroleum producing fields of Texas and Louisiana. Tangible evidence of petroleum was familiar to members of the First Nations and immigrant settlers, exemplified by natural gas seepages near Waterton and surface crude showings along the shores of the Athabasca River near Fort McMurray. Commercial natural gas production had begun late in the 1800s, after a water-directed well hit a shallow natural gas deposit near Medicine Hat, and at least three commercial petroleum booms had been stimulated by discoveries in the prolific Turner Valley field, southwest of Calgary. A small, shallow, Lower Cretaceous pool of wet gas, laced with light crude, was discovered just before World War I; in 1924, a deeper wet gas pool (in the Mississippian Rundle formation) came in; and a large, deeper oil pool in the Rundle formation was tapped in 1936. A provincial oil and gas conservation board was established in 1938 with regulatory powers over the industry. Oil output, mainly from Turner Valley, grew in the early years of the Second World War, peaking in 1942.

Other small exploratory successes occurred during the 1930s and 1940s, but Alberta's potential as a major oil producer was in doubt. Many communities, including Edmonton and Calgary, were connected to local natural gas supplies, but Alberta, by 1946, was a net importer of crude oil and refined petroleum products. Most of the major North American oil exploration companies were becoming disenchanted. Then, in February 1947, a crew drilling one of the last exploratory wells that Imperial Oil planned for Alberta struck

Table 2.1/continued

		1950	1960	1970	1980	1990	2000	2010
25	Marketed Gas Output 10^9 m^3	1.4	9.1	42.9	62.1	84.6	142.2	112.8
26	Field Operating Expenditures 10^6 $	10.5	55.5	119.4	950.7	3,604.6	4,832.4	10,055.1
27	Field Operating Expenditures 10^6 1990$	68.6	276.1	434.2	1,553.4	3,604.6	4,078.0	6,863.5
28	Total Operating Expenditures 10^6 $	16.5	92.0	219.5	1,679.2	4,445.4	5,751.2	12,097.6
29	Total Operating Expenditures 10^6 1990$	107.8	457.7	798.2	2,743.8	4,445.4	4,853.2	8,257.7
30	Oil R/P Ratio	35.4	24.9	23.2	11.3	9.6	6.7	8.9
31	Gas R/P Ratio	88.0	96.2	31.9	29.2	20.0	8.5	8.8
32	Conventional Crude Revenue 10^6 $	80.6	318.4	844.1	6,185.1	8,209.9	10,756.8	12,302.5
33	Gas revenue 10^6 $	2.9	31.3	243.9	5,121.1	4,667.0	23,091.1	15,564.7
34	Oil Sands Revenue 106 $	0	0	32.8	1,676.7	2,799.4	8,044.4	36,690.0
35	Oil Average Sales Revenue $/$m^3$	18.74	15.09	16.07	97.69	154.61	244.72	450.07
36	Oil Average Sales Revenue 1990$/$m^3$	122.48	75.07	58.44	159.62	154.61	208.68	307.48
37	Gas Average Sales Revenue $/$m^3$	2.04	3.46	5.69	82.51	55.18	162.34	137.98
38	Gas Average Sales Revenue 1990$/$m^3$	13.33	17.21	20.69	134.82	55.18	137.00	97.50
39	Gas Price Relative to Oil (energy content basis)	0.11	0.23	0.36	0.85	0.36	0.66	0.29

Sources and Notes:

*2009 is the last year for which data is available.

Row 2: Values are for the end of the fiscal year (March 31) except for 2000 and 2009, which are as of December 31. From *Annual Reports* of the relevant Alberta government department (Mines and Minerals, Energy and Natural Resources and Energy); 2000 and 2009 from Dianne Johnston, Department of Energy.

Rows 3, 4, 7, 9, 10, 22, 24, 25 26, 27, 29, 33, 34, 35, 36, 38: CAPP *Statistical Handbook*.

Rows 6, 13, 14, 15, 16, 17, 18, 19, 21: ERCB *Reserves Report* ST-98.

Rows not mentioned are calculated by the authors.

The Canadian GDP price deflator was used to calculate 1990 real dollar values.

oil in upper Devonian rock (the D-2 formation). A neighbouring 'stepout' well proved disappointing in the D-2 formation but was extended deeper in May, with good production flow from the D-3 formation. The Leduc success, following over 130 Imperial dry holes, marked the start of the Alberta oil boom.

3. Alberta's Upstream Petroleum Industry

A. Exploration

Exploratory activities, as discussed in Chapter One, include land (mineral rights) acquisition, geological and geophysical (G&G) prospecting, and exploratory drilling.

1. Land Acquisition

The legal right to explore for, develop, and produce petroleum must be acquired by petroleum companies from the owner of the mineral rights. The mineral rights on more than four-fifths (81%) of Alberta's area are held by the provincial government ('Crown land'); the federal government holds another 9 per cent, on Indian reserves and, primarily, in National Parks; the remaining 10 per cent is held by private individuals and companies as 'freehold' rights on land issued to the Hudson's Bay Company, railroads, or homesteaders prior to 1887 (Alberta Department of Mines and Minerals, April 1972). Petroleum companies typically lease below-the-ground mineral rights from owners of mineral rights for a specified period of time, or as long as petroleum production might occur. In return, the original mineral rights owner is compensated, often with a lump sum bonus payment when the deal is signed, an annual rental, and a royalty out of any petroleum revenues. Companies must also obtain the right to the use of the surface of the land for purposes of exploration and production facilities.

Table 2.1 shows companies' holdings of Alberta Crown leases in selected years since 1946. (To illustrate the development of the industry in Alberta, this chapter includes data on the industry at ten-year intervals, commencing in 1950). Oil companies must hold leases on the land before producing oil. In addition, a variety of reservations permits and licences have been issued to allow exploration and were partially or wholly convertible into production leases. By the 1980s most mineral rights were issued in the form of leases. The total amount of land held under some form of permit rose sharply after 1950 as unexplored parts of the province became of interest to companies and as leases above discoveries were retained to allow production. Companies allow rights to lapse on explored plots that do not appear to be economically productive, thereby saving rental payments. As the industry matures, the total area held in mineral rights can be expected to level out and eventually decline. This has not yet become apparent in Alberta, although land acquisition appears to have become more focused on natural gas and oil sands prospects than on conventional crude oil.

2. Geophysical and Geological (G&G) Surveys

As was discussed in Chapter One, G&G surveys precede costly exploratory drilling, in order to locate the most promising drill sites. Table 2.1 includes measures of such exploratory effort for select years.

Unfortunately, consistent data on crew months is available only to the 1960s. For 1950 and 1960 we report the number of months of effort during the year by petroleum survey teams in Western Canada (mostly in Alberta). There has been significant knowledge growth and technological change in G&G activities, especially from the mid-1980s, with the development of new computer techniques including 3-D and 4-D seismic surveys. As a result, a crew month in the year 2010 was more productive than in the year 1950. Surveys occur early in the life-cycle of industry activity, as low-cost information gathering. However, G&G work will continue, and even grow in later periods, as new areas or deeper formations become of interest, as new companies commence exploration and undertake their own surveys, and as growing scientific knowledge develops new G&G techniques or interpretations.

Table 2.1 also includes data on G&G expenditures in Alberta, both in current (nominal or 'as-spent') dollars and in dollars of 1990 general purchasing power ('real 1990 dollars'); the latter shows the size of expenditures after allowance is made for general inflation, therefore showing changes in expenditure by this sector in terms of general purchasing power in the economy. From an economic perspective, the real expenditures are of most interest, since an increase in expenditures simply because price levels have been rising would be quite misleading. Nominal G&G spending tended to increase over the years. Real spending does not show a clear trend across time, presumably reflecting factors such as varying real oil and gas prices, changing views about exploratory prospects and different government policies.

Note that choice of the 'correct' adjustment for inflation is difficult. There is some ambiguity about the meaning of the concept of 'real' expenditures. We may intend it to refer to the activity undertaken by this sector of the economy, or what we might call the 'quantity of effort expended.' In this case one would wish to adjust current dollar expenditures by a price index specific to the activities of this sector. However, such detailed price indices are not readily available, so it is generally necessary to rely on a broader price index. It would also be desirable to modify expenditures to reflect technological (quality) improvements, but this is hard to do. Alternatively, real expenditures might refer to purchasing power in the economy at large, that is, what quantity of goods 'in general' could be bought by the expenditures of this sector. In this case, the adjustment index should be a general price index such as the Canadian GDP price deflator. However, there are different inflation rates across regions, and differently defined 'reference bundles' of goods, so that a number of general price indices are available. In a relatively open economy such as Canada's, the longer-term trends in broadly defined price indices are much the same, but this is not necessarily true for narrow indices, such as for one specific economic activity.

3. Exploratory Drilling

In Chapter One we said that exploratory wells generate two main products: geological knowledge (which is normally generalizable beyond the specific drill site) and petroleum discoveries. Most exploratory wells, particularly 'wildcat' wells located some distance from previously discovered pools, are 'dry,' not recording a commercial find.

Knowledge is particularly important for the initial wells drilled in a geographical area or through particular geological formations. The first exploratory well to discover a large pool in an entirely new formation is particularly productive for both knowledge and petroleum; significant examples in Alberta include Imperial Oil Leduc No. 1 in 1947; which, as noted above, signalled the beginnings of Alberta as a major crude oil producer, Socony Seaboard Pembina No. 1 in 1953; and Banff-Acquitaine Rainbow West in 1965.

Each of these defined a new geological 'play,' and set off a major surge in exploratory activity and oil discoveries in Alberta.

Table 2.1 shows petroleum industry exploratory drilling activity, including the total number of exploratory wells drilled, the oil and gas reserves discovered, the exploration drilling success rate, and current and real (constant) dollar exploration expenditures. Another measure of exploratory activity is the total exploratory drilling footage in a particular year; we have not included this variable. In models and descriptions of exploration, exploratory drilling effort is variously measured as the number of wells drilled, the total exploratory footage drilled, and the real expenditures on exploratory drilling. These measures are correlated with one another, but not perfectly. Thus, technological improvements might allow the same number of wells or footage to be drilled at a lower real expenditure; fewer wells might be drilled, but footage and expenditures increase if the average depth of wells rises. From a modeling perspective, different results might be attained in the same empirical model, depending upon which measure of exploration effort is used.

Reserves are volumes of petroleum known with a relatively high degree of certainty to be recoverable under current economic and technological conditions. The reserves reported in Table 2.1 are estimates made in the year 2009 of the size of reserves discovered in past years, that is, initial reserves (before any production) as reported in the discovery year and as 'appreciated' or revised since then. In this appreciation process, reserve additions reported in any year for an oil pool or gas reservoir are credited back in time to the year in which the pool was discovered. As was reported in Chapter One, most reserves are credited due to development activities ('extension' or 'outpost' wells) in years after the pool is discovered. Pools discovered in 2000 and 2009 have had fewer years for such appreciation to occur, so reserves discovered may be understated relative to earlier years. The success rate is the proportion of exploratory wells that discovered oil or gas pools. As can be seen in Table 2.1, there are much higher success ratios for the years shown after 1980. This could reflect a number of factors, such as a fall in the proportion of wildcat wells, improved technology allowing increased efficiency in selecting drilling sites, and more emphasis on outpost drilling in natural gas pools. Table 2.1 shows that discoveries tended to shift towards natural gas over the fifty-year period, although relatively low gas prices near the end of the period led to renewed oil-directed exploration.

Year-to-year data show large fluctuations in the industry's exploratory drilling activity in Alberta. The general trend was upwards, at least until the early 1980s, as Table 2.1 suggests. Exploratory drilling fell off in the 1980s, but then picked up again by the year 2000, although real expenditures were still smaller than in 1980. The cyclical variations reflect mainly the succession of new petroleum plays, the variability of oil and gas price expectations and changes in government tax and other regulations. The incentive to drill and obtain general geologic information tends to be strongest in the early years of industry activity, as is true of G&G work. However, the number of specific drilling sites in Alberta is very large, so, for many years, new knowledge and any increases in price, or cost-reducing technological improvements, will tend to make a significant number of new potential drilling locations attractive. As Table 2.1 makes clear, these new drilling sites have become less and less productive, as shown by the decline in the volume of reserves discovered per exploratory well drilled, although the decline would be smaller if more recent finds were adjusted to allow for future reserve appreciation. This falling finding rate typifies what many economists call a 'stock' or 'degradation' effect: as the stock of undiscovered resources in a given geological play becomes smaller, new discoveries require more effort (i.e., tend to become more costly). In short, diminishing returns emerge. Table 2.1 suggests that the degradation effect has not operated in Alberta through deposits becoming harder to find (the success rate has not shown a persistent tendency to fall); rather, discoveries have been becoming much smaller on average. Table 2.1 also suggests that discoveries have tended to shift towards natural gas, as oil productivity has declined, and as natural gas markets have grown.

The reserves per well drilled in Table 2.1 are only roughly indicative because petroleum exploration is a joint-product process. The industry's exploratory drilling produces: (1) knowledge, (2) oil discoveries, and (3) gas discoveries. It is impossible to specify what proportion of the total wells drilled was necessary to produce, separately, any one of these three products. One of the three may have been dominant in the mind of the company drilling, but we rarely have access to this information.

B. Development

Recall, from Chapter One, that development activities by the petroleum industry are concerned with 'proving up' reserves (by demonstrating the existence beneath the surface of commercially recoverable petroleum volumes) and providing productive capacity that can be used to lift oil, that is by installing such capital equipment as completed wells, enhanced oil recovery (EOR) injection facilities, water disposal wells, and gathering and separation equipment.

While exploration is necessary to locate oil pools, most oil reserves (except in very small deposits) are added through development activities, particularly extension drilling and EOR investments. Historical experience in Alberta suggests that the amount eventually recovered from a typical oil pool (excluding the smallest ones) will be on the order of nine times the reserves estimated to be present on the basis of the discovery well (four times for gas). More accurately, ultimate oil recovery will be nine times the first year's estimate of recovery.

Table 2.1 shows how total initial Alberta petroleum reserves have grown over time. Initial reserves are all those that have been discovered in the province up to the date shown and consist of the remaining reserves in that year plus past production. Since production continually depletes reserves, remaining established reserves are less than initial reserves. If total ('gross') reserves additions in a year exceed production in that year, remaining established reserves will rise, showing positive net reserves additions. Conversely, if production exceeds gross additions, net additions will be negative and remaining established reserves will fall. It can be seen that gross additions exceeded production for natural gas through 1980. Remaining established reserves for oil went into decline earlier.

Table 2.1 also provides an historical review of development drilling and development expenditures in Alberta since the Leduc discovery. As would be expected, the success rate is much higher for development than for exploration wells. The dominance of oil in the early decades of this industry's growth is evident, as is the increased importance of natural gas since 1960. As with exploratory expenditures, there was a peak in real expenditures in the early 1980s, followed by a decline. However, in the 1990s nominal development expenditures rose markedly, and, unlike for exploration, real expenditures were higher in the years 2000 and 2010 than in 1980.

C. Lifting (Operation or Extraction)

Table 2.1 indicates how Alberta oil and gas production and operating expenditures have changed since 1950. As was discussed in Chapter One, output in any year reflects underlying natural conditions (e.g., volumes of oil in developed pools and reservoir characteristics), physical capital constraints (developed capacities of wells, gathering, separation and transmission equipment), market conditions (demand and prices) and government regulations (taxes, output controls, export restrictions).

Production has tended to follow levels of remaining established reserves, with natural gas output generally rising from 1950 to 2000, while conventional crude production increased to the mid-1970s, then levelled out and fell. Production changes have reflected both changes in the level of reserves and also the intensity with which reserves are used, as indicated by the R/P ratio, which shows end of year remaining reserves divided by annual production. As can be seen in Table 2.1, the R/P ratios for conventional oil and natural gas have both fallen. Oil sands production commenced in the late 1960s and has risen throughout the period, as plant expansion and new projects occurred. By 2010, oil sands and bitumen output significantly exceeded conventional oil production. Nominal and real operating expenditures have risen throughout the period, reflecting in part, output increases. In addition, unit costs have risen as oil discoveries have tended to become smaller, and output rates have fallen due to production decline in reservoirs. Also, in many pools the water to oil ratio rises over time, so water disposal costs increase.

The nominal sales value of Alberta petroleum output has risen dramatically over time, as illustrated in Table 2.1, partly due to output increases. In addition, beginning in the early 1970s petroleum prices increased very markedly. But, after the mid-1980s, prices declined again, especially for natural gas. During the 1990s, natural gas revenues surpassed conventional crude oil revenues for the first time, and by 2010 oil sands revenue exceeded that from natural gas. Alberta oil and gas prices (as shown by the average sales revenue figures in Table 2.1) have been affected most strongly by two factors: (1) international oil prices, which directly affect the value of oil everywhere in the world and influence the price of other energy products, such as natural gas, and (2) government regulations, particularly restrictions on international trade (which break the direct link with international prices) and direct price control regulations. Natural gas prices in Alberta have generally been well below oil prices on an energy-content basis, mainly reflecting the higher costs of shipping energy to markets in the form of natural gas. An exception to this was during the mid-1970s to mid-1980s, when oil and natural gas prices were fixed by Canadian governments; these regulations will be discussed in Chapters Nine and Twelve. In the mid-1980s, crude oil and natural gas markets were deregulated and became subject to the interplay of market forces, exhibiting significant instability, so there has been considerable year-to-year variation in industry revenue. The late 1990s saw the relative value of natural gas increasing again as Alberta became increasingly integrated in the North American natural gas market and demand rises for gas in North America began to exceed supply increases, but by 2010 the gas price had fallen again relative to oil.

D. Government Activities in the Crude Petroleum Industry

1. Government Objectives

Governments are vitally concerned with the operations of the petroleum industry. As a result, the economics of the industry must be seen within the context of a regulatory environment: an economic history of petroleum is a story of petropolitics. Since 1930, the Alberta provincial government has been owner of most of the province's mineral rights, and so it has an obvious interest in ensuring that it receives a fair return on petroleum leases transferred to companies. Moreover, as the representative of citizens, the government has responsibility for establishing a legal/regulatory environment that is consistent with the interests of Albertans. Returns to owners, taxation (royalties, rentals, and bonus bids), conservation, and macroeconomic impact are major concerns.

The crude petroleum industry can be a major revenue source for governments. High quality natural resources generate surpluses above the required expenditures to find, develop, and produce the resource. (In economic terms, 'expenditures' include the return required on capital investment.) This surplus (a profit or 'economic rent') can, in theory at least, be taken from the industry without affecting industry activity, hence proving to be an ideally neutral and efficient source of funding for government activities. The practical problem is to approach such ideal rent collection as closely as possible in a world in which the precise size of the rent surplus is uncertain and changing. Some use the term 'taxes' to refer to the payments the petroleum industry makes to governments; other analysts call this 'government take.' It includes payments made to the government in its capacity as owner of Crown mineral rights (e.g., bonus bids, rentals, and royalties) plus those more conventional taxes (e.g., corporate income taxes) which apply to the petroleum industry as to other industries. Some have argued that the two categories of payments to the government should be kept strictly separate. However, the government, unlike private mineral rights owners, can use its powers to unilaterally change the level of royalties it receives on previously issued mineral rights, just as it can change income and sales tax rates. Moreover, the levels of conventional taxes and payments of royalties and bonus bids are not independent of each other. Hence it is desirable to consider conventional taxes and petroleum-specific payments together. For industry, the 'bottom line,' after deduction of all payments to governments, is crucial. Chapter Eleven examines government rent collection in Alberta.

'Conservation' is, surely, an unexceptionable objective. Exactly what is meant, however, is often left vague. We will use the term quite generally as meaning 'good physical production practices' and note that it has implications both for current industry activities ('intragenerational') and for the future ('intergenerational'). Intragenerational concerns relate partly to the physical impacts of production, such as minimizing unnecessary surface damage during exploration and production; careful disposal of water and other waste material; handling abandoned wells; reducing the risks of pipeline leaks and well blowouts. Also important is the desire to minimize the loss to society of the petroleum resource itself (e.g., reducing the flaring of natural gas, and increasing the recovery factor, the percent of the resource volume physically available that is commercially producible). The intergenerational issue is that of balancing the legitimate needs and concerns of present and future citizens, given that the petroleum resource base is limited, at least in a geological sense. It is tempting to see only the costly side of this intergenerational problem – our production of petroleum must, by the very nature of physical depletion, reduce production possibilities in the future. Balanced against this, however, are any improvements in knowledge that reduce costs or increase recoverability in the future and the returns future citizens derive from investments undertaken today with petroleum revenues. An obvious question with respect to conservation is to ask why companies will not automatically incorporate good production practices, obviating the need for government regulations. In Chapter Ten we consider one such conservation issue in detail.

Macroeconomic impacts relate to the effects that petroleum activities have on population, investment, employment, and per capita income, since the industry is such a critical part of the province's economy. The Alberta petroleum industry can be seen as an example of an 'export base' industry, one where external market conditions governing exports from the province play a key role in determining the value of the industry's output. Fluctuations in the level of petroleum industry activities will tend to generate similar fluctuations in the province's economy, with obvious possibilities for 'boom-bust' cycles. Sub-regions of the province may be especially prone to such cycles. At the macroeconomic level, it is usually judged desirable to have a relatively high and stable rate of economic growth, with low rates of inflation. Chapter Thirteen delves into these macroeconomic issues.

The three government objectives – revenue, conservation, and stable macroeconomic growth – are not independent of one another but need not always be in conflict. Consider, for example, an 'attenuated' resource development scenario, where the government imposes some limits on the speed of development of the resource base (Scott, 1976); annual tax revenues collected will tend to be smaller, at least in the earlier years of industry activity, but intergenerational conservation and macroeconomic stability may be enhanced, and total lifetime tax revenue from the industry may actually increase.

Government treatment of the petroleum industry is particularly complicated in a confederation like Canada since both provincial (Alberta) and federal ('Ottawa') governments may wish to regulate the industry in the interests of their constituencies. Sometimes the concerns of the two levels of government may coincide. For example, both may wish to minimize local pollution and obtain fair export prices. In other instances, the two may disagree; for example, Ottawa will have more concern than Alberta with the impact of higher petroleum prices on consumers, and the two governments can easily disagree on how tax payments by the industry should be shared. Later chapters, especially Chapters Nine and Eleven, will detail a number of vociferous intergovernmental disagreements about Canadian petroleum policies, especially after the revolutionary increases in world oil prices in the 1970s.

2. Government Policies

After the Leduc discovery of 1947, oil companies dramatically increased exploratory activity in Alberta. Initially corporate interest centred on the Devonian reef formations in the centre of the province, but the new-found optimism about Alberta's oil potential led to increased G&G surveys and drilling in all areas, and other productive plays soon followed. The government is concerned with the nature of exploration activities, their pace, and the financial effects.

The industry's physical activities have been monitored and regulated by the Alberta government. The aim has been to prescribe safe and clean methods of exploration, to require that companies obtain a permit or licence before undertaking activity, and to watch over what companies do to ensure regulations are followed. Economists have frequently been critical of regulation by direct control of activities undertaken by firms; this approach may involve high costs of administration and lack sufficient flexibility to respond to the variety of conditions faced by the industry, and the rapidity with which our world changes. It has been necessary in Alberta, for instance, to recognize changes in the geographical interests of companies (e.g., from the central plains to the forested foothills and north to permafrost land) as well as changes in technique (e.g., new drilling fluids and larger all-terrain vehicles). The government of Alberta has transferred prime responsibility for regulatory administration of the petroleum industry to an independent body, funded by the province and a levy on petroleum company revenues.

The Petroleum and Natural Gas Conservation Board (PNGCB) was set up in 1938, initially to regulate the Turner Valley oil discovery (Breen, 1993). Some small producers were having difficulty finding markets for their oil, gas flaring was common, and output rates were so large that future productivity was in danger. The high production rates were a legacy of the provision of British Common Law known as the 'Rule of Capture,' which had long plagued the U.S. petroleum industry. Daintith (2010) provides a detailed review of the rule of capture in the U.S. and elsewhere (though not in Canada). He notes that it stems from provisions in Roman law and also operated under civil law in countries with no connection to British common law. Under the rule of capture, petroleum in a reservoir with divided interests is owned in common by all producers with access. It became the property of one company only when that company lifted it to the surface. In the circumstances, producers had strong incentives to drill many wells and produce at high output rates to capture the oil before their neighbours, even though this often damaged the recovery process in the reservoir. Turner Valley, prior to regulation, was a prime example.

In 1950, the PNGCB was given prime responsibility for administering the province's new *Oil and Gas Resources Conservation Act*. In 1957, with the *Oil and Gas Conservation Act*, it was renamed the Oil and Gas

Conservation Board (OGCB). In 1971 its responsibilities were extended over other energy products (coal and electricity), and it became the Energy Resources Conservation Board (ERCB). In 1994, the government announced that the activities of the ERCB would be combined with those of the Public Utilities Board in a new Energy and Utilities Board (EUB). Then, in 2008, the EUB was bifurcated and the ERCB resurrected.

The board has a large technical staff, which, among other things, monitors the industry to ensure that regulations are observed. It has been given quasi-judicial powers in certain areas to hold hearings and sanction or order changes in the actions of companies, and it frequently makes recommendations to the government about policy issues of concern, usually after hosting public hearings. Some economists and political scientists feel that one problem with agencies such as the ERCB is that they are 'captured' by the industry they are designed to regulate (Stigler, 1972). Many individuals have moved easily between the board and industry, but most observers regard the ERCB as an efficient and independent body in much of its regulation of the technical aspects of petroleum industry activities. In this book, we shall not investigate the administration by the board of regulations governing purely technical aspects of industry behaviour nor the more controversial issues related to public health and the environment. Later chapters will deal with several of the programs that have had a great impact on the economics of the industry, including market-demand prorationing schemes to regulate oil output rates from pools (Chapter Ten), and natural gas export sales restrictions (Chapter Thirteen).

In December 2012, the government of Alberta passed the *Responsible Energy Development Act*, which set up a new 'Alberta Energy Regulator'. As of final editing in April 2013, the regulator was still in the process of being established; it would be responsible for the regulatory functions previously handled by the ERCB including matters which require public hearings.

The pace of petroleum exploration is another matter of concern to governments. It can be influenced by licensing regulations for G&G surveys and exploratory drilling, and by government taxes and/or subsidies, depending on whether the government wishes to discourage or encourage activity. In Alberta, however, the most immediate influence on the level of exploration activity comes through the rate at which the province issues mineral rights on Crown land (Crommelin, 1975; Crommelin et al., 1976). There is little, if any, evidence of a specific government policy on the rate of issuance of petroleum and natural gas exploration and production rights. The main influence on the number, and the surface area, of rights issued in any period appears to have been requests for rights by the industry. Alberta has, on occasion, issued regulations making it less attractive for companies to hold leased rights inactive and unexplored, thereby encouraging more exploratory activities. The requirement that exploration rights be partially relinquished back to the government, on a checkerboard basis, has ensured that the government retains an ongoing interest in sub-areas of the province that become particularly attractive for petroleum exploration. In addition, beginning in the 1970s, a series of royalty-rebate and/or subsidy programs – both at the federal and provincial level – were introduced to stimulate additional exploration. These programs came at a time of high oil prices but also of higher royalties/taxes and of reduced average discovery size, especially for oil, and continued as oil prices became lower in the 1980s.

Fiscal effects on exploration by the Alberta government are primarily tied to the financial terms of Crown mineral right issues. As was discussed earlier, the government would like to capture as much as possible of the economic rent (profit) from petroleum production. Three broad classes of financial payment have been common to virtually all mineral rights:

(1) Bonus bids: the mineral rights are issued to the company that offers to pay most in a competitive, sealed-bid auction.
(2) Rentals: an annual rental per hectare is assessed on mineral rights held.
(3) Royalties: companies pay a portion of any petroleum sales revenue to the government.

Chapter Eleven will review these tax (rent collection) regulations in more detail. Provincial income tax also captures rent for the province, and the federal government has, of course, an interest in the petroleum industry as a tax base. From the government's perspective, it is very important to find the right mix and levels of payment. Table 2.2 gives some feel for the importance of the petroleum industry as a source of Alberta government revenue since 1950, excluding income tax. Considerable year-to-year variability in government revenues from the petroleum industry is apparent, particularly in bonus bids. Particularly high government revenue, in real terms, came in the early 1980s when world oil prices were at a peak, and total petroleum revenues amounted to over three quarters of total provincial government receipts. As

Table 2.2: Alberta Government Petroleum Revenues (10^6 $)*

	1950	*1960*	*1970*	*1980*	*1990*	*2000*	*2010*
Fees and Rentals	6.3	32.7	58.3	69.3	72.4	141	158
Royalties	3.6	27.3	143.7	3,456.1	2,118.8	3,970	6,533
Bonus Bids	23.2	81.3	26.5	1,057.7	389.1	743	1,165
Total	30.1	141.3	228.5	4,583.1	2,580.3	4,854	7,756
(% of Provincial Government Revenue)	(30.2)	(42.2)	(25.1)	(80.9)	(21.3)	(24.1)	(19.7)

* For the fiscal year ending March 31 of the year indicated.

