5-2012

Three essays on oil scarcity, global warming and energy prices

Matthew Riddle

University of Massachusetts Amherst, mriddle@econs.umass.edu

THREE ESSAYS ON OIL SCARCITY, GLOBAL WARMING AND
ENERGY PRICES

A Dissertation Presented
by

MATTHEW RIDDLE

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2012

Department of Economics

THREE ESSAYS ON OIL SCARCITY, GLOBAL WARMING AND ENERGY
PRICES

A Dissertation presented

by

MATTHEW RIDDLE

Approved as to style and content by:

_____

James K. Boyce, Chair

_____

Michael Ash, Member

_____

Erin Baker, Member

_____

Michael Ash, Department Chair
Economics

ACKNOWLEDGEMENTS

I would like to thank my dissertation chair, James Boyce for his guidance and support throughout my time at UMass, and my committee members Michael Ash and Erin Baker for their thoughtful suggestions. I would also like to thank my wife, Allison, for her ideas and support.

I also would like to thank the E3 Network and the Political Economy Research Institute for providing fellowships that helped support my dissertation research.

ABSTRACT

THREE ESSAYS ON OIL SCARCITY, GLOBAL WARMING AND ENERGY PRICES

MAY 2012

MATTHEW RIDDLE, B.A. CARLETON COLLEGE
Ph.D UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor James K. Boyce

This dissertation is composed of three essays. In the first essay, I construct a supply and demand model for crude oil markets. I then fit the model to historical price and quantity data to be able to project future oil prices. Ex-post forecasts using this model predict historical price trends more accurately than most oil forecasting models. The second essay incorporates the supply and demand model from the previous paper into a complex systems model that also includes oil futures markets. Adaptive-agent investors in futures markets choose from a set of rules for predicting future prices that includes the rational expectations equilibrium rule, as well as rules that rely on more short-term information. The set of available rules evolves following a genetic algorithm; agents choose which rules to follow based on their past performance. While outcomes vary depending on the specific assumptions made, under a plausible set of assumptions investors can fail to anticipate shortages properly, leading to significant price spikes that would not occur in the rational expectations equilibrium. The last essay addresses the impacts of carbon cap-and-trade policies on consumers. I calculate how higher carbon prices would affect the prices of different consumer goods, how consumers would respond to the price changes, and how the price changes, along with revenue recycling, would impact consumers of different income levels.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Energy in historical context

The ability to harness energy sources and put them toward productive use has played a crucial role in economic development worldwide. The industrial revolution in Europe was driven in part by the use of coal to power steam engines for rail and water transport, to facilitate iron smelting, and to power looms and other industrial equipment (Heinberg, 2003). Expanded use of easily accessible oil helped to fuel continued expansion in the 20th century. Agricultural production was transformed by the use of motorized farm equipment and petroleum-based fertilizers and pesticides. Cars, trucks and airplanes powered by oil products revolutionized the transportation of people and goods. Electricity, largely powered by coal and to a lesser extent oil and natural gas, contributed to the further automation of manufacturing and made possible the development of numerous electricity-powered technologies (Heinberg, 2003).

1.1.1 Problems with fossil fuel consumption

While fossil fuels have helped spur economic growth, the rapid consumption of fossil fuels has also contributed to environmental damage, and may lead to even greater costs in the future. The mining, drilling for and transportation of fossil fuels can lead to the removal of mountaintops, the contamination of ground water with methane, and oil spills that despoil oceans and beaches. The burning of fossil fuels can release chemicals that contribute to smog, acid rain and mercury contamination.

These immediate environmental costs are also accompanied by two prominent concerns about future costs. Carbon dioxide emissions from fossil fuel consumption are the main drivers of climate change, the effects of which are likely to become more and more severe as temperatures rise. The depletion of oil and other fossil resources leaves less available to future generations, and increases the likelihood of price spikes if demand outpaces supply. These concerns are difficult, if not impossible, to address without cutting fossil fuel use, since all fossil fuel combustion produces carbon dioxide, and fossil fuels, once consumed, cannot be re-generated.

1.2 Oil scarcity

Oil has several advantages over other fossil fuels: it is easily transportable and energy-dense, and when refined it is suitable for a wide variety of uses. It currently accounts for the largest share of world energy use of any source, slightly ahead of coal (BP, 2011). Changes in oil prices have been found to play an important role in predicting economic growth in the US (Hamilton, 2005a).

Considering the important role that oil plays in our economy, if persistent shortages were to emerge, the economic implications could be enormous. However, there is no consensus as to how seriously the threat of oil resource depletion should be taken. Some warn of a colossal societal collapse in the not-too-distant future, while others argue that technological progress will allow us to shift away from oil before resource depletion becomes an issue. How much of a problem oil depletion poses depends on the amount of oil that remains accessible at reasonable cost, and how quickly the development of alternatives allows the demand for oil to be reduced.

Chapter 2 provides one attempt to evaluate the level of the threat posed by oil depletion. The focus is on projecting oil prices, since these serve as a measure of scarcity as well as a mechanism by which scarcity will negatively impact society. Price projections are made by first constructing a demand-and-supply model and fitting it to price and quantity data. Demand and supply each evolve over time following a pattern that is based in historical data, while supply is also constrained by resource availability. The responsiveness of supply and demand to price changes in the short and long run is also carefully modeled, and prices are set to equilibriate supply and demand.

With this model, it is possible to project how prices and quantities would move far into the future if the same model continues to hold. The price path we project provides an indication of how high prices would have to get, and at what rate, to make demand drop fast enough to accommodate decreasing supplies driven by dwindling resource availability. We find that prices are projected to rise gradually but persistently throughout the simulation period. Concerns about drastic consequences in the near future appear to be overblown, but in the long run, as oil resources approach exhaustion, prices do reach extremely high levels.

These projections are clearly speculative, as any attempt to make forecasts far into the future is subject to concerns that the model used to fit past data may not be the best model of the potentially different conditions that may exist far into the future. It is still useful to make these projections, however, as they provide a best guess as to what might happen if no dramatic structural changes occur.

1.3 Market anticipation of future scarcity

While chapter two helps to establish that oil scarcity is a legitimate concern, chapter three addresses the ability of markets to anticipate future scarcity and make the transition as smooth as possible. If market players can anticipate that oil supplies will be scarce in the future, they should be able to make money by holding onto oil and selling it once prices rise. This should drive current prices up, which creates the right incentives for everyone to make the adjustments necessary to make the anticipated shortage as painless as possible.

For this process to work, market participants must be able to anticipate future shortages. It is not clear, however, that oil market players will look far enough into the future to account for long-run scarcity concerns when betting on short-run price changes. While there are clear benefits to accurately predicting short-run price movements, the benefits of factoring in long-run scarcity concerns into these calculations are not so clear. Chapter three aims to evaluate whether short-sighted or long-sighted rules are more likely to persist in an evolutionary setting where the most accurate prediction rules proliferate.

The model built in chapter three combines adaptive agent investors who choose from a set of long-sighted and short-sighted price prediction rules with the demand and supply model from the previous chapter. Model simulations are used to evaluate which types of prediction rules are favored by the evolutionary algorithm, and what oil price trajectories result from these prediction rules. We find that different model runs can produce widely different outcomes, but in some simulations short-sighted rules dominate leading eventually to extreme spikes in oil prices.

While we do not provide evidence as to how likely such a short-sighted outcome is, the fact that it could exist in a plausible setting with smart investors poses a significant challenge to standard economic theories of exhaustible resources. This modeling exercise goes a long way toward building a rigorous theoretical argument as to how market structures with investors responding to market incentives can lead to an outcome where markets do not properly anticipate future scarcity in spite of apparent incentives to do so.

1.4 Climate change policy

There is no mechanism for the market on its own to address concerns about climate change. However, if policies are put in place to price the costs of climate change into the price of fossil fuel consumption, then this should trigger market incentives that should lead efficiently\to the desired emission reductions. Policies that take this approach include carbon taxes and cap-and-permit programs.

One concern with this approach is that poor and middle-income families will be hurt by the higher price of fuels. Studies of the distributional impacts of these policies have shown a consistent pattern: absolute payments into the charge increase with the income of the household, but payments as a percentage of income are highest for low-income households. If the government revenues from the charge are returned to consumers on an equal per capita basis, the net effect of the policy would be a significant progressive redistribution of income, while if the revenues instead flow to polluting companies, the net effect would be regressive.

The fourth chapter of this dissertation expands on past studies of the distributional impacts of a carbon charge by including a more thorough analysis of producer and

consumer responses to a carbon charge, and how these responses could affect the distributional outcomes. The main conclusions of past work on this issue hold up to most of these changes in assumptions. The assumption that has the greatest impact on the incidence outcomes is the producers' assumed rate of pass-through of price increases. If the pass-through rate is low enough, producers could bear much of the burden of the charge, and this burden would be passed onto shareholders who are primarily in high-income groups. Other adjustments to the model provide interesting insights about the role of different assumptions in determining incidence outcomes, but none of the effects are large enough to alter the primary conclusions of past studies.

CHAPTER 2

MODELING OIL MARKETS AND FORECASTING OIL PRICES

2.1 Introduction

Oil plays a crucial role in our economy. It provides most of the energy to meet the world's transportation needs, including passenger travel and cargo transport. It provides fuel for home heating, electricity production, and to power industrial and agricultural equipment. It provides the source material for the construction of plastics, many fertilizers and pesticides, and many other industrial chemicals and materials. It is difficult to find any product that does not require the use of oil at some point in the production process.

The current rate of consumption can not be sustained forever, since oil supplies are finite. However, there is no consensus as to whether the depletion of oil resources will lead to any significant economic hardship, and if so, when. Saudi oil minister Sheikh Ahmed Zaki Yamani was famously quoted as saying, "the Stone Age didn't end for lack of stone, and the oil age will end long before the world runs out of oil." This quote reflects the view that the development of new technologies will lead to a shift away from oil consumption before oil resources are fully depleted.

On the other hand, a number of observers have warned that resource depletion, and dwindling oil resources in particular, could lead severe consequences in the not-too-distant future. There is still enough oil left in the ground for many decades of consumption at current rates, but as the easily available oil is depleted, it will become increasingly difficult to continue pumping oil at the rate at which it is currently being consumed. There has been an explosion of popular literature recently predicting that oil

production will peak soon, and that oil shortages will force us into major lifestyle changes in the near future – a good example of this is Heinberg (2003). The point at which world oil production reaches a peak and begins to decline permanently has been referred to as 'Peak Oil'. Predictions for when this will occur range from 2007 to 2025 (Hirsch, 2005).

There has been debate within the economics profession as to the dangers of resource depletion, but in recent times many economists have questioned the urgency of this threat. Tilton (2002), in a summary of the state of knowledge among economists of the threat of mineral resource depletion, summarizes that "during the next 50 to 100 years, we have found that mineral depletion is not likely to rank among the most pressing problems confronting society," (p. 119) and that "in the long run, should mineral depletion cause shortages, they are likely to emerge gradually, perhaps over decades, as the real prices and costs of mineral commodities rise slowly but persistently" (p. 76).

2.1.1 Predicting price paths

In a market environment, oil scarcity is felt through high prices. If supply is not sufficient to keep up with demand at current prices, prices will rise. This will have a range of negative effects. The most visible impact of crude oil prices is on the price of gasoline, which is closely followed by consumers and regularly reported in the news. Changes in oil prices can also influence the overall health of the worldwide economy as well as the relative strength of different national economies. Nine of the ten recessions between 1946 and 2005 were preceded by spikes in oil prices (Hamilton, 2005), and the latest recession followed the same pattern.

It is difficult to predict the path that oil prices will take. Many factors contribute to changes in oil prices, from technological developments to weather patterns to economic trends and geopolitical events. These factors can be difficult to model, and in some cases can be entirely unpredictable. It is probably futile to try to provide an accurate forecast of future price movements. Nevertheless, it is a useful exercise to evaluate what is likely to happen to oil prices if past trends continue. Given the importance of oil prices, and the vastly differing views that exist about the prospects for impending scarcity, it is valuable to have an idea as to which way price trends are most likely to go in the short run, and as to the likelihood of shortage-induced rises in prices in the long run, given the currently available information.

Economists take several approaches to addressing this question. Theoretical modeling of the depletion of exhaustible resources has provided some insights into how prices can be expected to behave as a scarce resource approaches depletion. Empirical studies have looked at how prices have evolved historically over time, and used the results to project what could happen in the future. Other structural models predict how demand and supply are likely to evolve, and use these projections to anticipate price movements. However there are some shortcomings of each of these approaches, and none has proven very effective at predicting prices.

2.1.1.1 Theoretical models

The seminal work in theoretical modeling of resource depletion was by Harold Hotelling (1931). Hotelling provided a framework for analyzing producer behavior when extracting a scarce, non-renewable resource. In this modeling framework, the owners of a resource choose the extraction rate so as to maximize the present value of the profit

9

they would receive over the life of the resource. The primary conclusion of Hotelling's

original model, in its simplest form, is that the price of the resource will rise over time at

the going interest rate. This conclusion has an intuitive explanation: if the price of a

resource went up faster than the rate of interest, producers would choose to put off

extracting and selling the resource until its price had increased, because they could make

more profit by putting off the sale. If the price were falling or rising more slowly than

the rate of interest, producers would choose to sell more of the resource early and invest

their returns at the going interest rate. The equilibrium occurs when prices are rising at

exactly the interest rate, making producers indifferent between producing now or holding

reserves for later.

The Hotelling model relies on some very restrictive assumptions. Refinements

and extensions of the model have been developed that relax many of these assumptions.

One simple refinement, which only slightly changes the conclusion of the basic model, is

to include a cost of extracting the resource. In this case, instead of the price rising at the

rate of interest, it is the difference between the price and the marginal extraction cost that

should rise at the rate of interest – a quantity that is referred to as the shadow price,

scarcity rent or *in situ* value of the resource. If the extraction cost is constant but

positive, then the price rises monotonically but at a rate lower than the rate of interest

(Krautkraemer, 1998, p. 2068). If the marginal extraction cost falls faster than the rise in

shadow price, the price of the resource could fall. Eventually, however, the shadow price

will dominate, necessarily leading to a rise in price as the resource approaches exhaustion

(Fattouh, 2007, p. 7).

The price path can be further adjusted with additional extensions of the original

model. One common refinement is to allow the cost of extraction to vary based on the

cumulative extraction of the resource (or conversely the amount of the resource left in the

ground), to reflect the fact that as the resource becomes more depleted, the resources that

remain get more costly to extract. This allows the shadow price to rise at less than the

rate of interest (Krautkraemer, 1998, p. 2069). Allowing the reserve level to increase

with new exploration can provide an explanation for decreasing extraction costs early in

the life of a resource, which causes prices to follow a U-shaped pattern over time

(Pindyck, 1978).

While these models show prices eventually turning up, it is also possible that they

never will. Perhaps the most important and difficult-to-model factor that can influence

the path of prices over time is the development of new technologies, both in reducing the

cost of extracting the resource, and in developing alternative energy sources and efficient

technologies that reduce the demand for the resource. If technological developments lead

to sufficient drops in demand, the long-run exhaustibility of the resource may never come

into play. In the words of Tilton, "the long run availability of mineral commodities

largely depends on a race between the cost-reducing effects of new technology and the

cost increasing effects of resource depletion" (Tilton, 2003, p. 63).

2.1.1.2 Statistical analyses of prices

Another approach taken by economists is to look at how oil prices have evolved

over time in the past, and use this to make suggestions about what will happen in the

future. One of the first studies to use long time series for natural resource prices to make

inferences about how resource scarcity evolves over time was by Barnet and Morse

11

(1963). They found that the prices of nonrenewable resources were generally lower in 1957 than in 1870, in spite of the depletion that occurred during that period. Later, Margaret Slade (1982) found a U-shaped relationship between mineral prices and time, with prices decreasing at first but later turning up. A more recent update by Berck and Roberts (1996) finds mixed results, depending on the regression technique used. With their favored technique, they find that there is no significant trend for resource prices over time.

The stochastic process used to model the path of energy prices plays an important role in these studies. It can affect the regression results, and therefore the prediction for future prices, as demonstrated in Berck and Roberts (1996). In addition, in projecting future price paths it is important to have projections about the types of price fluctuations that are likely to happen in addition to a best-guess price. For this purpose the type of process followed by the time series is important in its own right. Slade (1982) models prices as following a trend-stationary ARIMA (1,1,0) model. Berck and Roberts (1996) argue that using a trend stationary model is inappropriate, and use a difference stationary model instead. Pindyck (1999) suggests a more complicated model, with a stochastically moving trend-line, in addition to stochastic variation around the trend with gradual reversion towards the trend.

2.1.1.3 Modeling Supply and Demand

A final approach to projecting oil prices is to use structural models in which prices are set by the interaction of demand and supply. While there is a substantial body of work modeling demand, supply and price movements, there is surprisingly little research that uses this approach to provide forecasts of future oil prices. Groups like the

International Energy Agency (IEA) and the US Energy Information Administration (EIA) use complicated oil market models to project demand and supply into the future. These demand and supply projections are then used by analysts in the business world who require estimates of future oil prices in making financial decisions. The EIA model also produces its own projections for oil prices, but these prices are highly dependent on their assumption about future OPEC supply, which is required as an exogenous input into the model. OPEC supply numbers are determined based on "expert judgment and/or offline analysis" about anticipated OPEC "output and pricing behavior" (EIA, 2007, p. 10). In practice, OPEC production levels in earlier runs of the EIA model appear to have been chosen based on the assumption that OPEC members will produce enough to meet demand at a target price, so the price projections are based mainly on the assumption about the target prices that OPEC will try to reach (Gately, 2001). The IEA's World Energy Model does not attempt to generate price predictions, but takes prices as an exogenous variable (IEA, 2007, p. 6). It is not clear where the price projections they use come from.

Other models in the academic literature have used a similar approach, combining demand and supply models with a price adjustment rule to show how prices are set (Dees et al., 2007; Bacon, 1991). These models produce good fits with past oil price data, but are not used to produce future projections. Doing so would require projections for several variables that are used as inputs into the model, which the authors do not attempt.

2.2 The Model

The approach taken in this paper is a variation of the supply and demand approach, but incorporates elements of the other approaches as well. Demand and supply

are modeled as a function of time, current and past price levels, and the amount of oil remaining in the ground. The demand and supply equations are constructed to be consistent with results from the literature, and are fit to historical crude oil price and quantity data. Prices forecasts are produced by extending the model to future time periods, using the mean value for the stochastic term.

This approach has several advantages over other approaches that have been used in making price projections. It is a simple technique that allows important theoretical modeling considerations to be combined with historical data in a flexible framework. Theoretical models on their own have either been too restrictive in their assumptions, leading to unrealistic projections of price paths, or too broad to produce meaningful predictions. Empirical work that models how prices have evolved over time does a reasonable job of fitting past prices, but the shape of the time trends that are fitted to the data generally have little theoretical basis, and do not account for the exhaustible nature of the resource. As a result, using this approach to make long-run projections may be inappropriate. Structural models of demand and supply do a good job of using past data to show how demand and supply respond to different inputs, but few models have used this approach to forecast future prices. One reason may be that demand and supply models generally depend on several other variables that must themselves be projected into the future in order to produce demand and supply forecasts.[1] By modeling demand and supply only as a function of time, prices and cumulative production, the model in this paper limits the number of variables that need to be projected and the data requirements

---

[1] For instance, EIA (2007) requires projections for GDP and OPEC production levels as inputs into the model. Dees et al (2007) require projections for GDP, global natural gas liquids production, crude oil production by Russia and China, and OPEC capacity (p. 184).

in estimating the model. It combines the simplicity of models that focus solely on oil prices with some of the realism of structural models, while accounting for the exhaustible nature of the resource as in theoretical models.

2.2.1 Demand Function

Demand is modeled as a function of time and the price history. Other variables that are commonly included, such as income, are left out to minimize the data requirements of the model and make future projections easier. Instead, they are captured indirectly through shifts in demand over time.

2.2.1.1 Relationship with time

The log of demand is modeled as a quadratic function of time plus a price response term and a disturbance term:

$$\log(QD_t) = b_{0D} + b_{1D} \cdot t + b_{2D} \cdot t^2 + f(p_t, p_{t-1} \ldots, p_0) + e_t \tag{2-1}$$

The relationship with time (t), along with the disturbance term (e_t), captures all developments that could lead to shifts in demand (QD_t) that are not related to current and past price changes (p_t). The log-quadratic relationship captures trends in growth rates of demand over time, and the disturbance term captures random unanticipated movements. The underlying reasons for these movements in demand include changes in world income and how it is distributed, changes in population and in the taste of consumers, weather patterns that may affect the demand for heating and cooling, and any exogenous policies to curb demand, improvements in efficiency, or the development of substitutes for oil that are not driven by price changes.

This functional form is assumed to do a reasonable job of capturing factors that are evolving in a consistent pattern over time. For example, if income or population growth has caused demand to increase, but the percentage rates of increase are gradually declining, this functional form will capture that trend and project continued declines in the growth rate in the future. On the other hand, a more significant structural shift could lead to changes in demand that are not well captured by the model. If a technological breakthrough leads to rapid adoption of a substitute for oil, for example, this could cause a shift in demand that is not predicted by the model. The model instead assumes that trends in technological progress will proceed as they have in the past. If technological development has slowed the growth in demand in the past, that trend is projected to continue, and if the technological progress has become more rapid, that trend also is projected to continue, but any technological progress that leads to a divergence from past trends will not be anticipated.

Using this relationship to produce long-term forecasts entails more difficulties. If the data show a decreasing growth rate in demand over time, projections into the future will eventually show demand growth turning negative. This projection is highly dependent on the functional form chosen – a decreasing growth rate could also be modeled by a relationship in which growth is slowing but never turns down.

For these reasons, projections made with this model, especially long-run projections, should not be taken too literally. Any number of developments could lead actual demand to diverge from the model's projection. Still, it is useful to make projections as to how demand would move if past trends continue, based on a model that fits well with past data. A log-quadratic form produces a robust relationship with

16

historical demand data, and as good a projection as can be made without adding more complications to the model.

2.2.1.2 Demand response to prices

Demand responds to prices for a number of reasons. Individuals and companies that consume oil products may adjust their behavior immediately to consume less of those products. They may also invest in more efficient equipment, such as cars with higher MPG, which allows them to reduce their consumption further. Higher prices may also spur the development of new technologies that make further reductions possible. Finally, higher prices can motivate the development of new policies to help encourage reductions in demand.

One consistent finding in the literature on demand responsiveness to price changes has been that the price responsiveness of demand in the long-run is considerably greater than the short-run response (Dahl, 1993; Fattouh, 2007). This makes sense, since many of the ways that the demand can adjust to a change in prices are not likely to happen immediately. Behavioral changes can happen the most quickly, but it may still take some time to change old habits. Improvements in the efficiency of capital equipment, including cars and trucks, requires replacing the old capital stock, a process that will happen gradually over time. The implementation of new policies and the development of new technologies may take even longer to have an effect on demand.

One common approach to modeling the dynamic response of demand to changes in price is known as the Koyck model (Bohi, 1981, p. 18), which specifies demand as a function of current and lagged prices in the form:

$$QD_t = a_0 + a_1 \cdot \sum_{i=0}^{\infty} \lambda^i \cdot p_{t-i} + e_t \qquad\qquad (2\text{-}2)$$

In the Koyck model, e_t is an independent, randomly distributed error term. An

alternative specification known as the partial adjustment model produces the same

relationship between demand and prices, with a slight difference in the error term: instead

of being independent, the e_t would follow: $e_t = \lambda \cdot e_{t-1} + u_t$, with the u_t independent and

randomly distributed (Bohi, 1981, p. 19). In the partial adjustment model, demand moves

a fixed portion of the way in each period toward an equilibrium level that would be

achieved in the long run if prices remained constant.

Another important finding in the oil demand literature is that demand responds

asymmetrically to rises and falls in price, and also responds differently to increases that

represent price recoveries compared with new maximum prices (Gately, 1993; Gately and

Huntington, 2002; Hamilton, 2003).

In our model, demand responds gradually to changes in what I call the 'effective

price' in a similar manner to the partial adjustment model. The effective price is defined

in such a way that price increases, and particularly new maximums, can have a greater

impact than price decreases, to allow for asymmetry in the demand response.

For computer modeling purposes, it is useful to break the demand function down

into the short-run response to prices and long-run shifts in the demand curve over time.

The short-run demand response is given by the constant elasticity demand function:

$$QD_t = D1_t \cdot p_t^{SRE_D} \qquad\qquad (2\text{-}3a)$$

or equivalently:

$$\log(QD_t) = \log(D1_t) + SRE_D \cdot \log(p_t). \qquad\qquad (2\text{-}3b)$$

In these equations, QD_t is the quantity demanded at time t, $D1_t$ is a demand parameter that shifts over time, p_t is the price at time t, and SRE_D is the short-run price elasticity of demand.

In the long run, the demand parameter $D1_t$ moves stochastically over time, following a quadratic time trend while also responding to past prices:

$$D2_t = \log(D1_t) = b_{0D} + b_{1D} \cdot t + b_{2D} \cdot t^2 + LRA_D \cdot pMA_{t-1} + e_t \qquad (2\text{-}4)$$

Here, LRA_D is a long-run price adjustment parameter that combined with the short-run elasticity gives the long-run price elasticity of demand, and pMA_{t-1} is a composite of past prices and maximum prices designed to capture the delayed and asymmetric effects that prices can have. Consistent with the partial adjustment model, *pMA* is expressed as a moving average of past 'effective prices,' which will be defined later.

$$pMA_{t-1} = (1-\delta)\left(\sum_{i=1}^{\infty} \delta^{i-1} \cdot \log(pEff_{t-i})\right) \qquad (2\text{-}5)$$

The error term e_t follows an AR(1) process:

$$e_t = \theta_D \cdot e_{t-1} + u_t \qquad (2\text{-}6)$$

The term u_t is a normally distributed disturbance term with mean 0, constant variance σ_D^2 and is independent of all other independent variables and error terms from previous periods.[2] θ_D determines the extent to which demand shocks persist.

Combining the above equations, we get:

---

[2] One possible improvement would be to check if the variance of the disturbance term is constant over time, and allow for heteroskedasticity if it is not.

$$\log(QD_t) = b_{0D} + b_{1D} \cdot t + b_{2D} \cdot t^2 + LRA_D \cdot (1-\delta)(\sum_{i=1}^{\infty} \delta^{i-1} \cdot \log(pEff_{t-i}))$$

$$+ SRE_D \cdot \log(p_t) + \theta_D \cdot e_{t-1} + d_t \tag{2-7}$$

This is similar to the equations from the Koyck and partial adjustment models, with the only differences being: 1) the use of effective prices instead of actual prices; 2) an additional parameter that determines the short-run elasticity separately from the long-run elasticity and the adjustment rate; and 3) in the error term, which is similar to the error term in the partial adjustment model, the autoregressive term for the error, θ_D, need not equal the adjustment term for prices, δ.

Effective prices are here defined to capture the fact that demand responds asymmetrically to changes in price. When prices rise, this can spur the adoption of new, more efficient technologies, a process that will not be reversed if prices later fall. A recovery from an earlier drop in prices also is unlikely to have as much of an effect as a rise to a new maximum, because many of the efficient technologies and strategies that were implemented during the previous price rise will still be in place. Several approaches to modeling this feature of price response have been tried, but none has done a perfect job of capturing its underlying logic. One promising approach is to use the maximum price that has been achieved to date as a separate dependent variable, in addition to the current price (Gately, 1993; Gately and Huntington, 2002). This would capture the fact that both price decreases and price increases that are below an earlier maximum would have less of an effect than a price increase that leads to a new maximum. However, it doesn't recognize the fact that a maximum achieved in the past 10 years may be more relevant than a maximum achieved, say, over 100 years ago. With the price series used here, prices from the first two years of the series (1870-1871)

reached \$45.82 in real terms – a maximum that was only barely exceeded by the prices spike of the early 1980's, which reached \$53.14 in 1981. As a result, most of the price rises in the 1970's did not produce a new maximum, even though for practical purposes the price in 1871 was not likely to have been relevant 100 years later. One way to surmount this problem would be to look at the maximum from some fixed period, such as the past 30 years. But the choice of the number of years to pick would be arbitrary, and could lead to a sudden change in the variable when an old maximum price is no longer in the time period. A better approach is to use a weighted average of the maximums using different time-length periods. This is the approach taken in this model. Specifically:

$$pEff_t = (1-\phi)\left(\sum_{k=0}^{\infty}\phi^k \cdot \max\{p_t,...,p_{t-k}\}\right) \qquad (2\text{-}8)$$

Combining equations (2-3)-(2-6) and (2-8), the demand model is determined by the following set of equations:

$$QD_t = D1_t \cdot p_t^{SRE_D}$$

$$\log(D1_t) = b_{0D} + b_{1D} \cdot t + b_{2D} \cdot t^2 + LRA_D \cdot pMA_{t-1} + e_t$$

$$e_t = \theta_D \cdot e_{t-1} + u_t \qquad (2\text{-}9)$$

$$pMA_{t-1} = (1-\delta)(\sum_{i=1}^{\infty}\delta^{i-1} \cdot \log(pEff_{t-i}))$$

$$pEff_t = (1-\phi)\left(\sum_{k=0}^{\infty}\phi^k \cdot \max\{p_t,...,p_{t-k}\}\right)$$

2.2.2 Supply Function

There is even more disparity in how supply is modeled in different studies. A number of models predict oil supply based on purely geological considerations,

21

considering the amount of oil remaining, assumptions about the depletion rate of existing reserves and the rate of discovery of new reserves (Hubbert, 1962; Campbell and Laherrere, 1998). Others address how production responds to changes in oil prices. However, there is no consensus as to the best model to use. Supply from OPEC and non-OPEC countries are frequently separated. For non-OPEC countries a positive relationship between prices and supply quantities is generally assumed, usually with little theoretical justification for the form of the relationship (Cremer *et al*., 1991, p. 61). For OPEC countries there is no consensus even as to the direction of the effect that prices have on production levels (Kaufmann *et al.*, 2008; Ramcharran, 2002). Theoretical models of supplier behavior, such as the Hotelling model, focus on how suppliers choose when to bring oil to the market depending on how the current price relates to expectations of future prices.

The supply model in this paper incorporates aspects of both the geological curve-fitting model developed by Hubbert (1962) and economic models of price responsiveness. Supply levels respond to current and past market prices, and can move over time based on historical trends as well as geological constraints.

Expectations about future prices could also affect supply decisions, as suppliers could choose to hold onto resources if they expect that future scarcity will drive up prices. However, future price expectations do not enter into the supply equations in this paper. We leave a discussion of the role of price expectation to chapter 3 of this dissertation, as the formation of price expectations is the primary focus of that chapter.

2.2.2.1 Short-run supply function

I choose a supply function that responds positively to prices, subject to a maximum capacity constraint in the short run. In the long run, production capacity also adjusts in response to past price changes. This approach is consistent with the assumptions used for non-OPEC supply by the US Energy Information Administration in producing their International Energy Outlook (EIA, 2007, pp. 9-10).

In the short run, there is a limit on how much can be produced, no matter how strong the economic incentive, because time is required to install the new capital needed for production. When oil prices are high, production will be close to full capacity, and it will approach full capacity as prices approach infinity. For lower prices, production levels will gradually drop until, at a price of zero, production reaches zero, since there is no incentive to produce. A simple function that satisfies these properties is:

$$QS_t = (1 - e^{-Ap_t}) \cdot C_t \tag{2-10}$$

Here, C_t is the maximum capacity at time t, p_t is the price of oil at time t, and A is a constant that determines how quickly production approaches maximum capacity as prices rise.[3]

This short-run relationship between prices and production is a middle ground compared to the results of past theoretical and empirical studies. Most empirical studies have found very low short-run price elasticities of supply, and in some cases, negative

---

[3] It would perhaps be more realistic to model supply as a function of prices relative to some expectation of future prices, rather than absolute prices. This would require an assumption about how price expectations are formed, which we put off dealing with until chapter 3.

price elasticities have been found (Fattouh, 2007, p. 18-19). One explanation for negative price elasticities, known as target revenue theory, is that OPEC countries aim to meet budget requirements, and sell more oil when prices drop to keep revenues from dropping below target levels (Ramcharran, 2002).

On the other extreme, the assumptions behind the Hotelling model and other related models suggest that production could jump from zero to the maximum in response to a small change in price. Producers, in deciding when they should produce to maximize profits, should withhold production completely if prices are lower than expected future prices, while if prices are higher than expected future prices, they should sell all their oil immediately.

A compromise, with small but positive responses to price changes, is most realistic, and the best to use for long-run modeling purposes, for several reasons. While there appear to have been historical periods where the target revenue theory may have been valid, particularly the late 1980's and 1990's, when low prices compelled OPEC countries to increase their production beyond quotas due to revenue shortages, the long-run strategy of OPEC is to try to stabilize prices around target prices by increasing production when prices are too high, and decreasing production when prices are too low. This implies a positive relationship between prices and production.

For firms operating in a competitive market setting, one would also expect a positive price elasticity of supply – but not the sort of extreme response predicted in the Hotelling literature. Incorporating uncertainty about future prices into the Hotelling model leads to the conclusion that production could respond more gradually to changes in price than in the basic model. Extensions of the Hotelling model that account for the

costliness of capital investments needed to build production capacity also suggest a less dramatic response to a change in prices, as well as a limit on the amount that can be produced in a given year (Campbell, 1980). With empirical studies producing mixed results, it is appropriate to assume a positive but limited relationship, as this has the strongest theoretical foundation.

2.2.2.2 Long-run supply changes

In the long run, several factors can cause capacity levels to shift. Resource depletion decreases production capacity in mature fields. Past prices can motivate changes in the level of investment in new exploration and production capacity. As with demand, the price responses are asymmetric, with new maximums, decreases and recoveries having different effects. In addition, technological developments, new discoveries, weather, political changes and conflicts can lead to changes in production capacity over time. These are not modeled explicitly, but are captured by a quadratic time trend combined with an error term to capture random, unpredictable shifts.

These shifts in supply, whether caused by a time trend, past prices, or stochastic volatility, should not be expected to have a linear effect on capacity. No matter how much these variables shift, capacity can never exceed the total amount of oil in the ground, and in fact, it is unlikely to ever exceed some fixed percentage of total reserves. This is because for any given oil field it is inefficient, if not impossible, to pump all the remaining oil in the field in one year. The model therefore assumes that production capacity will never exceed c times the amount of ultimately recoverable reserves remaining in the ground, where c is a constant between zero and one. On the other end, production capacity cannot go below 0.

To fit these constraints, I use a logit transformation to convert *S2*, which is a function of time and prices, into C_t, the maximum capacity at time *t*, which ranges from 0 to $c \cdot R_t$, where R_t is the amount of ultimately recoverable reserves remaining in the ground at time *t*.

$$C_t = \frac{1}{(1 + e^{-S2_t})} \cdot c \cdot R_t \tag{2-11}$$

S2 evolves as a quadratic function of time, and also responds to a moving average of effective prices, along with a disturbance term. The disturbance term follows an AR(1) process, as with demand:

$$S2_t = b_{0D} + b_{1D} \cdot t + b_{2D} \cdot t^2 + LRA_S \cdot pMA_{t-1} + e_t \tag{2-12}$$

$$e_t = \theta_S \cdot e_{t-1} + u_t \tag{2-13}$$

The moving average of effective prices is defined as it was in equations (2-5) and (2-8). Supply adjusts gradually to price changes because building capacity requires capital, and it takes time for capital levels to adjust to a new price level. Also, as with demand, there is reason to believe that supply responses to price increases may not be reversed by subsequent decreases in price. I am not aware of any studies that address this issue, but the reasons for it are similar to demand: an increase in price to new highs will drive the development of new production technologies, new exploration and new capacity investment that will not be completely reversed if prices later fall. Therefore, the supply model uses the same effective price variable as the demand model.