Source: *Annual Reports* of the Department of Mines and Minerals, Department of Energy and Natural Resources, and Department of Energy, depending on the year.

will be discussed in Chapter Thirteen, the government questioned whether such high revenues from a depletable natural resource could be expected to continue indefinitely and whether some should be saved rather than spent immediately. Petroleum industry payments to the provincial government fell after 1980, and then rose again in the 1990s; however, after allowance is made for inflation, and the growing size of the provincial economy, the relative importance of petroleum revenues to the government is not as high as it was in the early 1980s.

As with exploration, the Alberta government has had responsibility for ensuring that development activities are carried out safely and in an environmentally sound manner. This book will not discuss details of these types of regulations, though we note that the possibility of well blowouts and other petroleum leakages has generated controversy, especially where population centres and sour (high sulphur) petroleum come together. There is also concern about clean-up costs of abandoned wells, particularly where wells were shut-in many years ago and where companies have gone out of business (Horner, 2011). There have been hearings before the EUB and its predecessor boards with respect to major EOR projects and natural gas processing plants. The EUB possesses some powers to force changes in development plans (e.g., compulsory EOR investment) if it is clearly in the interests of greater commercial oil recovery.

Government regulations with respect to well spacing (establishing minimum spacing requirements) have had a significant impact on the industry's development activity. The government has also had to examine how its policies affect the rate of development; high royalties, for instance, may discourage reserve development and lead to premature well abandonment unless they are mitigated. Output control schemes such as market-demand prorationing also affect industry development. These issues will be considered later in the book.

The board has also been given responsibility for assessing whether major oil sands ventures are 'in the public interest.' In Chapter Seven, we will look at how the board considered this with respect to the impact of oil sands production on the market for conventional Alberta crude oil. We will not, however, consider the board's treatment of the environmental (including health) impacts of oil sands investments. These issues have been quite different from the board's historic concern with oil and gas production techniques in the conventional industry; some critics have argued that the board has insufficient expertise in these broader environmental areas and is too sympathetic to the viewpoint of the industry.

As noted above, governments have an obvious interest in both the timing and value of petroleum output. Timing concerns relate to the depletable resource base, such that lifting today means foregoing future production. Governments may feel that output should be spread relatively evenly over time to help stabilize revenue and tax flows. Such timing concerns are clearly linked to prices. The provincial government may, for instance, prefer to restrain production if output rises would drive down market prices or to delay production if prices are expected to rise in the future. Once again, the question arises of why private producers would not themselves react in this manner, therefore making government action superfluous.

The two decades following Leduc were dominated by the desire to find markets for the province's rapidly expanding oil reserves; this search took place against a backdrop of falling real international oil prices. Two issues became the focal point of discussion, as discussed in detail in Chapter Nine. Should Alberta oil be reserved primarily for use in Canadian markets, even if this meant bypassing closer markets in the Midwestern United States and shipping the oil to Montreal where it was less competitive with offshore international supplies? Should Alberta oil be offered protection from declining international crude oil values? As it happens, Canadian policy, implemented in Ottawa as the National Oil Policy, encouraged exports of oil to the United States and restricted Canadian access to offshore oil, thereby allowing Canadian prices above international levels. The United States was debating the same issues; the resultant oil import quota program maintained high prices in the United States and quickly incorporated special treatment for oil produced in Canada but did not leave the U.S. market completely open to Alberta oil.

The 1970s brought rising international oil prices and international supply disruptions in connection with political events centred in the Middle East. The government of Alberta, and oil producers, viewed higher prices and buoyant U.S. markets with pleasure. However, the Canadian federal government felt grave concern about the cost and security of oil supplies – both immediately, for users in Quebec and the Atlantic provinces, and through the longer term for consumers in Ontario. Ottawa also worried about the impacts on consumers, and its own fiscal position, of sharply increased oil and natural gas prices. The resultant tension between Ottawa and the petroleum-producing provinces brought a decade of intergovernmental rancour, with export limitations on oil, price controls to keep domestic oil prices below international levels, extension of oil transmission facilities to Montreal, and new federal taxes on the crude petroleum industry. The provinces and a new federal government in Ottawa finally agreed, in 1985, on a policy of 'deregulation' of the petroleum industry; this coincided with a collapse of international oil prices.

Natural gas policies over this period exhibited similar tensions between domestic security of supply and export potential as well as between higher prices for producers and lower prices for consumers. Ottawa introduced price controls on natural gas and a new federal natural gas tax. These policies are reviewed in Chapter Twelve.

These instances of government involvement in the crude petroleum industry highlight an underlying uncertainty by some about the wisdom of relying on relatively open markets to allocate oil and natural gas supplies, given the importance of factors external to Canada in setting prices and the depletable nature of the resource base.

The provincial government has also been concerned with variations in the level of petroleum industry activity and in the revenues received by the industry because such changes will impact on the level of economic activity in the province. In Chapter Thirteen, we look at the macroeconomic impact of the industry on Alberta.

4. Alberta's Downstream Petroleum Industry

The focus of this book is on the Alberta crude petroleum industry. In this section of Chapter Two, we briefly review aspects of the downstream industry in Alberta and provide some comments on economic issues, which may be useful to an understanding of the upstream industry.

A. Transportation: Industry Activities

Chapter One argued that two different categories of crude petroleum transportation have been important in Alberta. First, a network of 'gathering' pipelines is needed to collect the lifted petroleum from various reservoirs in the province and move it to major collection points or local markets. Then large-diameter 'trunk' pipelines move volumes to major market destinations outside the province. The physical heterogeneity of crude oil raises some technical problems. As noted in Chapter One, most crude oils are mixed together (blended) as they are shipped, but it is sometimes desirable to separate a particular grade from others in the line (to ship by 'batch'), or to build a separate line to handle a particular product (for example, heavy oil, or bitumen or natural gas liquids, NGLs). While there are chemical differences between volumes of gas from different pools (e.g., in the presence of sulphur and NGLs), natural gas is more homogeneous than oil, especially after treatment in gas plants. Transmission of crude oil and natural gas has proceeded quite differently.

Crude oil pipelines in Alberta were usually built by oil companies themselves, often companies that both produce crude oil and refine it. Estimating the appropriate size of a line is difficult. Economies of scale imply that a pipeline should be as large as possible (up to limits that are significant relative to the volumes of oil shipped from Alberta). Therefore, oil companies normally try to anticipate the likely

Table 2.3: Sales of Alberta Petroleum by Region

	1950	1960	1970	1980	1990	2000	2010
OIL SALES (10^6 m^3)							
Alberta	4.3	4.1	6.4	16.1	19.3	23.9	26.6
B.C.	0	3.7	3.2	7.5	3.6	2.3	1.9
Sask/Manitoba	0	3.3	4.4	3.6	1.9	7.7	10.8
Ontario	0	6.4	13.4	25.8	20.5	1.3	7.7
Quebec	0	0	0	15.4	3.6	0	0
Total Canada	4.3	17.5	27.4	68.4	48.8	35.2	46.9
U.S.A.	0	4.1	32.5	8.7	29.8	55.6	79.3
GAS SALES (10^9 m^3)							
Alberta	2.0	4.2	6.8	13.5	17.3	23.1	32.1
B.C.	0	0.2	0.3	0.3	0.9	2.9	2.6
Sask/Manitoba	0	0.1	2.4	3.5	2.9	6.5	8.5
Ontario	0	0	11.5	18.0	17.3	29.7	17.1
Quebec	0	0	1.4	2.9	4.8	5.9	1.1
Total Canada	2.0	4.5	22.4	38.2	43.2	68.1	61.5
U.S.A.	0	1.1	17.4	19.4	35.7	66.1	49.6

Note: Includes bitumen, synthetic crude oil, pentanes plus and condensate.

Sources: 2000 and 2010: ERCB, *Alberta Energy Resource Industries Monthly Statistics* (ST-3).
1950–1990: ERCB and OGCB, *Alberta Oil and Gas Annual Statistics* (ST-17) and *Cumulative Annual Statistics of the Alberta Oil and Gas Industry*.

volumes forthcoming from future discoveries in the area, as well as volumes immediately available. The various gathering lines have usually been constructed by the first company or companies to generate significant crude oil discoveries in a particular region. The major gathering lines converge on Edmonton, which is the starting point for two trunk lines, both of which were initially built in the 1950s. The Interprovincial Pipe Line (now known as Enbridge) heads to major markets in the east, looping below the Great Lakes into the United States, as far as Toronto (with a link to Montreal built in the 1970s). The Trans Mountain Pipe Line (briefly known as Terrason and now as Kinder-Morgan) traverses the Rockies to Vancouver and the Pacific Northwest states. The gathering lines are wholly owned subsidiaries of various oil companies, while the trunk lines are shareholder-owned, with shares traded on the public stock exchanges (generally, large blocks of shares have been held by oil companies).

The first natural gas pipelines were built by local Alberta distributors (utilities) to bring gas to their customers. In 1954, the government of Alberta introduced legislation that set up a shareholder-owned, publicly traded company called Alberta Gas Trunk Line (AGTL), which would have sole responsibility for gathering natural gas in the province and moving it to the borders for ex-Alberta sales. (In 1980, AGTL became NOVA, an Alberta Corporation, then, in the late 1990s, it merged with TransCanada Pipe Line.) AGTL sold a transportation service to the owner of the gas. Until the 1980s, the gas was usually bought by and owned by the major trunk line transmission company, which took possession at the Alberta border. These gas-purchasing pipelines were not themselves natural gas producers but new investor-owned companies with publicly traded shares set up explicitly as gas transmission companies. The largest early buyers of gas were TransCanada Pipe Line (for sales to the East), Westcoast Transmission (for sales to the West), and Alberta and Southern (owned by Pacific Gas Transmission, for sales mainly to California). Beginning in the 1970s, a number of new marketers and shippers began to enter the Alberta market, generally purchasing gas for shipment to export markets in the United States. Deregulation in the late 1980s further increased the number of natural gas buyers, as will be discussed in Chapter Twelve.

Table 2.3 shows the destination of sales of Canadian crude oil and natural gas for various years since Leduc, clearly demonstrating the significance of markets external to the province.

B. Transportation: Government Activities

In addition to concerns about personal and environmental safety, governments have been interested in the price charged for transmission, access to facilities, and the route and destination of pipelines. As was discussed in Chapter One, pipeline tariffs are of concern largely because of the 'natural monopoly' nature of the service, since economies of scale usually mean that a single pipeline is the most efficient way to move the product. The government wishes to ensure that the transmission company does not take unfair advantage of its monopoly status by charging a tariff far above costs, or by buying petroleum itself at artificially low prices, or selling at artificially high ones. Access concerns relate to the possibility of the pipeline denying service to some potential users (e.g., refusing to move a competitor's oil or gas). A government may be concerned about pipeline routes because it does (or does not) wish to see a specific geographic market penetrated, or because security of supply or environmental risks dictate certain routes.

In Canada, there has been far less government attention to crude oil pipelines than natural gas, perhaps because the ownership of crude oil lines by the oil companies themselves was conducive to results that the industry found acceptable. In the early years of operation, the possibility of government regulations may also have helped persuade the owners of crude oil pipelines to keep access open to all potential users and to base tariffs on pipeline costs, following the approaches used by regulated pipelines in the United States (Lawrey and Watkins, 1982). In addition other government regulations on oil output, in the form of market-demand prorationing, ensured that all oil reserves holders were given the opportunity to produce; thus, the large oil producers could not use their ownership of pipelines to squeeze out other producers while increasing their own oil production. The *National Energy Board Act* of 1959 declared interprovincial trunk lines (and intraprovincial lines used by more than the owner) to be common carriers, requiring that pipeline capacity be equally accessible to all potential shippers. In addition, the act gave the board the power to regulate oil pipeline tariffs, although the NEB did not begin to exercise this power until 1977 (Lawrey and Watkins, 1982).

Variations in Alberta oil output also raised questions related to government permits for pipeline construction. In the early days, the question was that of determining pipeline sizes and destination markets when the potential for oil production in Alberta was unknown and just being established. As conventional oil reserves began to decline after 1970, there was concern about underutilization of facilities. Pipeline expansion became an issue again as the oil sands picked up in the new century, with discussion of whether pipelines would handle upgraded oil or bitumen and whether Alberta should continue to rely on traditional North American markets or look toward Asia.

Natural gas transmission has been more contentious from the start. For example, in the 1950s the federal government insisted that the TransCanada Pipeline be built entirely on Canadian territory, even though it would have been cheaper to follow the Interprovincial oil line through the northern tier of the United States, south of the Canadian Shield. The natural gas trunk lines' role as the main buyers of natural gas also raised potential problems. The pipelines – if they crossed provincial boundaries – were subject to cost of service rate regulation under the authority of Ottawa's National Energy Board, but petroleum companies frequently complained that the gas pipelines (especially TransCanada's) monopsony position as a buyer of natural gas led to artificially low prices. Exacerbating this problem was the prevalence of long-term natural gas purchase contracts, often of more than twenty year's duration, with relatively fixed prices. The issue came to a head in the 1970s when oil prices rose dramatically, increasing the value of competing fuels such as natural gas. The Alberta government began to use its export licensing requirements to force renegotiation upward of natural gas prices, and later began to cooperate with Ottawa in fixing natural gas prices in relation to oil prices.

Deregulation in the mid-1980s generated another series of public policy concerns. As a number of new pipeline proposals arose, drawing on Alberta natural gas, governments had to determine whether they were all compatible and in the public interest. Would there be costly duplication of facilities? Were markets strong enough to provide a fair return on Alberta gas (especially markets on the far east coast of the United States)? Could too many new facilities be constructed thereby providing so much additional Alberta gas to markets that prices would fall or increasing the risk of raising unit shipment costs? Readers will note that such questions all betray an anxiety about the operation of unregulated markets. The Alberta government was also under pressure to change its policies with respect to NOVA, and its method of handling gas movements within the province. NOVA was rate-regulated, but, for many years, applied a 'postage

Table 2.4: Alberta Oil Refining

	Number of Refineries	Crude Oil Refining Capacity (m^3 per calendar day)	Refinery Runs (m^3 per day)
1950	7	7,450	
1960	11	15,650	12,248
1970	8	27,810	17,909
1980	6	45,311	45,301
1990	6	63,200	55,612
2000	5	68,055	70,167
2010	5	72,135	69,498

Sources: CAPP *Statistical Handbook* (excludes oil sands upgraders).

Table 2.5: Alberta's Primary Energy Consumption by Fuel (Primary Energy Shares, %)

	Oil	Natural Gas	Coal	Hydro	Other
1966	41	52	5	0	2
1990	48	36	15	1	0

Source: For 1990, ERCB, *Energy Alberta* Reports (oil includes 24% bitumen); 1966 figures calculated from energy use tables in Appendix C of NEB (1969).

stamp' tariff, in which all gas would be assessed the same charge regardless of pick-up or delivery point, even though some gas obviously had lower transportation costs. Only in the late 1990s did the government require that NOVA (now TransCanada) abandon the postage stamp tariff

C. Refining and Marketing: Industry Activities

It was mentioned in Chapter One that it is cheaper to transport crude oil than refined petroleum products (RPPs); thus, regional refining capacity tends to be geared to the size of the local market. Table 2.4 shows the number and capacity of Alberta refineries in select years from 1950 to 2010. Refineries exhibit economies of scale; the reduction in the number of refineries since 1960, and the significant rise in total capacity and throughput, reflect a move to larger more efficient refineries. The main Alberta refineries are owned by large vertically integrated oil companies; some of the smaller refineries are former assets of the major oil companies that were purchased by employees or smaller companies (perhaps other gasoline marketers) when the majors were rationalizing facilities and disposing of less profitable assets.

Table 2.5 shows the relative importance of utilization of different energy products in Alberta in 1966 and 1990. The importance of petroleum is evident. Natural gas plays a particularly high role in Alberta in comparison with other parts of Canada. Coal's role has increased as a source for thermally generated electrical energy, and as the demand for electricity has risen more rapidly than the demand for other major secondary energy products.

D. Refining and Marketing: Government Activities

Government attention to petroleum refining and marketing has focused on competition policy, taxation, and conservation.

The importance of economies of scale in refining has meant that a relatively small market can support only a limited number of efficient refineries, raising the possibility of imperfectly competitive behaviour. In the oil industry, the oligopolistic nature of refining has combined with the vertically integrated nature of the industry to raise the possibility of restrictive competition from crude oil through to the marketing of oil products. For example, as will be discussed in Chapters Six and Ten, the process of refiners 'posting' prices that they would pay for crude oil, in conjunction with the government's 'market-demand prorationing' regulations, led to rigid prices for crude oil from 1950 through 1972. In another possible example, in Alberta (but not to the same extent in Ontario) motor gasoline prices have often exhibited a certain amount of rigidity, rather than the short-term variability common to some other commodities. Concerns have also been expressed about restrictive tied marketing arrangements imposed by refineries on retail distributors (for example, forcing them to handle a particular brand of motor oil or tires). These concerns have attracted a number of government studies.

There has not been much regulatory response, however, since clear evidence of anti-competitive behaviour is not strong (Watkins, 1981). This may, in part, reflect effective competition among the limited number of firms active in refining, and the greater number in marketing. Competition is also enhanced by the threat of new entry (Baumol et al., 1988), and the possibility of importing RPPs from more competitive external markets (e.g., the United States).

With respect to natural gas, distribution to consumers is efficiently done by a single company in any market area as a natural monopoly. Hence, gas utilities have been subject to cost of service rate regulation by the Alberta Public Utilities Board (PUB); after 1994, regulation was by the Energy and Utilities Board (EUB), and, after 2007, the Alberta Utilities Commission (AUC).

Some RPPs have been very attractive to governments as targets for taxation. For revenue purposes, a commodity is especially appealing for a sales (excise) tax if its consumption is relatively unresponsive to price changes. (Economists would say that the demand is 'inelastic'; this concept is set out in Chapter Four.) This is characteristic of a good that is viewed by many consumers as a necessity and for which few substitutes exist. Motor gasoline has been particularly appealing in this regard. In Alberta, the provincial tax on a litre of regular grade gasoline was 2.5 cents in 1962, 3 cents in 1970, zero cents in 1981, 5.0 cents in 1989, and 9 cents by 2011 (Canadian Tax Foundation, *Provincial and Municipal Finances*, various years). (The 1990 tax was equivalent to over eight dollars per barrel of oil.) In 2011 the federal gasoline tax was about 10 cents per litre (plus the 5% GST).

Conservation regulations include those designed explicitly to discourage current utilization of petroleum, as has been advocated by many as part of the response to man-made global warming. (Such environmental problems are beyond the scope of this book.) Higher taxes on RPPs, such as a carbon tax, can be useful here, since higher taxes discourage current use. Regulations may also prohibit the utilization of certain petroleum products for particular purposes. For example, from the mid-1970s until 1992, Alberta did not allow new thermal electricity plants burning natural gas. (The United States had similar regulations for much of the 1970s and 1980s.) The removal of these restrictions stemmed from apparent surpluses of natural gas even at falling prices and from recognition of the special environmental risks posed by coal and nuclear-powered generation facilities.

5. Conclusions

This chapter has provided a preliminary overview of major developments in the history of the Alberta petroleum industry and of issues that have attracted the attention of governments. Underlying both historical events and government policies is the operation of petroleum markets and the extent to which such markets adequately reflect society's interests. The remainder of this book includes more formal analysis of petroleum markets and major government regulations and policies associated with the Alberta crude petroleum industry. Before concluding this chapter, however, brief comments are made on three important topics.

First, Alberta's economic performance after 1947 was tightly bound to developments in the conventional petroleum industry. In purely physical terms, however, the conventional petroleum resource base is dwarfed by the volume of non-conventional petroleum resources. For example, the year 2011 Canadian Association of Petroleum Producers (CAPP) *Statistical Handbook* estimates the total volume of conventional liquid in place in the province (including past production) to be about 10.6 billion m^3, while non-conventional heavy oil and bitumen deposits are estimated to hold over 400 billion m^3. Since at least the 1920s, private companies and government bodies such as the Alberta Research Council have experimented with ways to produce at low cost from these non-conventional deposits. In 2013 there were five operating oil sands mining companies, two having commenced more than thirty years previously – the Suncor plant (commenced in 1967) and Syncrude (commenced in 1978); in 2011, upgraded synthetic crude made up about 31 per cent of Alberta's liquid hydrocarbons. Several large-scale and a number of smaller *in situ* heavy oil projects were also in operation; in 2011, bitumen from such ventures amounted to about 45 per cent of liquid hydrocarbon production. While the physical potential for large volumes of oil from non-conventional sources is high, actual development hinges on perceptions of expected future economic conditions, the scope of technological innovations and the regulatory environment. Similarly, new techniques or improved economic conditions have been necessary to stimulate exploitation of Alberta's natural gas volumes held in very 'tight' (relatively non-permeable) formations or trapped as methane in the province's coal seams. The problems and potential of these non-conventional resources differ considerably from those of the conventional crude petroleum industry. Chapter Seven deals with Alberta's non-conventional oil, and we touch on non-conventional natural gas in Chapter Twelve.

Second, this chapter has not yet raised a controversial public policy issue – foreign investment in the

Canadian petroleum industry. Foreign ownership has been significant almost from the beginning, at least since Standard Oil acquired control of Imperial Oil back in 1898. By way of illustration, in 1989, the Federal Petroleum Monitoring Agency reported that 44 per cent of the revenues in the Canadian crude petroleum industry were foreign owned; if all aspects of petroleum industry activity were considered, the foreign ownership percentage rose to 46 per cent of revenues, 49 per cent of assets, and 52 per cent of expenditures. Pervasive and persistent foreign ownership in the petroleum industry touches the nerves of many Canadians. Opponents of foreign investment see foreign owners capturing jobs and profits that would otherwise go to Canadians. And large volumes of a scarce resource are seen as siphoned away from Canadian users to consumers south of the border. Foreign – read 'American' – values and mores are argued to be imported by the ex-Canada owners and imposed over traditional Canadian social values, transforming Alberta into a pseudo-Texas. Proponents of foreign investment see the financial capital and technical expertise of the multinational oil companies generating employment opportunities and income gains for Canadians that would not otherwise occur. The higher per capita living standard that results makes it easier for Albertans and Canadians generally to provide those private and social goods that define our society. Such deep-seated and conflicting views cannot be reconciled by economic analysis alone, but, in Chapter Six, we briefly consider some of these views.

Third, there is no discussion in this book of significant issues in environmental economics, even though a number of controversial issues have attracted much public attention in Alberta, including the hazards of well blowouts (which generated an extensive public hearing after the Lodgepole blow-out in the 1980s); the health effects of petroleum production, especially sour natural gas; global warming; and the very significant environmental concerns associated with expanded oil sands production, especially from gigantic strip-mining ventures.

CHAPTER THREE

Alberta and World Petroleum Markets

Readers' Guide: Alberta is not an isolated economy. This chapter provides an overview of the world oil market, what it is, and how it has developed. Since the 1960s, this has been, to a considerable extent, the story of the Organization of Petroleum Exporting Countries (OPEC): the role of the OPEC governments gives obvious meaning to the term 'petropolitics.' We look at Canada's position in the world oil market and argue that it has been such a small player that changes in production and consumption here have minimal effects on world oil prices. In economic terms, Alberta has been a 'price-taker' in the world market, and the international price of oil sets the value of Alberta oil.

1. Why World Markets Matter

It is commonplace to remark that global interdependence has been growing. Rapid, low-cost transportation systems easily move people and goods across the globe, and falling costs of communication have led to almost instantaneous exchange of information. Glimpses of other lifestyles have fuelled consumer tastes everywhere, including the material expectations of the world's poor. Financial capital moves with ease, changing location and form at a moment's notice, its price in all parts of the world responding to the latest news or rumour. Branches of large multinational corporations reach into the world's most distant corners, part cause and part symptom of the shrinking global village. For the world at large, international trade has been increasing faster than purely domestic trade, and governments throughout the world have been negotiating regional and global agreements that impact on trade and investment.

International trade in crude oil provided one of the earliest of truly global markets. There are a variety of reasons for this. First, energy is a necessary input to economic activity throughout the world, and oil provides a particularly convenient (and, in internal combustion engines, essential) form of energy. Second, oil is a relatively homogeneous commodity in the sense that crude oils from different deposits around the world can substitute for one another. Third, crude oil can be transported readily and at relatively low cost. Today, for instance, a charge of several dollars per barrel or less would normally be sufficient to cover the cost of moving oil up to 3,500 kilometres through a large-diameter pipeline, or the 7,000 kilometres between Saudi Arabia and Japan in an ocean-going tanker.

Crude oil is, and has been for decades, the most highly valued commodity moving in international trade, as measured by the total value of shipments. The ease of moving oil between markets means that, in the absence of any government regulation hindering the flows of oil between countries or imposing special import or export taxes, the price of crude oil within any one country is very closely tied to international oil prices. If the domestic price were significantly lower, those who own domestic oil (whether companies or consumers who have purchased it) have an economic incentive to move the oil into the higher-priced

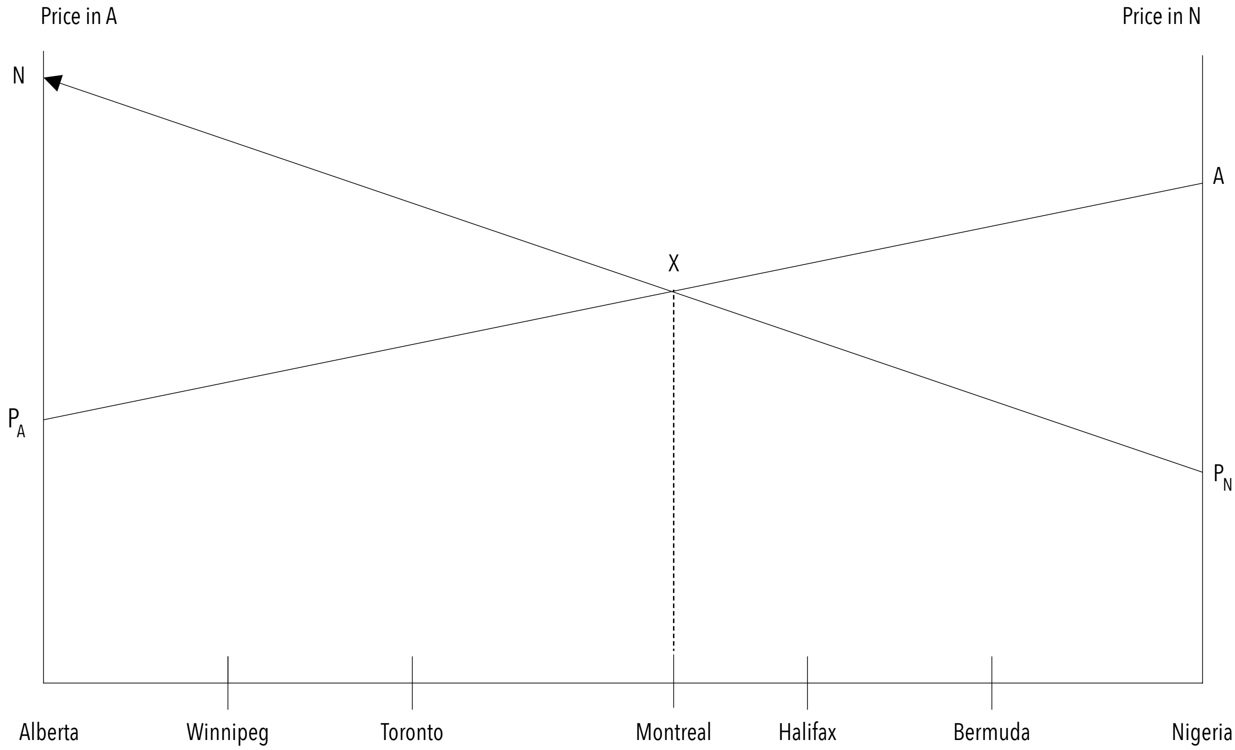


Figure 3.1 International Oil Flows and Prices

international market, and the reduced domestic supply would drive home prices up. Conversely, if domestic prices were above world prices, there would be strong incentives to purchase in the international market, and the reduced demand for domestic oil would drive the price down.

If oil could be instantaneously transported, and if the world oil market had immediate and perfect information flows to all actual and potential market participants, and if buyers and sellers were free to contract with one another and sellers could not discriminate amongst buyers, then oil prices in different parts of the world would differ, at most, by the marginal transportation cost between locations. Economists often refer to this as the 'Law of One Price,' and argue that the exploitation of profitable trading opportunities ('arbitrage') will ensure this result. More accurately, there would be a structure of crude oil prices. Prices would vary across different qualities of crude oil, and, for any particular grade of crude oil, prices between any two locations would differ at most by transportation costs between the two. Crude oil from any large producing region would satisfy demands in that immediate location, with the delivered price of oil rising by the incremental shipment cost as sales occur in progressively more distant markets; at some location, oil from this region would be higher priced in the more distant market than oil from another producing centre, and no sales would occur.