The reserve variable R_t represents the level of ultimately recoverable resources in the ground. It begins at time zero with a fixed value $\overline{R_0}$ that is chosen based on estimates from the literature. After this it is depleted as oil is produced, but otherwise

26

does not change. Many analysts argue against using a fixed quantity of ultimately

recoverable resources, arguing that reserves are variable over time depending on

technologies and prices (Lynch, 2002, p. 378; Fattouh, 2007, p. 7). Some reserves may

not be recoverable with current technologies, but may become recoverable with future

technologies – others may be too expensive to be worth extracting now, but may become

economically feasible if prices rise sufficiently.

There is still a finite amount of oil available for extraction, however, which limits

the amount that can be extracted no matter the price or the rate of technological

advancement. It is possible to allow prices and technological developments to influence

how much oil is ultimately used while still imposing a limit on the amount of oil

available. Including variable reserve levels complicates the model with little clear

benefit.

Putting all this together, the supply function is given by the following set of

equations:

$$QS_t = (1 - e^{-Ap_t}) \cdot C_t$$

$$C_t = \frac{1}{(1 + e^{-S2_t})} \cdot c \cdot R_t$$

$$R_0 = \overline{R_0}$$

$$R_t = R_{t-1} - QS_{t-1}$$

$$S2_t = b_{0D} + b_{1D} \cdot t + b_{2D} \cdot t^2 + LRA_S \cdot pMA_{t-1} + e_t \qquad (2\text{-}14)$$

$$e_t = \theta_S \cdot e_{t-1} + u_t$$

$$pMA_{t-1} = (1 - \delta)(\sum_{i=1}^{\infty} \delta^{i-1} \cdot \log(pEff_{t-i}))$$

$$pEff_t = (1-\phi)\left(\sum_{k=0}^{\infty}\phi^k \cdot \max\{\,p_t,...,p_{t-k}\,\}\right)$$

2.2.2.3 Hubbert's curve

Although this model is motivated by some simple intuitive rules for determining how supply would evolve over time, it produces a supply function of a form that is related to another common approach to modeling oil supply developed by M. King Hubbert (1962). Hubbert proposed that oil production over time would evolve according to the function:

$$Q_S(t) = h * \frac{e^{-\frac{(t-t_{\max})}{w}}}{\left(1 + e^{-\frac{(t-t_{\max})}{w}}\right)^2} \tag{2-15}$$

where t_{max} is the time at which the peak of the distribution occurs, and w and h are parameters that determine the width and height of the distribution. It is the derivative of the logit transformation function, and is shaped like a bell curve – though it is slightly different from a normal distribution.

In a limiting case, the model in this paper can produce a supply function over time that is precisely equal to Hubbert's curve. This occurs when there is no random variation ($\sigma_S^2 = 0$), when prices are constant at $\overline{p}$, when there is no quadratic time trend ($b_{2S} = 0$) and when $b_{1S} = c \cdot (1 - e^{-A\overline{p}})$. However, this model adds some flexibility by allowing these parameters to vary, addressing some of the common criticisms of Hubbert's approach.

One criticism is that Hubbert's approach fails to take into consideration how economic factors can affect production. The present model addresses this by allowing both production capacity and capacity utilization rates to respond to price changes – within the constraints imposed by geological factors.

A second criticism addressed in this model is that Hubbert's peak predicts that the curve will be exactly symmetric over time, with the rate of increase on the way up equaling the rate of decrease on the way down, and the peak occurring when exactly half of the total resource has been exhausted. Several observers have noted that in places that are past their peak, such as the US, the down-slope has proven to be less steep than the upslope, with a fatter tail. If this is true in general, models that assume a symmetric form overestimate the rate of depletion after the peak is reached (Fattouh, 2007, p. 16; Lynch, 2002, p. 380).

The model used in this paper, on the other hand, will accommodate any rate of increase and any rate of decrease, creating the flexibility to allow for forms with a rate of depletion that is either faster or slower than the rate of increase. If prices are constant at $\bar{p}$, and if there is no quadratic term ($b_{2S} = 0$), the variable b_1 determines the rate of increase on the way up, and the rate of decrease on the way down is set by the quantity $c \cdot (1 - e^{-A\bar{p}})$. For further discussion and mathematical derivation of these results, see Appendix A. Including a quadratic term adds additional flexibility.

2.2.3 Determining parameters of the model

The model described above provides a flexible framework for modeling oil demand and supply, but to implement it and use it to make projections, specific parameter

values are needed. A mix of different approaches is used in choosing these values. The goal is to produce a good fit with historical data on prices and quantities, using parameter values that are logically sensible and consistent with estimates of similar values from elsewhere in the literature.

As described below, some parameter values are estimated using formal regression techniques, minimizing the sum of the squares of the values of the error term u_t, once the rest of the parameters have been set. The remaining parameters are chosen to be consistent with estimates from other studies when possible, or with basic intuition when such estimates do not exist.

2.2.3.1 Demand parameters

Two key parameter choices for the demand function are the long-run and short-run price elasticities. There is an extensive body of literature that estimates elasticities of demand for oil and oil products. Most of these studies look at demand for specific products such as gasoline or heating oil, where panel data is more available and aggregation problems are less of a concern. Studies of the elasticity of demand for crude oil are less common, and rely on more limited data with less spatial variation that makes it difficult to separate demand from supply and obtain significant results. Fattouh (2007) summarizes the results of several studies showing a range of 0.001 to -0.11 for short run elasticities and 0.038 to -0.56 for long run elasticities. Dahl (1993), in an extensive review of the literature on energy demand elasticities, suggests -0.1 and -0.5 as reasonable estimates for the short and long-run elasticities (Dahl, 1993, pp. 112, 114).

The price elasticity of demand for crude oil can also be approximated based on the results of studies of demand for oil products such as gasoline diesel and heating oil.

These studies are more common, and the estimates are more solidly established. However, inferring an overall demand elasticity for crude oil based on these results can be tricky. The price elasticity of demand for crude oil is likely to be lower than for oil products for two reasons. The first is that since crude oil is only one of the inputs that determines the price of oil products such as gasoline, an increase in the price of crude oil is likely to cause a smaller percentage increase in the price of oil products, which will make the elasticity of oil demand with respect to crude oil prices lower than elasticities with respect to retail prices such as the price of gasoline (Dahl, 1993). The extent of this difference varies depending on the price of crude oil relative to other inputs, but as a rough approximation for an average price level, we might expect elasticities to be about twice as high for gasoline as for crude oil.[4]

The second reason that the elasticity for crude oil could be lower is that if there is any substitution between the different oil products, an increase in the price of all oil products might have less of an impact on the demand for each oil product than if the product prices each rose separately. This is not likely to be a major concern, however, since in the short run there is little possibility of substitution between products, while in the long run, prices of different oil products generally move together so an increase in the price of one product will be accompanied by an increase in other prices as well. The long-run elasticities for oil products reported in the literature generally do not effectively control for the price of other oil products.

A comprehensive survey of price elasticities for different oil products was conducted by Dahl (1993). More recent updates have not significantly changed the

---

[4] This is based on the results of a regression of gasoline prices against oil prices in the US over the last 30 years.

conclusions (Graham and Glaister, 2002). The best estimates for the price elasticity of demand for gasoline from Dahl (1993) are -0.26 for the short run and a -0.86 for the long run (Dahl, 1993, p. 143). For residential fuel oil, the best estimates are between -0.2 and -0.26 in the short run, and from -0.75 to -1.0 in the long run (Dahl, 1993, pp. 111, 114-115). Industrial demand for oil products has a price elasticity around -0.2 in the short run and -0.8 in the long run (Dahl, 1993, p. 120).[5]

A weighted average for all oil products would come out between -0.2 and -0.25 in the short run, and -0.8 to -1.0 in the long run. Converting to crude oil prices, this comes close to the estimates of -0.1 and -0.5 that were found in studies that looked directly at the demand for crude oil.

For the model in this paper, I use a short-run elasticity of -0.15 and a long-run elasticity of -0.75. The long-run adjustment parameter LRA_D that is input into the model is the difference between these, -0.6.

The short-run elasticity is chosen to be slightly higher than has generally been found, to help avoid major price fluctuations in response to a small shift in demand or supply. In the real world, above-ground storage capacity can help to dampen these fluctuations since demand and supply need not be exactly equal in each period, leaving some more time for demand and supply to come back into balance. Because this model forces demand to equal supply in each period, slightly higher short-run elasticities are useful to keep supply and demand in balance without major price fluctuations.

---

[5] For gasoline, this is the conclusion of the most comprehensive recent review, Dahl and Sterner (1990, 1991a, 1991b). For fuel oil, the range is from a rough average of the new studies tables (C31 & C32) to the average of the old studies table (p. 111). For industrial oil demand, the best estimate numbers were quoted in Dahl (p. 120).

I also chose a long-run elasticity that was slightly higher than the average for the literature, for a different reason. It represents the elasticity of demand in response to a change in effective prices, which is only equivalent to a change in actual prices if the new price represents a new maximum. The effect of a change in price to a new maximum should be higher than the estimates from the literature, which capture the effect of an arbitrary change in price. This is supported by past studies of asymmetric price effects. Dahl (1993, pp. 121-122) reports that a study by Gately found a long-run price elasticity of demand for crude oil for non-transportation uses of -0.7 for increases in price that are above the maximum. A more recent study by Gately & Huntington (2002, p. 39) finds a long-run price elasticity of OECD demand for oil to be -0.64 using an asymmetric model.

To estimate the demand equation, I also need to choose parameter values for the constant ϕ that determines how the effective price is defined, and δ, which determines how fast demand converges to the long run equilibrium in response to a price change. The value I chose for ϕ is 0.95. This means that the one-year maximum received a 5% weight, the two-year maximum received a weight of 4.75%, and the weight applied to each additional year decreased by 5% per year. With this value, 37% of the weight went to maxima over periods of less than 10 years, 25% to periods of between 10 and 20 years, and 38% to periods of 20 years or more. This choice was arbitrary, but it felt intuitively like a reasonable amount of weight to give to different time periods.

For the rate of adjustment to the long-run equilibrium in response to a price change, I chose a value of $\delta = 0.85$. This is similar to the value for θ_D that will be calculated later, which makes sense since both represent the rate at which demand responds to a change; in one case via a change in prices, in the other via a shift in the

curve due to any other cause that is not explicitly modeled. In a partial adjustment model, these two parameters will be the same.

In addition, the short-run elasticity, long-run elasticity and adjustment rate are related in the partial adjustment model by the equation $SRE_D = (1 - \delta) \cdot LRE_D$. With the parameter values I have chosen, the short-run elasticity is slightly higher than the value that would satisfy this equation. This makes sense, since some additional adjustment in the first period can be expected beyond the adjustment that occurs as the first year of the long-run adjustment process. Some forms of consumption require converting fixed capital assets to make them more efficient, a process for which the partial adjustment model is appropriate. However, some behaviors, such as the number of miles driven, can be adjusted immediately, suggesting that there should be some additional adjustment in the first period.

The remaining demand variables, b_{0D}, b_{1D}, b_{2D} and θ_D, are estimated using time series linear regression techniques and will be discussed in more detail below.

2.2.3.2 Supply parameters

The short-run supply function requires a parameter *A* that determines how quickly production approaches full capacity as prices rise. I use a value of 0.1 for *A* in this model. It is chosen to make capacity utilization levels reasonable at normal price levels. With this choice of *A*, production will be at 63% capacity when the price is $10 a barrel, 86% of capacity when it is $20 a barrel, 95% when it is $30 a barrel, and 98% when it is $40 a barrel.

The short-run price elasticities of supply at each of these prices are 0.58, 0.31, 0.16 and 0.07 respectively. At average prices, these reflect a higher price elasticity than has generally been found in empirical studies. As with our choice for the short-run demand elasticity, this makes it easier for demand and supply to reach equilibrium without any change in above-ground storage.

The long-run supply adjustment term LRA_S determines how much supply adjusts in the long run to an extended period of higher or lower prices. If production capacity is a small portion of the maximum possible capacity, the long-run price elasticity is approximately equal to the sum of the long-run adjustment parameter and the short-run supply elasticity. As capacity approaches the maximum proportion of remaining reserves, the elasticity becomes lower than this number, and eventually approaches 0.

Our choice of 0.4 for LRA_S produces long-run elasticities that are at the high end of estimates from the literature. Using prices and production capacities from the past thirty years, this choice produces elasticities that range from 0.42 to 1.01, with a mean of 0.69. The higher elasticities in this range are driven by high short-run elasticities caused by low prices. This can be compared to estimates of long-run supply elasticities in other studies summarized in Fattouh (2007), which range from 0.08 to 0.58 (p. 19). As with demand, it is appropriate to have an elasticity that is higher than the average in the literature because the effective price variable captures changes in the maximum price, which should have more impact on long-run supply than the average change in price.

Another important parameter in the supply model is $\overline{R_0}$, the estimate for the total amount of oil available in the ground before any is extracted. For this, we use an

estimate from the USGS (2000) for ultimate recovery of crude oil of 3003 billion barrels of oil (EIA, 2004).

This is at the high end of all estimates of ultimately recoverable resources. MacKenzie (2000) summarizes earlier estimates from the previous 25 years, and finds that they had not changed much over time, and consistently ranged from 1.8 to 2.4 trillion barrels, compared with the USGS estimate of 3.0 trillion barrels. Some authors have criticized the USGS approach, and argued that the discovery rates implied by the USGS projections are unrealistically high and out of line with recent history (Heinberg, 2003). It is the most widely cited estimate in the recent literature, however. Moreover, a high-end estimate suits my purposes, since I am trying to capture all oil that could possibly be recovered taking into account the technological progress and greater economic incentives that are likely to occur. Lower-end estimates often fail to include oil that is currently not economical to recover with current technologies. In addition, using a high-end estimate helps to guard against the criticism that our results are driven by an unrealistically low estimate for ultimately recoverable reserves.

The parameter c determines the maximum that the yearly production capacity can reach as a proportion of total remaining ultimately recoverable reserves. This parameter is set at 0.2. This is somewhat higher than the depletion rates that have generally been observed for fields in decline. EIA (2004, p. 3), for example, uses a production-to-reserves ratio of 0.1 after production reaches its peak, based on data from existing fields that are in decline. I choose a higher parameter of 0.2 to capture the absolute maximum that could be achieved – a maximum that is in practice never reached in the model.

Based on the Hubbert's peak discussion earlier, this parameter influences the rate of decrease of the curve as the resource approaches exhaustion. If supply were modeled as a linear function of time, this choice of c would imply a rate of decrease that is faster than the initial rate of increase. However, including the quadratic term leads to decreasing supply and demand coefficients over time, and causes a slower rate of decrease.

For the rate of adjustment parameter δ, and the asymmetric price effect parameter ϕ, I use the same numbers as for demand. In principle they could be different, but I saw no theoretical or empirical basis for choosing either one to be higher than the other. The remaining supply variables, b_{0S}, b_{1S}, b_{2S} and θ_S, are estimated using regression techniques as discussed below.

2.2.3.3 Estimation of remaining parameters

Once the parameters discussed above have been set, it is possible to set up linear regressions for demand and supply to estimate values for the remaining parameters. Rearranging equations (2-4) and (2-12), we get:

$$D2_t - LRA_D \cdot pMA_{t-1} = b_{0D} + b_{1D}t + b_{2D}t^2 + e_t \qquad (2\text{-}16)$$

$$e_t = \theta_D \cdot e_{t-1} + u_t$$

$$S2_t - LRA_S \cdot pMA_{t-1} = b_{0S} + b_{1S}t + b_{2S}t^2 + e_t \qquad (2\text{-}17)$$

$$e_t = \theta_S \cdot e_{t-1} + u_t$$

The left side of these equations can be calculated using time series data on prices and quantities of crude oil production, along with the equations and parameter values discussed above.

2.2.3.3.1 Data

The data used in the regression are a long time series of annual prices and quantities for crude oil that goes back to 1870, near the time when oil was first produced and marketed. This allows me to check that the model is consistent with the full history of the use of the commodity, rather than fitting just one historical period. The best source for long time series for oil prices is Manthy (1978), in an update of earlier work by Potter and Christy (1962). This series runs from 1870 through 1973, and provides the average wellhead prices for the major US oil fields. After 1973, I use a US average wellhead acquisition price by first purchasers from EIA (2008). I use inflation-adjusted prices. Before 1947, I use real prices provided by Manthy, which are adjusted by the 1967 US wholesale price index, as the producer price index was called before 1978. From 1947 on, I use the producer price index series provided by the BLS to deflate the nominal price data from Manthy and the EIA, and to convert to 2006 prices.

The prices I use are average prices levels for US producers. It would be more appropriate to have a world average, since I am looking at worldwide consumption of crude oil. Unfortunately, no similar worldwide average price series exists that goes back as far, so I use the US data as a proxy for the world price.

To construct a long time series for oil quantities, I use results from Marland *et al*. (2007), who use data on liquid fuel production from Etemad *et al*. (1991), along with more recent updates from the United Nations Energy Statistics Yearbook, to construct a

time series going back to 1870 for CO_2 emissions from liquid fuels. They provide the conversion factors that they use for the rate of CO_2 emissions per ton of oil consumed, which can be used to back out the oil production levels in barrels per day from these data.

2.2.3.3.2 Setting up the regressions

From these prices and quantities, $D2_t$ and $S2_t$ can be calculated by rearranging equations (2-3b), (2-10) and (2-11), and using the quantity data q_t for both QD_t and QS_t:

$$D2_t = \log(D1_t) = \log(q_t) - SRE_D \cdot \log(p_t) \tag{2-18}$$

$$S2_t = -\log(\frac{c \cdot R_t}{C_t} - 1) \tag{2-19}$$

$$C_t = \frac{q_t}{(1 - e^{-Ap_t})} \tag{2-20}$$

Once the adjusted demand and supply variables have been calculated, the estimation technique is identical for demand and supply, so for simplicity I will refer to the estimation parameters as b_{0i}, b_{1i}, b_{2i} and θ_i, where i can be either D or S. To estimate the model, I use an iterative process in which values for the b_{ji}'s and for the θ_i's are estimated alternately until they converge. I begin with an arbitrary initial estimate for θ_i (0.8), and estimate b_{0i}, b_{1i} and b_{2i} using a Cochrane-Orcutt transformation by subtracting θ_i times the lagged value of the dependent variable, making appropriate adjustments in the interpretation of the constant, linear and quadratic terms (Cochrane and Orcutt, 1948). The values for b_{0i}, b_{1i} and b_{2i} in the original, untransformed equation can be calculated from the estimated coefficients of the transformed equation. The error term e_t in each time period is then regressed against its

lagged value to get an estimate for θ_i. This whole process is then repeated with the new

estimate for θ_i until the value for θ_i converges (that is, until it changes by $< 10^{-5}$ from

one run to the next).

This approach is appropriate if the process is stationary, which will be true if θ_D

and θ_S are less than 1. If $0 < \theta_i < 1$, then a random disturbance in one period carries

over partly but not completely to the next period, so that demand and supply revert

gradually toward a trend, with θ_i determining how quickly it converges toward the trend.

If $\theta_i = 1$, on the other hand, then the series is non-stationary, and the error term follows a

random walk, with no reversion toward a trend.

I find that θ_D converges to 0.84 and θ_S converges to 0.81. To test that this is

significantly different from one, I use a Dickey-Fuller test on the error term from the final

repetition of the regression. For both demand and supply, I find that it is different from 1

at the 1% significance level.[6] It is therefore appropriate to model this as a trend

stationary process, with gradual reversion toward the trend at a rate determined by the

estimated values for theta.

The estimates for b_{0i}, b_{1i}, b_{2i} and θ_i in the demand and supply equations are

shown in Table 2.1. All coefficients are significant at the 1% level.

---

[6] This is for a Dickey-Fuller test with the constant term suppressed, since there should be no drift in the error term. Stata doesn't provide p-values for this test, but for a Dickey-Fuller test with drift, the estimated p-values are 0.0003 and 0.0002 for demand and supply respectively.

Table 2.1. Demand and supply regression results

	Demand[a]	Supply[a]
Const	5.80	-12.37
	(0.15)	(0.17)
T	0.090	0.109
	(0.005)	(0.006)
t^2	-0.00021	-0.00034
	(0.00004)	(0.00004)
Θ	0.84	0.81

Standard errors in parentheses. A Cochrane-Orcutt transformation was used in the estimation, but coefficients and standard errors reported are for the coefficients in the original, non-transformed equation.

[a]The dependent variable is adjusted demand and supply variables as defined in the text in equations (14) and (15).

One important thing to notice is that the coefficient for t^2 is negative and highly significant. There is no clear theoretical basis for deciding whether or not to include a quadratic term in the estimations, but the significance of the coefficient of the quadratic terms suggests that it is important to include it.

Figures 2.1 and 2.2 show how closely the model fits the actual values for the adjusted demand and supply variables that serve as the dependent variables in the regressions. The fit for demand is closer than the fit for supply, which has more random variation that is unexplained by the model. This makes sense because of the difficulty predicting OPEC behavior, the presence of conflicts that affect supply, and other random supply shocks.

Figure 2.1. Fit of demand regression
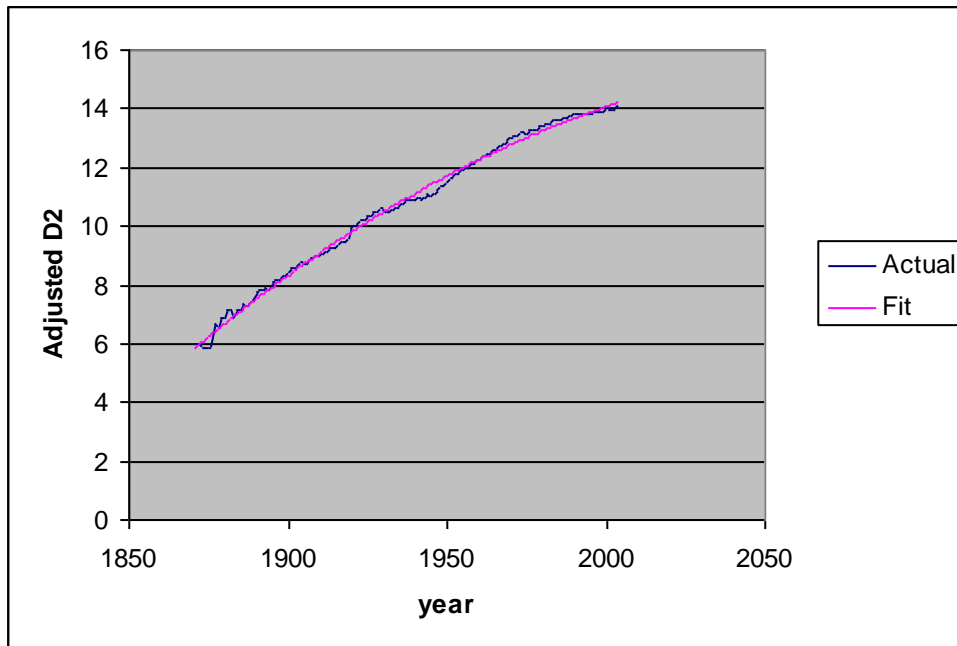
Figure 2.2. Fit of supply regression
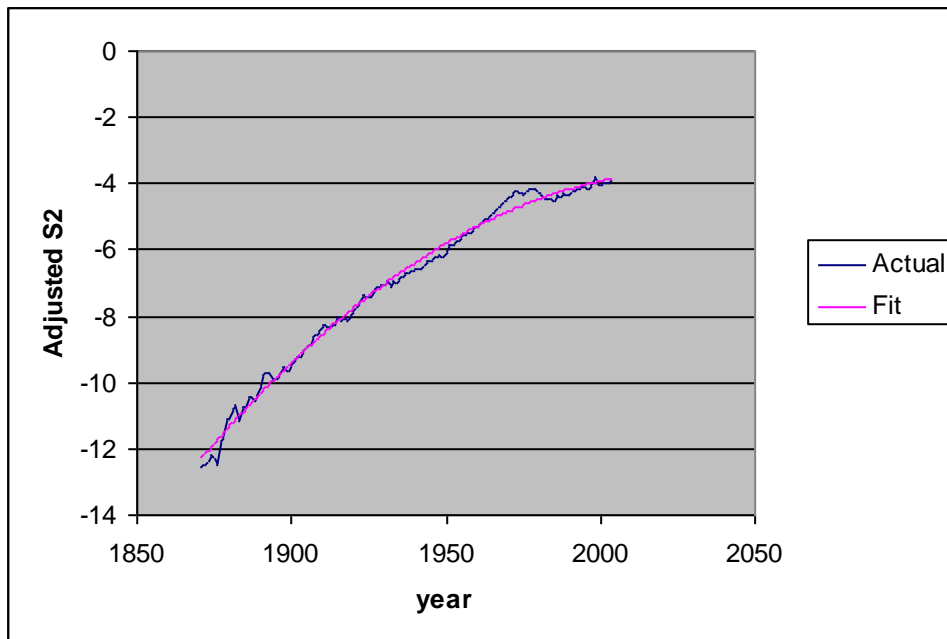
Figures 2.3 and 2.4 show how closely the model fits the original demand and

supply variables, $D1_t$ and C_t, with long-run price adjustments and supply responses to

reserve depletion taken into account. In the case of demand, the model captures well the

steady increase in demand through the end of the 1970s, the fall in demand in the 1980s, and the resumption of demand growth afterwards. One minor shortcoming is that after the dip in the 1980s, demand has not resumed growing as fast as predicted by the model.

The supply model does not come as close to capturing the movements in supply that have occurred over time. Again, this can be explained by the fact that many of the factors that influence supply are not included in the model, but are modeled as random supply shocks.

Figure 2.3. Model fit with demand
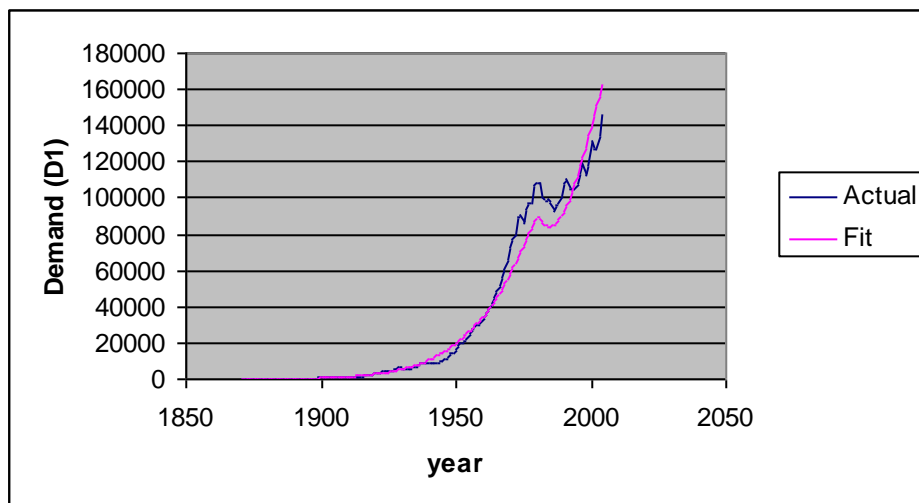
Figure 2.4. Model fit with supply
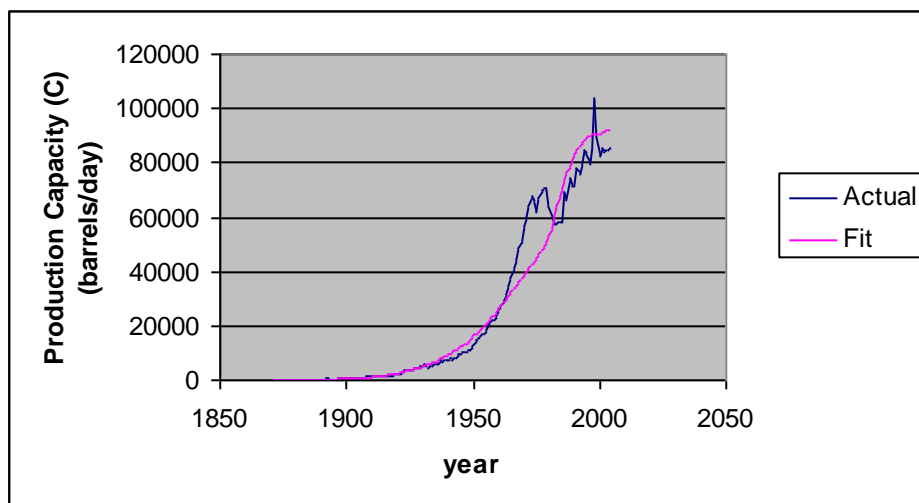
2.2.4 Price Setting

In each period, prices are chosen to equate the quantity demanded to the quantity supplied. Setting prices to equilibriate demand and supply is a common assumption of economic models, but in models of oil markets it is not universally adopted. In reality, supply need not equal demand in each period because above ground stocks can be built or drawn down to accommodate a temporary imbalance between demand and supply. Some models address this by including a price setting rule in which prices respond to the amount of oil stored above ground (Dees *et al.*, 2007). Others argue that significant imbalances cannot persist for long, so using an equilibrium model is appropriate for long-run forecasting (EIA, 2007). For simplicity, I stick with an equilibrium model in this chapter, and put off addressing the role of above-ground storage until chapter 3.

The equilibrium price is calculated numerically by testing different prices until an equilibrium is reached between the short-run demand and supply functions $QD_t = D1_t \cdot p_t^{SRE_D}$ and $QS_t = (1 - e^{-Ap_t}) \cdot C_t$. The demand and supply functions are constructed in such a way that there will always be a unique equilibrium price $p > 0$ as long as there are any oil reserves in the ground. This is guaranteed by the combination of three facts: 1) the short-run demand curve is monotonically decreasing while the short-run supply curve is monotonically increasing; 2) at $p = 0$, supply is always less than demand (in fact supply is zero while demand approaches infinity); and 3) as p approaches infinity, supply eventually exceeds demand. The first two facts are clear from the short-run demand and supply equations. The third fact is true because as p approaches ∞, demand approaches 0, while supply approaches C_t, which is given by $\dfrac{1}{(1 + e^{-S2_t})} \cdot c \cdot R_t$.

Each term in this expression must be positive as long as R_t is positive, which will always

be true since production in each period can never exceed $c \cdot R_t$.

Figure 2.5 shows the prices predicted by the model during the sample period, and

how they compare to actual prices. It replicates the price rise in the early 1980's, though

the timing of the spike doesn't quite match perfectly with the data.

Figure 2.5. Model fit with price
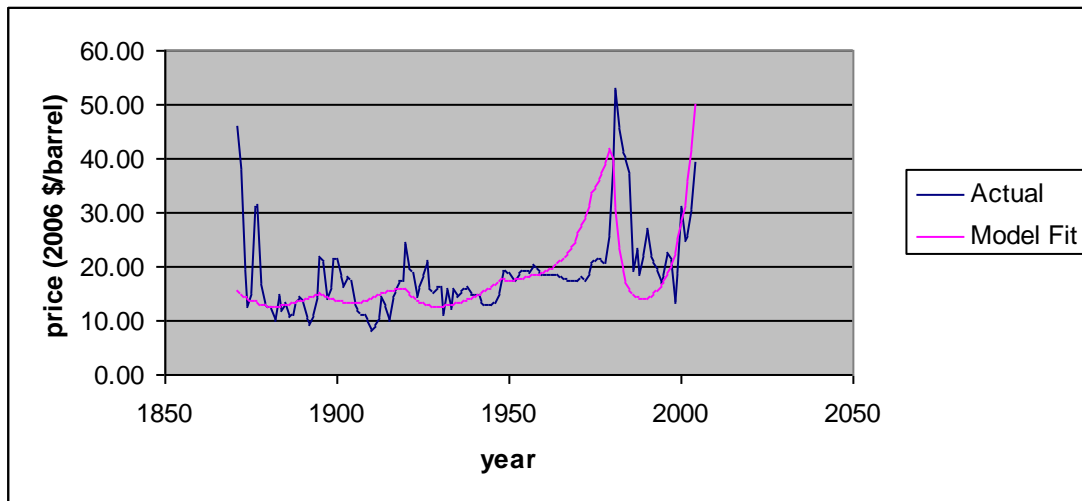
The fit for oil production and consumption levels is shown in Figure 2.6. The

model shows a slight slowdown in the growth of production in the 1980's at the same

time as actual production dips, but it doesn't come close to matching the extent of the

downturn. A greater demand response to the high prices of the early 1980's might be

needed to replicate this better.

45

Figure 2.6. Model fit with quantity
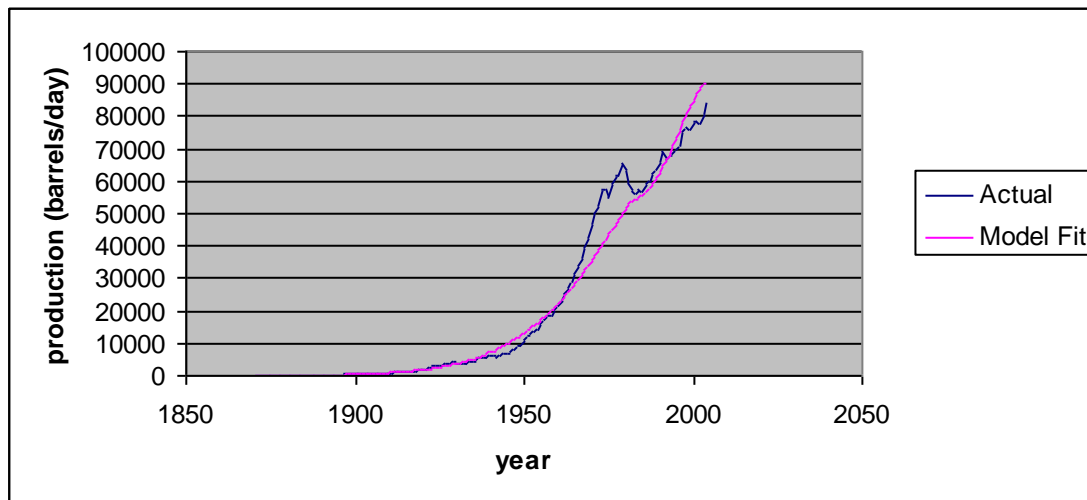
2.3 Model output

2.3.1 Projections

Projecting prices and quantities into the future is a straightforward exercise with this model, since everything is modeled directly or indirectly as a function of time. The projections can be broken down into an extension of the trend line along with a projection of the error term, starting with the error term in the final period of the sample, and assuming that the random part of the error term takes on its mean value of zero.

2.3.1.1 Quantity Forecast

The projections for the quantity of oil produced and consumed are shown in Figure 2.7. The shape looks similar to a Hubbert's curve, with production levels projected to peak in the year 2017, after which they begin to fall. The drop in the quantity supplied comes largely because resource depletion constrains supply more and more as time goes on. In addition, the time trends for supply and demand have negative

46

quadratic terms, which leads to additional drops in the growth rate of quantity supplied over time.

Figure 2.7. Quantity forecast
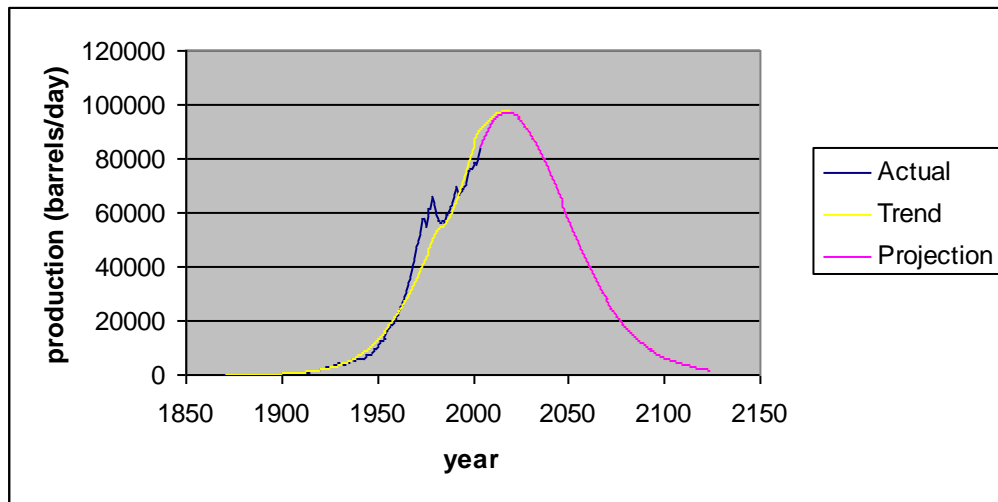
2.3.1.2 Price Forecast

The projections for the price of oil (adjusted for inflation) are shown in Figure 2.8. They show that prices are likely to increase persistently, though not too dramatically in the near term. In 2010, prices are projected to be at $60.44 per barrel, rising further to $93.72 in 2020 and $155.45 in 2030. Looking farther ahead, projected price levels become more extreme, reaching $400 a barrel in 2050. Beyond the years shown in Figure 2.7, prices are projected to continue to rise, eventually surpassing $10,000 a barrel in the year 2104, and $30,000 a barrel by 2124.