Figure 3.1 is a simple illustration of this effect for two producing regions (Alberta and Nigeria). The vertical axis shows oil prices, initially P_A in Alberta and P_N in Nigeria, while the horizontal axis shows locations between the two producing centres (Winnipeg, Toronto, Montreal, Halifax, and Bermuda). The line P_AA rising from left to right shows the delivered cost of Alberta oil in various markets, while P_NN shows the delivered cost of Nigerian oil. The two lines meet at point X, in this case Montreal, with Alberta oil being chosen by consumers in Winnipeg and Toronto, and Nigerian oil chosen in Halifax and Bermuda. Prices in various markets are given by line P_AXP_N, and it can be seen that the price difference between any two regions may be equal to transportation costs between the markets (compare Montreal to either Alberta or Nigeria) or less than such transportation costs would be (if no oil moves between the markets; compare Alberta to Bermuda). The market where crudes from the two regions are priced equally (i.e., Montreal) is often called the 'watershed' market, or the 'competitive interface.' It can readily be seen that if the price of crude oil fell in one of the producing regions (e.g., P_N declined) with nothing else changing (i.e., neither P_A nor shipment costs) the watershed would move further away from that market (e.g., to Toronto) and sales of the other region's oil would decline. Such a reduction in the demand for Alberta's oil might lead to a price fall in Alberta, with a new watershed somewhere between Toronto and Montreal and lower prices in all markets.

In an interconnected world market of this sort, the exact level of oil prices, and the specific geographic pattern of prices (and location of watershed markets), depends upon the entire global set of demand, supply, and transportation cost components. In general, any change in a major demand, supply, or shipment cost component will lead to changed prices, production, and consumption everywhere in the world. Consider, for example, the following stylized example of how a major oil conservation scheme in Tokyo might reduce the price of crude from the Alberta Pembina pool. Mandated improvements in transportation fuel economy in Japan reduce the Japanese demand for crude, which drives down the price of Indonesian crude oil sold in Japan; but a reduced Indonesian price cuts the price of Indonesian oil in India, and Indian consumers switch from Middle Eastern to Indonesian oil. The reduced Indian demand for Middle Eastern oil drives down its price in Western Europe so consumers there switch from Nigerian oil to Middle Eastern oil; in turn Nigerian oil prices fall, consumers in Ontario switch away from Alberta oil, and the reduced demand for Alberta crude oil drives down prices in Pembina.

Most economists have accepted this general depiction of interconnected world oil markets as accurate, or, as Adelman (1984b) noted, "the world oil market, like the world ocean, is one giant pool." Casual observation of the world oil market suggests that the Law of One Price is true, at least to a first order of approximation: any large price change for oil internationally has carried into oil markets in all regions. On the other hand, Weiner (1991) argues that the immediate and perfect interconnectedness implied by our simple model does not appear to be empirically valid. Price differentials between regions of the world tend to change over time, as market conditions vary, rather than moving strictly together as the "one pool" analogy would suggest. Weiner's results are somewhat difficult to interpret. They could reflect, as he suggests, varying degrees of market power in different regions, which allow sellers to exercise some degree of price discrimination as market conditions change. However, short-term inflexibilities in transportation systems and unique local demand and supply conditions for specific grades of crude may also affect the results, as may varying lags in shipment time, differences in the effectiveness of arbitrage responses in different markets, and differences in contractual terms governing price adjustments. More recent research (Gulen, 1999; Kleit, 2001; Fattouh, 2010) has questioned Weiner's conclusion, finding that world oil markets exhibited a large degree of integration. Overall, we are willing to accept the simple model of an interconnected world oil market as a reasonable way to depict the general structure of world oil prices while accepting Weiner's caution that it may not properly capture very short-term market adjustments for specific crude oil grades or locations. By way of example, in 2010, the price of WTI oil at Cushing, Oklahoma (and also the price of Alberta crude oil which is linked to Cushing) fell significantly below that of North Sea Brent oil for the first extended period of time; this price differential still existed at the time of final editing of this volume (March 2013). The reason lies in a relative surplus of supply at Cushing and a shortage of pipeline capacity to connect Cushing with the rest of the oil market. The former reflects increasing oil production from the Alberta oil sands and the mid-west United States. The latter results from the lags in building new pipeline capacity to handle the increased output; this, in turn, may result from the difficulties in planning pipeline capital investments when oil production levels are uncertain, but it also comes from unexpectedly long lags in obtaining regulatory approval for pipeline construction.

Highly interconnected international gas markets have not developed, in large part because natural gas is so much more costly to ship, per unit of energy content, than oil. This is true for movement by pipeline but is particularly so for ocean-going tankers, which require facilities to liquefy and regasify the gas at either end of the shipment route. In addition, there are no essential energy needs that require natural gas specifically, and a high-cost pipeline distribution system is necessary to move natural gas to consumers, so that many regions of the world do not use much or any natural gas. The result has been that, up to now, natural gas markets have been regional rather than international. Therefore, it may make sense to speak of a North American natural gas market, and to expect the price of Alberta natural gas to be influenced by supply conditions in Texas and demand conditions in California, among other locations. However, the concept of a world natural gas market, with Alberta natural gas prices directly tied to Algerian gas supplies according to the Law of One Price, is not particularly useful. Of course, Algerian natural gas developments could affect the Alberta natural gas market. However, in the absence of low-cost transportation for gas bound from Algeria to North America, any interconnections are likely to be more indirect. For example, a higher supply of Algerian natural gas to Europe by pipeline under the Mediterranean means lower European gas prices, thereby reducing the European demand for oil; this in turn might lead to lower world oil prices, which would lower oil prices in North America and lead to a reduced demand for natural gas and lower Alberta natural gas prices. But the mechanism here is the world oil market and regional energy markets for products such as natural gas, which are tied to local conditions in the oil market, not a world natural gas market. Of course, natural gas prices in North America could rise to a level high enough that large-scale imports of LNG from Africa and the Middle East become economic, so that the natural gas market also becomes a world market. Significant cost reductions in the movement of natural gas across oceans would further this globalization.

Recently there appear to be supply possibilities for large-scale 'non-conventional' natural gas production within North America (such as 'tight' gas, gas from shale formations, and coal bed methane) at prices lower than those required for large-scale LNG imports. It is not yet clear whether the additional supplies are sufficient to keep North American gas prices low enough to enable exports of LNG from North America; if such exports became feasible, the North American natural gas market would become globalized with prices tied to those in the export markets.

2. Alberta's Role in the World Oil Market

Alberta is a relatively small player in world oil market. Table 3.1 shows that Canada's oil output in 2011 was 4.3 per cent of the world total. Canada's share of proved reserves of oil was 10.6 per cent. (This percentage includes non-conventional bitumen and synthetic crude. This is the first year in which the BP source has included large oil sands volumes in Canada's oil reserves.) Similarly, Alberta's oil consumption, even all Canada's, is a small part of total world oil demand. In 2011 Canada consumed about 2.5 per cent of the world's oil (BP, 2012). For this reason, variations in Canada's supply and demand for oil tend to be relatively insignificant for the world oil market. It is common, in fact, to assume that Canada is essentially a 'price taker' in international oil markets. Certainly, the actions of any single Canadian oil producer or consumer are so small that they will have no discernable effect on the world crude oil price; more generally, plausible variations in the Canadian oil supply (certainly Canadian conventional oil supply) or Canadian demand for refined petroleum products (RPPs) are small enough that they would generate barely noticeable effects upon *world* oil prices. The effect upon *Alberta* oil prices will tend to be small, but not necessarily negligible, depending upon the size of the regional markets in which Alberta oil sells, as well as the responsiveness of demand to price changes. Consider, for example, an increase in Alberta oil supply in the framework of the simple market depicted in Figure 3.1. If imports of oil into the watershed market (Montreal) are large enough, some incremental Alberta oil supply may be absorbed there with no price change by backing out Nigerian oil. However, if imports of Nigerian oil are not large enough, then the price of Alberta oil will have to fall slightly, thereby encouraging more consumption of Alberta oil in all markets up to Montreal, and shifting the watershed slightly to the east. The 'price taking' assumption simply says that reduced Nigerian oil sales to Canada are such a small part of total world oil supply that no noticeable change is needed in the world (Nigerian) oil price. Hence, price changes for Alberta oil due to changes in the location of the competitive interface will be small; ripple effects in the 'world oil pool' will be minimal.

As was noted at the end of the previous chapter, Alberta's non-conventional oil resources, in the form of tar sands and bitumen deposits, are very large in comparison to current world oil reserves, with estimates of potential reserves of over 350 billion barrels (Alberta Energy Resources Conservation Board [ERCB], *Reserves Report*, ST-98, 2010). This is over one third of current estimated world reserves, and 32 per cent higher than Saudi Arabia's 2010 reserves. One can appreciate that any technological change, or crude oil price rise, which would be significant enough to make Alberta's non-conventional oil resources economic, would be of tremendous importance, not only to Alberta, but to the world oil market generally. In fact, OPEC spokesmen have expressed awareness of the importance of pricing their oil below the cost of such large-volume alternatives as tar sands and oil shale hydrocarbons. An implicit assumption underlying these concerns is the belief that non-conventional oil output will prove to be very supply-elastic above some critical price; that is, large quantities could be

Table 3.1: World Oil Reserves and Production, 2011

	Share of World Reserves (%)	*Share of World Production (%)*	*R/P Ratio (Reserves/ Annual Production)*
Canada	10.6	4.3	136.5
United States	1.9	8.8	10.8
Mexico	0.7	3.6	10.6
North America	13.2	16.8	41.7
Ecuador*	0.4	0.7	33.2
Venezuela*	17.9	3.5	298.6
South and Central America	19.7	9.5	120.8
Norway	0.4	2.3	9.2
Russia	5.3	12.8	23.5
United Kingdom	0.2	1.3	7.0
Europe and Eurasia	8.5	21.0	22.3
Algeria*	0.7	1.9	19.3
Angola*	0.9	2.1	21.2
Libya*	2.9	0.6	269.4
Nigeria*	2.3	2.9	41.5
Africa	8.0	10.4	41.2
Iran*	9.1	5.2	95.8
Iraq*	8.7	3.4	140.1
Kuwait*	6.1	3.5	97.0
Qatar*	1.5	1.8	39.3
Saudi Arabia*	16.1	13.2	65.2
United Arab Emirates*	5.9	3.8	80.7
Middle East	48.1	32.6	78.7
China	0.9	5.1	9.9
Indonesia*	0.2	1.1	11.8
Asia Pacific	2.5	9.7	14.0
OPEC Total	72.4	42.4	91.5
World Total	100.0	100.0	54.2

Notes: Reserves estimates are for December 31, 2011, and are "proved reserves" defined as "those quantities which geologic and engineering information indicate with reasonable certainty can be recovered in the future from known reservoirs under existing economic and operating conditions." Reserves estimates for non-conventional oil in Canada ("under active development") and Venezuela (the Orinoco Belt) are included. Libya's production is unusually low reflecting the 2011 political instability associated with the Arab Spring.

* Member of OPEC (except for Indonesia which has had a suspended membership since 2009).

Source: BP *Statistical Review of World Energy*, June 2012.

forthcoming with very little increase in cost. If this were true, then non-conventional oil might help play a role as a part of the eventual backstop technology for crude oil. A 'backstop technology,' strictly speaking, is an energy source available in essentially unlimited quantity, at constant cost, which is a perfect substitute for conventional crude oil (Nordhaus, 1973). Such a backstop would set a ceiling price for crude oil.

Table 3.1 makes clear the very uneven geographic distribution of conventional crude oil reserves, particularly their concentration in the phenomenally productive, low-cost oil fields of the Middle East and North Africa, which held over 50 per cent of world reserves at the end of 2011. The political instability of this region raises concerns about the security of supply of much of the oil moving in international trade, although it is easy to overstate security risks. Most of these Middle East reserves, plus those of a number of other countries, amounting in total to over 70 per cent of world reserves, are held by members of the Organization of Petroleum Exporting Countries (OPEC).

OPEC is an intergovernmental group of 'third world' nations that rely upon crude oil exports for the majority of their foreign exchange earnings. OPEC was formed in 1960 by Venezuela, Iran, Iraq, Kuwait, and Saudi Arabia, and, by the end of 2012, had grown to the thirteen members listed in Table 3.1. (OPEC membership has changed somewhat over the years. Ecuador suspended its membership in 1992 but indicated intent to rejoin when domestic economic conditions improved; it re-entered in October 2007. Gabon departed in the mid-1990s. In 2007 Angola joined. Indonesia suspended its membership in 2009.) OPEC's share of world oil output in 2011 was just over 42 per cent, much smaller than its 72 per cent share of oil reserves. Clearly OPEC nations in general are using their proved reserves much less intensively than other countries. The ratio of end-of-year reserves to annual output (R/P) shown in the last column of Table 3.1 confirms this. In 2011 the United States, for example, had an R/P ratio of 10.8 in comparison to OPEC's 72.4. Most economists would attribute OPEC's output restraint to a desire to maintain cartel-like high oil prices, rather than to a lack of 'need' for revenue or to a conservationist concern about the long-term availability of energy supplies for consumers. Some OPEC members exhibit much higher R/P ratios than others, but only Algeria and Indonesia had ratios smaller than 20 in 2011.

It is interesting to note that, despite growing production, remaining proved reserves of oil in the world have increased over the past four decades. Remaining

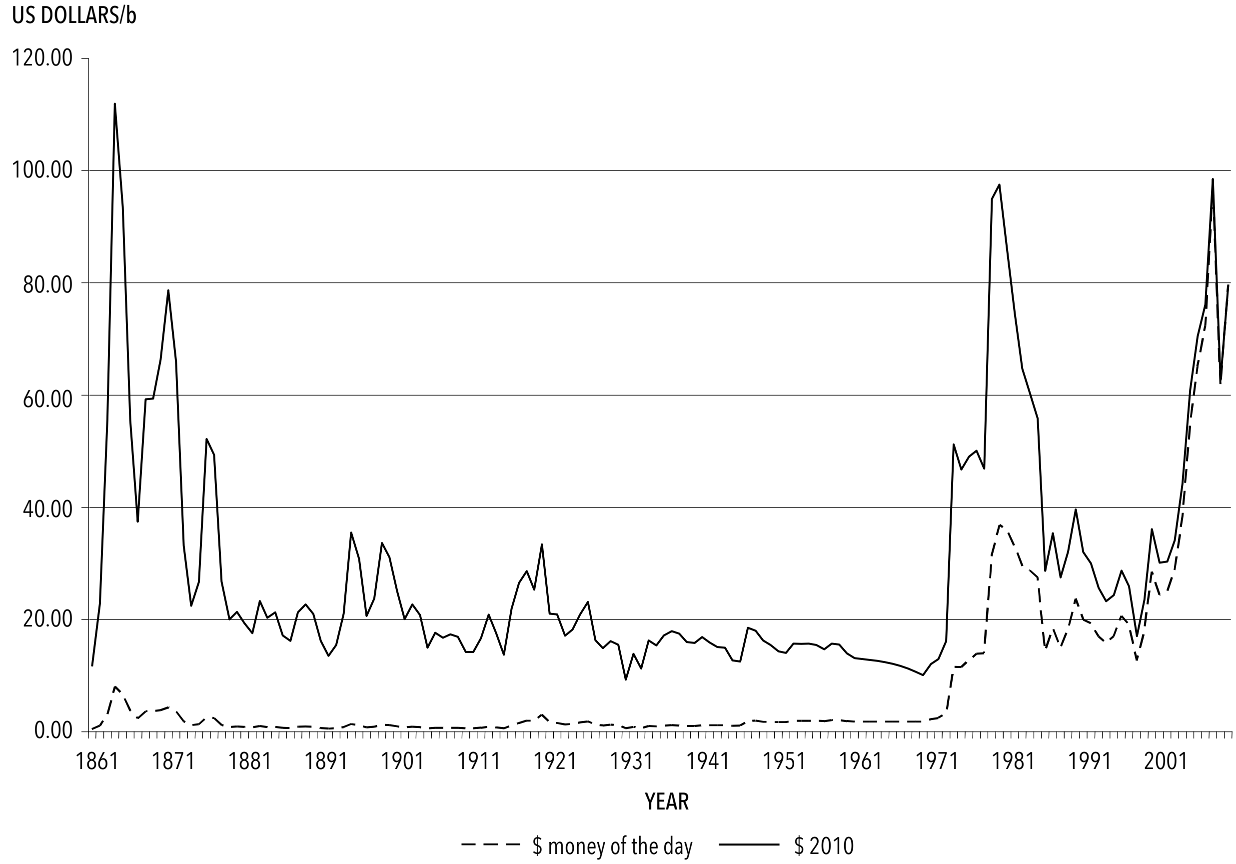


Figure 3.2 World Crude Oil Prices, 1861–2010

Source: BP *Statistical Review of World Energy*, June 2011

reserves of about 80 billion barrels in 1948 increased to about 350 billion barrels in 1965, doubled to just over 700 billion in 1985, and rose to over 1 trillion barrels by 1990, staying relatively constant at this level despite continuing production, then rising more recently to over 1600 billion barrels, partly as a result of the inclusion of non-conventional oil reserves in Canada and Venezuela (Dahl, 1991; BP, 2012). OPEC's share of remaining reserves has also been increasing, particularly in the late 1980s when Venezuela, Iran, Iraq, Saudi Arabia, and the United Arab Emirates (UAE) all added appreciably to their reserves.

3. Determination of World Oil Prices: History

We have argued that world oil prices are of critical importance to Alberta and Canada but that Alberta is such a minor player in the world oil market that these prices are determined independently of any events in that province. What has been the pattern of international oil prices, and how have they been determined?

Figure 3.2 (from BP, 2011) provides an overview of the historical development of world crude oil prices, since the beginnings of the world oil industry in about 1860. The higher of the two lines on the figure shows the price that a reasonably well-informed buyer might expect to pay for crude oil at a major world production location, expressed in real U.S. dollars of 2010 purchasing power. The lower line gives the current dollar price of oil as actually paid in any year before 2010 (at the prices of that year); it is lower than the real 2010 dollar price due to inflation in the world economy over most of this period. The real price of oil provides a measure of the price of crude relative to commodities in general, and so is a more valid indicator of oil price changes than a nominal series showing actual price quotations in each year. The figure shows an average field price for U.S. crude up to 1944, the posted price for Saudi Arabian 'light' at the Persian Gulf from 1945 to 1985, and the Brent spot price since then. (Brent refers to a blend of North Sea light crudes. The changes in pricing location imply slight inconsistencies in the series in 1945, with Saudi prices slightly lower than U.S. prices, and in 1986 with Brent prices slightly higher than Saudi Light.) The pattern of movements shown would hold for other grades of crude oil at other locations, though there would be small differences in values reflecting quality differentials and transportation tariffs. Price trends would look similar for purchasers outside the United States, although, again, changes in real crude oil prices would differ somewhat to the extent that other countries experienced different inflation rates than the United States and/or currency fluctuations vis-à-vis the U.S. dollar. These two effects tend to offset one another to some extent, since inflation rates markedly higher (lower) than in the United States tend to be accompanied by depreciation (appreciation) of the local currency relative to the U.S. dollar. Most refined petroleum product prices would also show the same broad trends as crude oil prices; for some products, however, changes in government consumption taxes have played a major role in changing final retail prices over time.

There is no simple long-term trend in crude oil prices. This casts doubt on the hypothesis that over time increasing physical scarcity of a depletable, non-renewable natural resource such as crude oil will necessarily bring increasing economic scarcity (in the form of real price increases). While each year-to-year price change will have its own unique set of underlying stimuli, it is useful to distinguish three broad periods of time: 1860–1930, 1930–70, and 1970–present. (Many authors have reviewed the history of international oil prices; see, for example, Adelman, 1972, 1995; Blair, 1976; Danielsen, 1982; Frank, 1966; Hamilton, 2011; Hartshorn, 1967, 1993; Jacoby, 1974; Odell, 1986; Penrose, 1968; Yergin, 1990.) Dvir and Rogoff (2009) provide a statistical analysis of world crude oil prices that finds three historical sub-periods similar to those noted here, though they date the end of the first period at 1878, while finding the years 1878–1934 to be more unstable than their first period.

Period 1, 1860–1930. International crude oil prices were dominated by events in the United States, which was far and away the world's largest oil producer. Apart from two years at the turn of the century when Russia ranked first, the United States produced more oil than any other country every year until the early 1970s and was a net exporter until the late 1940s. By the turn of the century, a pricing pattern for oil had emerged, which was called 'Gulf Plus' pricing; the delivered price of oil anywhere in the world outside of North America was equal to the price of oil at the U.S. Gulf of Mexico (i.e., Texas) plus the cost of shipment from the Gulf of Mexico to that market. Gulf Plus pricing generally held up to World War II, apart from occasional local price wars and excepting markets that deliberately isolated themselves from the world market (e.g., Mexico after nationalization in 1938 and Russia after 1917).

Until the 1930s, U.S. crude prices showed instability occasioned mainly by the operation of the 'rule of capture' in conjunction with erratic reserve additions (McDonald, 1971; Daintith, 2010). As discussed in Chapter Two, the rule of capture is a legal provision that ownership of a mobile or fugacious natural resource (such as wildlife and water, or petroleum in an underground reservoir) belongs to the party that captures (lifts) the resource. Typically, in the United States, the mineral rights that landowners issued to oil companies covered small areas so that more than one potential producer would have access to any oil pool. The oil beneath the surface was essentially unowned; that is, it was 'common property' to the producers with access to the pool. It is as if keys to your storeroom were owned by your competitors (Adelman, 1964). As a result, there was a powerful incentive for each producer to attempt to lift the oil early, rather than risking its capture by a neighbouring competitor, and despite the damage this might cause to the reservoir's natural drive and future producibility.

Combined with the unpredictability of discoveries, the rule of capture induced pronounced cycles of production, with resultant price instability. Discovery of a major new oil play attracted eager explorers and there was a rush of discoveries and rapid output increases under the stimulus of the rule of capture; market prices would plummet. However, before long, rapid production decline in reservoirs, occasioned by the high initial output rates, would reduce output and force prices up again until the next major oil discovery in a new geologic play. Refined product prices exhibited less price instability over this period, in part because unstable crude oil prices made up only a part of refined product prices. In addition, for much of this early period the refining sector of the oil industry was subject to strong anti-competitive corporate control, as exemplified by the Standard Oil Trust from the

1870s through to the 1911 U.S. Supreme Court decision that forced the break-up of Standard Oil.

Since the 1930s, the rule of capture has ceased to operate strongly as a factor influencing world oil prices, partly because the U.S. dominance of production began to disappear. Also of importance were government regulations in North America that offset wild production swings (McDonald, 1971; Daintith, 2010). Of particular significance were: (1) market-demand prorationing regulations adopted by several U.S. states (and Alberta), which restricted output to the amount the market could absorb at current prices, and (2) various incentives to companies to 'unitize' oil pools and run them as a single producing operation so that production took place cooperatively rather than competitively. In addition, the rule of capture did not operate in many parts of the world that were of increasing importance after 1930 because British common law did not apply and/or because single production companies (like the Arabian Oil Company in Saudi Arabia) held vast tracts of land so reservoirs were not shared among producers.

Period 2, 1930–70. The period from the early 1930s to the late 1960s is characterized by an unusual degree of price stability, relative to the earlier and later periods. Over these years, oil output in the United States grew, but at a slower rate than output elsewhere in the world, especially in less economically developed nations such as Venezuela, and some Middle Eastern and African countries. The majority of these new oil supplies were controlled by a small number of large, vertically integrated oil companies, the international 'majors,' often referred to as the 'Seven Sisters.' Three of the companies derived from Standard Oil – Standard Oil of New Jersey, now known as Exxon; Standard Oil of California or Chevron; and Standard Oil of New York, or Mobil. There were two other U.S. companies – Gulf and Texaco – plus one British/Dutch company long active in the United States – Shell – plus one company with significant British government ownership – the Anglo Persian Oil Company, now known as BP (privatized in the 1980s).

Despite occasional disagreements and price wars in some local markets, the majors maintained relatively tight oligopoly control in the world oil market outside of North America and the USSR. (As will be discussed in more detail in Chapter Four, an oligopoly is a market in which some producers are large enough that they are able to affect the market price significantly. Thus they have an incentive to restrict production below what a price-taking producer might do, thereby increasing the market price and generating higher profits for themselves.) In earlier years, the companies entered into a number of formal agreements. After World War II interactions between the major producers involved their joint operating agreements in many countries along with more informal (tacit) understandings. Oligopoly markets, dominated by a few large sellers, exhibit an inherent tension between collusive (cooperative) and competitive tendencies. The large producers are motivated to cooperate, by restricting sales, to generate higher prices and profits. However, as a result of the high prices, any one of these collusive producers has potential additional production available with costs lower than price. It is tempting to produce this output, in violation of the formal or tacit group understanding, but such production will push the price down. These contradictory behavioural impulses clearly serve as a source of potential price instability in oligopolized markets. In such an uncertain situation, producers in a reasonably well-functioning oligopoly may tend to adopt a pricing and marketing strategy of relative inaction, to avoid sending frightening signals to competitors and to maintain the status quo. Such oligopolistic rigidity seems to fit the international oil market quite well for the period from about 1930 through to the mid-1950s, with the prices for Middle Eastern crude tending to follow prices in the U.S. domestic market; U.S. prices, in turn, were quite inflexible under the domination of market-demand prorationing regulations.

As was described above, prior to the mid-1940s, international oil prices generally followed a system known as 'Gulf Coast Plus,' in which the delivered price of oil anywhere in the world was the price in the Gulf of Mexico plus transportation charges between the Gulf of Mexico and the delivery point. This was true whether or not oil was actually shipped from the Gulf of Mexico to that market. As a pricing system, this made eminent sense when the United States was by far the world's largest producer of oil, the main source of incremental supply, and when most markets did in fact rely on shipments from the United States. However, Gulf Coast Plus pricing became increasingly inappropriate as regions outside of the United States became major oil suppliers. By the late 1940s, the United States had become a net crude oil importer and was beginning to draw supplies from the Middle East so that pricing oil throughout the world as if it came from the United States no longer made sense. The first major break in Gulf Coast Plus pricing had come in 1942 when, at the insistence of the UK Navy, the price of Middle Eastern crude was set equal to the

U.S. price (rather than the U.S. price plus the transportation charge from the U.S. to the Middle East). With the removal of price controls after World War II, U.S. oil prices rose more than Persian Gulf prices. Still, in the early 1950s, Persian Gulf Prices, while lower than those in Texas, followed any changes in U.S. crude prices.

The late 1950s saw the decoupling of North American and international oil prices. International prices fell under the pressure of increased competition, stimulated by the entry into the international oil market of new crude oil suppliers such as the USSR and a number of 'independents.' At the same time, the U.S. Oil Import Quota Program (1957) and the Canadian National Oil Policy (1961) allowed prices for North American-produced crude to remain significantly above the international level (U.S. Cabinet Task Force on Oil Import Control, 1970; Watkins, 1989). From 1959 to 1970 the selling price of international crude fell gradually from $2.08/barrel (U.S. dollars for Saudi Arabia 34° crude f.o.b. Ras Tanura) to about $1.20/barrel. (The 'posted' price, in fact, remained constant at $1.80/barrel from 1960 through 1970 and served as a tax reference price, but this price was widely discounted in sales of oil.) The per barrel 'tax-paid cost' of oil to companies in Saudi Arabia included about $0.10 of production costs and $0.85 to $0.95 in taxes throughout the 1960s, so that corporate profits per barrel fell drastically as the selling price of Middle East oil fell. Falling world oil prices meant that the consumption of oil rose very rapidly, with expanded Middle East and North African production providing much of the increased production.

Period 3, 1970–present. The third historical period, dating from 1970 to the present, is the OPEC-era (see Evans, 1986; Skeet, 1988; and Parra, 2004, amongst others). OPEC was founded in 1960 by Venezuela, Saudi Arabia, Iran, Iraq, and Kuwait, in response to the declines in oil prices in 1959 and 1960. (Lower sales prices meant lower tax revenues to the government.) After 1960, nine other members were admitted: Ecuador, Algeria, Gabon, Libya, Nigeria, Qatar, the United Arab Emirates (U.A.E.), Indonesia, and (in 2007) Angola; as noted above, Ecuador and Gabon left OPEC in the 1990s, with Ecuador rejoining in 2007.

OPEC had no dramatic effect on international oil markets during its first decade. It consolidated its existence and administrative structure, expanded to include new members, encouraged the standardization of the oil tax regimes of the members, and was, through its continued existence and periodic protests, instrumental in persuading companies to accept royalty/tax assessments based upon the unchanging posted price of crude rather than the declining actual sales price. OPEC failed completely in its announced objective of forcing increases in international oil prices, in establishing an output control (prorationing) scheme, and in substantially increasing its overall 'take' per barrel. But it did defend this take in the face of declining international oil prices.