Figure 2.8. Price forecast
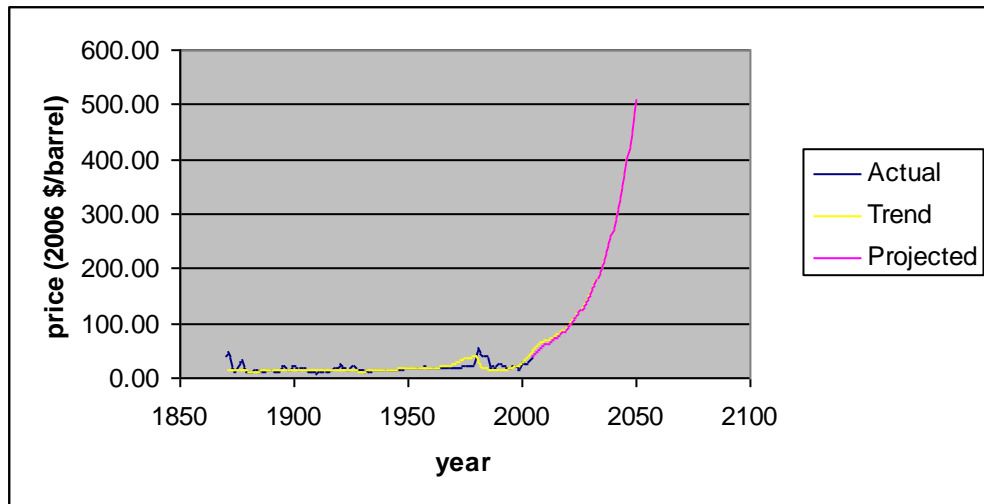
These results reflect the fact that resource exhaustion eventually leads to significant scarcity in this model. In the race between technological developments to reduce demand and the depletion of resources, resource exhaustion comes first. This outcome is not inevitable with the model assumptions we use. Since demand is projected to turn down eventually, continued reductions in demand could be fast enough to keep up with drops in supply due to resource exhaustion without the need for high prices. The fact that prices eventually rise to extreme level indicates that under the assumptions of the model, if past trends continue, the reductions in demand will not occur fast enough to avoid shortages without significant price increases.

This result is dependent on the assumptions of the model holding throughout the period for which the forecasts are made. The results could change if ultimately recoverable resources turn out to be higher than in the model, or if innovations can help spur more rapid demand reductions than could be forecast from past trends. Prices also may not rise as high as they do in these forecasts if demand responds more strongly to price increases once they reach a certain level.

Sensitivity analysis of how robust the results are to changes in these assumptions could help to address these concerns. Alternative assumptions about the shape of the demand relationship with time and price are more difficult to formulate, and I do not attempt to do so in this paper. Adjusting the resource level is much more straightforward, and provides some indication of how sensitive the results are to changes in the assumptions that drive the long-run forecasts.

2.3.1.3 Sensitivity analysis

As a test of the sensitivity of our results to the assumption about resource levels, we try two alternative scenarios, one with remaining resources in 2004 half as high as in the baseline model, and one with remaining resources doubled. The results are shown in Figures 2.9 to 2.12. Prices rise more quickly when resources are lower, and less quickly when resources are higher. By 2020, prices are projected to reach $149 per barrel in the low reserves case and $79 in the high reserves case, compared with $94 in the base case. Peak oil also occurs sooner (2009) with a lower resource estimate and later (2024) with a higher resource estimate.

Even in the high resource scenario, prices eventually rise to over $3000 / barrel by 2100. In other words, even if we double the amount of remaining resources from one of the highest estimates currently available, demand is not projected to decrease fast enough to avoid future scarcity. A more dramatic change in demand is therefore needed to avoid significant increases in the price of crude oil in the long run.

Figure 2.9. Price forecast with low reserves
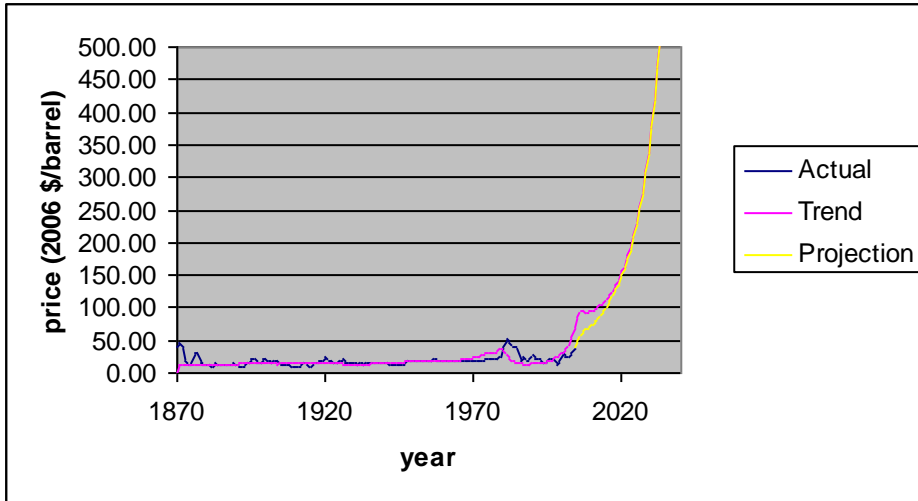
Figure 2.10. Price forecast with high reserves
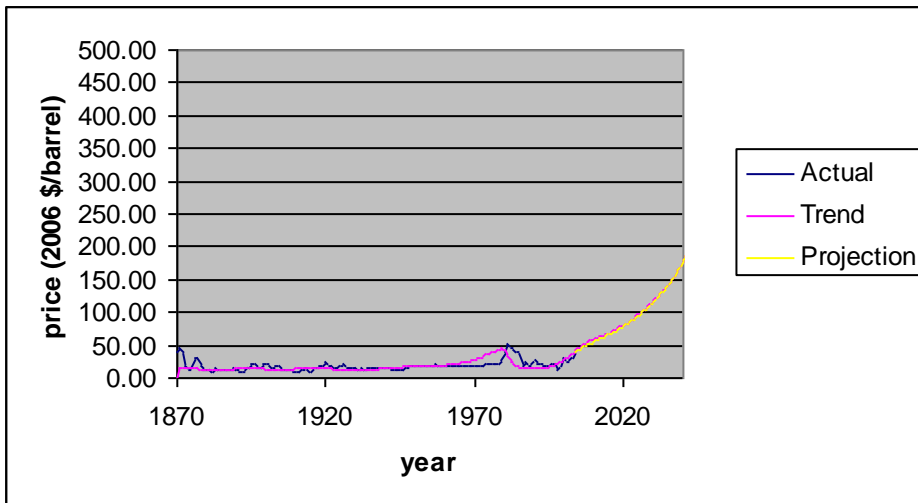
Figure 2.11. Quantity forecast with low reserves
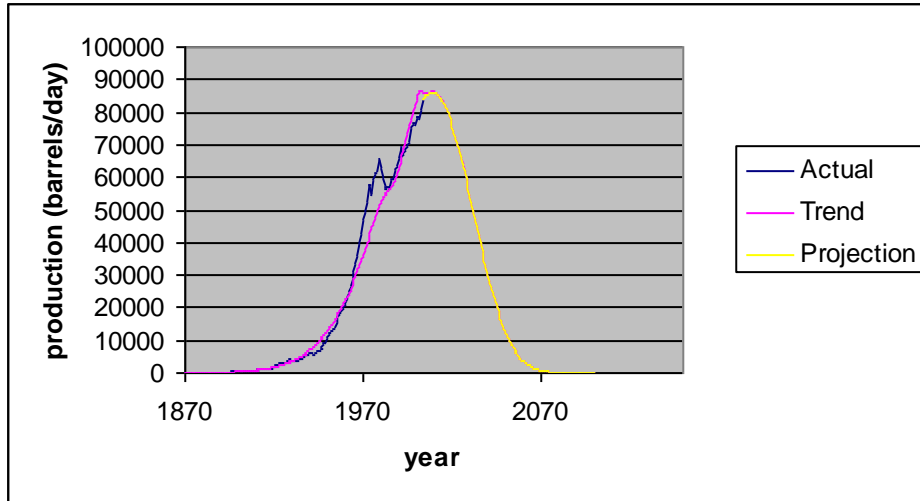
Figure 2.12. Quantity forecast with high reserves
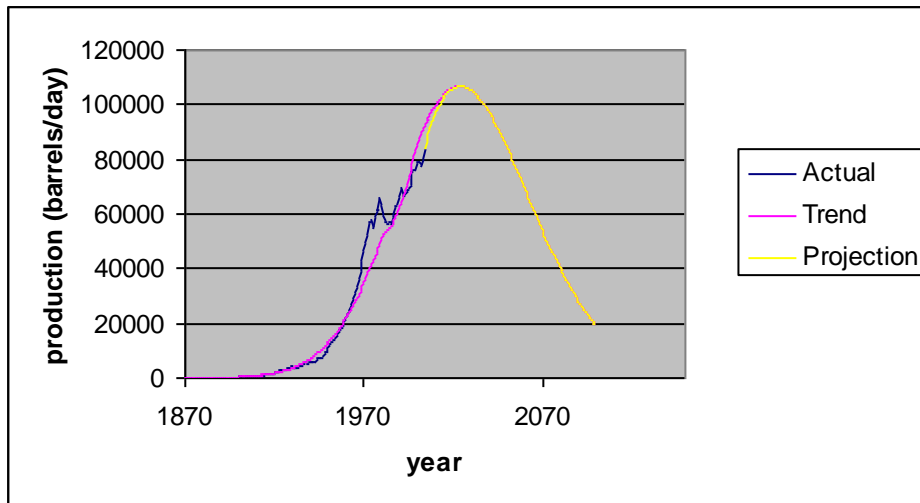
2.3.2 Simulations

In addition to projecting a best estimate for the expected price in each coming year, it can also be useful to see a simulation of how prices and quantities could behave, with random fluctuations included. This allows us to get a sense of the likely volatility of the series.

51

I estimate the variances σ_d^2 and σ_S^2 of the disturbance term u_t to be 0.009 for demand and 0.016 for supply.[7] I also test whether the disturbance terms of the two series are correlated. This is important in order to accurately capture the volatility of the price and quantity series. Positively correlated movements in demand and supply will have less effect on price than negatively correlated or uncorrelated movements.

I find that the correlation coefficient for the disturbance terms for demand and supply is 0.62. It is not clear if this correlation suggests that there is a legitimate reason for demand and supply to be correlated (possibly because of technological developments that affect both demand and supply), or if it simply a product of how our demand and supply variables were constructed. In either case, it is important to include it in our simulations, to produce as accurate as possible a simulation of prices and quantities.

Using a randomly generated normally distributed disturbance term, I construct simulated demand and supply variables, and use them to calculate the associated equilibrium prices and quantities. The simulation results for price and quantity are shown in Figures 2.13 and 2.14 respectively, along with actual price and quantity data. During the sample period, the simulated patterns look similar to the actual patterns – the ups and downs do not occur at the same time, but the size of the spikes and the length of time that they persist looks similar. As the simulation is projected into the future, it can be seen that the price projections presented earlier do not tell the whole story. In addition to the general increasing price trend, the simulated price variable shows some significant fluctuations in price, with a peak of $87 per barrel occurring in the year 2006 and $160 in 2017.

---

[7] If the variance of the disturbance term changed over time, this should be accounted for in projecting the disturbance term into the future.

Figure 2.13. Price simulation
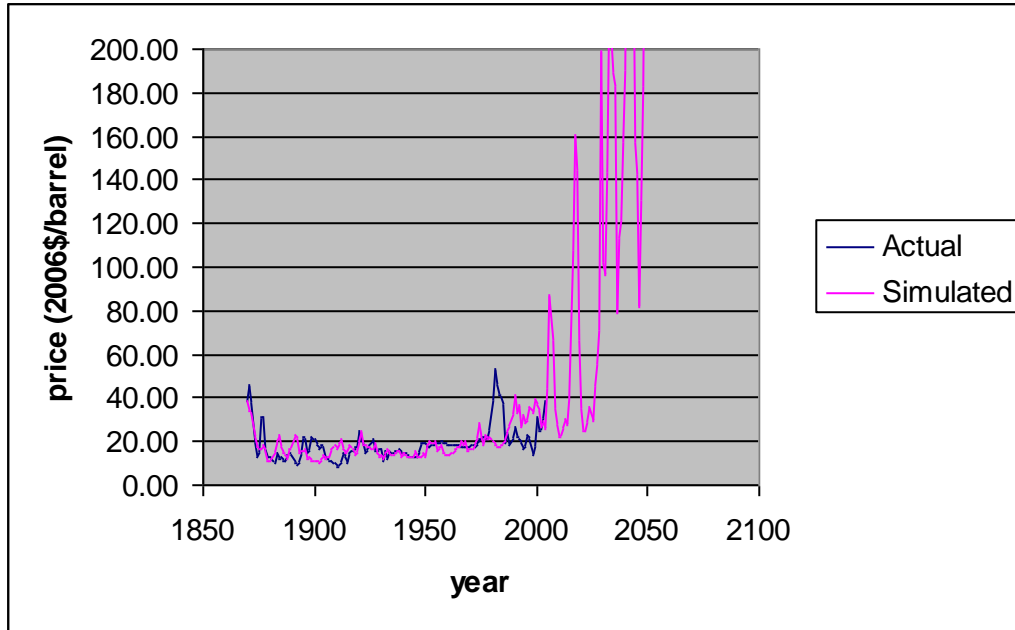
Figure 2.14. Quantity simulation
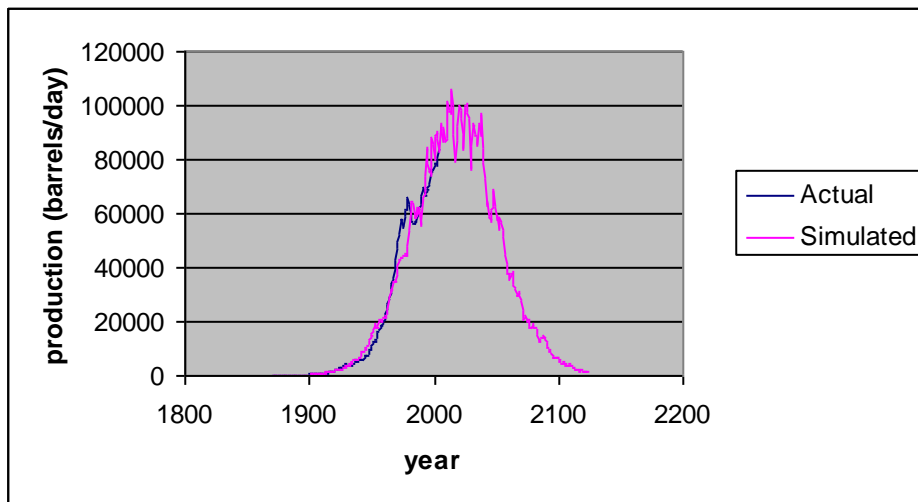
2.3.3 Testing effectiveness of forecasts

To test the effectiveness of the forecasts, I run the model on data from part of the sample period, and use it to predict prices for the remainder of the period. These ex-post forecasts can then be compared with actual prices to see how effective they were. Figure

2.15 shows the results of eight different ex post forecasts, each of which uses data from before a given base year to project prices after that year. The base years range from 1964 to 1999 in five-year increments.

Figure 2.15. Ex-post forecasts
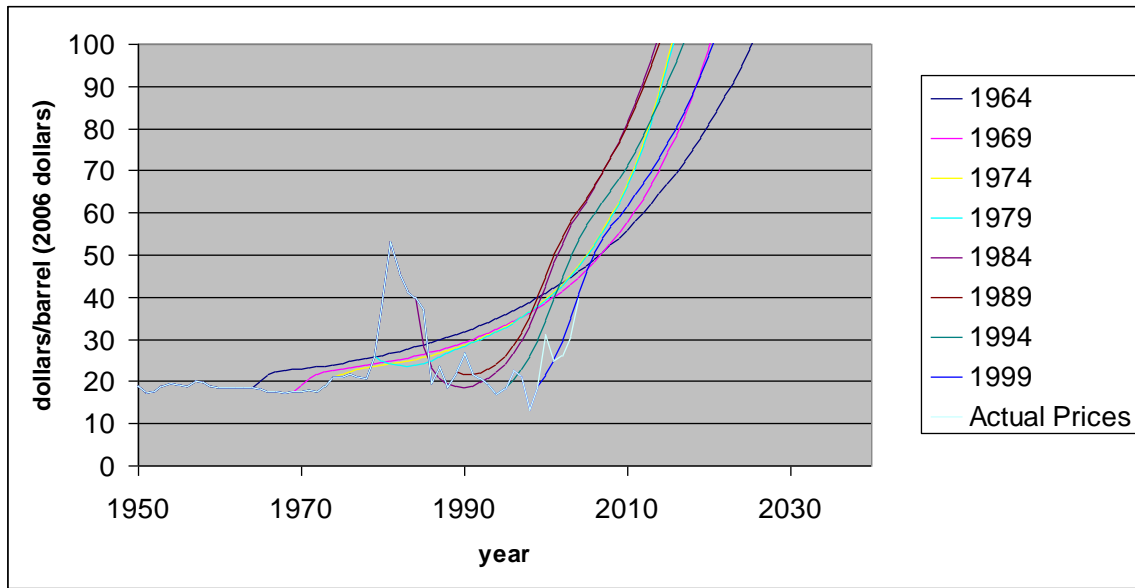
The forecasts that result are not perfect, but they compare favorably with the results of most other attempts at forecasting oil prices. They do not predict the spike in prices of the early 1980's, but this is understandable since it was driven largely by geopolitical developments that were hard to anticipate. The drop in prices in the end of the 1980's is different; it was driven largely by an imbalance between demand and supply due to adjustments in demand and in non-OPEC supply that occurred in response to the high prices of the early 1980s. Our model anticipates this well: the prediction from 1984 correctly predicts the drop in prices, followed eventually by a price rebound. The price rise beginning in 2000 is also predicted, though the timing of the predictions is not perfect.

To compare these forecasts with other forecasts that were made over this period, I look at two studies that summarize the results of other forecasts: Huntington (1994) and Lynch (2002). Huntington focuses on ten models that were collected as part of the Energy Modeling Forum's World Oil Study conducted in 1980 and 1981. Huntington summarizes the results of the projections for 1990 using each of the different models and compares them with the actual data. The predictions for crude oil prices in 1990 range from 133% to 301% higher than the actual price (Huntington 1994, p. 5). The model in this paper performs much better. Its prediction for the price in 1990 using data through 1980 is actually 29% lower than the actual price. For the years immediately before and after 1990, when prices were lower, the model projection comes even closer to the true value.

Lynch (2002) looks at a series of forecasts made by the US Energy Information Administration between 1978 and 2001. As with Huntington (1994), he finds that forecasts from the 1980's give price projections that were significantly too high, while more recent projections kept being revised downwards until the projections for 2000 and 2001 predicted that prices would remain near $20 per barrel through 2020. The earlier forecasts failed to anticipate the fall in prices in the 1980's and the continued low prices of the 1990s, while the later forecasts failed to anticipate the even greater rise in prices that has occurred in the 2000s. The model in this paper performs significantly better at anticipating both these shifts.

In making ex-post forecasts, I have one advantage over studies that were conducted in 1980 and 1981. In choosing the form of the model and some of the parameters I was able to benefit from experience and research that has occurred since

55

1981. For example, my estimate for ultimately recoverable reserves was based on a USGS study that was conducted in 2000.

It is impossible to adjust the model to make it entirely based on information that was available before 1980, since the general form of the relationships used in this model drew on studies conducted since 1980, but we can adjust some key parameters to be based only on data that were available then. Predictions of ultimately recoverable resource level that were made in the 1970's were notably lower than the USGS estimate from 2000 that was used in this paper. The elasticities of demand for oil can also be adjusted based on differences between estimates available in 1980 and those used in this paper. Specifically, the estimates of price elasticity of demand for gasoline from the review in Bohi (1981) are slightly lower than the estimates in Dahl (1993) that contribute to the elasticities chosen for this paper. To estimate what parameters I would have chosen given the information available in 1980, I revise the demand elasticities downward based on the ratio of the gasoline estimates from Bohi (1981) and Dahl (1993), and replace the USGS resource estimate of 3.003 trillion barrels with the average prediction from the 1970s of 2.283 trillion barrels as reported in Horn (2007).

The price forecasts from 1980 with this adjusted model are shown in Figure 2.16. The estimated price in 1990 increases slightly, but still remains 17% lower than the actual price. Looking farther forward, the model performs somewhat less well, predicting that prices would rise again sooner than actually occurred.

Figure 2.16. Ex-post forecast from 1980
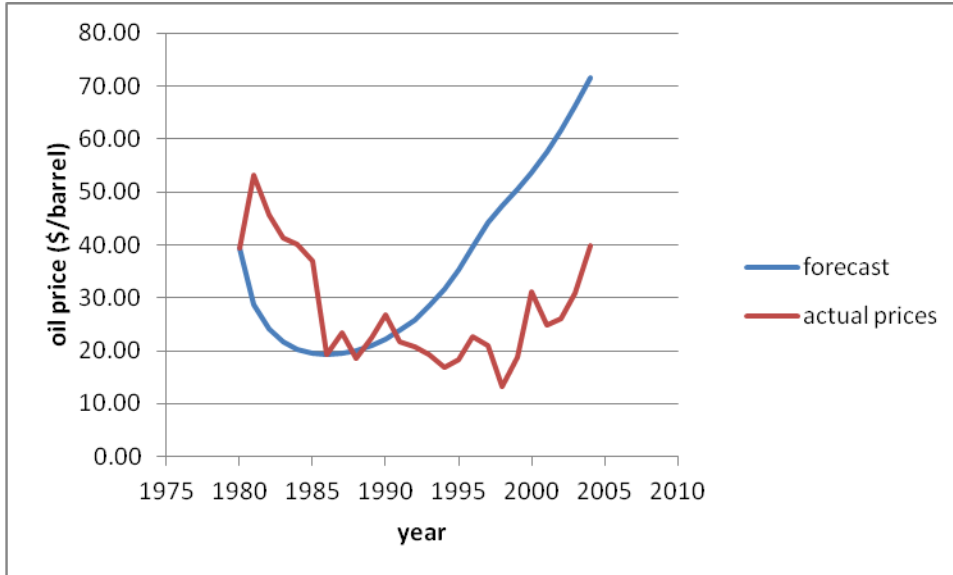
A final check on the forecasting model is to look at how it has performed at predicting prices in the years since 2004. Extending the data through 2010, we see that the predictions for recent years from the forecasting model using all data through 2004 have been fairly accurate (see Figure 2.17). The price spike in 2008 was higher than anticipated, but prices in 2009 and 2010 returned to levels very near the model forecast.

Figure 2.17. Extended price data

57
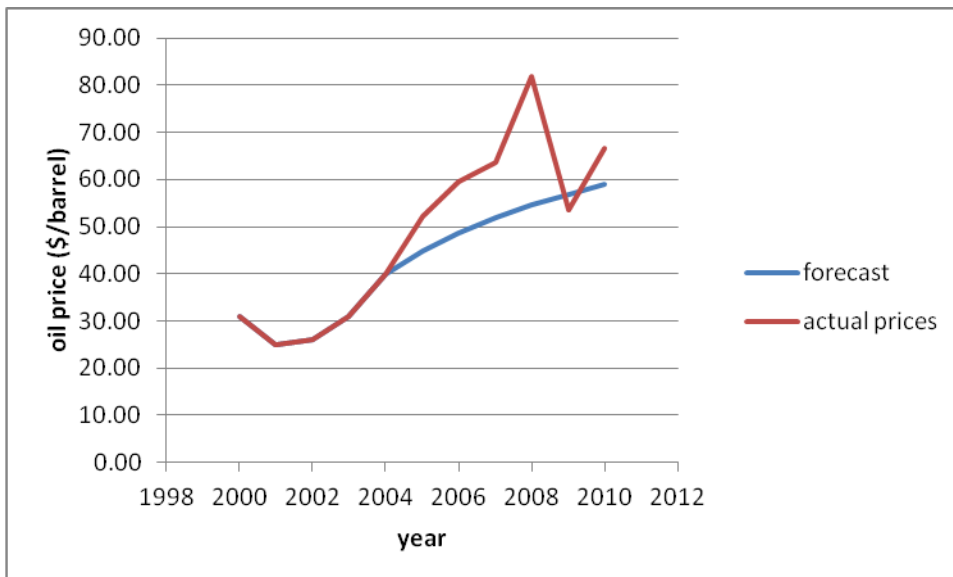
2.4 Conclusion

This paper presents a straightforward way of combining an analysis of actual data with simple theoretical considerations to produce a reasonable forecast of future prices and quantities. It combines curve-fitting techniques with some basic theory to ensure that the model being fitted is consistent with economic and geological constraints. Its performance, when tested through ex post forecasting, compares favorably to forecasting attempts from other studies.

The results provide some evidence that oil depletion is a legitimate concern. Under the assumptions of the model, prices rise gradually in the short run, but continue to rise long into the future, eventually reaching extreme levels. The price increases are reduced, but not eliminated, if a more optimistic assumption about resource levels is used. The results do not support warnings by some in the 'peak oil' camp about extreme hardships in the relatively near term, but they do point to oil scarcity eventually being a significant problem if demand is not reduced beyond levels suggested by past trends.

While any forecasts made this far into the future are tenuous at best, and heavily dependent on the functional forms used in the model, it is still useful to evaluate competing claims about resource scarcity with a model that captures current trends as well as possible. Considering the vast differences in beliefs that exist about resource scarcity and its impacts, and the societal importance of these questions, it is surprising that more rigorous modeling has not been done to attempt to evaluate these claims. By combining trend-fitting models of supply and demand with information from studies on resource availability and responses to price changes, this paper provides an important contribution to this literature.

There is also considerable room for improvement in this model. It provides a framework for modeling oil markets that can be built on in a number of different ways. The supply and demand functions can be made more realistic by introducing more complexity. More variables could be added, consumers and producers could be disaggregated into different groups with different supply and demand functions, and the functional forms could be refined further to capture actual behavior more accurately. The effectiveness of this model demonstrates that even with a relatively simple model, some good results can be produced.

The supply and demand functions from this model can also be combined with a more realistic price setting rule to analyze some even more interesting questions. In chapter 3, I look at how prices are set more closely, considering the role of futures markets and how the behavior of investors in futures markets can affect the path of oil prices. In a correctly functioning market, standard economic theory predicts that futures markets could help to smooth out price series by allowing investors to anticipate future shortages or gluts. On the other hand, a popular theory has emerged that speculators have created bubbles in futures markets, leading prices to diverge from fundamental values. Incorporating the role of futures markets and above-ground storage into the model can help evaluate each of these theories, and provide some insight into the conditions under which each of these theories might hold. The supply and demand model described in this paper provides a good basis that can be used in that analysis.

CHAPTER 3

A MODEL OF INVESTOR BEHAVIOR IN CRUDE OIL FUTURES MARKETS

3.1 Introduction

Chapter 2 laid out a model of producer and consumer behavior in oil markets, and how they adjust over time. One shortcoming of this model is that there is no mechanism for market participants to anticipate future developments in the market, and adjust their current behavior accordingly. This chapter addresses this shortcoming by introducing oil futures markets in which participants try to predict how oil prices will move. These predictions then influence real markets by determining when holders of crude oil choose to buy and sell their oil.

3.1.1 Anticipating future shortages

The model in chapter 2 suggested that if prices are set to equilibriate current demand and supply, with no anticipation of future developments, oil shortages are likely to occur at some point. Simulations based on the model show temporary shortages at times throughout the model run, in the form of sharp price spikes, with the highest price spikes occurring late in the model as the resource approaches exhaustion.

If participants in oil markets are able to anticipate these shortages, sudden price spikes like these should not occur. Anyone who can correctly anticipate that a rise in prices is likely can buy oil before the price spike, store it, and sell it when the price goes up. If enough people do this, it should drive up prices before the price spike, and keep them from ever getting as high as they would have otherwise.

James Hamilton, a prominent economist who has written about oil markets, has posted frequently to the EconBrowser blog about oil market issues from an economist's perspective. In one post "How to talk to an economist about peak oil" (July 11, 2005), he outlines the economic argument as to why a predictable oil shortage should be able to be anticipated by the market in advance, raising current prices. He explores a hypothetical scenario in which an anticipated 30% drop in oil production would cause prices to rise from $60 a barrel to $200 a barrel in two years.

> Anybody who pumps a barrel out of a reservoir today to sell at $60 could make three times as much money if they just left it in the ground another two years before pumping it out. The same is true for anybody with above-ground storage facilities—they're throwing way money, and lots of it, for every barrel they sell at $60 that they could have instead stored for two years and sold at $200. If oil producers did respond to these very strong incentives by holding back oil from today's market, the effect would be to drive today's price up.

He then goes on to explain the effects that an immediate rise in prices to $180 a barrel would have:

> For one thing, it would be a very powerful incentive to force today's users of oil to reduce their consumption immediately. It would likewise be a very powerful incentive for investing heavily in oil sands and alternative technologies. And of course, it would leave us more oil in the future to keep the economy going as we make the needed transitions. In other words, the consequences of oil producers trying to sell their oil for the highest price would be to help move society immediately and powerfully in the direction that we earlier determined it ought to move in anticipation of what is going to happen in the future.

In other words, if future oil shortages are anticipated properly, the profit motives of participants in oil markets will be sufficient to generate socially desirable outcomes, with smooth price paths and as graceful a transition as possible. This argument helps to explain why many economists are not too concerned about resource depletion, and feel that, in the absence of other market imperfections, resource scarcity on its own is not a reason for government intervention into oil markets.

3.1.1.1 Questions about ability of market to anticipate future shortage

A number of observers of oil markets are skeptical that economic factors could do much to alleviate oil shortages. Geologists in the 'peak oil' camp, such as Colin Campbell, argue that the geology of oil reservoirs, combined with the difficulty of finding substitutes for oil, will trump any economic or political factors (Campbell, 2002).

Some more substantive arguments why markets may not respond sufficiently to impending shortages were made by anonymous posters in response to the blog posting by James Hamilton in EconBrowser:

> "Markets thrive on uncertainty, but they function best when there is quantifiable uncertainty -- that is, when probabilities can be estimated, based on past experience. When a market enters completely new territory, it's difficult to make intelligent bets. Right now, the players are still mostly betting that things will continue as they are at least for several years. Put yourself in their place. Wouldn't you have a hard time facing your shareholders if you didn't lay down the majority of your chips down on the square labeled "status quo"? That's called fear overwhelming greed, and it happens all the time." (posted by Ralph on July 11, 2005)

> "The result [of lack of information] is that the markets become nervous and unstable, easily swayed by rumors, everyone looking at everyone else to try to get a sense which way things will break. Prices often become metastable, sticking to one trading range for no particular reason and then suddenly switching to a new price range, seemingly on a whim." (posted by Hal on July 11, 2005)

> "Markets are just a community of humans. Humans are very emotional in their decision-making. They cannot really credit that the future will be sharply different than the past until they have experienced the emotional consequences of a new regime. This is especially true of humans in groups, whose individual judgement gets subjugated to that of the herd, until the situation become so obviously untenable that everyone starts to change (as in the sudden changes of prices across a bubble-crash sequence). We create fire brigades only after major fires, earthquake codes only after major earthquakes, departments of homeland security only after 9/11, and we will figure out how to price post-peak oil only after the peak." (posted by Stuart Staniford on July 11, 2005)

These comments raise important questions about how market participants form their expectations as they predict what will happen with future oil markets. It may

difficult for many players in oil markets to make purely rational calculations in their predictions. However, these arguments also raise questions about why sub-optimal investment strategies would persist over time. Would the participants who did a better job of predicting market developments be rewarded financially, and therefore be copied by other investors, moving markets toward the optimal outcome? More research is needed to evaluate these questions.

3.1.2 Futures markets

Oil futures markets play an important role in forming expectations as to what will happen to oil prices over the next several years. Futures markets for crude oil are markets that allow investors to exchange contracts to buy and sell crude oil meeting certain specifications, at specified location and date. Futures markets are set up to allow investors of all sorts to participate in oil markets, without having to physically engage in storing and transporting oil. Anyone with information that could affect oil prices movements can bet on oil prices rising or falling by taking a long position (entering contracts to buy oil) or short position (entering contracts to sell oil) in futures markets. They can then zero out their position before the contract comes due so that they can profit or lose money without ever having to physically engage in buying or selling crude oil. The result is that the prices of crude oil futures at different time periods represent the market's aggregate 'best guess' as to how oil prices are likely to move over the next several years.

In US media reports on the movement of oil prices, the price most commonly reported is the price of the nearest term oil futures contract traded on the New York Mercantile Exchange (NYMEX). This represents the price of futures contracts for light

sweet crude oil in Cushing, Oklahoma, on the last business day before the 25th of the coming month. The actual spot price paid in Cushing for immediate delivery of crude oil usually does not vary much from these near term futures prices. There are also longer term futures markets for contracts to buy and sell crude oil farther into the future. There are futures markets that come due every month up to 30 months, and every year up to 7 years from the current time.

Many factors are reported in news reports to cause changes in oil prices: supply disruptions from wars and other political factors, from hurricanes or from other random events; changes in demand expectations based on economic reports, financial market events or even weather in areas where heating oil is used; changes in the value of the dollar; reports of inventory levels, etc. These are all factors that will affect the supply and demand for oil, but they often influence the price of oil before they have any direct impact on the current supply or demand for crude oil. The information affects oil prices immediately because investors in futures market use this information to change their assessment of what oil prices should be and how they are likely to change. Futures prices then affect spot prices because if there are significant differences between spot and futures prices, anyone who has flexibility as to when they can buy or sell oil supplies will choose to buy when prices are lowest and sell them when prices are highest. If spot prices are higher than futures prices, they will sell oil now, driving spot prices down closer to futures prices. If spot prices are lower, they will buy oil now at spot prices, and wait to sell it at the higher prices of futures contracts.

3.1.2.1 Role of futures markets in aiding price prediction

Properly functioning futures markets should help market participants to anticipate future developments in oil markets more easily by allowing more people to use diverse sources of information to predict future price movements, allowing more information to be brought into the price discovery process. The price discovery role of futures markets is seen as an important economic benefit of futures markets (US CTFC, 2011).

While allowing more people to participate in the process of price formation should, in principle, make price predictions more accurate, there is also a risk that allowing too many people to participate in future markets could lead markets to be dominated by psychological factors that could lead to bubbles and busts that are unrelated to market fundamentals. This idea has gained attention recently as several observers have argued that the participation of institutional investors in futures markets led oil prices to diverge from their fundamental value during the rise in oil prices that peaked in 2008. The sharp price spike and the sudden drop in prices that followed certainly had the appearance of a speculative bubble bursting. A report by a US Senate subcommittee (US Senate, 2006) identified excessive speculation on future price increases as a key reason for the rise in prices that had already occurred up to that point. Between 2003 and 2008, there was a large increase in the allocation of investment funds into commodity index trading strategies that took long positions in futures markets for a number of different commodities, including oil (Masters, 2008). Once commodity prices started falling in 2008, many fund managers dropped those positions (Masters and White, 2009). The claim is that this led to a speculative boom and bust, with prices diverging from what

fundamentals of supply and demand would suggest that they should be (Masters and White, 2009).