However, since 1971, OPEC has been a dominant influence on world crude oil prices, although the degree of dominance has waxed and waned. The crumbling control of the large multinational oil companies was replaced by OPEC as an even more effective price-determining oligopoly. Initially OPEC exercised its power in the oil market by actions designed to increase the taxes paid by the major oil companies; as the tax-paid cost of oil to the companies increased, they were forced to raise the selling price of the crude oil. However, by the end of the 1970s, the OPEC governments had nationalized the oil reserves of the major oil companies, and the production and pricing of oil was the responsibility of the government-owned 'national' oil companies (NOCs). (Jaffe and Soligo, 2007, provide a recent survey of the relative significance of NOCs and shareholder-owned private oil companies in the world market.) Figure 3.2 demonstrates two ways in which this new (OPEC) oligopoly differed from the old (corporate) one: prices were increased to levels far above those of the 1930–70 period, and international crude oil price instability became pronounced once again. Competing explanations have been offered for the price changes over this period (Gately, 1984, 1986; Griffin and Teece, 1982; Mead, 1979). Our brief survey draws largely on the work of Morris Adelman, who has argued convincingly that the most critical factor is OPEC's role as an effective cartel (Adelman, 1980, 1984a, 1986, 1989, 1990, 1993a, 1995, 2002, amongst many others).

The major price rises after 1970 suggest that OPEC has been a much more effective cartel than were the major oil companies. A partial explanation must lie in OPEC's lack of inhibition in fully exploiting its powers in the market; OPEC is not, for instance, subject to corporate anti-combines (anti-monopoly) laws. However, in its search for higher profits, OPEC remains subject to the overall discipline of the market. Higher prices will induce conservation of energy and substitution away from oil to other energy products, as well as increased output by non-cartel suppliers. As a 'residual supplier' in the market, OPEC must beware lest declining sales more than offset rising prices.

Further, the quest for a price that maximizes OPEC's profits necessarily has certain similarities to a game of blind man's bluff. The OPEC cartel can see the current market situation but has only imperfect information about two key factors: what will happen to sales at prices significantly different from today's, especially over the longer run, and what changes in behaviour, technology, and government regulation may occur in the future (induced, perhaps, by current OPEC decisions). As a result, OPEC's search for the optimal price takes place in an environment of uncertainty in which the optimal price is best seen as a moving target.

OPEC's rise as a successful cartel was precipitated by actions taken independently by Libya and Algeria (both OPEC members) in 1970. Following the closure of the Suez Canal in 1967 and of the Trans Arabian pipeline from Saudi Arabia to the Mediterranean in Spring 1970, profits on North African oil increased. Shipment costs for oil from the Persian Gulf to Europe (now forced to travel around the Cape of Good Hope) increased, thereby raising the delivered price of oil in Western Europe and generating a higher value there for North African oil, which did not face a large rise in transport costs. Under the tax/royalty regime of the time, based on unchanging posted prices (that is, 'tax reference prices,' which were the price levels used in calculating royalty and income tax payments to the government), all the profit increase on North African oil went to the oil companies and none to the government. Libya responded by demanding higher posted (tax reference) prices and a rise in the income tax rate.

The bargaining position of the governments of the oil-producing countries was generally thought to be weak in circumstances like this. For instance, when Iran had nationalized the assets of the Anglo-Iranian Oil Company (AIOC, now BP) in 1951, the majors boycotted Iranian oil and increased oil output elsewhere in the Middle East, particularly from AIOC holdings in Kuwait. Iran found that it was unable to sell any appreciable amounts of its oil and was forced to negotiate a new agreement with AIOC and other international oil companies in order to get back into the world oil market. Libya's situation, however, differed in two important respects from that of the Persian Gulf OPEC members. First, rather than facing a single consortium producing virtually all the nation's oil, Libya had issued oil permits (concessions) to a number of companies. Second, some of these companies were 'independents,' like Occidental Petroleum, which had international oil output only in Libya, unlike the 'majors,' which had productive concessions elsewhere as well. Libya strengthened its bargaining position by negotiating with the companies separately, by focusing initially on the independents, and by ordering production cutbacks for those companies. After the first independent agreed to Libyan terms, the other oil companies soon followed. Algeria, whose oil was produced by French companies, attained similar agreements.

Libya's success in generating more revenue per barrel for the government, through higher posted prices and tax rates, galvanized other OPEC members, whether or not they possessed Libya's temporary locational advantage. OPEC focused its attention initially on oil shipped from the Persian Gulf (which came from Iran, Iraq, Kuwait, Saudi Arabia, Abu Dhabi, and Qatar), demanding changes similar to those agreed to in Libya. The oil companies had always refused to recognize OPEC as a legitimate bargaining agent, insisting on separate single company-single government negotiations. However, this had not worked to their advantage in Libya and Algeria. In late 1970, meetings began between representatives of countries with oil moving from the Persian Gulf and representatives of all the companies producing that oil. On February 15, 1971, the Teheran Agreement was signed by the governments and companies, pretty well agreeing to the OPEC demands. Later in the spring, a similar agreement was signed at Tripoli covering oil that left OPEC from Mediterranean ports, and corresponding agreements quickly followed covering the rest of OPEC oil (i.e., Venezuela, Nigeria, and Indonesia). The Teheran-Tripoli agreements were to apply for five years, bringing 'stability' to the world oil market by providing for an increase in OPEC tax rates and a schedule of moderate posted price rises. We can see now that their real significance was in demonstrating to OPEC that it possessed the ability to increase the price of oil and that neither the oil companies nor the world's major oil importing countries would offer any effective resistance to this.

It is of some importance to note that the rises in posted price under the Teheran-Tripoli Agreements did not directly increase world crude oil prices because, as was noted earlier, oil was not in fact sold by the major oil companies at the posted price. However, the Agreements did raise oil prices in an indirect manner. Since the posted price served as a tax reference price, higher posted prices raised the 'tax-paid cost' of oil to the oil companies, who in turn were pushed to increase the actual selling price of crude to recover these higher costs. OPEC members began to nationalize the crude oil producers beginning in 1972, and the posted pricing system became less important.

Increasing volumes of oil were sold directly by the government oil companies at a price set by the government. Thus, in the early to mid-1970s, the major influence on international oil prices shifted from OPEC's determination of the posted price and tax rates (which affected the per barrel payments the companies made to the governments) to OPEC's direct establishment of a selling price for crude.

Opinions differ on why no effective opposition to OPEC appeared. It might be tied to the basic logistic and coordination difficulties involved in obtaining responses from the oil companies and consumer governments; that is, this side of the bargaining table was effectively competitive. However, others (Adelman, 1972) have argued that in 1971 a number of the other main players also desired higher oil prices. The major oil companies, for instance, had clearly lost much of their power over the oil market in the 1960s. If crude oil prices rose, they stood to benefit by higher profits, which would follow on non-OPEC oil, and even on OPEC oil, if they could increase the selling price of oil by more than increased taxes to OPEC. U.S. oil had become increasingly non-competitive throughout the 1960s, and higher oil prices would both make the relatively high U.S. oil prices more politically tolerable and bring oil costs in the rest of the world more in line with those faced by U.S. businesses. Moreover, most of the international oil companies were U.S.-based, so that higher international oil profits benefited U.S. citizens and the U.S. balance of payments. Other governments with important domestic energy production sectors (Canada's fossil fuels, coal in the UK, Belgium, and West Germany, etc.) may also have seen higher oil prices as desirable. In addition, the rapid growth in oil consumption in the 1960s, spurred by low and falling prices, had brought to an end most of the excess production capacity in North America and had increased the world's reliance on OPEC oil.

The promised stability of the Teheran-Tripoli agreements proved illusory, as some observers forecast (Adelman, 1972). Once OPEC began to receive the fruits of higher prices, why would its members settle back quietly for five years? Later in 1971 and again in early 1973, OPEC insisted that negotiations be reopened, and, in agreements with the oil companies signed in Geneva, posted prices and tax payments per barrel were increased again. Signs of growing disarray in the international oil market led to attempts by participants to stake out positions with a firm foundation. OPEC spoke of the priority of a "Principle of Changing Circumstances," to which responsible governments must respond in the interest of the citizens they represent, while companies increasingly emphasized the "Principle of the Sanctity of the Contract." OPEC's Principle prevailed as the Organization moved to consolidate its power in world oil markets.

The result was an explosion in oil prices in the decade after 1972. This occurred with two separate price eruptions in 1973–74 and in 1979–81. Each followed a four-stage process (Adelman, 1995):

(1) A disruption in international oil supplies (the 1973 Arab-Israeli War, or 1978 Revolution in Iran) put strong upward pressure on oil prices in spot sales. OPEC members not involved in the disruption faced contradictory incentives. On the one hand was a willingness to see oil prices rise, with a corresponding reluctance to immediately make up for all the supplies lost as a result of the supply disruptions. On the other hand, there was an incentive to increase production to generate more revenue at the high spot price.
(2) Other participants in the crude oil market – oil producers outside OPEC, refiners, marketers, oil consumers – had no certain knowledge about how long the crisis might last or how high oil prices might rise. They reacted by stockpiling more oil for their own use or subsequent sale. But the resultant decreased supply and increased demand on the spot market drove the spot price even higher. Data on oil production during the crises suggests that the reduced output during the crisis was quickly replaced by higher output, mainly from other OPEC members, so that the main stimulus to higher prices during the crisis came from the inventory (stockpiling) effects.
(3) Members of the cartel found that the high spot prices largely benefit companies that had purchased oil at the cartel's lower official selling prices. OPEC members individually or collectively acted to raise official prices, ensuring that they captured the incremental oil profits.
(4) After the crisis passed, OPEC kept prices at the new higher level, adjusting its output so as to maintain the price.

Several dimensions of this four-stage process deserve further comment. First, from the mid-1970s to the early 1980s, OPEC operated as a 'price-fixing cartel.' At meetings of the OPEC conference, the group would decide upon a price for a 'marker' or reference crude oil (34° Saudi oil at the Persian Gulf). Individual members were then responsible for fixing appropriate prices for their specific crude oils. Presumably, this was a price (relative to the reference grade) that would generate a demand for this country's oil such that its share of OPEC production was at a level recognized as appropriate by other cartel members. Prior to 1982, OPEC did not publicly announce any agreed-on output levels or quotas for individual members, but if one country set its price too low, so that its sales rose 'too high,' other cartel members would accuse it of 'cheating.' In other circumstances, during supply crises, for instance, with fixed OPEC prices, increases in spot market prices conveyed a message of 'foregone profits' to OPEC members.

Second, OPEC operates under a consensual method of decision-making, rather than through a majority voting procedure; on matters of any import to the group, a policy change can occur only if all members agree to it. More accurately, none must disagree, and it is possible that they may agree to allow different actions by different members. This method of decision-making brings rigidity to OPEC behaviour.

Third, uncertainties in the market meant that, while all OPEC members saw significant gains accruing from high oil prices in the early 1970s, no one knew how high the price should go in order to maximize group profits. Reference was made to the possible costs of alternative energy supplies such as non-conventional oil, but agreement on exactly how far to raise the price was hard to attain, especially if all members had to agree on the new price. Under these conditions, it is easy to see why spot prices may have been taken as an indication of how much consumers would be willing to pay for oil.

Fourth, in the short term, neither consumption of oil nor production by non-OPEC suppliers is very responsive to a price rise. Therefore, OPEC sales declined relatively little immediately after a price rise, even a large one, and higher prices initially raised group profits. The lack of near-term response by non-OPEC producers and consumers was accentuated by government policies in some non-OPEC countries, such as price controls, to limit oil price rises, and higher production taxes. Price controls keep oil consumption higher, and along with higher production taxes keep oil production lower, thereby reducing the fall in OPEC sales as OPEC raises the price of oil.

These factors help explain the peculiar course of the international crude oil price in the 1970s. Instead of a steady price rise as OPEC gained strength, a 'ratchet' process is apparent with periods of relative stability alternating with sharp increases. However, as time passed, longer-run market adjustments began to occur, and OPEC was forced to reassess its strategy.

The fall in prices as the 1980s progressed illustrates the problems inherent in cartel management when responses to its actions are gradual and cumulative rather than abrupt. The Saudi Arabia reference crude had risen in price from about $14.00/bbl in 1977 to as high as $34.00/bbl in 1981 and was still at $28.00/bbl in 1985. By 1985 OPEC was producing at only 50 per cent of the 1979 level. Thus, despite the doubling of nominal prices, members were in about the same earnings position as 1979 and worse off after allowance for inflation. Saudi Arabia was bearing a disproportionately large share of the burden of cutting output (falling from over 10 million barrels per day (b/d) in 1979 to under 3 million by early 1985). This reflected the Saudi adoption of a 'balance wheel' (or 'swing producer') role in 1982, where it agreed to vary production above or below its quota amount as necessary to maintain the OPEC price. Continuing declines in OPEC's share of the market were in prospect at these prices; OPEC had overshot the mark with its price increases of 1978 to 1981. Saudi Arabia precipitated the required market adjustment by increasing its output sharply, driving spot prices below $10/barrel and forcing a new OPEC agreement. Revival of the OPEC cartel involved two components: agreement on the desirability of a significantly lower price, more compatible with the longer-term realities of the oil market; and a shift in group strategy from price-fixing to quantity-fixing, as OPEC instituted a system of production quotas.

The change in OPEC's method of operation can be seen as part of what some authors have called 'commoditization' of the world petroleum market (Verleger, 1982). The nationalization of the assets of the major oil companies by OPEC governments in the 1970s meant that the majors no longer controlled large volumes of petroleum within vertically integrated channels. The majors were transformed into large buyers of crude oil, broadening the crude market dramatically. Further, the spot market for oil, and in the 1980s futures and options 'derivative' markets, has grown rapidly. In the late 1960s, about 5 per cent of international crude oil sales were in the spot market, the rest being either intercorporate transfers or sales under longer-term contracts. By the early 1980s, over 50 per cent of sales were in the spot market. These new spot and derivatives markets have many active buyers and sellers, with flexible prices; entry is relatively easy, so that spot and future prices are very reactive to day-to-day perceptions about variations in demand and supply whether occasioned by real events or rumour. Of course, the play of spot and derivatives markets is against the all-critical backdrop of OPEC's basic output decisions. By switching from a price-fixing to a quantity-fixing strategy, however, OPEC has moved to tie its exercise of oligopoly power directly to the volatile spot market, with the prices paid to the country reflecting current market values.

The importance of these changes can be seen in the similarities and differences between the 1990 Gulf War in the Middle East and the two earlier political crises. With the Iraqi invasion of Kuwait in August 1990, spot prices of oil rose markedly, as in previous crises, from under $15/barrel in June 1990 for Saudi Light crude to near $40/barrel in late September. The market reacted sharply to the decreased supplies from Iraq and Kuwait (even though these were largely replaced by increases in output elsewhere, especially in Saudi Arabia), and the levels of stockpiles desired rose sharply. However, unlike the earlier crises, the price of oil fell sharply back to lower levels, reaching around $18/barrel by February 1991. The differences reflect five interrelated factors. First, OPEC had learned the longer-term dangers of pushing oil prices up too high. Second, even though Iraqi and Kuwaiti oil was completely lost to the market in 1991, other OPEC members – especially Saudi Arabia – were more than willing to make up the shortfall (as was necessary if continuing high prices were to be avoided). Third, as a quantity-fixing cartel, with sales prices for these quantities tied to spot prices, OPEC producers automatically benefited from the rise in spot prices during the crisis, unlike the earlier two crises where OPEC-fixed prices lagged behind the spot market and OPEC had to raise the official government selling price to profit from the crisis. Fourth, strategic petroleum reserves (SPRs) were higher in OECD countries, with the United States actually releasing small amounts from its SPR in September. The use of SPRs helped to dampen prices, and knowledge of their potential for a moderating effect may have reduced the incentive to build up inventories. Fifth, the development of oil derivatives markets meant that companies could reduce their exposure to possible price increases during the crises by engaging in futures markets transactions instead of building up inventories. This reduced the pressure on spot prices during the crises.

Since OPEC adopted a quantity-fixing approach to the oil market in 1987, almost all crude oil sales take place at fluctuating spot prices. The spot prices are highly variable, reflecting current market conditions and expectations. On an annual basis, from 1987 to 2002, the price of Saudi Arabia light oil was as low as $12.16/bbl (1998) and as high as $26.24/bbl (2000). On a daily basis, the price was as high as above $40/bbl (during the Gulf War and the uncertainties of early 2003 leading up the U.S. invasion of Iraq); it fell below $10.00/bbl. when world oil consumption declined during the economic crisis in the Asia Pacific in 1998. Participants in the world oil market have had to learn to live with price variability. There was no obvious trend in price over this time period, although prices from 1987 to 1999 were largely in the teens, followed by about four years with prices more frequently in the twenties. This was reflected in OPEC's 'target' price, which was $18.00/bbl. in the 1990s (and normally not attained!), but was increased to a range of $22–$28/bbl in 2000. (The target price is the average price for a 'basket' of seven specific crude oils.)

However, beginning in the year 2004, as shown in Figure 3.2, international crude oil prices rose dramatically, in a manner much welcomed by OPEC. (Smith, 2009, provides a good discussion of world oil prices with an emphasis on events from 2000 to 2009; see also Hamilton, 2009.) For example, Saudi Light oil was priced at $27.08/b as of the start of 2004, $31.86/b in 2005, $50.86/b in 2006, and $55.94/b in 2007. Midway through 2007, prices increased markedly again, so that Saudi Light was priced at over $88/b by January of 2008, and still on an upward path, hitting a high of $136.02/b for the first week of July 2008, well above the previous real price maximum in 1980. (It should be noted that the real price refers to international prices expressed in terms of real dollars of U.S. purchasing power. For the many countries, like Canada, the UK, Australia, and those in the Euro in Western Europe, the real oil price was still significantly below the heights of 1980 because their currencies had been appreciating relative to the U.S. dollar.) These high prices generated tremendous revenue increases for the OPEC nations. OPEC's March 2006 *Long-term Strategy* document re-stated OPEC's verbal commitment to helping maintain 'stability' in oil markets but included no mention of a long-term target price (OPEC, 2006).

However, the high prices of mid-2008 were short-lived, as prices fell steadily. Saudi Arabia Light averaged $38.35/b in the last week of December 2008! International crude oil prices then resumed a generally upward trend, reaching $71.58/b (for Saudi Light) in the first week of August 2009. From then until the time of final revision of this book (Spring 2013), prices have fluctuated but in a narrower range than that from 2004 to mid-2009, with a low weekly average price for

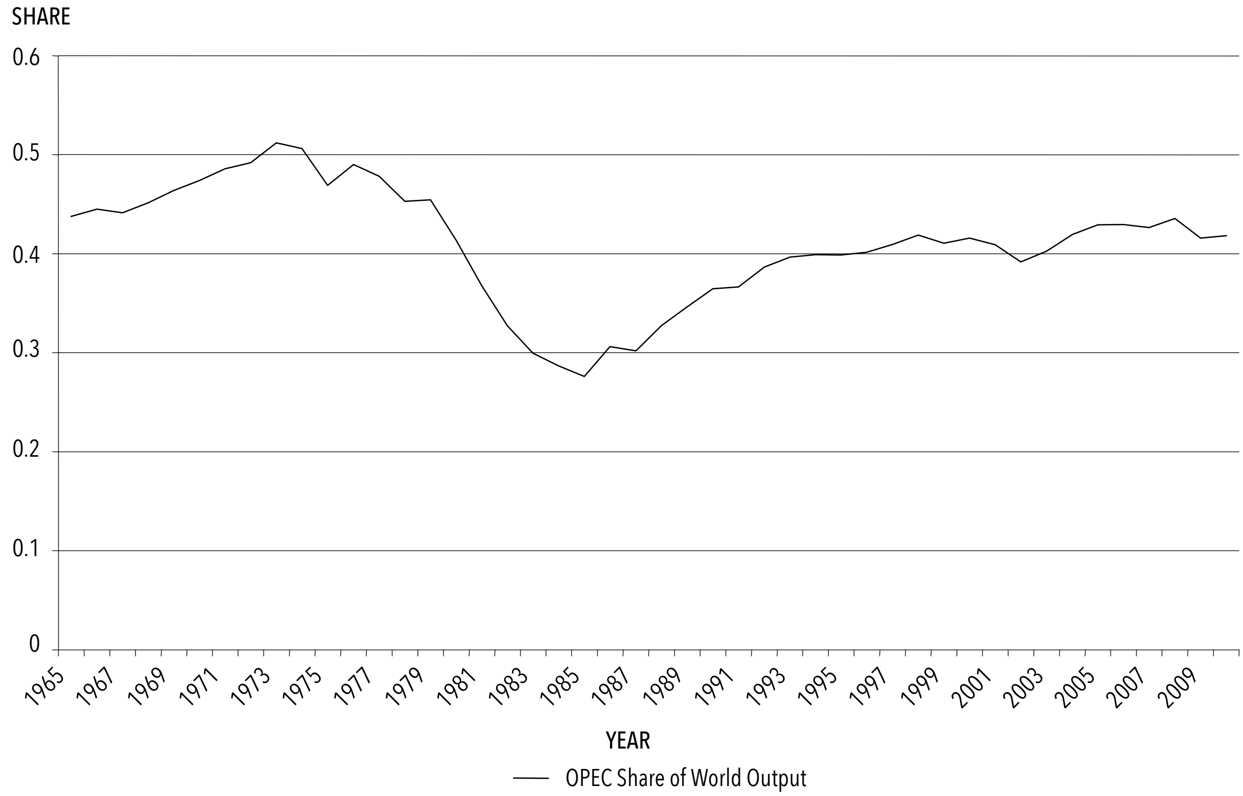


Figure 3.3 OPEC Share of World Crude Oil Output, 1965–2010

Saudi Light of $66.33/b (mid-September 2009) and a high of over $100.00/b (by the end of February 2011).

We now turn from this brief historical review of world crude oil prices to a more general discussion of the major factors that determine the level of oil prices in the world market.

4. Major Determinants of World Oil Prices

It is tempting to explain the evolution of oil prices by the physical 'realities' of petroleum as a scarce natural resource. For example, the inevitable depletion of the resource base may be called upon as justification for forecasts of higher prices. However, as mentioned above, over a century of history does not show a consistent tendency for higher real oil prices. Others have suggested that the geographic distribution of world oil necessitates increasing dependence on OPEC, especially Middle Eastern oil. Figure 3.3 shows that the world did become increasingly reliant on OPEC oil from 1965 to 1973, but that dependency fell after that. A missing element in these physical explanations is the operation of the market, specifically the slowdown in total oil consumption (and, hence, production) and the increase in non-OPEC production (especially outside the United States and the USSR) after OPEC forced oil prices up in the 1970s. Physical realities must underlie what occurs, but petroleum prices reflect economic factors.

It is sometimes argued that OPEC producers maintain high R/P (reserves to production) ratios because oil left in the ground has a higher value to them than oil produced today; in terms that will be set out more fully in Chapter Four, oil has a relatively high 'scarcity value' or 'user cost.' In strict economic terms, based on oil prices, the argument makes little sense. A simple numerical example illustrates this. If we abstract from production costs (which are very low for most OPEC producers) and consider the use to which current oil revenue might be put, $60/b, derived from the sale of one barrel of crude oil produced in the year 2010, would have a value of $638 in eighty years if invested at a real rate of return of only 3 per cent per annum. Eighty years is used to approximate the length of time that would pass before a country with a high R/P ratio would be required to draw on any oil left in the ground today. (If the long-term inflation rate were 2% per year, the nominal value of the investment, in dollars of year 2090 purchasing power, would be $3112.) It would make sense to leave oil in the ground as an investment only if oil prices by 2090 were expected to be at these very high levels. This seems unlikely, given many analysts' estimates for such potential backstop technologies as non-conventional oil and various solar/electric/hydrogen alternatives. A more plausible explanation for the high OPEC R/P ratio is a cartel (oligopoly) reason: the group's desire to generate higher current prices and profits by restricting output. (However, Hamilton, 2009, argues that the higher oil prices after 2006 may incorporate a significant pure scarcity value, and, even if they do not now, rising world demand, especially in the developing world, is likely soon to generate such a scarcity value. While this would support competitive world oil prices higher than the production costs in the main OPEC nations, we do not believe it would justify prices of $60/b or more.)

However, the high R/P ratios also point out a major source of potential instability in the oil market, since most OPEC members would have little technological difficulty in expanding oil output significantly. (At an R/P ratio of 25, still far above the 2010 non-OPEC ratio of 15.1, OPEC alone would have produced over 110 million barrels per day in 2010. Total world oil utilization in 2010 was about 82 million!)

Adelman (1990, 1993a) notes another sense in which the potential for rapid oil output increases is greater now, in the OPEC era, than it was when the multinational majors dominated the industry. When the Seven Sisters managed the market prior to 1970, their productive capacity was held quite close to planned output levels. The relative rigidity in pricing and market shares extended to restraint in the installation of new capacity, an accommodation that was, no doubt, fostered by the prevalence of joint production agreements amongst the companies. OPEC, however, has been operating in many years since early 1970s with significant excess producing capacity. In part, the excess capacity reflected the major fall in OPEC output after 1979, but, in addition, a number of members were active in increasing capacity. There has been little formal economic analysis of why countries might undertake expenditures for development that is not utilized. Among the reasons that suggest themselves are: forecasting errors, in the sense that they had expected to need the capacity but subsequently discovered it was not needed; preservation of the ability to capitalize rapidly on increased sales if OPEC should collapse; holding spare capacity in reserve for use during any international supply disruption, like the 1990/91 crisis; using spare capacity as a potential threat to other OPEC members, thereby gaining greater influence over OPEC pricing policy and output quota allocations. Whatever the reasons, an overhang of spare capacity provides a clear threat of potential downward price instability to the market. The loss of Iraq's and Kuwait's 6.5 million b/d or so from the market in late 1990 dented the excess capacity considerably for several years. In addition, production decline as oil pools are depleted will gradually eat up spare capacity, though there is disagreement amongst experts on how rapidly production decline is occurring in the large oil reservoirs of the Middle East. The U.S. Energy Information Administration (EIA, 2013) reports relatively low OPEC spare capacity in the mid-2000s (from 1.0 to 2.0 million b/d in the 5 years from 2004 to 2008) but an increase to 5 million b/d by 2010, followed by reduced levels (2.8 million b/d by February 2013).

While the *potential* for a significant collapse of the world oil price is clear in a market with an overhang of spare capacity and high R/P values for key OPEC producers, the *likelihood* of the scenario is more difficult to estimate. The major inhibiting factor is the widespread recognition by OPEC members that they would all be worse off in this case, particularly if non-OPEC producing countries provide governmental support for their domestic oil production. On the other hand, prices of oil far above marginal production costs may tempt one or more OPEC members to exceed quotas; this behaviour could draw other members into similar action, trying to increase sales before prices plummet. Given the short-run unresponsiveness of world consumption and non-OPEC production to lower prices, even relatively small output increases could generate a significant price reduction and induce responses from other OPEC members. If the short-run elasticity of demand for OPEC oil were −0.2, and total demand for OPEC oil were 22 million barrels/day, an output increase of only 400,000 barrels/day by any single cartel member would be sufficient to reduce price by almost 10 per cent. And no OPEC member would find a 10 per cent decline in revenue to be negligible. (An elasticity of demand of −0.2 means that if the oil price were to rise by 1%, the quantity of oil demanded would fall by 0.2%. Chapter Four provides discussion of the elasticity concept.)

While many OPEC members have often held significant spare production capacity, it is only fair to note that there are disincentives to investing in unused capacity, most importantly the foregone opportunity costs of the required expenditures. This is particularly so given the highly politicized atmosphere in which most OPEC government oil companies operate, where governments would prefer to see oil revenues utilized on programs that generate political capital amongst the population at large. In the early years of the new millennium, OPEC spare productive capacity fell to historically low levels. In such a market, the short-run unresponsiveness (inelasticity) of production and consumption behaviour means that the predominant instabilities in oil prices are in the upward direction in the face of unexpectedly large demand increases or supply disruptions.

From an economic perspective, the crucial element in the determination of international crude oil prices is not the pressure of declining supplies of an exhaustible natural resource, but the oligopoly structure of the international oil market, and the opposing pressures on cartel members to collude for higher group profits or compete (cheat) for higher market shares. As a result, one can build plausible scenarios for international crude oil prices over the next several decades that range from under $10/barrel to over $100/barrel, as well as mixed scenarios that move from lower to higher prices as OPEC exercises more or less production restraint in the face of changing market circumstances, and as periodic political crises occur in the Middle East. Pindyck (1999) argues that no single empirical model of OPEC is likely to prove adequate. This is a realistic message but not a comforting one for producers, consumers, and governments who cannot avoid making policy and investment decisions that hinge on long-term oil prices.

Some oil market analysts find it useful to distinguish between day-to-day or month-to-month shorter-term market fluctuations and the underlying longer-term phenomena. Given the commoditization of the crude oil market discussed above, the short-term unresponsiveness of consumption and production to price changes, and the feasibility of short-term storage, prices in spot and futures markets may change rapidly and significantly. As a result, commodity trading activities, including the extensive trading in derivatives, have become critical for oil companies.