A number of authors have attempted to address the question of whether the price movements of 2008 were driven by market fundamentals, or were a speculative bubble. Phillips and Yu (2010) use statistical modeling of price patterns to find evidence of a speculative bubble during the price spike in 2008. Lagi et al (2011) use a behavioral model to address the question of whether the spike in grain prices was driven by speculation or by market fundamentals. They break down price changes into those driven by fundamentals and those driven by speculation using a model of trend following with reversion to mean that leads to speculative oscillations, along with fundamental factors such as demand growth and ethanol production. They find that most of the price movements of 2008 fit better with the trend following speculative model then they do with the fundamentals-based model.

Other authors have argued against the speculative nature of the price movements on the basis that inventories did not adjust as would be expected in the presence of a speculative bubble. Hamilton (2009a) and Krugman (2008) argue that if speculation caused prices to diverge much from fundamentals, demand and supply would no longer be in equilibrium, leading to changes in inventories that were not observed. In other work, Hamilton also finds that the changes in oil prices between 2005 and 2011 were fairly consistent with what might have been expected due to stagnation of oil production, using reasonable elasticity estimates (Hamilton, 2011). Ederington et al (2011) provide a good overview of research on this question.

3.1.3 Modeling investor behavior in futures markets

The discussion in the previous section suggests two very different views of the role of futures markets. They can reduce volatility of oil prices by allowing market participants to draw on the full knowledge base of futures market in predicting future price movements, allowing markets to anticipate shortages better and smoothing prices over time. Or they can increase volatility by allowing more uninformed speculators to get involved, generating bubbles that lead to price spikes that would not have otherwise happened. These divergent outcomes are driven by different views about investor behavior.

While the authors in the last section focused on evaluating the causes of recent price movements, the goal for this paper is to explore what could happen to oil prices if oil resources become more scarce and approach exhaustion. The behavior of investors also has significant implications for this question, as it determines how well the market anticipates the scarcity so it can make adjustments in advance.

It is difficult to answer this question using an empirical analysis based on historical data. There is no real historical precedent for the exhaustion of a resource such as oil. Many historical examples of resource exhaustion are for renewable resources such as forests or fisheries. No non-renewable resources have been exhausted that play nearly as important a role in the economy as oil does.

Analyses of oil markets up to this point also do not give a great idea as to what would happen if oil ever approached exhaustion. While there is certainly much to be learned from analyzing past price movements, there is no guarantee that as a resource moves closer to being exhausted the patterns observed up to that point will persist.

Hamilton (2009b) finds that scarcity rent contributed little to oil price in 1997, but could now play more of a role. The short-run concerns of developing supplies fast enough to keep up with rapidly rising demand have dominated price setting in oil markets through most of their history, but this could be less true moving forward.

Since there is little relevant data to use as a basis for empirical analysis, this paper focuses on developing a theoretical model of oil markets that can be used to explore potential scenarios as to how prices might evolve with different assumptions about investor behavior. The model is set up so that attempts to anticipate future shortages could help lead to smoother price paths, while also allowing for the possibility that investors are not perfectly rational in their predictions of future prices.

This modeling exercise will not provide any definitive answers about how investors will behave or what will happen to oil prices, but it should help to clarify our thinking about the problem. Patterns observed from running model simulations should help to highlight important issues that need to be considered when developing approaches to deal with possible resource scarcity. It could also provide a basis for doing more empirical analysis in the future, using the model as a basis to make comparisons between historical data and model outputs under different assumptions to evaluate what scenario we might currently be in.

There have been no modeling efforts to date, that I am aware of, that incorporate the ability of market participants to anticipate future shortages in setting oil prices, while also allowing for the possibility that human investors may not properly anticipate shortages, or even create periods of speculative booms and busts. There is a large literature of exhaustible resource modeling in economics in which the anticipation of

68

future scarcity plays a major role. However, these models are entirely based on rational actors who, by assumption, cannot behave in ways that would lead to a speculative bubble. There is also a large literature that looks at behavioral models of different financial markets, with investors following boundedly rational behavior such as adaptive experimentation or herding. However, the few models that look at oil futures markets do not combine the behavioral models with a model of oil markets that takes into account the exhaustible nature of the resource.

3.1.3.1 Rational actor models

The classic model of non-renewable resource price setting in economics is the Hotelling model (Hotelling, 1931). In this model, applied to oil, owners of oil reserves decide when to extract the oil and sell it to consumers in order to maximize the present value of their profit, given the changing price of oil over time. The conclusion is that oil prices must satisfy an equilibrium condition known as the Hotelling rule: the *in situ* value of the oil (the price minus the extraction cost) must rise at the rate of interest. The extraction path over time in this model matches the optimal path to maximize discounted societal benefits.

A key assumption of this model is that the actors are able to properly anticipate what prices will be. This assumption is a version of the rational expectations assumption commonly made in economics. The same assumption is also the basis for a whole set of more sophisticated models that have been developed based on the idea of the Hotelling model but with less restrictive assumptions. While these models add uncertainty about resource discoveries, changes in demand and extraction cost, they all use the rational

expectations assumption that market participants are able to make the best calculation possible of the path prices will follow, given the information available at the time.

However, this assumption requires a very strong notion of rationality that is not likely to exist in the real world. In order for actors to correctly anticipate what prices will be, they must not only being able to calculate correctly the expected value of many future variables that will affect demand and supply, but they must also assume that all other actors will make the exact same calculation and invest on the same basis. If, however, other investors act differently, prices will not follow the Hotelling-optimal path, so the choices made will not be optimal. With heterogeneous agents, the assumption that all actors will follow the logic needed to reach the rational expectations equilibrium seems unreasonable (Arthur *et al.*, 1996, pp. 4-6).

In such situations – where there is little rational basis for behavior without knowing how others will act – an alternative approach is to use evolutionary models in which investors try out different investment rules, compare their effectiveness with other strategies, and stick with those rules that prove to be most effective. This approach has been used to model investor behavior in stock markets and other asset markets in the behavioral finance literature.

3.1.3.2 Behavioral models of stock markets

Behavioral models of stock markets have grown out of concern that the efficient markets hypothesis (EMH) advanced by Fama and others (e.g. Fama, 1965; Fama, 1970; Fama, 1997; Rubenstein, 2001) does not explain some market phenomena.

It is hard to justify many of the historically large price swings that have occurred entirely based on fundamentals. One particular instance was the 1987 stock market crash

in the US, where the Dow Jones Industrial Average dropped by a record 22.6% in one

day, with no significant news to justify such a large price change (Shleifer, 2000). Shiller

(1981) also demonstrates that stock prices are considerably more volatile than can be

explained based on variation of market fundamentals. Another challenge to the EMH is

that many studies have found evidence of market predictability. Low market-to-book

ratios (De Bondt and Thaler, 1987) low price-to-earnings ratios (Shiller, 2005) and an

extended history of bad returns (De Bondt and Thaler, 1985) have all been shown to

predict higher future returns in the long run. On the other hand, Jegadeesh and Titman

(1993) show that a shorter history of price losses can predict that those losses will

continue in the future (Shleifer, 2000). Some additional puzzles that are difficult to

explain based on the EMH are surveyed in LeBaron (2006). The volume of trading is

much higher that would be expected with purely rational agents, and high volumes tend

to persist for many periods. The same applies to market volatility as well: in addition to

the surprisingly high levels of volatility found by Shiller, there are swings between

periods of high volatility and periods of more stability.

A number of authors have put forward models that show how boundedly rational

behavior can lead to market outcomes that differ from the rational markets equilibrium.

Grossman and Stiglitz (1980) considered the evolutionary outcome if there is a cost to

obtaining accurate information about fundamentals. They find that in this situation, the

equilibrium prices can't perfectly reflect fundamental values. DeLong, Shleifer,

Summers and Waldmann (1990a) developed a model that showed that even if

information is not costly, the presence of some noise traders who make arbitrary trading

decisions could keep prices from reaching their fundamental value. In a model with

noise traders and rational investors, the rational investors have to be aware that noise traders could cause prices to diverge further from their fundamental value, which adds a level of risk to their investment decisions. As a result, rational investors bring prices only part way toward their fundamental values.

Some models have attempted to provide an explanation for specific empirical findings that rational market models can't explain. Barberis *et al.* (1998) and Daniel *et al.* (1997) present two alternative behavioral models to explain why the market tends to under-react in the short run to a single bit of new information, but overreacts to longer term, repeated information. Lux and Marchesi (2000) reproduce higher volatility than would be found in the efficient market hypothesis, as well as clustered volatility, with switches between more and less volatile periods, using an evolutionary model with three types of agents: fundamental traders, optimistic chartists and pessimistic chartists.

More complicated simulation models have been developed based on agents selecting strategies from a wide variety of different types of rules. A classic example of this type of model is the Santa Fe Artificial Stock Market (*Arthur et al.*, 1996), in which investors choose from a set of available rules that adapts over time following a genetic algorithm, with rules that are used successfully being more likely to survive. These models have also replicated some features of real world markets that are not explained by rational market theory, such as high volatility and volatility clustering. This approach comes closest to the type of modeling that is applied to oil markets in this paper.

3.1.3.3 Behavioral models of oil markets?

There have been very few behavioral models that have focused on oil markets. Spyrou (2006) looks at evidence of overreaction or underreaction in futures markets, with

mixed results, but does not include a model of why this may be happening. Ellen and Zwinkels (2010) introduce a behavioral model of oil futures markets, with investors who choose between fundamentalist and chartist rules for predicting oil price movements. Both rules focus on past price patterns to predict future movements, with fundamentalists assuming that prices will revert to past mean levels, while chartists follow trends in prices. There is no mechanism for investors look forward at the possibility of future shortages and take that information into account in their investment decisions.

None of these papers provides much direction as to how to design a behavioral model that will help to address the question of how predictions about future price movements can affect the long run trajectory of oil prices in the face of future scarcity. The model in this paper draws on the literature for parts of the model. It uses ideas from behavioral economics for the process by which investors select their price prediction rules from the set of possible rules. It also draws on rational actor models of oil markets for the relationship between above-ground storage, spot and futures prices, and to aid in constructing the set of possible price prediction rules. The types of behavioral differences assigned to investors in the model, however, are not taken from the academic literature, but are based on my own evaluation of what are the most important behavioral differences for the understanding of oil markets and how well investors might be able to anticipate future scarcity.

The two key behavioral differences that are highlighted in the model are the beliefs that players have about the abundance of the resource – specifically the rate of discovery of new reserves – and how far investors look into the future in making their calculations. How much oil there is in the ground is a hotly contested question. Some

observers have warned that we are near the point of using half of the oil resources that were initially in the ground and recoverable before we started consuming oil, based on estimates of ultimately recoverable reserves (URR) of around 2 trillion barrels (Campbell and Laherrere, 1998; Heinberg, 2003). A more optimistic view has come from the USGS, whose estimate for URR of 3 trillion barrels we used in the last chapter (USGS, 2000). Others observers have argued that even the USGS estimate is too low, because new technological developments are likely to make oil available that is not currently feasible to recover (Lynch, 2002). The answer to this question has important implications for the long-run trajectory of oil prices.

The other question is how far into the future market players look to determine what prices should be. There is plenty of anecdotal evidence that most market players focus on short to medium-term developments. News stories about oil price movements focus on developments that are likely to affect oil markets in the next few months, and rarely mention new discoveries that won't come to market for years to come. Another indication of how far into the future oil market participants tend to think is that crude oil futures traded on NYMEX only go up to seven years into the future, with the near-term markets generating by far the most activity. Even the US Energy Information Association's longest projections for oil market developments only look about 25 years into the future (EIA, 2011). Yet in the Hotelling model, actors may have to look hundreds of years into the future, or as long as needed for the resource to be exhausted, to determine what the current market price should be. How far investors look into the future could have significant implication for their calculation of what prices should be.

We focus on these two factors because of their prominence in discussions about oil scarcity, and their importance for the ability of market players to correctly anticipate future scarcity. There are other factors that may also play an important role that I do not address in this model. Most notably, differences in beliefs about the availability of alternative technologies that can substitute for oil, and about how much these technologies will progress, could also play a key role in determining whether oil supplies will become scarce in the long term. This could be an important area for future research. However, I believe that some of the dynamics associated with differences in these beliefs would be similar to the dynamics of this model based on differences in beliefs about resource abundance, since both differences primarily affect the prospects for long run oil scarcity.

3.1.4 Overview of remaining sections

Section two describes how the model of supply, demand, storage, and futures market investment is constructed. This model is implemented in java, using the Repast java library for constructing the agent-based model. Section three provides an overview of results from model simulations, and section four concludes with thoughts about the contributions of the model and areas for further work.

3.2 The model

The model used in this chapter builds on the demand and supply model from Chapter 2. A behavioral model of speculation in oil futures markets is added in which investors try to predict the price of oil, and invest in futures markets based on their prediction. Futures prices are related to spot prices through the addition of above-ground

storage markets. In addition, the supply model is adjusted to add some uncertainty about resource levels through an exploration and discovery process.

3.2.1 Roles

The key roles played by participants in the model are: demand, supply, storage, and speculation.

In the Hotelling model and many others based on it, suppliers of oil predict future prices and make decisions about how much to produce and how much to store for later based on that prediction. The model in this paper divides that process into three steps: futures market investors make price predictions and choose to buy or sell futures on the basis of their prediction – this determines futures prices. Holders of above-ground storage then adjust their storage levels based on the levels of futures prices relative to spot prices. Producers then use spot prices in choosing how much to supply in this period. The decisions of speculators, storage holders, producers and consumers are modeled separately.

Although the roles are modeled separately, the same people or companies could be playing more than one role. An oil producer can also invest in futures market and hold above-ground stocks of oil. However their decisions as to what to do in each case can be decided separately.

Under reasonable assumptions, a utility maximizing producer, consumer or holder of oil can use the price of futures as the basis for their timing decisions. If their own prediction about how oil prices should change differs from the futures price, they can best take advantage of this by investing in futures markets while still basing their supply and

demand timing decisions on the price of futures. This is demonstrated to be true for a

fairly general form of optimization problem in Appendix B.

3.2.1.1 Demand

The demand model we use in this model is the same as in the model from chapter

2, and is specified by equations (3-1):

$$QD_t = D1_t \cdot p_t^{SRE_D}$$

$$\log(D1_t) = b_{0D} + b_{1D} \cdot t + b_{2D} \cdot t^2 + LRA_D \cdot pMA_{t-1} + e_t$$

$$e_t = \theta_D \cdot e_{t-1} + u_t \tag{3-1}$$

$$pMA_{t-1} = (1 - \delta)(\sum_{i=1}^{\infty} \delta^{i-1} \cdot \log(pEff_{t-i}))$$

$$pEff_t = (1 - \phi)\left(\sum_{k=0}^{\infty} \phi^k \cdot \max\{p_t, ..., p_{t-k}\}\right)$$

3.2.1.2 Supply

3.2.1.2.1 Supply from previous essay

The supply model we use is similar to that in equations (2-14) from the model in

chapter 2.

$$QS_t = (1 - e^{-Ap_t}) \cdot C_t$$

$$C_t = \frac{1}{(1 + e^{-S2_t})} \cdot c \cdot R_t$$

$$R_0 = \overline{R_0}$$

$$R_t = R_{t-1} - QS_{t-1}$$

$$S2_t = b_{0D} + b_{1D} \cdot t + b_{2D} \cdot t^2 + LRA_S \cdot pMA_{t-1} + e_t \tag{3-2}$$

$$e_t = \theta_S \cdot e_{t-1} + u_t$$

$$pMA_{t-1} = (1 - \delta)(\sum_{i=1}^{\infty} \delta^{i-1} \cdot \log(pEff_{t-i}))$$

$$pEff_t = (1 - \phi)\left(\sum_{k=0}^{\infty} \phi^k \cdot \max\{ p_t, ..., p_{t-k} \} \right)$$

The parameter values for both the supply and demand equations taken from chapter 2 are also the same as those used in chapter 2. It should be noted that since these were calculated by fitting the chapter 2 model to data, the expanded model in this chapter will not be properly fit to data. While the supply and demand models used in this paper are nearly the same as in chapter 2, the ability of investors to anticipate future developments in oil markets may help to smooth price paths, leading to less price variation in this model than in the model from chapter 2. Since the demand and supply parameters were chosen so that the chapter 2 model would replicate the amount of price variation in the real-world data, there may not be enough random variation in demand and supply to produce the correct historical levels of price volatility once the anticipation of future movements is taken into account. Therefore these model runs should not be seen as fitted projections of how prices are likely to behave, but as an illustration of how changing particular behavioral assumptions affect the results, under a reasonably realistic, but not perfect, demand and supply scenario.

3.2.1.2.2 Uncertainty about resources

Since uncertainty about resource abundance is a key component of the behavioral model, we expand the supply model from chapter 2 to allow for some uncertainty. To

implement uncertain resource levels, we distinguish between proven reserves, which have been discovered and are available, and ultimately recoverable resources, the level of which is unknown. Exploration is required to turn resources into reserves. Production capacity is limited by proven reserves, so we replace the production capacity equation

$$C_t = \frac{1}{(1 + e^{-S2_t})} \cdot c \cdot R_t \text{ with } C_t = \frac{1}{(1 + e^{-S2_t})} \cdot P_t, \text{ where } P_t \text{ represents proven reserves. To}$$

make the model as similar as possible to the one used in chapter 2, enough exploration is done in each period to bring proven reserves at least as high as $c \cdot E[R_t]$. The amount of exploration needed to reach this level depends on how successful the exploration is.

Our model of the distribution of deposits follows Lasserre (1984). Exploration produces deposits according to a Poisson process with rate parameter 1 (we can set the rate parameter as desired by adjusting the units of land). There is a fixed total amount of land to be explored, all of which is equally likely to produce deposits, and all deposits are of equal size, r. The total number of deposits follows a Poisson distribution:

$$p(D = n) = \frac{(\gamma)^n e^{\gamma}}{n!} \tag{3-3}$$

The parameter γ is the expected number of deposits, which is equal to the number of units of land, X_0, available for exploration. This is set to be 500. The number of deposits affects the amount of uncertainty that can exist as to the total resource levels. A small number of large deposits leads to more variation in the distribution of total resources than a large number of small deposits. The size of each deposit, r, is set to be one 500th of 3003 billion barrels, so that the expected level of ultimately recoverable resources is equal to the starting resource level from Chapter 2.