It is common to argue, however, that the price instability in the spot market is tied to, but relatively independent from, longer-term 'market fundamentals.' The hypothesis is that the *average* level of spot and futures prices is tied to longer-term factors, but that most changes in spot prices are largely independent of the longer-term underlying determinants of price. The *trading* departments of oil companies are vitally concerned with spot market prices that respond to seasonal factors, crises in the Middle East, breakdowns in facilities, changes in patterns of inventory behaviour, and all-and-sundry news reports about possible changes in government oil policies, OPEC solidarity, new technologies, etc. But the *investment* divisions of companies are vitally concerned with the longer-term fundamentals. It is important to note that the hypothesized independence of short- and long-term influences on oil prices was not true during the supply crises of the 1970s, when OPEC used spot prices as a signal for their longer-term pricing strategies. As discussed earlier, the shift by OPEC to a quota-fixing strategy, and recognition of the unreliability of spot prices as an indicator of sustainable long-term prices, make a repetition of the crisis-driven price responses of the 1970s less likely. However, we should note that governments may become very quickly attached to the rapid rise in petroleum revenues from an oil price increase. (Revenue rises since, in the short term, sales fall very little as price increases.) However, such governments run the risk of very substantial sales declines in the longer run as consumers and other producers adjust their behaviour to the higher prices.

The longer-term course of world oil prices hinges critically on the cohesion of the OPEC cartel in light of the usual (though difficult to predict) evolution of economic growth, technological change, discoveries of new oil plays, developments in mature areas, changes in consumer tastes, and government policies.

As noted, a 'cartel' is an endeavour by producers to generate profits by withholding output from the market to induce higher prices. A cartel is subject to high tension between this collusive tactic and an opposing temptation to expand output and cheat on the group agreement. OPEC members are all sovereign states, not subject to any binding international law, so that the ultimate oil output and pricing decisions are those of the individual members. In effect, each OPEC member has the following three decisions it must make:

(1) The individual member of OPEC has some proposal for the actions of the group in its entirely; i.e., some *target price and quantity (quota) for OPEC as a whole* which is consistent with current market conditions.
(2) The individual member has some implicit or explicit proposal of how the group's total output will be allocated amongst the 13 members; e.g.,

proposed output levels (quotas) for the thirteen which add up to the target output.

(3) The individual member must determine actual output or price levels for the various types of oil it produces; these may be consistent with the OPEC group decision or may involve 'cheating' by the individual member.

Numerous factors complicate these three decisions. While any one member will tend to prefer a group decision that favours it, the consensual group decision-making procedure means that none of the other members must find the proposed group decision unacceptable. Since conditions in the oil market are uncertain, the group decisions cannot fix both the price of OPEC oil and the quantity produced. One of the two must be flexible to allow oil markets to find their economic equilibrium. Since 1986 OPEC has functioned as a quantity-fixing cartel, so price is flexible, although the quotas have ostensibly been set with a target price in mind. (Starting around 2000, OPEC briefly flirted with a policy to automatically raise output if the price exceeds the target for a set period of time and to cut production if price falls below the target level. But this policy was not consistently applied.) While OPEC procedures require full agreement, it seems reasonable to suppose that some OPEC members have more ability to achieve their objectives than others. For example, small producers (like Algeria) may be able to attain particularly high output rates (relative to reserves) and, perhaps, successfully cheat because they have such a limited effect on the market. Conversely, the largest producers, especially Saudi Arabia, must be accorded great weight in group decision-making, simply by virtue of their huge reserves and ability to vary output significantly to help attain any proposal favoured and to frustrate any plan not liked. But these are generalizations, and need not be true at any particular time, especially since the high responsiveness of oil prices to short-term output fluctuations gives even middle-sized producers the ability to have a noticeable impact on the oil market.

It is difficult to know exactly when an OPEC member is cheating by overproducing. For one thing, output statistics are not available immediately, nor are they perfectly reliable, and OPEC's auditing procedures have been singularly unsuccessful. Also, one must assume that the quota represents an average to be met over the period of agreement (e.g., half year); but the quota may well be exceeded for some weeks or even months, without any implication of cheating. Another possible ambiguity lies in the definition of what constitutes oil – does it include condensate, NGLs, or very heavy oil (bitumen)? Apparently OPEC interprets the quotas as applying to the production of crude oil in member countries, excluding liquid petroleum derived from natural gas production.

The inescapable conclusion is an awkward one: decision-makers in the oil industry must base their actions on expectations about the future course of international oil prices, but the outlook brackets a wide range of possibilities.

5. Conclusion

The impact of the OPEC governments on world oil prices illustrates 'petropolitics' writ large. To label OPEC a 'cartel,' as do most economists, conveys some economic information, but not a great deal. It implies that members of OPEC cooperate to restrict output and generate higher world oil prices. But cartels come in all shapes and sizes and change form over time. Adelman (1980, 1989) characterizes OPEC as a 'loosely cooperating oligopoly' which has tended to move back and forth between two different modes of operation: a 'full cartel' mode in which all members operate to vary output together to control the market, and a 'residual supplier' mode in which only certain producers (especially Saudi Arabia) take responsibility for controlling the market by playing a balance wheel role. (Hansen and Lindholt, 2008, provide a statistical analysis of the world oil market that seems consistent with Adelman's characterization.) It is easy to be misled by the size of Saudi Arabia's output and reserves and to assume that it is the only OPEC member that really matters. In fact, empirical investigations of OPEC's behaviour over the years since 1973 suggest that almost all OPEC members have been willing to cooperate to some extent to share fluctuations in the demand for the group's oil (Griffin, 1985; Jones, 1990; Smith, 2005). Nevertheless, a cartel always walks a knife edge between collusion and cheating. And the rest of the market must live with the resultant price uncertainty.

What is the relevance of this to the Alberta petroleum industry?

A country such as Canada can have little if any direct impact on world oil prices, at least until large non-conventional oil reserves come into play. Two concluding comments, therefore, are in order. First, if Alberta, or Canada, does wish to address concerns about OPEC's control of international oil prices, or security of supply risks for international oil, the most effective action will involve cooperation with other similarly concerned industrialized nations. Second, apart from such cooperative action, Canadians must accept international oil prices as determining the commercial value of domestically produced oil. These international oil prices are necessarily subject to uncertainty. The short-term instability of oil markets and the political instability of the Middle East mean that very high prices may occur, and a cartel may maintain such prices for some period of time. There is also potential instability on the downside. Cartel members maintain prices by holding output lower than one might expect under more competitive conditions, as demonstrated by the high reserves to production ratios of most OPEC members. However, weakening of cartel resolve may lead to widespread cheating and large price declines. Adelman (1989) notes that this possibility may be made more likely if attempts to gain intra-cartel bargaining power lead members to install excess capacity, since production increases can then occur very quickly.

The difficulties in forecasting international oil prices became all too apparent after the year 2000. For the previous fifteen years, oil prices, except during political crises in the Middle East, were rarely over $20/b. Then prices crept over $20/b and OPEC, its members pleased by the increased revenue, raised its 'target price.' By 2003, the prevailing expectation seems to have been that OPEC would likely be successful in maintaining real crude oil prices at these levels (in the mid-20's per barrel) through the rest of the decade. Instead prices rose dramatically (to over $130/b) by mid-2008. Oil analysts were left grasping for explanations. (For a discussion see Smith, 2009.)

Some saw the rise as largely temporary, reflecting a shortage of spare capacity in the market when faced by unusually large consumption increases and continuing political uncertainty in the Middle East (the war in Iraq, continued Palestinian-Israeli violence, and fears over Iran's nuclear intentions). From this point of view, the high prices include a significant, and presumably temporary, 'security premium.' At some time in the not-too-distant future, prices would come down again, although many expected that OPEC, happy with the revenue increase the higher prices brought, would now try to defend a price in the range of $40 to $60/b. Presumably OPEC would monitor world consumption and non-OPEC production to ensure that its market share did not plummet, as had happened with the price increases of the 1970s.

Other analysts suggested that the large price rise from 2004 to mid-2008 reflected a permanent change in the oil market, with increasing demand, driven by high growth in countries such as China and India, pressing against resource limits for conventional oil. Often those making this argument suggested that OPEC members have overstated their reserves and understated the production decline problems in existing producing fields. Even were this not true, OPEC may simply prefer very high prices, nearer $100/b, so long as there is no evidence of the rapid decline in sales which occurred in the 1980s.

Proponents of these opposing views have relevant criticisms of the other side. The lower-price advocates remind us that, back in the 1973–80 period, it took many years for significant supply and demand responses to occur and that behavioural adjustments are further inhibited if governments (like China) are slow to pass through price increases. Those anticipating continued high prices note that oil inventories did not increase as dramatically as would be expected if the price was as high as it was largely due to a 'security premium,' so the high prices must reflect 'real' factors.

As was shown above, oil prices fell drastically after mid-2008, as forecast by the first of these two lines of argument. However, they quickly rose again and stayed in the $70/b to $80/b range for most of the next two years, then rose up over $100/b. As Figure 3.2, shows, this is higher than prices have been for any time in the industry's history, apart from brief periods in the 1860s and from 1980 to 1984. This might be taken to support the second line of argument but only if prices remain at this level over the longer term.

The history of international oil prices suggests that high, low, or medium prices are quite possible for either brief or more extended periods. Canadian decision-makers – producers, consumers, and governments – have no choice but to live with this uncertain situation.

CHAPTER FOUR

Economic Analysis and Petroleum Production

Readers' Guide: Chapter Four reviews the major microeconomic tools and concepts utilized by economists to analyze the operation of the market for a specific product and shows how they can be applied to the petroleum industry, focusing on the market for crude oil. The tools are applied to a number of specific policy issues in oil economics with specific reference to the objective of economic efficiency. Readers who are well acquainted with the vocabulary and tools of economics may wish to skim this chapter.

1. Introduction

Is petroleum too important to be left to the marketplace? Can it be rational to allow people to purchase scarce oil for use in a third snowmobile? If petroleum were an 'irreplaceable asset' how could we rely upon the decisions of profit-maximizing corporations? Questions such as these are frequently asked and betray a common concern: oil and natural gas – the main energy sources in our modern world – are finite in a physical sense. With these physical limits, how can we trust an allocation procedure based upon selfish economic valuations? To view petroleum as just another economic good is often seen as perverse.

It is not surprising to find that most economists disagree with this thesis!

Most people take for granted the ready availability of goods and services without thinking of the complicated coordination problems involved. In Chapter One, we summarized the many specific tasks that make up the modern petroleum industry. Obviously countless individuals are involved in the production and consumption of petroleum products, and each of these individuals is driven by complicated personal goals. How is it possible to bring together these many interests in an efficient manner? For most economists, the essence of the answer is clear: through the mechanism of the 'market,' which is potentially open to all.

Much economic analysis is concerned with the analysis of the operation of markets as a way of handling society's complex production and consumption decisions. Amongst the goals of economic analysis are the following:

(a) to help us understand the physical and social world in which we live;
(b) to provide useful input into the decision-making processes of individuals, companies, and governments.

Economists have developed a number of tools of analysis to aid attainment of these goals. This chapter includes an introduction to the most important of them, and illustrates how they can be applied to several policy issues in the petroleum industry. Attention will focus upon the concepts of 'supply' and 'demand,' and the precise meaning economists attach to such words as 'competition' and 'monopoly.'

2. Supply and Demand

A. Introduction

Supply and demand are the most ubiquitous of the modern economist's analytical tools. Transferring goods or services from one party to another involves some explicit or implicit exchange ratio. The exchange ratio is approximated by the price of the product, and this price is a major variable in reconciling the interests of buyers and sellers. Attendant conditions of exchange, beyond the market price, may also be a part of the 'true' exchange ratio; for example, service guarantees, credit terms, delivery arrangements, pleasantness and promptness of service, etc. The following discussion abstracts from these considerations. Economic exchange in a market is a voluntary activity, so for exchanges to occur the price of the good or service must settle at a level that is acceptable to both buyers and sellers. Hence economists' tools of supply and demand focus upon market price as the key element in exchange.

The basic concepts of 'supply,' 'demand,' and 'market equilibrium' will be reviewed with specific reference to the wellhead price for crude oil. The term 'supply' (or 'demand') is most frequently utilized by economists to refer to a curve (or schedule) that shows a hypothetical relationship between (1) the quantities of a product that would be supplied (or demanded) in the market place over a particular period of time and (2) various possible market prices of the product. The relationship is hypothetical in that it is not a description of what actually does happen but of what would happen if a certain price were to prevail. For example, a supply schedule for oil might indicate that, at a price of \$70.00 per cubic metre (\$70/m^3), Canadian producers would be willing to bring 300,000 cubic metres of oil a day (300,000 m^3/d) to the market, whereas at \$30/m^3 they would supply only 100,000 m^3/d. It is not the intention of economists to argue that only the price of the product affects the quantity supplied (or demanded). Rather, a particular supply (or demand) curve is defined for a specific and fixed set of underlying variables other than price, and the curve shows the quantities supplied (or demanded) at various alternative prices, given the values of those other variables. A well-functioning market will tend to an equilibrium price at which it clears; that is, the quantity willingly demanded by buyers is exactly matched by the quantity willingly supplied by sellers, and neither buyers nor sellers wish to change their behaviour (at that price).

Prices and costs for these demand and supply functions should be thought of as real (constant dollar) values that show prices relative to other goods in the economy. General inflation in the economy will not ordinarily shift supply and demand curves, since real dollar values are unchanged. What is relevant is a greater or lesser change in input costs and other values for this industry than for the economy in general.

B. Supply

1. The Supply Curve

One could postulate a 'supply function' that shows the relationship between the quantities of oil (in m^3/d) that producers would be willing and able to supply at the wellhead and all the major variables that determine that quantity. We have not said whether buyers are there to take the quantities concerned: this is the meaning of the independence of demand and supply. The reader can probably make a long list of factors that might influence the quantity of oil the producer would be willing to supply. Economists typically handle a complex problem such as this by building a simplified analytical model that is assumed to be a reasonable depiction of the behaviour under study. In the case of crude oil supply, for instance, it is commonly assumed that producers wish to maximize profits, so that the following factors would influence supply.

(a) the **price of oil** (the higher the price, the greater the quantity supplied, everything else affecting supply held fixed);
(b) the **technological conditions** of production, including the nature of the equipment used and the physical characteristics of the reservoir (the better the technology, or more amenable the reservoir, the greater the quantity supplied);
(c) the **costs of inputs** such as labour, materials and supplies, land and capital, including the minimum profit the operator must receive in order to continue his activity, i.e., the 'normal profit' (the lower the costs of inputs the greater the quantity supplied);
(d) the **price of natural gas**, natural gas liquids, and other products produced in conjunction with oil;
(e) other financial charges incurred with production, e.g., royalties and income **taxes** (the lower the taxes the greater the quantity supplied); and

(f) **expectations about the future** values of all these variables (generally, the less favourable to profits the operator expects these to be, the greater the quantity supplied now out of available reserves, but the less attractive are additions to reserves). In addition, in light of uncertainty, production closes off the option of waiting until more information is available; an opportunity cost of foregone anticipated profits may be associated with this (Dixit and Pindyck, 1994). (Kellogg, 2010, finds evidence of such an effect on oil investments in Texas.)

If all factors except the first (the price of oil) were assumed to be fixed at some level, then we arrive at a hypothetical supply schedule of the type described earlier. The supply curve is the locus of points of maximum quantities that would be supplied to the market at various prices. At any given price, suppliers will be willing to supply less, but they can't be induced to supply more. That is, the curve tracks the minimum prices that will induce suppliers to place the various quantities on the market. Suppliers will be happy to accept a higher price for a given quantity but will not supply that quantity for a lower price. The minimum price necessary to entice the supply is called the 'supply price.' To be profitable, this minimum price must cover the cost of producing the unit of oil in question, so the general condition for supply is that price equals marginal cost ($P = MC$) for the last unit of oil the producer is willing to bring to market. ('Marginal' is the term economists use to refer to the individual unit.)

We expect the supply curve to slope upward to the right, showing a greater quantity of oil willingly supplied at a higher price (curve *S* in Figure 4.1), since a higher price will cover the higher-cost units of petroleum that were not profitable to producers at the lower price. Such a curve could be imagined for any unit of production in the industry (well, pool, field, or company). Horizontal summation of the quantities across all units of one type (e.g., oil pools) at each price would generate a market supply curve for the region. This supply curve would have a positive slope both because a higher price may induce more production from any one unit (e.g., pool of oil) and because a higher price makes higher cost units (e.g., low productivity pools) attractive to produce.

We have been speaking of the wellhead supply of oil. Oil is non-homogeneous in two important respects, in a regional producing market. Firstly, pools differ in location, with pools closer to the market in a

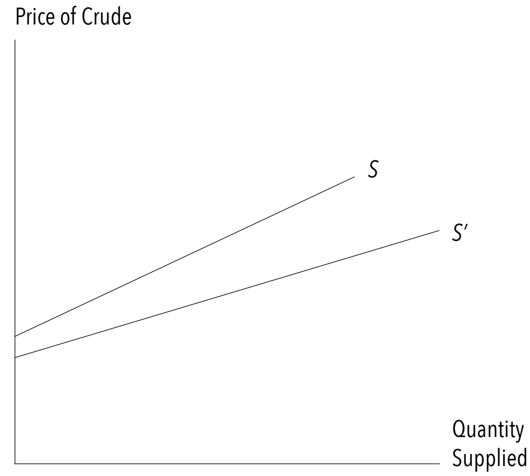


Figure 4.1 The Supply Curve for Crude Oil

preferred position. Secondly, crudes differ in quality, yielding differing arrays of product when refined and possessing undesirable impurities (e.g., sulphur) to varying degrees. Because refined products differ in price, this affects valuation of the crude; also different quality crudes incur different costs of transportation and processing. Hence it is best to think of the market supply of crude oil in a producing region as defined for one specific grade of crude oil at a central gathering point in the region. In Alberta, for instance, the reference oil might be light (42°), low sulphur crude oil at the Edmonton terminal of the Enbridge pipeline. The actual wellhead price for any specific barrel of crude oil will be higher (or lower) than this reference price as: (1) the transportation cost to the central gathering point is lower (or higher) than that for the reference crude, (2) the market value of the array of refined products obtained from the crude is higher (or lower) than for the reference crude, and (3) the cost of moving or refining the crude is lower (or higher) than for the reference crude.

There are, obviously, an infinite number of hypothetical market supply curves derived from the infinite number of possible assumptions about variables other than price that underlie the supply curve. However, at any particular time, one set of underlying variables will be extant and one of the hypothetical supply curves will exist. (This is subject to qualification about 'short-run' versus 'long-run' curves, as will be discussed below.) If the specific value of one of the underlying variables should change, then a new supply curve would be generated. For instance, in Figure 4.1, the curve *S'* shows an increase in supply relative to curve *S*; that is, at every hypothetical price the quantity supplied is greater on curve *S'* than on curve *S*. This increase in supply might result from successful new exploration, improved technology, or reduced royalties or input costs, among other possible causes. Similarly a shift to the left in the supply curve, for instance as a result of higher input costs, represents a decrease in supply. As time passes, decreases in the oil supply curve for a particular well are expected, at least in the later years of operation, reflecting the phenomenon of production decline in oil reservoirs: with the depletion of oil reserves, there is less oil available and the internal production drive of the reservoir falls.

2. Supply and Costs

The supply curve is usually taken to show the marginal cost (incremental cost) of the additional unit of output indicated on the horizontal (quantity) axis. This recognizes that the operator will be willing to produce an additional unit of output when, but only when, the price of the product is as high as the incremental cost involved in providing that unit. It is common (e.g., Davidson, 1963; McDonald, 1971; Watkins, 1970) to divide crude oil marginal costs into three components: (1) variable input costs including labour costs, equipment and material costs, rent, and normal profits; (2) production taxes (e.g., royalties); and (3) the present value (discounted value in today's dollars) of any future profit foregone by producing the cubic metre of oil now instead of leaving it in the ground for later production. A present value is an expected future dollar value multiplied by a discount factor that allows for the return foregone on those funds by having to wait for them rather than having them available now. If r is the relevant annual rate of interest, and the future value would occur in T years, the discount factor is $1/(1+r)^T$.

This third cost element is called the 'marginal user cost' of production, and derives from the consideration a profit-maximizing operator gives to possible conservation of a depletable natural resource like petroleum. The user cost of an oil pool reflects two ways in which current lifting of oil reduces future profit possibilities. The first is a pure 'timing' effect, which reflects the depletable nature of oil deposits so that one cubic metre produced today is simply not available for future lifting. The second is a 'stock' or 'degradation' effect, which captures the internal reservoir dynamics of an oil pool and measures the increase in future production costs caused by the reduction in reservoir pressure as a result of producing the cubic metre today (Bohi and Toman, 1984). Obviously expectations of future prices, costs, and taxes influence the marginal user cost. The higher expected future prices, the higher will be the marginal user cost, and the lower current supply. Producers would be induced to wait for the better market in the future. Such conservation will operate only to the extent that producers are able to make reasonable predictions about the future and expect to control the oil pool then. It is also affected by government regulations that influence output rates, and by any contractual obligations the producer may have undertaken.

The user cost concept is particularly helpful in understanding the development of pools. Why, for example, would a company refrain from drilling a low-cost infill well that produces significant amounts of oil? Typically it is because the infill well has a high cost of foregone future profits because it reduces the later production from adjacent established wells; in essence, the infill well simply accelerates production, but the producer might find the future profits if he does not drill the well more attractive.

In analyzing oil supply it is necessary to distinguish between the supply of produced (lifted) crude oil and the supply of discovered oil in the ground (i.e., reserves) (Uhler, 1981). Exploration, and development activities, such as outpost (or extension) drilling and enhanced oil recovery (EOR), are concerned with the supply of oil reserves. Higher oil prices will cover higher cost reserve additions. Figure 4.2 shows, for a given set of underlying factors (input costs, geological and technical knowledge, future expectations, taxes, etc.) an upward sloping supply curve (S_{RA}) for reserves additions as the price of reserves additions rises. The 'price of reserves additions' is a sales value for oil reserves in the ground (an *in situ* price); it will be greater the higher the price of produced or lifted crude oil but will tend to be lower than the price of produced crude since the operator must still pay lifting costs for the oil and must wait into the future before he is able to recover all of the oil reserves.

Uhler (1976, 1977) argues that the supply curve of reserves additions in a given basin is subject to two contrary dynamic influences: technological improvements and new geological knowledge tend to increase the supply of reserves additions (shift it rightward to *S'*), while the depletion over time of the stock of undiscovered reserves tends to make new additions more costly and shifts the supply curve to the left, to *S"*. Uhler has hypothesized that the first effect tends to dominate early in the history of reserves additions

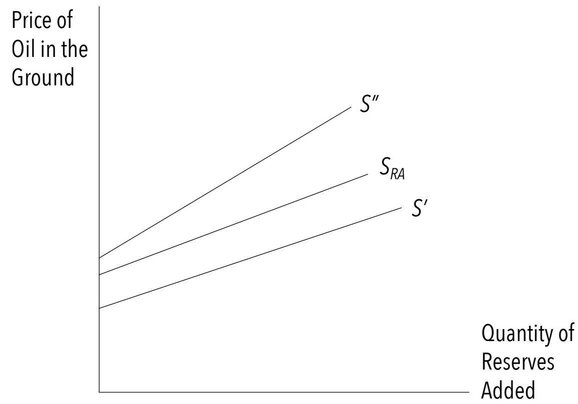


Figure 4.2 The Supply Curve for Reserves Additions

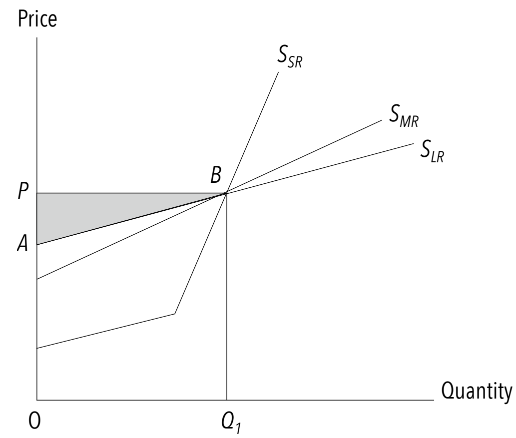


Figure 4.3 Short-, Medium-, and Long-run Supply Curves for Crude Oil

from a particular geological play (or formation), while the second effect dominates later on, thereby giving a tendency to first rising then falling reserves additions per unit of exploratory effort. Over the long run, it is reserves additions that support continued crude oil production.

In this book, we will emphasize the supply of oil production (lifted crude) rather than the supply of reserves in the ground.

3. The Analytical Time Dimension

Economists commonly distinguish between short-run and long-run supply curves. The short-run supply curve describes the relationship between the price of the product and the quantity that operators are willing to supply when some of the factors of production (e.g., major pieces of capital equipment, like the number of wells in a pool) are fixed and cannot be changed. The long-run supply curve describes the relationship between price and quantity when the operator has enough time to vary all inputs. Some writers further differentiate the short-run (when all capital facilities, including the number of wells, are fixed), the medium-run (when already discovered pools can be drilled more intensively and EOR schemes put in place, but no new pools can be brought on stream) and the long-run (when operating, development, and exploration activities are all variable) (McDonald, 1971). These analytical distinctions provide a convenient bridge from physical to economic descriptions of petroleum industry activity. Figure 4.3 shows short-run (S_{SR}), medium-run (S_{MR}) and long-run (S_{LR}) supply curves of lifted crude oil from a region for some particular year. Recall that these are 'hypothetical' curves illustrating what would happen under certain assumed conditions. Figure 4.3 illustrates the variations in supply if capital investment decisions possess a specified flexibility: complete inflexibility in the short-run so only existing equipment is used; flexible development capital but no new exploration in the medium-run; and both development and exploration possible in the long-run. Other factors influencing the supply of oil are assumed to remain at fixed levels.

Consider an initial price such as *P*, which is assumed to apply to identical short-, medium-, and long-run outputs (Q_1). The short-run supply curve is very steep (inelastic) since higher prices can draw forth very little additional output given that no new capital is installed; a higher price does serve to cover the marginal costs of some higher-cost (and, usually, low-output) wells, which would otherwise be abandoned or shut-in. The greater quantity responsiveness of the medium- and long-run supply curves is due to a higher price covering increased marginal costs of lifting and development in previously discovered pools (medium-run) and of lifting, development, and exploration in newly discovered pools (long-run). The shape of the short-run supply curve (S_{SR}) indicates that the price of oil would have to fall very low before it failed to cover the operating and user costs of those wells currently in operation (see for Canada, Edwards, 1972; for the United States, Griffin and Jones, 1986, and Adelman, 1992). The vertical distance between the S_{SR} and S_{MR} curves, for output levels below Q_1, represents the (sunk) development expenditures undertaken in the past to support output up to level Q_1, while the vertical distance between S_{MR} and S_{LR} represents sunk exploration costs.

An increasing cost industry (i.e., one with rising supply curves), such as the crude petroleum industry, will earn unit revenues (prices) higher than marginal costs. For price *P*, and output Q_1, this excess dollar profit (usually called 'economic rent,' or 'producers' surplus'), can be represented by shaded area *APB* in Figure 4.3; this rent is also known as a 'differential rent' or 'Ricardian rent' and measures the difference in cost between the highest cost unit produced (at Q_1) and lower cost units. Economic rent is usually defined to include the user cost component of the supply (marginal cost) curves since this represents an expected future profit, not an expenditure on production or required return on capital; this component of economic rent is sometimes labelled a 'scarcity rent,' deriving from the exhaustible nature of an individual oil deposit.

It is important to note that the differences between the three supply curves relate to an *analytical* time distinction (the time required to invest capital) rather than a *calendar* time distinction. A price rise for oil that was expected to be sustained would immediately induce short-, medium-, and long-run supply responses. Some of these might take place very rapidly (a new shallow exploratory or development well), while others might take many years (a large EOR scheme or frontier exploration program). Therefore, observed real world supply behaviour would include a mixture of (i) oil supply decisions as indicated by supply curves and (ii) the process of adjustment between different supply curves due to shifts in those curves and the process of capital investment.

Between any two calendar time periods the three supply curves of produced oil can be expected to change position. These dynamic adjustments can be credited to two somewhat different forces. First, a number of the factors that are assumed constant in an initial time period may change; examples would include changes in technology and knowledge, in input costs, in taxes, in expectations about the future. Second, the potential supply of lifted oil is affected by the supply of reserve additions. If at current prices (of both oil as produced and oil in the ground) additions to reserves just equal production, then the supply curves for output would tend to remain unchanged. However, if production exceeds reserves additions, then the process of production decline in existing pools would reduce supply (shift the supply curves to the left). On the other hand, if reserves additions exceed current production, then the supply of lifted oil would tend to rise, with the supply curves shifting to the right.

The underlying physical dimensions of the crude oil production decision, and the related economic supply concepts, may be made somewhat clearer by the slightly more extended graphical treatment of Figure 4.4, which examines a single oil pool in more detail. Panel A illustrates the capacity output path over time from an oil pool with a fixed amount of capital equipment in place; this is a short-run situation and the output path, q_t, shows falling production due to the depletion of reservoir energy as cumulative output rises. The dashed line, q_a, shows, for each year, the minimum output level acceptable to the producer. It is the output level that would generate just enough revenue to cover operating costs, and it is equal to the operating costs (*OC*) of the wells divided by the price (*P*) of oil. (If $P_t \cdot q_a = OC_t$, then $q_a = OC_t/P_t$.) Production decline means that actual output is steadily pushed down towards the minimum acceptable output rate; typically this defines a time of abandonment, T_A, at which the oil pool will be shut down. The variable tax component of the short-run production decision is treated sometimes as a component of operating costs (*OC*) and sometimes as a deduction from price (*P*) to yield a net (after-tax) price.

Panel B translates this pool's output path into economic short-run supply curves (abstracting from complications of the user cost component). Given that capital costs are sunk, the variable operating costs of a pool would typically yield a marginal cost curve like MC_o. The decision to operate (produce even one cubic metre of oil) requires an expenditure to maintain the equipment for the year. But once that is done, the incremental lifting cost is very low (e.g., pumping costs only) until capacity of the equipment is reached under the pool's current reservoir conditions (capacity output q_o, where the marginal cost curve becomes extremely steep). Curve AVC_o shows the average variable cost of various possible output levels; the first cubic metre has the same average and marginal cost and, thereafter, the low cost incremental units reduce the average cost, up to output level q_O. Under these conditions the producer will decide either: (1) to produce at capacity, where price, $P_O = MC_O$, so long as average revenue covers average operating costs i.e., $P_O \geq AVC_O$) or (2) to shut down (if $P_O < AVC_o$). Therefore the short-run supply curve (S_O) for the oil pool is the marginal cost curve above the minimum point of the average variable cost curve. Production

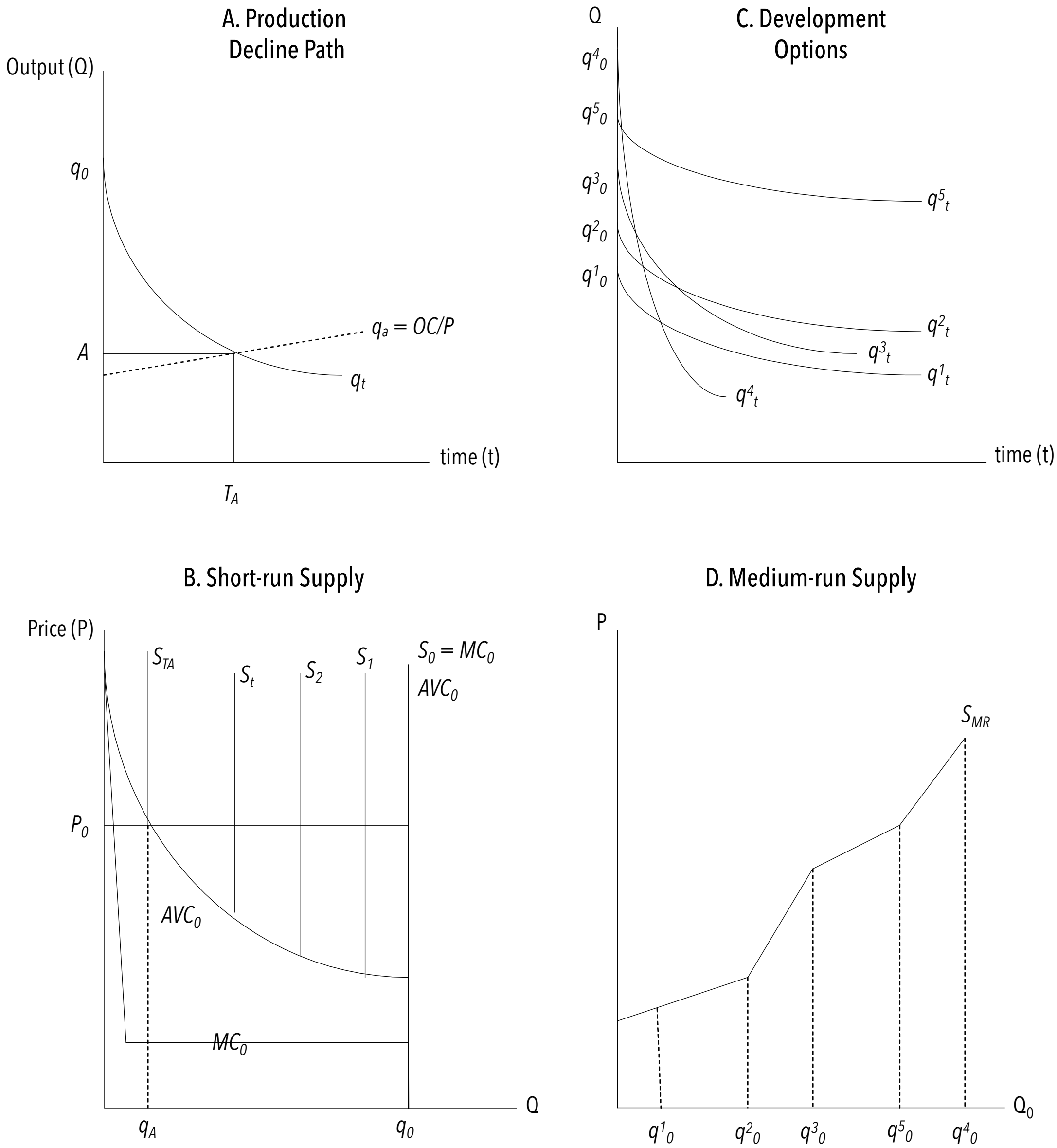


Figure 4.4 Crude Oil Supply: One Pool

decline means that the capacity output level from the fixed capital becomes smaller and smaller, so that the short-run supply curve shifts back to the left-over time (to S_1, S_2, …, S_t) until the year of abandonment (S_{TA}, if the real price of oil is fixed at level P_o).

Panel C of Figure 4.4 shows changes in the time-path of output from the pool as more wells are drilled. Apart from occasional dry holes, additional wells will increase the initial output rate from the pool (q_o). Panel C shows five possible developmental options with progressively higher initial output rates from q^1_o to q^5_o, illustrating different types of medium-run pool development. From an initial development plan (q^1_t), output path q^2_t shows extension (or outpost drilling) which brings more oil reserves into production and allows higher output levels in all years. Output path q^3_t is an example of infill drilling that produces the same reserves more quickly, so that $q^3 > q^2$ in early periods, but $q^3 < q^2$ later. Curves q^2_t and q^3_t illustrate 'pure' cases of extension and infill drilling; in most oil pools, incremental development will involve a gradual transition from extension to infill activities. Output path q^4_t shows a somewhat perverse case of infill drilling in which the high initial output rates significantly damage reservoir flow rates so that the production decline becomes very high. Most petroleum reservoirs have a 'maximum efficient rate' (MER) for oil wells or for the reservoir as a whole, which, if exceeded, results in high output decline rate and a loss in total recoverable reserves. Profit-maximizing producers ordinarily have no incentive to exceed the MER since it means a large loss of future profits (i.e., user costs are very high above MER). A significant exception is when several companies jointly produce from an oil reservoir, have no agreement on sharing output and profits, and operate under the legal convention known as the 'rule of capture' in which ownership of the oil belongs to the party that brings it to the surface. Here companies may exceed the MER since they believe that any oil left unproduced today will be captured by the other companies in the pool. Government agencies like the Alberta Energy Resources Conservation Board (ERCB) often impose regulations that restrict output to levels less than the MER. (See Chapter Ten.)

Finally, output path q^5_t, in Figure 4.4, Panel C, shows the impact of a successful EOR project that raises the recovery factor, thereby bringing new reserves into production, allowing more output in all periods. Panel D of Figure 4.4 shows the medium-run supply curve for crude oil produced this year from the pool shown in Panel C. Pure extension drilling adds output at a cost very close to the initial wells (q^1_o to q^2_o); infill drilling generates higher per unit costs due to well interference and rising user costs as current production reduces future output and profits (q^2_o to q^3_o); EOR schemes are typically more costly than primary recovery schemes ($q^3_o$to q^5_o); output in excess of the MER has very higher user costs of forgone future profits so is very high in cost (q^5_o to q^4_o).

The way in which various development options might be sequenced is very reservoir-specific.

4. Elasticity of Supply

Economists frequently use the term 'elasticity of supply' to describe the shape and position of the supply curve. The concept of 'elasticity' is important: it is the relative (percentage) change in one variable divided by the associated percentage change in a related variable, and shows the responsiveness of one variable to change in another. 'Own-price elasticity of supply' (usually called, simply, 'elasticity of supply') is the percentage change in the quantity of a product supplied divided by the percentage change in the price of the product, all else being equal; that is, it describes movement along the supply curve. Where the symbol 'Δ' is used to represent a change:

$$E_s = (\Delta Q_s/Q_s)/(\Delta P/P) = (\Delta Q_s/\Delta P)(P/Q_s).$$

The own-price elasticity of supply (E_s) is equal to the reciprocal of the slope of the supply curve ($\Delta Q_s/\Delta P$) multiplied by the ratio of price to quantity. The slope of the supply curve by itself is not a satisfactory measure of supply responsiveness. Exactly the same supply information can be conveyed using a number of different quantity measures; i.e., b/d; b/year; tons/year, m^3/month, etc. The same supply curve would yield quite different numerical values for its slope in each of the cases, even though the relative changes in price and quantity were the same. It is desirable to measure elasticity as a pure number, independent of the particular units of measurement chosen, and using percentage changes does this.

A higher price elasticity of supply means that any given percentage rise in price yields a higher percentage rise in quantity supplied. Thus, for instance, the long-run price elasticity of supply for crude oil will tend to be higher than the short-run elasticity as was seen in Figure 4.2.

It is customary to refer to a value of E_s greater than unity as showing 'elastic' supply and a value of E_s less than unity as showing 'inelastic' supply. Empirical work by economists has clearly demonstrated that the elasticity of a supply for crude oil is significantly greater than zero (e.g., Uhler, 1977; Bradley, 1989). This gives lie to one of the common beliefs in energy analysis – that fossil fuel availability is determined solely by nature, and that economics is essentially irrelevant to energy policy: price does affect availability.

C. Demand

Our treatment of demand is analogous to our discussion of supply. One can postulate a 'demand function' that shows the relationship between the quantities (in m^3/d) that purchasers are willing and able to buy and the major factors that influence that desire. Amongst the important factors affecting demand for crude oil at the wellhead would be:

(1) the **price of oil** (the lower the price, the greater the quantity demanded);
(2) conditions in the markets for refined oil products, for example,

(i) tastes of consumers (Do they prefer compacts or larger cars? Do they like really warm houses or cooler ones? Is it an austere, puritanical society, or a conspicuous-consumption-oriented one?);
(ii) **incomes** of consumers;
(iii) **population** in the market;
(iv) **prices of substitute products** (e.g., natural gas, insulation) and of complementary products (e.g., automobiles);
(v) **technological and cost conditions in industries that use petroleum** products in their production processes;

(3) **conditions in the supply industries between the oil field and the final consumer**, e.g., pipe line and tanker systems and oil refineries (the higher the costs of moving or refining the oil, the lower the quantity demanded at the wellhead); and
(4) **current expectations of the future values** for these variables and the price of oil (higher expected prices in the very near-term tend to induce higher current demand to capitalize on today's lower prices, while higher expected prices in the far-term tend to induce lower current demand as consumers purchase non-oil-using capital equipment today).

A demand curve shows the hypothetical relationship between the quantity demanded and various prices of the product, with all the other factors that might affect demand assumed to be constant. The demand curve is the locus of points representing the maximum rate of purchase at the given price; equivalently, it is the maximum price that would be paid for the given quantities. It should exhibit a downward slope since a reduced price for crude oil will both free income for more consumption (including more crude oil) and induce the consumer to substitute crude oil products for products whose price has not fallen (e.g., other fuels). If any of the other factors underlying demand should change, then a new demand curve would appear. In Figure 4.5, for instance, an increase in demand from curve D_{SR} to curve D' can be seen (at every price the quantity demanded is higher on curve D'). This might result from a rise in the price of natural gas, an increase in consumers' income, a rise in population, a fall in refining costs, or the like. In addition, it is useful to distinguish between the short-run demand for crude oil, when the ability of purchasers

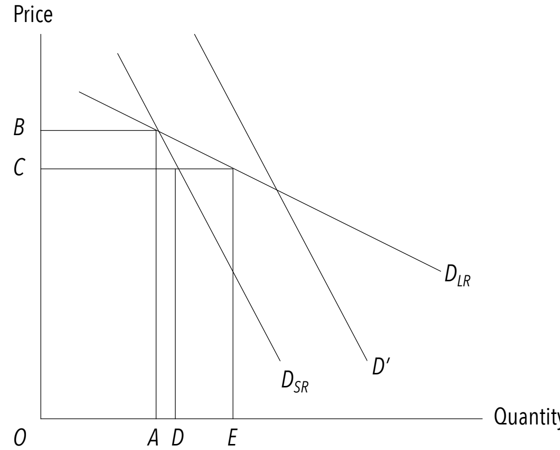


Figure 4.5 Crude Oil Demand

to adjust their capital equipment is assumed to be non-existent, and the long-run demand, shown by D_{LR}, when the stock of energy-using capital equipment can be adjusted in a way that is optimal for that oil price. For example, in Figure 4.5, a fall in price from *OB* to *OC* would lead to an increase in quantity demanded of only *AD* in the short-run, but of *AE* in the long-run.

The (own-) price elasticity of demand (E_d, usually called the elasticity of demand) measures the percentage change in the quantity demanded divided by the percentage change in price of the product, other things affecting demand held unchanged; i.e.

$$E_d = (\Delta Q_d/Q_d)/(\Delta P/P) = (\Delta Q_d/\Delta P)(P/Q_d).$$

The elasticity of demand will normally be negative since a price rise means a fall in the quantity demanded. (It is wise to be aware that some economists drop the negative sign, utilizing the absolute value of the elasticity.) A "large" elasticity of demand (e.g., an elastic demand, where $|E_d| > 1$) implies that a given percentage change in price gives a relatively large change in quantity demanded. The long-run elasticity will exceed the short-run elasticity. Empirical evidence clearly demonstrates that price changes do affect petroleum consumption (e.g., Berndt, 1977; Berndt and Greenberg, 1989; Berndt et al., 1981; Watkins, 1991c). Hughes et al. (2008), however, find that the short-run elasticity of demand for motor gasoline appears significantly lower after 2000 than it was in the 1970s and 1980s.

An interesting relationship exists between the price elasticity of demand and changes in consumer expenditures as a result of price changes. With an elastic demand ($|ED| > 1$), a rise in price will generate a fall in the amount consumers spend for oil; conversely, a price fall will give increased expenditures. If the demand curve is inelastic ($|ED| < 1$), a price rise (fall) gives increased (decreased) expenditures. With unitary elasticity of demand ($|ED| = 1$), expenditures are constant as prices change. By way of example, consider a price rise along a downward sloping demand curve. If all else were equal, an x% rise in price would imply that expenditures on oil rise by x%. However, the price rise generates a reduction in the quantity purchased; if quantity fell by y%, all else being equal, then expenditures would fall by y%. The change in total expenditures clearly depends on which of these two effects is larger. (That is, which is larger, x or y?) Recall the definition of the price elasticity of demand: percentage changes in quantity along the demand curve (i.e., y) divided by percentage change in price (i.e., x). If the curve is elastic ($|ED| > 1$), then the numerator (y) must exceed the denominator (x), so that a price rise implies reduced expenditures after the consumer adjusts his purchases to the higher price.

This price elasticity/expenditure relationship helps explain some economic phenomena. Consider, for example, the effect on oil-importing regions of the OPEC-generated oil price rises of the 1970s. Recall that in the short-run the demand curve for oil is quite inelastic, as low as –0.1 or –0.2 according to some studies. Sharp increases in the price of oil, then, generate large increases in expenditures on oil, at least in the short term. This helps us understand OPEC's desire for higher crude oil prices. Or, to put the issue somewhat more generally, large price rises are not simply a result of the establishment of effective oligopoly power by sellers but also depend on conditions on the demand side of the market (i.e., the elasticity of demand).

We might also point out some macroeconomic implications of the elasticity/expenditure connection. For oil-importing regions, the oil price rises of the 1970s generated larger expenditures on imported oil because of the short-run demand inelasticity. Therefore consumers had less money to spend on other goods and services, generating a deflationary effect. In the absence of some countervailing change to increase the economy's aggregate demand (e.g., expansionary monetary or fiscal policy, or large increases in investment by domestic energy industries or substantially increased exports of goods and services to OPEC), the oil price rise was recessionary, giving lower growth and higher unemployment.

3. Market Equilibrium in Perfect Competition

Supply and demand provide a ready framework to explain market price when the selling and buying sides of the market are completely independent of one another. This condition is met under 'perfect competition,' in which a great many fully informed buyers and a great many fully informed sellers buy and sell exactly identical units of some product with no collusion whatsoever; it is assumed that new buyers and sellers can enter the market, and old buyers and sellers leave, with no impediment. Perfect competition is an idealized market form. In the real world the hypothetical perfectly competitive results may be approximated under conditions of 'effective' or 'workable' competition.

Under such competitive market conditions, the price will tend toward an equilibrium point at which demand is equal to supply, i.e., the market clears at price P_E and quantity Q_E in Figure 4.6, Panel A. A price higher than this (for instance, a price of OA) will mean that the quantity sellers are willing to supply (OC) exceeds the quantity purchasers are wishing to buy (OB); this excess supply (BC) can be expected to put downward pressure on market price until the equilibrium point is reached. For purposes of illustration, it is convenient to assume that the market always adjusts instantaneously to the equilibrium point. This avoids the difficulty of illustrating the nature of the adjustment process, including any lags and the accumulation and disposal of unwanted inventories. Of course, in the real world, this adjustment process might proceed in a variety of ways and is a matter of major concern to participants in the oil market. Trading departments in oil companies and assorted middlemen (including speculators) operate in large part in response to current disequilibria. And in a well-informed trading environment, their actions help to move the oil market to equilibrium.

The analytical advantages of the demand and supply tools become evident when attention is turned to the impact on market equilibrium of changing circumstances. The major factors underlying both demand and supply were noted earlier; changes in any of those factors will lead to changes in the equilibrium price and quantity. Figure 4.6, Panel B, for instance, shows the short- and long-run effects of an increase in the demand (from D to D') for oil at the wellhead (for example, as a result of increases in natural gas prices, rising population, etc.). In the short run, when supply responsiveness (elasticity) is low, there is a large rise

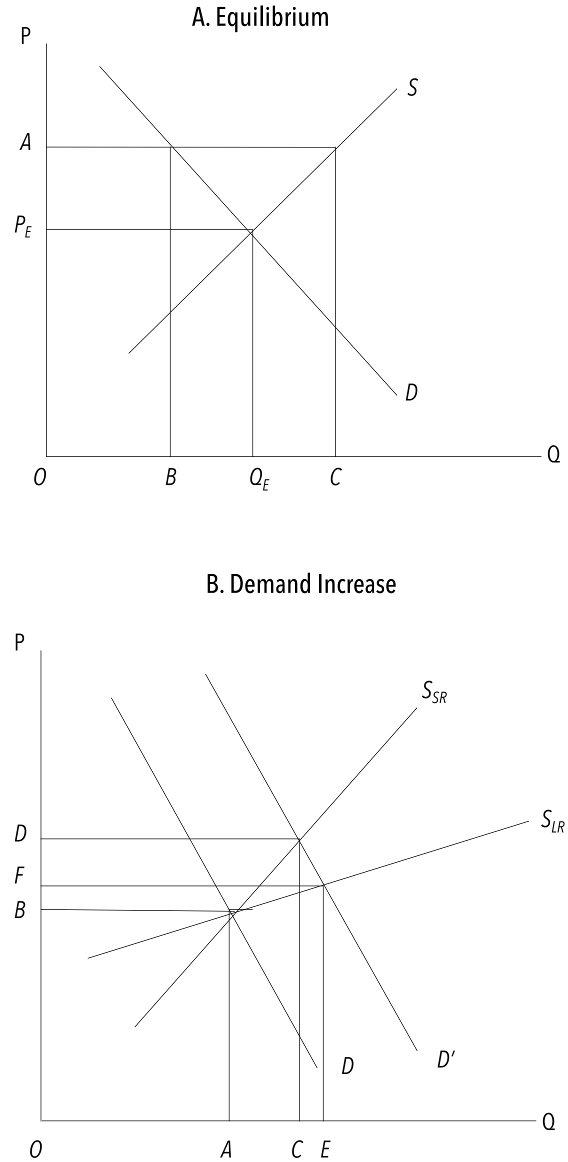


Figure 4.6 Market Equilibrium

in price from *OB* to *OD* and a relatively small rise in quantity from *OA* to *OC*. In the long run when producers are fully able to adjust to the price rise, the equilibrium will be established at a lower price (i.e., *OF*) and higher quantity (i.e., *OE*) than in the short-run.

It is easy to examine the impact of other shifts in demand and/or supply and the associated changes in equilibrium price and quantity: a number of cases will be considered later in this chapter. Remember that the more inelastic the demand curve (and/or the supply curve) is the greater the extent to which adjustments occur in the price of the product rather than the quantity. It is generally conceded that the short-run supply and demand elasticities for petroleum are low; less certainty exists about the long-run elasticities, except that they exceed short-run elasticities. It is also true that the more narrowly a market is defined, the more elastic at least one of the curves tends to become. For example, the demand for 36° oil from the Redwater pool would be very elastic because a wide variety of other crudes substitute very easily for this grade of oil.

We make two final comments. First, at any particular time only one market price and quantity will be observed. Under the simplifying assumptions made here, this is the equilibrium result, a point on both the demand curve and the supply curve. All other points on the two curves are hypothetical (i.e., unobservable in the market): they tell what buyers (or sellers) would do if the price were at some level other than the equilibrium one. The fact that most points on the demand and supply curves are hypothetical makes empirical derivation of the curves extremely difficult. Economists speak of the 'identification problem,' meaning that the plotting of price/quantity combinations as observed over time will not generally trace out either a supply curve or a demand curve. If all observations come from various intersections of the two curves, this information alone cannot be used to derive the parameters of one of the curves. More sophisticated techniques are needed that take account of the many other changing factors that affect supply and demand, and different modellers will generally estimate somewhat different supply and demand functions. Second, the terms 'demand' and 'supply' are frequently used in a much looser way than as defined here. Sometimes 'demand' is used to mean consumption and 'supply' is used to mean production (or availability), but the words then refer to observed market results, not hypothetical price-quantity schedules. For the economist bound by the conventions of current economic terminology, the equivalence of 'consumption' and the 'demand schedule' (or 'production' and the 'supply schedule') only make sense if the quantities of oil demanded (or supplied) are completely unresponsive to (perfectly inelastic with respect to) price and other variables discussed above. To say that the 'demand' for oil is higher because consumption has increased need not mean that there has been a rise in the demand schedule: it may simply be that the price of oil has decreased, generating a movement along the demand curve. The distinction between movements along a curve and shifts in a curve is important.

4. Normative Aspects

Thus far, we have dealt with several analytical tools of value in describing why a particular result occurs in the market for oil. Society is interested not only in why changes occur but in whether they should occur. This involves a change in focus from 'positive' (or descriptive) analysis to 'normative' analysis and means that the analysis of the economics of the petroleum industry is an exercise in 'petropolitics' rather than pure economics. Is the price of natural gas too low? Should Canada allow the price of crude oil to rise? Is a gasoline rationing system desirable? Should we impose a depletion tax on non-renewable energy resources? These are issues of social policy. What can economics contribute to their solution? Some economists feel that the tools of demand and supply are useful not only because they help us describe the operation of the economy but because they tell us something about the desirability or 'efficiency,' of alternative economic positions. As a general criterion for policy, economic efficiency is concerned with maximizing the benefits attainable from society's scarce resources.

That argument has been popular since an early version was propounded by Adam Smith in 1776. It relies strongly on the acceptance of what might be called 'individualistic liberalism' as a preferred social ethic: it assumes that the individual is the best judge of his or her own well-being, and that, if one individual is moved to a preferred position without harming anyone else, society is better off. Since most changes, however much they may benefit some people in society, involve moving others to a less-preferred position, it is necessary to further supplement the assumptions of individualistic liberalism. Specifically, it is often assumed that a common additive measuring stick can be applied equally to all people, that measuring stick being the monetary value individuals associate with changes, and that there is a social gain (rise in efficiency) if the sum of (dollar) benefits exceeds the sum of costs. This involves abstraction from all 'equity' considerations: how costs and benefits are distributed across different individuals. More accurately, it is usually suggested that equity considerations must be considered as well but can be considered independently of the efficiency criteria outlined so far. We shall return to this point shortly.

It might be noted that 'efficiency' can be treated as a purely descriptive concept: that is, efficiency rises if aggregate dollar benefits from a policy exceed aggregate dollar costs. However, to say that an increase in efficiency is desirable does necessitate the acceptance (explicitly or implicitly) of the value system outlined above. This is an extreme version of what Sen calls 'welfarism.' (For a discussion of this and other ethical premises for social policy see MacRae, 1979, and Sen, 1987.) That this is an extreme version is suggested by the observation that even where acceptance of this general system of beliefs has been most common, it is not usually applied to all segments of society (e.g., to children, lunatics, and criminals). It has also been suggested that the emphasis upon the individual fails to give adequate weight to humans as social animals. And some critics have questioned whether the individual's choices can be taken as representing the individual's best interests.

There are other ethical premises for valuing social policy. Some alternate value systems are 'paternalistic' in the sense that individual preferences are overridden by some other criterion of the social good. Still other value systems may give equity considerations over-riding importance: for example, a change is never acceptable if it moves poorer people in society to a less-preferred position. Other value systems may be strictly libertarian, arguing that a person should never be forced by government policies to move to a less-preferred position. Despite these alternatives, many economists feel that the concept of economic efficiency provides a workable and plausible basis for evaluating policy alternatives and feel comfortable enough with its underlying individualistic premises.

If the efficiency criterion is accepted as socially desirable, one can quite easily see why a well-functioning perfectly competitive free market economy is frequently characterized as preferable to alternate forms of economic organization. It is, after all, based upon the expression of individual desires through demand and supply and does not involve individual producers or consumers exercising control over market prices to their advantage.

The argument may be advanced in slightly more formal terms. Effectively competitive markets are assumed. Moreover, it is necessary to assume that all the real gains and losses from economic transactions are felt only by individuals actually participating in the market transactions. In formal economic terms, this means that there are no 'externalities' of either a positive sort (e.g., one firm benefiting by adopting, without appropriate charge, another firm's innovation) or a negative sort (e.g., harmful pollution, which a firm does not take into account as a cost of its production). Under these circumstances, the supply curve can be interpreted as representing costs for the whole society (i.e., social costs) as well as showing marginal costs to the individual private operator. In other words, it measures the amount society must pay inputs in order to achieve that particular unit of production. Both input costs and user costs would be accepted as valid marginal social costs. Royalties and taxes normally would not, since they are a transfer of profits from the private decision-maker to the landowner or government, rather than a cost to society. An exception is where the government devises a 'perfect' tax scheme to capture the marginal user costs and institutes a royalty equal to what these costs would be without the royalty. (With the royalty the user costs are zero, since the operator cannot capture any future profits, but the royalty is equivalent to the user costs. The games economists play!)

In the absence of externalities, the demand curve can be interpreted as a marginal social benefit curve, in that it approximates the amount that the individual is willing to pay to obtain that particular unit of output and is thus a measure of the value of that unit

Market equilibrium occurs where the marginal social benefit curve (*D* curve) intersects the marginal social cost curve (*S* curve). (See Figure 4.6A, again.) This quantity of output, with the associated market clearing price, is preferred to any other. Why? Any additional unit of output has a social cost greater than the social benefit (i.e., $S > D$, for that unit, as measured on the value or price axis), so it should not be produced. But all the previous units of output have a (marginal) social benefit greater than (marginal) social cost (i.e., $D > S$, for each unit), therefore each contributed a net benefit to society and should be produced. The conclusion? Competitive free markets yield the most desirable economic results for society: less production reduces net social benefit, as does more production. It is important to remember that this result depends on all benefits and costs being 'internalized' into the oil market.

It is useful to elaborate slightly on this idea of economic efficiency in effectively competitive markets, and on the concepts of 'producers' surplus' and 'consumers' surplus.' Recall that producers' surplus (economic rent) is the excess of market revenue above aggregate costs. In an efficiently competitive market (Figure 4.6A), it is the area between the price line and the supply curve. Consumers' surplus is the difference between the (maximum) amount a consumer would be willing to pay for a unit of output and the amount actually paid. In the market equilibrium in Figure 4.6A, it is the area between the demand curve and the price line. The producers' and consumers' surpluses represent net gains to market participants – profits (the excess of price above marginal cost) for producers, and net consumption gains (the excess of the value of consumption above price) for consumers. It is easy to see in Figure 4.6A that the price and quantity where supply intersects demand maximizes the sum of producers' and consumers' surpluses.

We would suggest that there are several normative interpretations of the result. Suppose, for instance, one were to look at the possible output levels in the market for a product, as illustrated in Figure 4.6A, and to ask what is the most desirable output level for society. Even without a clearly expressed general normative criterion, a plausible choice is the output level that maximizes the sum of producers' and consumers' surpluses. The fact that this is exactly the response that economic efficiency suggests can be taken as offering support to efficiency as a useful normative objective. A second normative interpretation is that which initially generated this discussion. If efficiency is the normative goal, then free and effectively competitive markets are socially desirable precisely because they do tend to generate maximum efficiency.

There are, of course, complications, including the concepts of 'second-best' and externalities. The failure of the economy as a whole to function in a perfectly competitive manner poses special problems in determining the value of inputs drawn from other sectors and the value of output in this sector. Costs and prices may no longer measure marginal social values. This is what economists label as the 'Problem of the Second-Best.' This book abstracts from these problems, although it has been argued that they dominate economic activity (Blackorby, 1990; Blackorby and Donaldson, 1990).

Externalities in the petroleum industry can, in theory, be incorporated within the normative theory: quantities can be evaluated at hypothetical costs and prices ('shadow prices') that include the dollar value of the externality. For example, we could find the maximum payment an individual would be willing to make in order to avoid industry-generated pollution and include this as an additional cost to society of the industry's activities. In this case the private industry would not be expected to produce at the efficient level unless forced to recognize ('internalize') the pollution cost.

Most market adjustments for a single product have relatively minor effects upon the overall distribution of income in the society. If the effects are not minor, they can usually be overcome by some general tax or subsidy arrangement. Hence, for many policy changes, an acceptance of the ethic of individualistic liberalism suggests the desirability of improvements in efficiency, regardless of the equity effects. Critics note that several major problems arise from the decision to separate efficiency and equity. The efficiency rule is based upon current prices of goods and services. To separate efficiency and equity suggests that these prices are independent of the distribution of income. If our equity judgment tells us that income differences between the average North American and the average resident of South East Asia or Ethiopia are morally indefensible, can we accept the price structure the current world income distribution gives us? The second problem relates to the weight given to future generations. They have no direct vote (can have no direct vote) in the efficiency rule: is the role given to them through our expectations about their demand for goods and services an adequate one? On this issue, see Page (1977).

Acceptance of the normative aspects of the efficiency criterion is obviously not universal. Some feel the concentration upon dollars as measured in the market is too restrictive, ignoring as it does historical processes, the non-market aspects of individual and social life, and pervasive second-best problems. Others simply do not accept an individualistic approach. Such criticisms argue that individual preferences do not equate with social welfare. However, a majority of North Americans seem inclined to an individualistic ethic; hence, in the remainder of this study, the normative model will be used on occasion. Readers who are sceptical of the appropriateness of the normative efficiency criterion might give careful consideration to the two following arguments:

(1) One of the major advantages of the efficiency objective is its broad applicability, with the associated virtue of consistency in policy recommendations. There is a tendency – especially strong if the economic system is viewed from the perspective of a specific interest group – to reject the policy conclusions of the efficiency rule, but to replace it with a series of *ad hoc* judgments. The resulting policy recommendations often turn out to contain contradictions. Two examples might be cited:

a. The efficiency argument suggests that North American domestic oil prices should equal international (OPEC) levels, since we import OPEC oil. Therefore, OPEC is a marginal supply source for Canada and the United States, and one extra barrel of oil consumption costs the OPEC price. But, it is often argued, higher prices reduce the standard of living of oil consumers. Therefore, oil prices within the country should be kept low. However, as the economic reasoning outlined earlier makes clear, lower oil prices both induce greater consumption and inhibit new supply additions, therefore increasing the dependence on high cost imported oil and generating even larger transfers of funds to oil producers in the OPEC world. Thus the viewpoint of current oil consumers is far too narrow a basis for consistent long-run policy analysis. The point is not that the effects upon oil consumers are unimportant. Rather, the point is that oil producers and taxpayers who subsidize imports from OPEC are also members of society and it is desirable to have a policy guide that incorporates the effects on them as well. The efficiency rule provides such a guide.

b. A second example relates to exports of energy products. It is frequently suggested that, since oil is a depletable natural resource, we must save all we can for future generations of Canadians. Therefore, all current exports should cease. What is less frequently suggested, but should surely follow as strongly, is that current generations of Canadians are also consuming too much of this scarce resource, at the expense of future Canadians. Limiting exports, without additional regulation, will increase availability for the domestic market and lower prices thereby tending to generate more energy consumption by current consumers in Canada. How does this meet the obligation to future generations?

It is not our intent to argue that Canada must export oil or that the efficiency argument is the only correct basis for public policy. However, what happens far too frequently is that policy recommendations are made on the basis of very limited or expeditious judgments. What is required is some broad overview of public welfare: the normative efficiency rule is one such view and therefore ranks among useful criteria for economic policy evaluation.

(2) Many critics of the efficiency criterion would suggest that the clearest proof of its failure lies in the area of energy policy, for reasons suggested in the first paragraph of this chapter: how can policy based upon the satisfaction of selfish desires of today's oil producers correctly allocate the fixed stock of petroleum that time and chance have bequeathed? Supporters of the

efficiency rule have two lines of argument in support of their position:

a. The first is to point to the historical record. For the past century or more, in the industrialized world, economic production and consumption have been based in large part on individual market decisions. This does not guarantee economic efficiency, but, as we noted above, the concept of economic efficiency does give strong support to the institution of free markets. Over this period we have witnessed sustained improvements in our physical well-being (measured both by consumption and health statistics). Moreover, most depletable natural resources seem, if anything, to have been more available (i.e., cheaper) in the post-World War II period than before. Thus, while admitting that there have been instances of monopoly power and agreeing that some externalities require correction, the historical record can be interpreted to suggest that free markets have not led to overly rapid exploitation of the resource base. Several simulations of the long-term future conditions in energy markets have reached the conclusion that energy needs can likely be met through the indefinite future, if the efficiency rule is followed. (For example, see Nordhaus, 1973, and Manne, 1976.) Of course, the argument that supplies of energy are likely sufficient for the indefinite future does not preclude the possibility that we are over consuming energy today because markets fail to recognize negative externalities such as environmental costs.

b. The second line of argument stresses the theoretical model of resource markets. In particular, it notes that efficient natural resource producers do take account of the expected future values of resources and will be quite willing to save resources for consumers in the future, if it appears profitable. The argument that the efficiency rule should be abandoned because it includes consideration only of immediate profitability is incorrect.

Once again, it has not been our intent to argue that the efficiency rule is the only feasible one for social decision-making. One can, for instance argue, that it gives insufficient weight to future generations because discount rates used by private decision-makers are too high (Lind, 1982). However, the rule is not as limited as many critics suggest, and it is reasonable to suggest that efficiency is one criterion that is useful for evaluating economic policies.

5. Market Equilibrium in Imperfectly Competitive Markets

Many markets do not have independent buyers and sellers: individuals (or small coalitions) are able to affect the market price. Supply and demand cannot, then, operate independently to establish price since one side or other of the market (or both, in some cases) is able to determine, in part, what the other side will do – in short some participants possess 'market power,' and competition is 'imperfect.'

It is important to note that market power is not conveyed to a firm simply because it is large. What is relevant is the size of the firm compared with the market and ease of entry. A large firm (measured by assets or sales) in a large industry may have no market power, whereas a small firm in a small market may have significant power. Moreover, it is not only current but potential participation that determines the extent to which markets may be effectively competitive. One approach emphasizes potential entry into (the 'contestability' of) markets (Baumol, 1986). This raises the contentious issue of the formal definition of a market. Obviously, it must be rather arbitrary, an extreme view suggesting that there are no markets, only buyers and sellers. Should we speak of the market for non-leaded premium motor gasoline in Don Mills, Ontario? Or the market for gasoline in Ontario? Or the market for oil in Western Canada? Or the market for energy in Canada? The most critical requirement of a 'market' would seem to be that it includes all individual buyers who show high willingness to move from one seller to another; similarly, all sellers must be very willing to move between buyers. A market typically covers a relatively homogeneous product and has good information flows among market participants.

Five cases of market power will be examined:

A. Monopoly

There is only one seller, so the monopolist faces the entire market demand for the product and can select where on that curve to operate. (The case of a

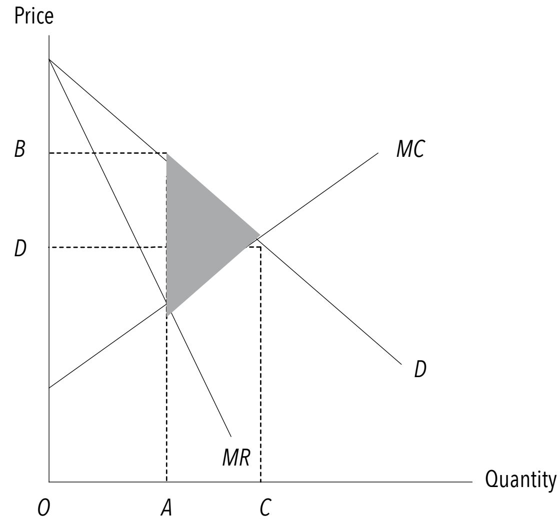


Figure 4.7 Monopoly in the Crude Oil Market

'price-discriminating monopolist,' which charges different prices for different units sold, will be discussed below.) In this case a company may profit by restricting sales in order to charge a higher price. Economists have formalized this in a simple but useful model. The firm will maximize profits if it produces output up to that unit which has a marginal cost (*MC*) equal to the marginal (incremental) revenue (*MR*) (Figure 4.7) and charges the price shown by the demand curve for that unit of output. The marginal cost of production for a monopoly is equivalent to what was called the 'supply curve' for a competitive firm. But the monopolist's marginal revenue curve lies below the market demand curve. The monopolist can sell an additional unit of the product only by cutting price and accepting a lower price on all sales. Thus to calculate marginal revenue it is necessary to deduct, from the revenue on the new unit sold, the reduction in revenue on all units that were previously sold at the higher price. For example, if 150,000 m^3/d were sold at \$35/m^3, total revenue would be \$5.25 million/d. Demand conditions might be such that 160,000 M^3/d could be sold if the price were reduced to \$34/m^3; revenue would be \$5.44 million. Marginal revenue (per additional cubic metre) would be \$190,000/10,000m^3 = \$19/m^3, which is less than the market price of \$34. The \$190,000 consists of \$340,000 on the new sales (10,000 m^3 at \$34 each) less \$150,000 (\$1 less on each of the 150,000 m^3, which used to be sold at \$35).

Contrast this with an effectively competitive market. Price is determined where market supply intersects market demand. Individual firms have no ability to influence the price of the product. If a firm attempts to charge more than the market price, no one will buy its product. It has no need to charge less than the market price because its small size means it can sell as much as it wishes at the market price. As far as the individual firm is concerned, the demand curve is perfectly elastic (i.e., flat) at the market price and this price will be the marginal revenue from each extra unit sold. In order to maximize profits, each firm in a perfectly competitive market produces where *MR* = *MC*, but this means it produces where price equals marginal cost.

In Figure 4.7 the monopolist would charge *OB* and sell *OA* units. An effectively competitive market with the same costs, including user costs, would yield price *OD* and quantity *OC*. If the normative efficiency criterion is applied, it is evident that there is a welfare loss (efficiency loss) under monopolistic conditions. The monopolist values additional units of sale (as indicated by the *MR* curve) less highly than does society as a whole (as indicated by the demand curve, *D*). The welfare loss of the monopoly is given by the shaded area: it is a dollar sum equal to the amount consumers would be willing to pay for the units of output the monopolist does not produce, less the cost the firm would incur in producing them, up to the unit of output at which the marginal benefit (demand) equals the marginal cost. Expressed slightly differently, by restricting output in order to maximize its own profits, the monopolist generates a market result that no longer maximizes the sum of producers' and consumers' surpluses.

For most changes in underlying market conditions – that is, if the demand and/or marginal cost curves shift – a monopoly market will react in the same directional manner as an effectively competitive one. Thus, the tools of competitive demand and supply are often useful for analyzing the direction of changes in equilibrium price and quantity, even though the market may not be perfectly competitive.

The discussion thus far has assumed that a monopolist charges a single price to all purchasers. Another possibility is a price-discriminating monopoly that charges different prices to different customers, or even to the same customer on different units. In this manner, the monopolist can capture more of the consumers' surplus. For such price discrimination to be possible, the purchasers in different price classes must be effectively segregated. For example, Buyer 'A' who is charged a lower price must not be able to increase purchases and resell them to Buyer 'B' to whom the monopolist is quoting a higher price. Geographical distance may provide a segregating factor, as when Buyer 'A' has to pay a shipment cost to move the product to Buyer 'B.' The presence of price discrimination is sometimes difficult to determine (Phlips, 1983). It includes, for instance, the obvious case of different prices for identical units of output. However, it also includes price differences not justified by differences in the characteristics of the output. For instance, to charge the *same* price for delivered output to customers in *different regions* would involve price discrimination because the prices do not reflect differences in the cost of delivery to the regions. It is interesting to note that a perfectly effective price discriminating monopolist would produce at the efficient output and price (just as under perfect competition, where the demand curve cuts the marginal cost curve). Each unit of output would be sold at the price indicated on the demand curve; all potential consumers' surplus, as well as producers' surplus, would be captured by the monopolist. Monopolies, let alone price-discriminating monopolies, are relatively rare. Their existence in the long-run depends upon an ability to exclude new firms from the industry, even in the face of substantial monopoly profit.

Four significant types of barriers to entry, and their possible application to the petroleum industry, will be discussed:

1. **Economies of Scale (Natural Monopoly).** Technical conditions of production may mean that the average cost of production for a firm decreases with increased output, up to the full extent of market demand. In this case, efficiency requires only one facility in the market so that production costs are at a minimum. On the other hand, efficiency requires a pricing policy other than the monopoly one. Small firms find entry difficult since costs are so high, and large firms will not wish to enter because the market is not large enough for them as well as the established monopoly. Significant economies of scale exist in the petroleum industry in pipeline transportation and natural gas distribution, though there is a size of pipeline above which no further economies can be realized. If the market were small enough, economies of scale in refining could also prove significant, although they would be circumscribed by diseconomies of distribution, depending on location in relation to markets – one giant refinery might serve all the market with the lowest unit cost of refining but would incur less than optimal distribution costs. Economies of scale may also be significant in 'frontier' areas in exploration and development. It should be noted that economies of scale serve as a barrier to entry only if they hold up to a level of output that is large relative to the size of the market demand.
2. **Government Monopoly.** Legislation or government decree may bring about monopoly conditions. The favoured company may or may not be government-owned. Often it is coupled with an enforced pricing policy that is designed to avoid the welfare costs of the monopoly.
3. **Absolute Cost Advantage.** The monopolist may have such a significant cost advantage over competitors that, even at the monopoly price, and with monopoly profits, new firms are not attracted to enter. This refers to the case in which the monopolist has lower levels of cost at all possible output levels, in contrast to the natural monopoly case where all firms have a lower average cost at higher output levels. Such cost advantages may arise, for example, from special knowledge, natural resource scarcity (one firm controls the lowest cost resource), vertical integration with a monopoly in upstream activities (the monopolist can charge high prices for essential inputs that it produces itself), monopoly in essential processes (generally supported by patent rights), access to unusually low cost capital, etc. In the petroleum industry, the major companies have been vertically integrated, often with a high degree of concentration in one or more levels of industry activity (e.g., the Standard Oil Trust in refining in the late 1800s, Middle Eastern crude oil reserves, refining in countries with small markets, pipelines). This opens up the possibility of input or output pricing practices that discourage entry at other levels of industry activity.
4. **Brand Loyalty.** On the demand side of the market, a particular firm may command such loyalty from consumers that new entrants are unable to break into the market place, even though the monopolist earns great profits. Usually this requires some legal protection (e.g., a patent or copyright) so that a competitor cannot produce an identical product. It may involve an absolute cost advantage, if a new entrant must undertake a major 'educational' expenditure to make buyers aware of its product.

These cases illustrate that the advantages of effective competition are not as clear as was suggested above. If there were significant economies of scale, or absolute cost advantages not due to monopoly restrictions of input supply or process technology, then real cost advantages accrue to society as a result of the monopoly. However, the monopolist's profit-maximizing position is not optimal for society since he will underproduce and overcharge. Usually government regulation or ownership is advocated by economists, such as the rate regulation that has been pervasive in the pipeline industry. Further, on the basis of dynamic considerations, some economists, like Joseph Schumpeter, have emphasized the importance of monopolistic markets in a technologically innovative society: the inventor of a new technology that benefits society also generates market power for itself, at least temporarily, so some degree of monopolization may be a price we pay for progress.

In general, unless entry barriers are extremely low or government regulations very effective, monopolies charge higher prices than effectively competitive industries and can generate efficiency losses.

B. Monopsony

There is one firm buying the product. The buyer may be able to price discriminate by paying less for some units purchased than for others, in effect picking up different units along the supply curve. In the absence of price discrimination, the result is analogous to the monopoly case (with a similar welfare cost), except that the reduced quantity produced is purchased at a price lower than with effective competition. The monopsonist buyer benefits by buying at the low price, but, as an upward-sloping supply curve tells us, at the lower price producers will make less available. If the monopsonist consumes the product directly, it has benefited by obtaining units at a price lower than the competitive one. If the monopsonist uses the product to produce something else, it benefits through the extra profits it gains from purchasing inputs at a low price. For the non-price-discriminating monopsonist, additional units of the product cost more than the price. This is because an offer to increase price in order to cover the increased marginal cost of the next unit of output also involves a higher payment for all units it was willing to purchase at the lower price.

A more complicated case arises if the monopsony buyer is a rate-regulated natural monopoly. For example, a natural gas pipeline company may be the sole gas buyer in a region, and the sole seller in a market area, but subject to rate regulation. In this instance, the monopsonist has a direct profit motive to restrict purchases (and sales) by paying less for the product only to the extent that he is able to disguise extra profits as allowable costs in the rate base or is allowed a rate of return in excess of the normal profit rate. A more indirect motive comes from the increase in the size of the monopsonist's system if costs are lower so demand for the monopsonist's product is higher, but this raises obvious problems in obtaining the input supplies necessary to service customers since the low price paid to the input suppliers inhibits their willingness to produce. Some analysts have suggested that a monopsony of this sort existed in the 1950s and 1960s in the export market for Alberta natural gas. Large volumes of gas had been found as a result of the search for crude oil. While the market for this gas was being built up, a monopsonistic natural gas purchaser (TransCanada Pipeline, TCPL) was able to purchase gas at exceptionally low prices and rapidly expand its market east of Alberta. However, one would expect that, eventually, a market imbalance would become apparent, where more gas was being demanded at these low prices than TCPL could contract from natural gas producers.

C. Oligopoly

As noted in Chapter Three on OPEC, an oligopoly is a market situation with more than one selling firm but few enough firms (or with one or more of the firms large enough) that companies are able to affect the market price and must consider competitors' reactions to their decisions. Oligopoly market structures are prevalent in our society: the Canadian refining industry is an example, as is OPEC. Economists have no single model of oligopoly, since what occurs depends upon how companies perceive one another, and exactly how they interact. Preservation of a profitable oligopoly with a small number of firms requires barriers to entry of the type noted above. 'Profitable' here means continuing profits above and beyond the minimum return needed to keep firms in the industry. Sometimes the word 'cartel' is used interchangeably with 'oligopoly,' but we prefer to apply the term 'cartel' to an oligopoly market structure in which the oligopolists are in direct communication with one another.

An oligopoly market might generate a result identical to the monopoly one. This would be true of a 'strict' cartel, in which the companies act as if they were a single firm – that is they rank output from the lowest cost unit to the highest cost unit, regardless of which firm produces it, and restrict output to the point where marginal cost equals marginal revenue. Investment in new facilities must also occur in the lowest cost manner. One can easily appreciate the difficulty in forming a strict cartel, particularly with regard to such matters as scheduling output shares, setting 'fair' profit shares, and determining the role of potential new entrants. Furthermore, there is a temptation for each individual producer to attempt to increase its share of industry output, especially if it feels that its share of industry profits is unfair.

In most countries, it is illegal for firms to enter into formal agreements to limit competition. However, a result close to monopoly may occur in oligopoly markets as a result of tacit (informal) collusion among firms. This is more likely if there are relatively few firms, producing virtually the same product, in an industry with high barriers to entry, and under relatively stable market conditions (e.g., steady growth in market demand). When companies all have the same information, they may collude by acting in a parallel fashion without direct communication. Frequently a pattern of price leadership evolves, in which one firm (often the largest) takes the initiative in changing price and other firms follow.

At the other extreme, an oligopoly situation may generate strong price competition among firms, with a result approaching the perfectly competitive one. This is the effective or workable competition model. If the good or service produced can be transported easily at low cost, effective competition may be attained with only a few domestic producers, given foreign competition.

Between the two extremes lie an infinite number of possible cases, some closer to the strict cartel end, some to the competition end. In comparison with the perfectly competitive case, all partake to some extent of the welfare criticism made of monopoly: the market result involves restricted production and higher prices. One type of intermediate position is thought to be very common in our society. Firms may compete in a non-price manner by means such as advertising, packaging, and changes in product quality, in order to differentiate their product. This raises costs, leading to a monopoly-type result (a higher price and lower quantity than with effective competition), but with the monopoly profits largely competed away. The absence of excess profits serves as a barrier to entry.

Generally speaking, the greater the number of firms in an industry and the smaller the largest firm relative to total industry activity, the more effectively competitive the industry is likely to be. With a very large number of firms, or a moderate number of firms of roughly equal size, it will be very difficult for one firm to lead the industry or for all firms to reach a collusive agreement.

D. Oligopsony

Oligopsony is when there are few enough, or large enough, buyers that they must take into account their interactions with one another. As with oligopoly, the market equilibrium may lie between two extremes, in this case the monopsony and perfectly competitive cases. Aspects of non-price competition occur in oligopsonies as in oligopolies; examples include prepayment agreements (where an initial payment is made prior to delivery), inclusion of take-or-pay provisions in the contract (where the buyer must pay whether or not it takes delivery), absorption by the buyer of certain gathering or distribution costs, etc.

E. Bilateral Monopoly

This is the case where a monopolist sells to a monopsonist. The actual market result depends on the relative bargaining strengths of the two parties. The equilibrium quantity may lie anywhere between the competitive quantity and the extreme, restricted, non-perfectly competitive quantity. The price may be anywhere between the low monopsony one and the high monopolist one. The efficiently competitive result may occur in a bilateral monopoly case (e.g., if the monopolist and monopsonist are 'perfectly' matched, and will not compromise), but it is unlikely.

In general, imperfect competition involves restricted market quantities relative to perfect competition and (if the normative model is accepted) some welfare loss.

Little has been made in the preceding discussion of the existence of vertical integration, which is corporate activity at successive stages of processing in an industry. Vertical integration has been prevalent in the petroleum industry. Economic policy views horizontal concentration, at a single stage of the petroleum industry, as the more critical problem. Vertical integration is compatible with degrees of market concentration all the way from monopoly to effective competition. However, as was suggested above, there may be interactions between vertical integration and horizontal concentration (often called 'economies of scope'). Vertical integration may generate significant economies (reduced costs). These are seen largely in economies in the information process, greater flexibility in scheduling operations, and absolute cost economies or economies of scale due to large size. However, the existence of vertical integration does not necessarily imply that such economies are important. For example, in the international petroleum industry, a major inducement to vertical integration historically was to avoid the market power exercised by existing firms at the refining or crude oil production level.

6. Applications to the Petroleum Industry

The purpose of this section is to apply the tools of economic analysis to several public policy issues in the petroleum industry. For ease of presentation, we shall assume that the petroleum market is effectively competitive. Empirical analysis of several of these issues is included in later chapters.

A. Interrelations among Markets

The petroleum industry is international; no part of the world oil market is immune from the impact of changes elsewhere. However, any relatively small producing region or consuming area may, by itself, have very little impact upon the basic price level of oil. As discussed in Chapter Three, to a great extent this has been true for Canada. Field prices of oil and natural gas have always been very heavily influenced by prices in the United States, which both produces and consumes much more oil than Alberta or Western Canada. And both Eastern Canada and the United States have been closely connected to overseas oil producers.

From the viewpoint of a small producing region, the price in the larger market ('the international price') might be seen as a price ceiling. Analytically, this can be handled in one of two ways: either (1) the demand curve may be viewed as perfectly elastic (flat) at the left since if producers in this region tried to charge a price higher than the international price, sales would fall to zero, or (2) the supply curve might be viewed as perfectly elastic (flat) to the right, since at the international price large quantities become available from other sources. There is no real reason to prefer one of

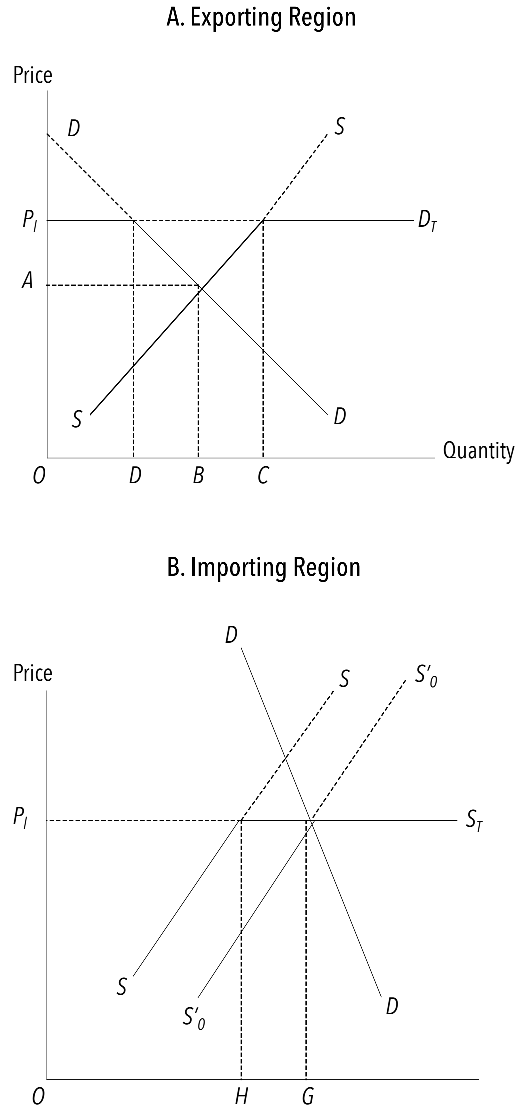


Figure 4.8 The Small Region in International Trade

these approaches to the other, so both are shown, but for somewhat different market positions, in Figure 4.8. In this figure, the price in the larger market is at level P_I, *SS* is the supply curve for the producing region, and DD is the demand curve in the domestic consuming market to which this region's oil moves. The approach is simplified by the exclusion of transportation costs within the domestic market.

In Panel A, the effective domestic demand curve, as viewed by the producers, is the kinked curve P_ID: even though consumers would (hypothetically) be willing to pay more than OP_1 for some oil, the producing region cannot charge more than the international price. This illustrates the situation in Central and Western Canadian markets since about 1950. It is assumed that regional supply is large enough relative to demand that there is no need for purchases from the larger world market. If producers are restricted from selling in the larger market they would produce quantity *OB* and obtain a price of *OA*. However, if producers are free to move into the larger market, they will view the total demand curve for their oil as curve $P_I D_T$ and produce *OC* units for sale at the international market price P_I. Of this quantity, *OD* will be sold in the domestic market and *DC* will be exported. Given the international price, increases in supply, domestic demand unchanged, would mean increases in exports. Increases in domestic demand, supply unchanged, would mean reduced exports.

In Panel B of Figure 4.8, domestic demand is large enough that it cannot be satisfied by regional supply at a price below the international level, as has sometimes been true of Canada as a whole. In this case a total supply curve can be drawn as viewed by the consuming portion of the market; this would be identical to the domestic supply (*SS*) at prices lower than P_I but show unlimited quantities available (i.e., from the world market) at that price (curve SS_T). Given the local demand (*DD*), price would be P_I and the quantity demanded would be *OG*, of which *OH* comes from domestic producers and *HG* is imported. An increase in demand (regional supply unchanged) would mean more imports but unchanged domestic production. A rise in domestic supply (demand unchanged) would mean reduced imports, but no reduction in price until domestic supply is greater than that shown by curve $S'_O S'_O$. (A supply greater than $S'_O S'_O$ yields the situation in Panel A.)

The history of Canadian oil pricing and production is discussed in detail in Chapter Six. Chapter Nine looks at government regulations that interfered with the free movement of production, consumption, or price.

B. Royalties

The term 'royalty' is generally applied to a payment made by the producer of petroleum to the resource owner on the basis of the amount of petroleum lifted. In Canada, it happens that most petroleum rights are owned by governments, so the royalty has been viewed as a tax by some observers. Its purpose is to allow the landowner (government) to share in the profits (economic rent) from petroleum production. It may also represent compensation for any reduction in the value of land transferred from an alternate use into oil production, though the resource owner may assess a rental payment for this purpose.

Royalties are viewed as a cost of production by the petroleum operator and therefore enter into the supply (private marginal cost) curve. It is evident that an increase in royalties will generate a reduced supply (leftward shift in the supply curve). Some aspects of royalties are shown in Figure 4.9. Panel A shows a 'specific' royalty, set at a fixed dollar value per cubic metre regardless of the market value of the oil (e.g., $\$3/m^3$). *SS* is the supply curve without royalty and *DD* the demand curve, with price *OA* and quantity *OC*. With the royalty, the new supply curve is *S'S'*, the royalty being the fixed vertical distance *SS'* per cubic metre. The new equilibrium price is at price *OD* and quantity *OG*: price is higher (but not by the full amount of the royalty, since $AD < EF$), and quantity supplied is lower. The reduced quantity reflects reduced capital investment in exploration and development and earlier abandonment of existing equipment due to higher operating costs for the company. The government collects revenue equal to *SS'* × *OG* = Area *SS'EF*. In the normative model, unless the royalty is set to internalize an external cost, there is a net welfare loss to society equal to area *EFB* (the shaded area), for the *GC* units now not produced. This equals the excess of their marginal value (shown by the demand curve) over their marginal costs (shown by the supply curve, *SS*). Formally, the royalty involves a loss of consumers' surplus and of producers' surplus (economic rent). The consumers' surplus loss is *EHB* for the unproduced units; it is the excess of the value of the marginal units to consumers (as given by the demand curve) over the price the consumer formerly paid (i.e., *OA*). The lost producers' surplus is *FHB*: it is the excess of the price formerly paid for the marginal units (i.e., *OA*) over the cost of production (as given by the supply curve). Taxes almost invariably involve some efficiency loss. One objective of taxation is to raise revenue with a minimum of such losses.

Owners of subsurface rights usually assess an '*ad valorem*' royalty instead of a specific royalty. An *ad valorem* royalty is some percentage of the market price of the product. Panel B of Figure 4.9 compares specific and *ad valorem* royalties. The initial demand is *DD*. If *SS* is the supply curve including a specific royalty, the market would clear with output *OB* at price *OA*. Consider now an *ad valorem* royalty that would generate exactly the same per unit revenue as the specific royalty when the price is *OA*. The supply curve with an *ad valorem* royalty would be *S'S'*, with greater

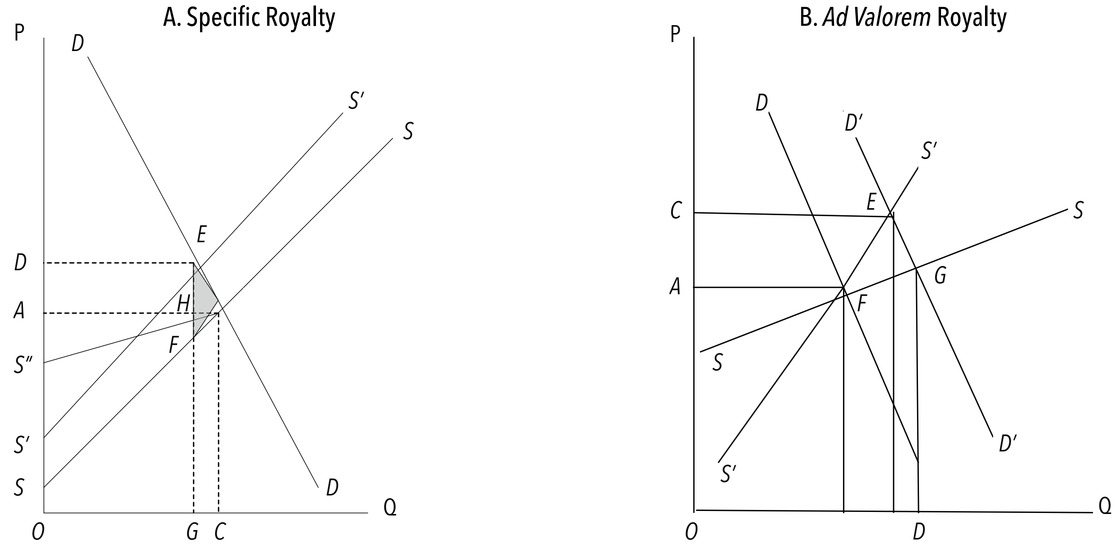


Figure 4.9 Crude Oil Royalties

supply than the specific royalty at lower prices and less supply at higher prices. If market demand were to increase, the new equilibrium would be at point *E* with an *ad valorem* royalty as compared to point *G* (a lower price) for a specific royalty. In the normative model, area *EFG* is the extra welfare cost of an *ad valorem* royalty as compared to a specific royalty, but the *ad valorem* royalty generates more revenue for the government or landowner than the specific royalty on the incremental units produced. (If the demand decreased, however, and price fell, there would be an increase in supply under an *ad valorem* royalty relative to a specific royalty, and a welfare gain.) Why are *ad valorem* royalties common? They have the advantage of allowing both the operator and the mineral rights owner to share in changing market conditions, since royalty payments vary directly with market prices.

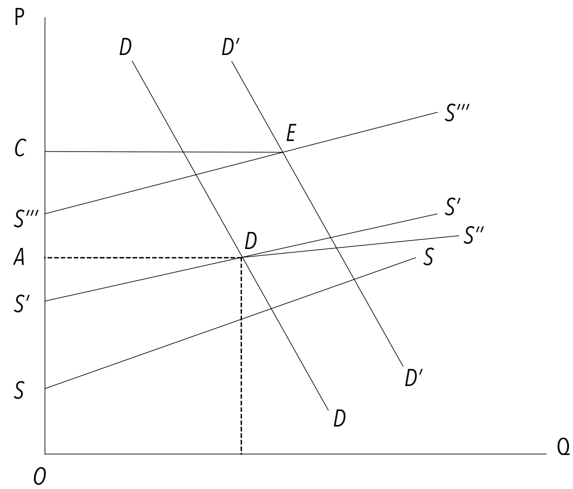
Some attempts have been made to generate more revenue from the royalty, while lessening the welfare costs, by the use of 'sliding-scale' royalties that attempt to assess a higher rate on the least costly (i.e., most profitable) production. Look at curve *S"BS* in Figure 4.9, Panel A. This involves a high royalty on the lowest cost units of output with royalties finally falling to zero on the OC^{th}, and all more costly, units. So long as the price is *OA*, curve *S"BS* is in effect a net royalty, as opposed to a gross royalty; that is, the royalty is based, not on the price alone, but the price less marginal cost. Unless the royalty is actually set up as a net profit tax, however, such an ideal scheme can only be approximated since (i) it would be too costly administratively to set a separate rate for each cubic metre of output, and (ii) the rates would have to change each time market price or costs changed, and it would be impossible to have an automatic formula that did so perfectly.

Frequently a sliding-scale royalty is based upon the well production rate, with lower royalties for wells with lower output rates, on the assumption that there is a significant negative correlation between costs per cubic metre and output per well. The assumption has some justification, since many costs are specific to the existence of a well, and invariant with respect to the production rate of the well, so that higher output means lower average costs. However, the correlation is not perfect, and some high output wells have high costs. This may hold for some, but not all, EOR schemes, and for very deep wells and those in hostile environments. As a result, a sliding-scale royalty will, like the royalties discussed above, generate reduced rates of oil production in a region, due to reduced investment and earlier abandonment dates for wells.

Sliding-scale royalties have also been tied to the price of petroleum, with higher rates the higher the price. Royalties have also been related to 'vintage,' generally the date of discovery of reserves, with higher royalties assessed on 'older' oil. This has been particularly popular if companies were willing to establish reserves in the past at lower prices, and subsequently prices rise substantially.

Panel C of Figure 4.9 illustrates some aspects of sliding-scale royalties. Curve *SS* is the industry supply

C. Sliding-scale Royalty



curve before royalties are assessed. An output-based sliding-scale royalty (assuming a negative correlation between cost and output levels) would shift the supply curve to S'S', with price *OA* and quantity *OB*. A vintage dimension, which reduced royalties on 'new' oil, would give a supply curve such as *S'DS"*, since some of the incremental long-run production would now be assessed a lower royalty. (The precise definition of 'new oil' is of vital concern to companies: newly discovered reserves normally qualify and frequently so do reserves from new EOR schemes; reserves added through extension drilling are often more problematic, and higher output from existing reserves due to accelerated depletion [infill drilling] is usually labelled 'old oil.')

The impact of a royalty scale that slides with price is harder to depict since the royalty payment (and hence the supply curve) will depend upon the equilibrium price. Thus, for instance, the supply curve might be *S'S'* if the price were *OA*, but *S'"S'"* if demand rose to *D'D'* with the equilibrium price rising to *OC*. (The *ad valorem* supply curve of Panel B [curve *S'S'*] showed marginal costs including the royalty on the assumption that the last unit produced was the equilibrium unit at which supply equalled demand. In effect, it traced out equilibrium points, like *D* and *E*, along varying supply curves like *S'S'* and *S'"S'"* of Panel C.)

In general, royalties can be used to raise revenues for the mineral rights owner and/or government, but they also change the market equilibrium (price and/or output) and generate welfare losses. More elaborate royalty schemes may minimize these changes and losses but cannot do so entirely and run the risk of becoming very complex to administer. For this reason, economists have not tended to favour royalties as the sole method of collecting economic rent from the petroleum industry.

Alberta's royalty and tax regime is discussed in Chapter Eleven of this book.

C. Production Controls

It has been common for government to impose limitations upon levels and methods of production from petroleum pools. Some of these regulations reflect safety and general conservation principles (like requiring that gas be reinjected rather than flared). In North America, the most important controls were largely dictated by the rule of capture. Consequently a number of governments (including Alberta) introduced prorationing (production control) restrictions, along with well-spacing regulations, with the avowed goals of (i) reducing the 'waste' in production associated with the rule of capture and (ii) protecting correlative property rights (i.e., the rights of access to the pool by adjacent property owners).

In Figure 4.10, Panel A, *DD* is the market demand curve in the region, and *SS* represents what the market supply curve would be if the rule of capture did not operate. This would be the case if the oil pools were 'unitized' – each pool produced by one operator only, although that operator might represent several companies. Equilibrium price would be *OA* and quantity *OB*. If the rule of capture (with shared oil pools) were suddenly to come into effect, the supply curve would shift to the right, to *S'S'*, with a corresponding fall in market price to *OC* and rise in quantity to *OD*. If supply and demand were inelastic, price could be much lower under the rule of capture. The increased supply (reduced marginal costs) reflects the tendency to ignore user costs under the rule of capture. Davidson (1963) argues that the rule of capture generates a negative user cost, which offsets the usual positive user costs. See also Watkins (1970) and McDonald (1971). Why would a company pay attention to the possible future profits that a cubic metre of reserves might generate if its competitor were likely to capture those reserves?

In theory, a well administered production control scheme could be imposed to limit production to quantity *OB*. In effect the supply curve would look

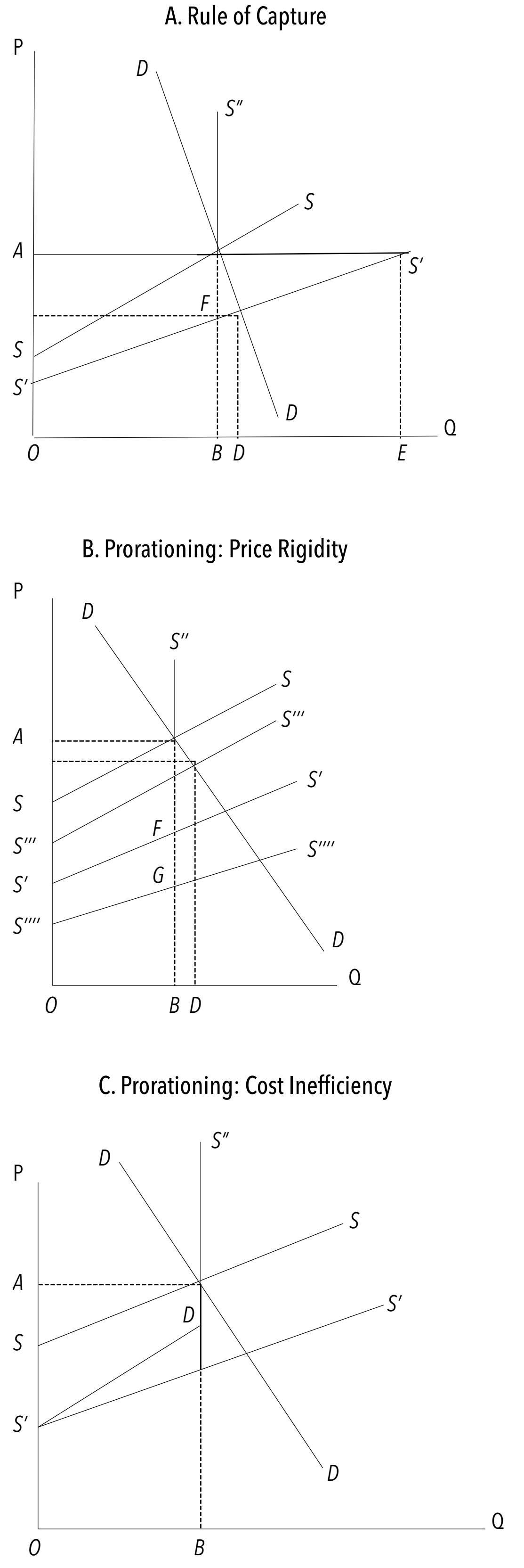


Figure 4.10 The Rule of Capture and Prorationing

like curve *S’FS”* (perfectly inelastic at quantity *OB*) and market equilibrium would correspond to that expected if the rule of capture did not operate. In practice, such a scheme is administratively infeasible since: (1) it would require restricting production to the lowest cost units of oil, even though at the price *OA* operators would be willing to supply substantially more (i.e., desired supply, without production control, is *OE* at price (*OA*)); and (2) with every change in demand (*DD*) or supply (*SS*), the regulations would have to change so that the only barrels produced would be those now corresponding to the hypothetical equilibrium without the rule of capture. It is unlikely that the administrators of the program would know exactly where the ‘true’ supply curve (*SS*) lies and which units of potential output have the lowest cost. Hence production controls are likely to give a market price higher or lower than the ‘desired’ price, *OA*.

Some critics of North American market-demand prorationing schemes accused the government regulators of administering the schemes in such a way that their prime effect was to fix petroleum prices at artificially high levels, thereby generating high consumer costs and higher petroleum profits than would otherwise have existed.

Figure 4.10, Panel B, illustrates this contention. Initially, assume the existence of a well-functioning prorationing scheme, such that production, *OB*, and price, *OA*, under the scheme (at the intersection of the demand, *DD*, and prorationed supply, *S’FS”*) correspond to the competitive equilibrium levels without the rule of capture (where *DD* and *SS* intersect). Suppose several major new discoveries are made, so that the basic market supply (excluding rule of capture considerations) shifts to *S’”S’”* (and the supply curve with the impact of the rule of capture moves to *S””S””*). The equilibrium price without the rule of capture would fall to *OC* and the quantity rise to *OD*. But regulatory authorities may continue to hold production at level *OB* (and price at *OA*); the prorationing equilibrium is where *DD* intersects *S””GS”*. The new supply addition is not allowed to affect the market, and there is a rise in ‘excess capacity’ (i.e., the amount producers would like to bring to the market at the existing price, but are not allowed to). In the normative model, there is a welfare loss equal to the excess of social benefit over social cost on the barrels that the regulatory authorities do not allow to be produced.

It is difficult to assess the extent to which prorationing schemes in North America generated welfare losses of this sort. The price of oil was relatively stable under these schemes from 1950 through 1970,

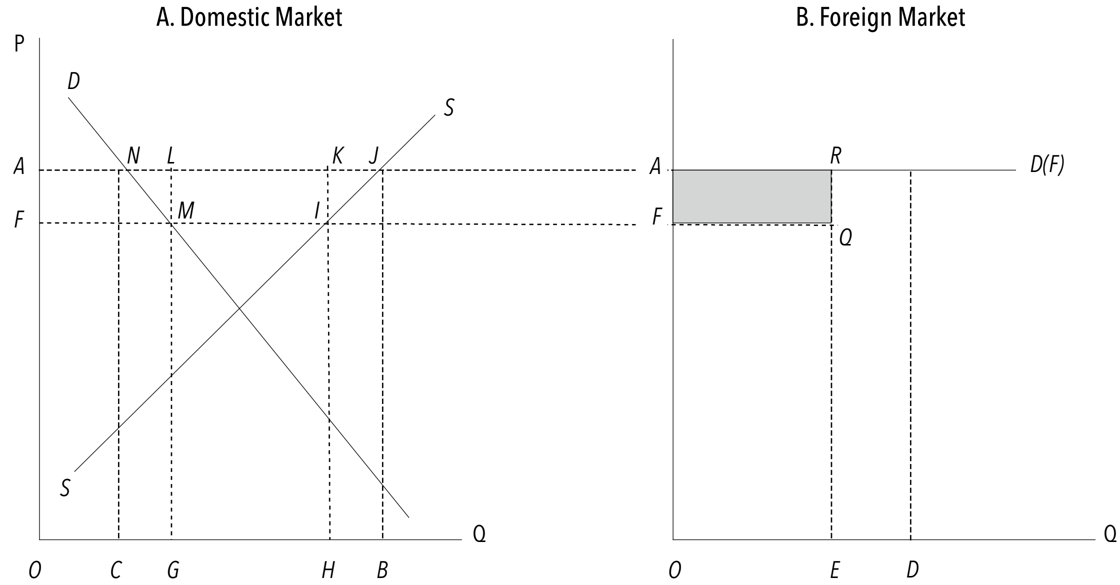


Figure 4.11 Oil Export Controls

even in the face of changing supply and demand conditions, while the amounts of excess capacity held off the market fluctuated more than price. It will be appreciated that under a prorationing scheme an individual producer has little incentive to cut price since it will gain little of any increase in quantity demanded. The increased production to meet a rise in quantity demanded will be divided amongst among all producers.

Prorationing schemes have also been criticized for inefficiencies due to their particular administrative structure. Typically (although not invariably) (i) some production from high-cost wells has displaced output from low-cost wells; and (ii) the schemes have induced the drilling of more development wells than is necessary to support allowable production. This situation can also be depicted by demand and supply curves, as in Figure 4.10, Panel C. In this figure, the curves *DD*, *SS*, and *S'S'* are the same as in the earlier diagrams: i.e., they represent, respectively, market demand, market supply with no rule of capture effects, and market supply with the rule of capture but without regulations. We assume a prorationing scheme is introduced that attains the 'desired' equilibrium price and quantity (i.e., *OA* and *OB*). The two administrative inefficiencies would lead to a supply curve under prorationing like curve *S'DS"*, instead of the *S'CS"* associated with perfect regulation. The area *S'CD* is the welfare cost associated with this imperfect production control scheme; it represents the dollar cost of oil production in excess of the lowest cost production pattern available. Empirical research has suggested that such extra costs have been substantial (Adelman, 1964; Watkins, 1971, 1977c).

Chapter Ten looks in detail at Alberta's conservation regulations.

D. Export Controls

The movement of oil between two separate jurisdictions increases possibilities for government intervention in the functioning of the market. For instance, at various times, both the Canadian and the U.S. governments have introduced measures to control the volume of crude oil, refined products, and natural gas flowing from Canada to the United States. Some useful conclusions about the economic impact of such measures can (of course!) be gained from the supply-demand apparatus.

Figure 4.11, Panel A shows a domestic producing region in which demand (*DD*) is entirely satisfied by domestic sources (domestic supply is *SS*). A much larger export market is available so that external demand is virtually unlimited as far as domestic producers are concerned, at a market price of *OA*. In the absence of trade controls, the equilibrium price would be *OA* and the equilibrium quantity of production *OB*; of this, *OC* is consumed domestically and exports are *CB* (equal to *OD* in Figure 4.11, Panel B, showing the export market). Suppose that the domestic government limits the volume of exports to *OE* (< *OD* = *CB*). (Or we could suppose that the foreign government limits the volume of imports.) The reduction in the demand for domestic oil will mean a fall in the domestic price to level *OF*, at which the quantity supplied (*OH*) is equal to the quantity demanded (*OH*, of which *OG* is domestic demand and *GH* = *OE* is a foreign demand). Domestic consumers benefit by such a policy on the part of the government since the oil price is reduced, but domestic producers suffer. There is also a benefit to those in the foreign market who are fortunate enough to obtain oil from the domestic region: the market price in the foreign market is *OA*, but foreigners can purchase these barrels at price OF, thereby gaining to the extent of the shaded area. In the domestic market, efficiency (welfare) losses equal to *NMIJ* occur: losses in revenue on exports (*LMIK* = *AFQR*), losses of producers' surplus on the reduced output (*IJK*) and the payments foreigners used to make in excess of the value of the extra domestic consumption (*NLM*). The program, therefore, involves income transfers from domestic producers to both domestic and foreign consumers.

If only select domestic producers are allowed access to export markets, those producers may be able to gain extra revenue by selling oil in the foreign market at price *OA*. This is an example of price discrimination, where foreign and domestic customers are charged different prices. The difference can exist only because the foreign customers are not allowed open access to competing Canadian suppliers.

Another possibility, an export tax, will be discussed below.

It must be noted that analysis is further complicated if the producing region operates a prorationing scheme. In this case, the imposition of export controls may simply mean reduced production without corresponding price changes.

E. Price Ceilings

In some instances, the imposition of a price ceiling at a level lower than the market clearing price is identical in impact to export control, as analyzed above. One might interpret Figure 4.11, for instance, as showing the impact of a ceiling price in the domestic market at level *OF*. The ceiling price implies an excess demand in the market. Domestic consumers are assumed to buy all they wish, but foreign consumers cannot obtain all they would like. In Figure 4.11, the excess demand does not pose severe adjustment problems since foreigners can obtain what they want from other sources at a price of *OA*.

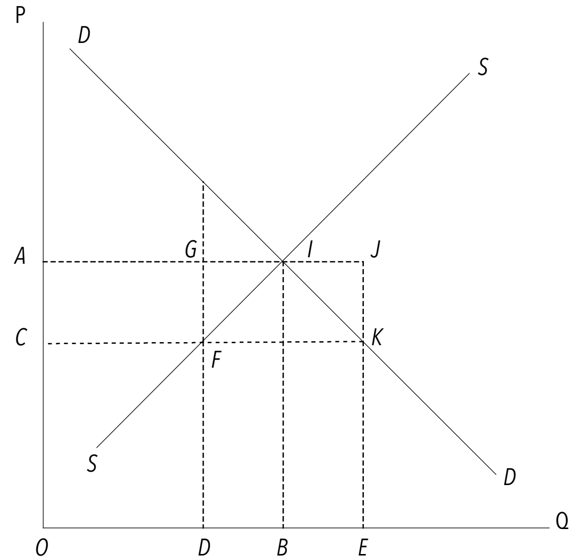


Figure 4.12 Oil Price Control

Once domestic demand is high enough to prevent exports, however, the price ceiling implies an excess of quantity demanded over quantity supplied within the domestic market. Some type of rationing scheme, deliberate or *ad hoc*, is necessary to allocate the shortage amongst users. Furthermore, a price ceiling, by definition, prevents the types of responses that are most natural in a market economy. For example, in response to impending shortages, there would be no price-induced restrictions in use and increases in investment and production.

Figure 4.12 illustrates the economic features of a price-control scheme that holds the quantity supplied domestically below the quantity demanded. In the absence of price controls, the equilibrium price would be *OA* and the equilibrium quantity would be *OB*. The imposition of a price ceiling at level *OC* (< *OA*) means a reduction in the quantity supplied (by *BD*) and an increase in the quantity demanded (by *BE*): a shortage (equal to *DE*) appears in the market. This shortage might be met by rationing the oil amongst buyers. If the scheme is perfectly efficient, in the sense that the petroleum goes to those areas that are willing to pay most for it, the net welfare loss to society is given by the area *FHI*. This is the excess of social benefit over social cost for those units of output that are now not produced. In addition, the program involves a transfer of purchasing power, equal to area *CFGA*, from producers to consumers. Producers receive less revenue on each of the *OD* units sold than they would without the price ceiling, while consumers receive the same benefit as before from consuming even though the units cost them less. In other words, an amount equal to *CFGA* is transferred from producers' surplus to consumers' surplus.

Alternatively, the domestic shortage might be met by imports at the international price (say *OA* in Figure 4.12), with the government subsidizing the imports to the amount of *CA* per unit consumed. In this case, in the normative model, there is a net social loss of *FGI* + *IJK* relative to an absence of price controls. This loss consists of: (1) the excess (on the extra oil consumption) of the cost of imports over the value to consumers of that oil (i.e., area *IJK*) and (2) the excess of the cost of imports over the production cost of domestic oil on the extra units domestic producers would willingly bring to market at the higher price (i.e., area *FGI*).

F. Export Tax

The previous two sections showed that the imposition of export controls and the imposition of a price ceiling could lead to equivalent results, when crude oil is sold both at home and in a foreign market. A third alternative could also bring this result: the imposition of an export tax. In Figure 4.11, suppose that the government has imposed a tax per barrel on exports equal to *AF*. Given the foreign price (*OA*), the netback to domestic producers will be *OF* (i.e., *OA* – *AF*), at which price they see virtually unlimited demand in the foreign market. So long as the quantity supplied at this price exceeds the quantity demanded domestically, this will be the price for the product. In contrast to the other two programs, however, the shaded area in Figure 4.11 represents the export tax payment to the domestic government, not a subsidy to foreign users.

In a sense, then, for a country that exports oil, any one of the three programs outlined (export controls, a price ceiling, or an export tax) could be used to achieve the same effect upon domestic prices and quantities. In a dynamic sense, they differ, however: that is, changes in demand and supply lead to different results under the schemes. For instance, if the foreign price rises, so does the domestic price under any given export tax, whereas with a price ceiling or an export quota there is no change in the domestic market. The dynamic differences, and the differing distributional effects, help explain why the Canadian federal government imposed all three types of schemes in 1973. At the same time, it must not be forgotten that Canada is a federal state, and that the government in Ottawa had to consider the ability of unhappy provincial governments to frustrate its policy goals. For instance, had the federal government imposed export controls alone and had Alberta prorationing authorities reduced the market allowable by an equivalent amount, there would have been no reduced price effect in Canada.

Chapter Nine will provide more detail on Canadian oil policies, and Chapter Twelve on natural gas, in the tempestuous 1970s and 1980s.

7. Conclusion

This chapter has been designed to demonstrate the usefulness of economists' analytical tools in describing the operation of petroleum markets. The tools are abstractions that aid in the understanding of real-world movements of market prices and quantities. They are also useful in describing the impact of changing market conditions or alternative regulatory policies upon different market participants: oil producers, domestic consumers, foreign consumers, and governments.

It has also been argued that the tools are useful in assessing the desirability of various alternative policies. Policy analysis should involve four stages:

(1) a clear statement of policy objectives;
(2) determination of alternative ways in which the policy objective might be achieved;
(3) assessment of the likely effectiveness of these alternative policy tools, and selection of that thought best; and
(4) assessment of the actual effectiveness of the policy chosen, after it has been applied.

This is a complicated process. For example, some of the policy objectives may conflict with others, and trade-offs must be considered.

The concepts set out in this chapter are of use in policy analysis in part for their descriptive value. However, beyond this, they have particular applicability to the assessment of one possible social policy goal – economic efficiency. This refers to the maximization of the difference between the dollar value of benefits received by members of the society and the dollar value of the costs of obtaining those benefits. Under certain conditions, the attainment of this maximum net benefit can be equated with effectively competitive market-clearing prices and quantities. Unless there is a strong reason for supposing otherwise, many economists judge this outcome to be a desirable one for society. What reasons might be raised for not preferring such a position?

(1) There are cases in which demand and supply forces do not accurately reflect benefits and costs to society. For example, there may be pollution costs not reflected in the supply curve. Or, profits on production may accrue to persons from outside the region. Or, monopoly power may exist. In this case, governmental action may be required in order that efficiency is achieved.

(2) The maximum efficiency outcome may conflict with the attainment of some other social goal. For instance, it may involve extreme hardship on the poorest members of society, may be felt to transfer undue political influence to the capitalist sector, may be felt to treat vested property rights unfairly, etc. In this case, it may be judged desirable to accept some inefficiency in order that another goal be more nearly satisfied.

Normative policy analysis is a critical issue in any society. At the same time, the establishment of goals is a difficult task: no universal agreement on policy goals is to be expected, or, many would argue, desired. This book will concentrate upon the goal of economic efficiency but will note the implications of various policies for the attainment of other possible social goals.