The amount of exploration needed to produce a deposit is a random variable E_λ that follows an exponential distribution, with density function:

$$d(x) = \begin{cases} e^{-x} & \text{if } x >= 0 \\ 0 \text{ if } x < 0 \end{cases}$$ (3-4)

In each period exploration continues until the level of proven reserves is at least as large as c times the expected value of ultimately recoverable resources. If a total of n deposits are needed to reach this level, the amount of exploration it will take to reach this is the sum of n independent exponential distributions, or the gamma distribution $\Gamma(n,1)$. Some additional exploration may occur without a discovery if the drop in expected resources from the lack of discovery makes the condition be satisfied before a discovery occurs.

3.2.1.3 Storage

3.2.1.3.1 Price smoothing

An important part of our model is the mechanism by which futures price affect spot prices. As discussed in the opening section, if investors anticipate a future rise in prices, this should help drive current prices up.

There are two mechanisms by which this could happen. One is that suppliers of oil would delay pumping oil out of the ground, and instead wait until the price rises to produce the oil. The other is that anyone with storage facilities could purchase oil at the lower price, hold onto the oil, and sell at the higher price.

In the Hotelling model, it is suppliers who adjust their production patterns based on the relative value of current and future prices. Producers have a fixed reservoir of oil

which they can extract at any time. They choose when to extract it based on the relative prices at different times.

This model is not very realistic however – in the real world, oil suppliers do not have so much flexibility as to when they pump their oil. They have a limited production capacity based on the wells they have drilled, and draining a reservoir too quickly can damage the reservoir, reducing the amount of oil that can ultimately be recovered (Banks, 2000). In addition, if producers adjusted their extraction levels significantly over time in response to oil prices, they would on average be producing well below capacity, which would not be the most efficient use of their capacity investments. As a result most wells, once drilled, operate at full capacity most of the time, regardless of the current price of oil.

Storage facilities can be much more flexible about when they move their oil to the market, and therefore respond most directly to the relative value of current and future prices. A number of economic models of exhaustible resources use a price smoothing rule based on equilibrium in the above ground storage market, rather than supplier decisions. In the industry standard oil pricing model, the difference between spot and futures prices depends on the equilibrium in the above ground storage market (French, 2005, p. 2, 3-4).

The flexibility of above-ground storage makes it the logical way of exploiting a short-term differential between spot and futures prices. However if a longer-term trend of increasing prices is expected, then putting off developing new fields until prices are higher may be a better strategy, since the amount of storage necessary to offset a long-run trend like this would become very expensive.

However, finding the optimal choice of investment in capacity and capacity utilization in the face of an anticipated path of future prices would be a very difficult optimization problem, and would be even more difficult to work into the model in this paper. I therefore focus on above ground storage for all price smoothing purposes. To allow long-run price smoothing, I keep the cost of above ground storage low enough that huge build-ups of above ground storage are possible when necessary. This approach can lead to levels of above ground storage that would be unlikely to occur in the real world, but it provides an alternative to allowing suppliers to hold resources underground as a method of long-term price smoothing.

3.2.1.3.2 Storage equilibrium

Stocks of oil, or oil that is held in storage above ground, are useful for several reasons. They may allow suppliers to provide a smooth supply of oil to customers (French 2005, p. 1, refs to Kaldor 1939 and Working 1948). Large consumers may also hold reserves to ensure that they have oil available when they need it. There is also an option value of holding resources, as it gives you the option of exploiting price movements (French 2005, ref to Dixit and Pindyck 1994). For these reasons, there is a 'convenience yield' to holding reserves. There is also a cost of physical storage of oil.

Holders of oil stocks can also command a profit or loss on their storage depending on the relative levels of spot and futures prices. We can use current futures prices instead of the actual price in the next period because owners of storage facilities can minimize their risks by buying or selling futures to lock in a price they will buy/sell their oil at, as shown in section 3.2.1. However, futures prices must be discounted, since they will not be received until the oil is sold at the end of the period. We use a real interest rate, r, of

2%, which may seem low by historical standards, but is not too low for an inflation adjusted risk free interest rate. The investment should be risk free since futures markets offset any risk from unforeseen price movements during the period.

Owners of oil storage will choose to hold enough oil so that the marginal convenience yield from holding an additional barrel of oil minus the marginal storage costs of storing the oil equals the loss of value from holding oil, as measured by the difference between spot prices and discounted futures prices.

The convenience yield and storage cost functions are given below:

$$CY(x) = a \cdot D \cdot \ln(x) \tag{3-5}$$

$$C(x) = c \cdot x \tag{3-6}$$

D is the demand for oil, and a and c are constants. The shapes of these curves are generally consistent with those suggested by French (2005).

Setting marginal benefits equal to marginal costs, we have:

$$\frac{a \cdot D}{x} - c = p - \frac{f}{1+r} \tag{3-7}$$

$$St(p, f) = x = \frac{a \cdot D}{p - \dfrac{f}{1+r} + c} \tag{3-8}$$

If there is no gain or loss of value from holding oil, then the optimal level of stocks will be equal to $\dfrac{a \cdot D}{c}$. I set set a/c = 0.2, so that the optimal stock level is 20% of demand – or that there are 73 days of forward cover. This is slightly higher than is common in the current market – forward cover has ranged from 51 to 61 days from 2006-2010 (IEA, 2011). C is set to be 2.0, which means the cost of storing oil is $2/barrel per year. True costs vary depending on how the oil is being stored. When storage needs are

high, companies can store oil in tankers at sea at costs of \$0.70 – \$1.00/barrel per month according to one estimate, or \$8.40-\$12.00/barrel per year (Saul and Johnson, 2010). However, storage costs could be reduced if there is a long-term need for large amounts of storage. Since we are using above-ground storage as a substitute for withholding production as a way to anticipate long-term shortages, we need to keep storage costs low enough to make it reasonable for there to be large amounts of above-ground storage held for long time periods.

3.2.1.4 Speculation

Speculators determine prices in futures markets. All speculators follow the same general approach to choosing what position to take. Each year, they make an evaluation of what prices should be based on a set of fundamental information and a belief about how prices are set. They also use information from the positions other investors take – if their own evaluation is that the price should be p, and the market value for futures prices is f, they assume that the price will actually be $0.1 \cdot p + 0.9 \cdot f$. This adjustment is made based on the belief that whatever led everyone else to predict prices differently from their prediction may still be true in a year. Since everyone is doing this, it doesn't affect the value of futures prices, but it does affect how predictions are evaluated. The most accurate predictions are more likely to be copied by others, and without this adjustment, predictions that are in the right direction but differ significantly from the market average would be unlikely to be rewarded.

If their final prediction is higher than the market price of futures, they take a long position in futures markets; if it is lower, they take a short position. The amount they take is chosen to maximize their expected utility, using the constant absolute risk

aversion utility function $u(x) = -e^{-\alpha \cdot x}$, and the belief that prices could take on a value

that will be drawn from a normal distribution with variance σ^2 centered around their

price prediction, $E[p_{t+1}]$. The result is that the position they would take in the futures

market is $c = \dfrac{2(E[p_{t+1}] - p_f)}{\alpha \cdot \sigma^2}$. (Arthur *et al.*, 1996). If c is negative, this means a short

position, if it is positive this means a long position.

Two things affect how strong a position investors take: the rate of risk aversion,

α, which may depend on the amount of money available, and the confidence that the

person has in their estimate for p_{t+1}, as measured by the variance of the distribution, σ^2.

I assume for now that all investors have the same amount of risk aversion, and the same

confidence in their estimate, so that the predictions of all investors are given equal weight

in determining the price of futures. A promising area for future research would be to

explore how variation in these factors could affect the results.

3.2.2 Equilibria

Equilibrium spot prices and futures prices are set simultaneously each period.

Futures market equilibria are set based on the price predictions of investors, and spot

market equilibria are set to equilibrate demand and supply from producers, consumers

and holders of above-ground storage, taking futures prices into account.

3.2.2.1 Futures markets

The futures market we focus on is for crude oil for delivery one year into the

future. One year is the shortest time period we can look at because we only update the

model once a year. Real-world markets of course are updated much more regularly, but

yearly updating should be enough for our purposes of looking at long run trends in prices.
More frequent updating would slow down model runs, making it more difficult to look at
results from large numbers of runs.

The model does not include markets for oil futures more than one year into the
future because the price of oil in the next period is the only thing that should affect
behavior in our model. Above-ground storage is the only market that is affected by
future prices. If the price next period is lower than the current price, expectations of price
increases further into the future should not lead to any additional storage today, because
holders of storage would prefer to wait until the price drops to buy the oil.

Even though we only look at the price of futures one year away, people may still
have to look farther into the future to determine what that price should be. In a rational
expectations model such as the Hotelling model, long-run calculations are needed to
determine what the current price of oil should be, as well as the price one year away.

Equilibrium in the futures market is easy to calculate: it is the average of the price
predictions of the investors. This is because all investors use linear investment functions
with the same risk aversion and confidence parameters, so all predictions get the same
weight.

3.2.2.2 Spot market

Spot prices are set so that the supply from producers equals the demand from
consumers plus the net demand from storage, given the value of futures prices.

$$D_t(p_t) = S_t(p_t) + St(p_t, f_t) \qquad (3.9)$$

A unique equilibrium will exist as long as oil has not been 100% used up. The
equilibrium value is calculated numerically.

3.2.3 Possible prediction rules

The prediction rules used by investors in our model can be broken down into two

categories: long time horizon or long-run rules and limited time horizon or short-run

rules.

3.2.3.1 Long-run price prediction rules

Investors using long-run rules use all the information available at a point in time

to make a long-run calculation of how demand, supply, storage and speculators will

behave over time, assuming that all other speculators behave the same as them. They are

able to project supply and demand arbitrarily far into the future based on full knowledge

of the supply and demand equations in the model as well as the current values of the

supply and demand constants, proven reserves, and the amount of land left available for

exploration. However, they are not able to predict how the stochastic terms in the supply

and demand equations will come out – instead they assume that they will always take on

their mean value of zero. They also assume that new discoveries will be made at a

constant rate, but they can have different beliefs about what that rate is.

In principle, long-run investors should look infinitely far into the future.

However, our supply and demand model leads to near exhaustion of the resource and

very low demand 250 years from the start of the model run. For programming ease we

therefore cut off these long-run evaluations, as well as our model runs, at that time.

To predict what prices should be, long-run investors need to calculate a price for

oil in each period that would lead to equilibriums in the spot market every year until the

end of the run, given their assumptions about supply and demand. Futures prices are

assumed to correctly predict prices for the next period. At the end of the run, there should be no oil left in above ground storage.

The equilibrium problem can be expressed as:

Choose $p_{t_0}, \ldots, p_{250}$ to satisfy:

$$\tilde{D}_t(p_t) = \tilde{S}_t(p_t) + St_t \quad t \in \{t_0, \ldots, 250\}$$

$$St_t = St(p_t, p_{t+1}) \quad t \in \{t_0, \ldots, 249\} \tag{3.10}$$

$$St_{250} = 0$$

Where $\tilde{D}_t$ and $\tilde{S}_t$ are projected demand and supply at time t based on demand and supply updating rules with stochastic terms set to mean value, and discovery rate set to belief of investor.

These long-run rules are similar to what the rational expectations rule would be for this model. The only differences are that the long-run rules may have different beliefs about discovery rates, and they do not properly account for random variation in the model, which could affect the optimal strategy. Doing the full stochastic optimization problem is not computationally feasible in model run time, so these long-run price prediction rules (with the correct beliefs about the most likely discovery rate) are the closest we can come to a rational expectations rule. When all investors use a long-run rule with the correct belief about the most likely discovery rate, the model output looks similar to what we would expect in a rational expectations equilibrium.

Figure 3.1. Price simulation with long-run investors only
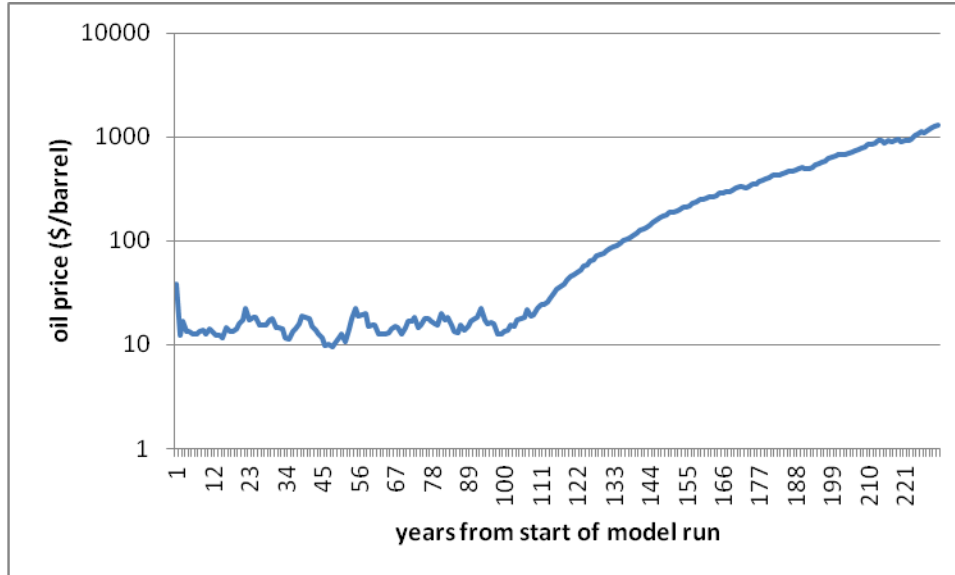
The y-axis is a logarithmic scale, so a straight line increase represents a constant percentage increase in price each year. There seem to be two distinct regimes, one for the first half of the model run, and one for the second half.

For the first half of the model run, short-run concerns dominate, and keep prices higher than they would need to be to anticipate long-run shortages. Stochastic variations in the demand and supply constants lead to some volatility.

In the second half of the model run, the long-run scarcity dominates, and investors anticipate this, building up above ground storage levels while prices rise gradually. The rate of increase depends on interest rates and storage costs. It is a little faster at first because storage costs are a higher percentage of prices. The high levels of above-ground storage also make it easier for markets to adjust smoothly to short run variations in demand and supply constants, smoothing out some of the volatility from the first half.

The relationship between model years and actual years should not be taken too literally since is based on the fit to real-world data taken from chapter 2, but it is still

interesting to note that, based on this fit, the transition to the long-run scarcity regime occurs around 1980. However, the timing of this transition is highly dependent on our assumptions about the interest rate and storage costs – a higher interest rate or storage cost would put off the transition and lead to more rapid increase in prices after it occurs.

3.2.3.2 Short-run price prediction rules

Short-run rules are similar to long-run rules, except that they have a limited time horizon for which they do their analysis. They are formed by looking at projected supply and demand a fixed number of periods into the future. As with long-run rules, we assume that in projecting supply and demand, speculators know the predictable parts of the demand and supply function, but not the stochastic parts.

Based on their projections of supply and demand, they make the same calculation as long-run investors make up to their time horizon. To complete the calculation of what the equilibrium price should be, they need to make an assumption about what state they expect above ground stocks to be in at the end of the period they are evaluating. Investors can have different beliefs about the number of days of forward cover there should be at that point - that is, how long stocks would last if demand remained at current levels with no new supply. The evolutionary algorithm can help this assumption adjust depending on what is proving most accurate. Beliefs that stocks should increase will lead to higher price predictions, and beliefs that stocks should fall will lead to lower price predictions.

3.2.3.2.1 One year time horizon

If we restrict the model to allow only short-run investors with one-period time horizons who have the same assumption about forward cover at the end of the period, we get a good approximation of the results from the model in chapter 2. This is because with very little change in stock levels over the course of the year, prices are being set to equilibriate supply and demand during that year. There are slight differences between this model and the one in chapter 2. The amount of above-ground storage can change slightly as demand for oil increases. There is also variation in the discovery rate from new exploration, which leads to some additional variation in supply over time that did not exist in chapter 2. However, these changes are small, and the results do not differ much from those in chapter 2, except perhaps at the very end when each discovery starts to have more of an effect. The results of one randomly chosen simulation are given below. There are price spikes that go as high as $57,000/barrel as the resource nears exhaustion.

Figure 3.2. Price simulation with short-run investors only
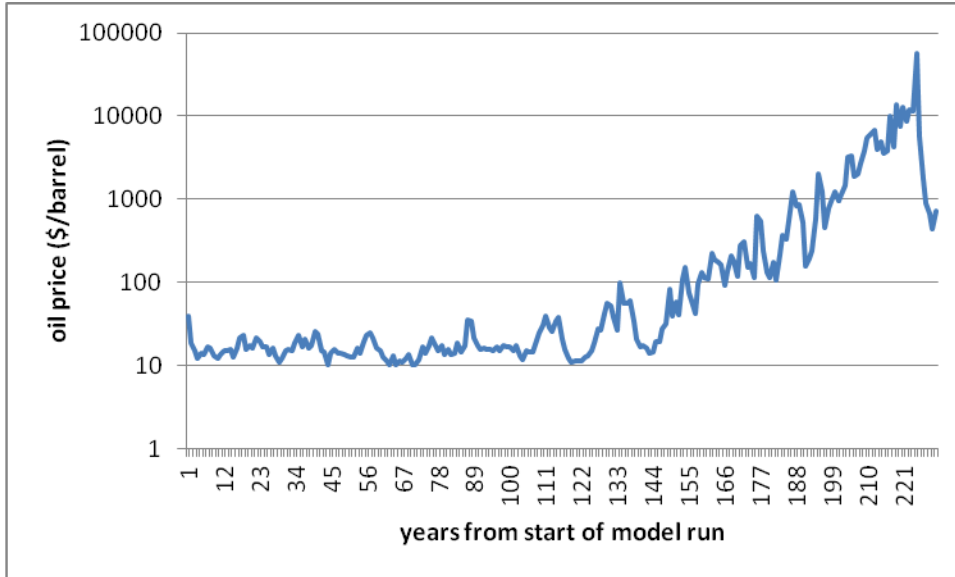
Figure 3.3. Comparison between short-run simulation and chapter 2
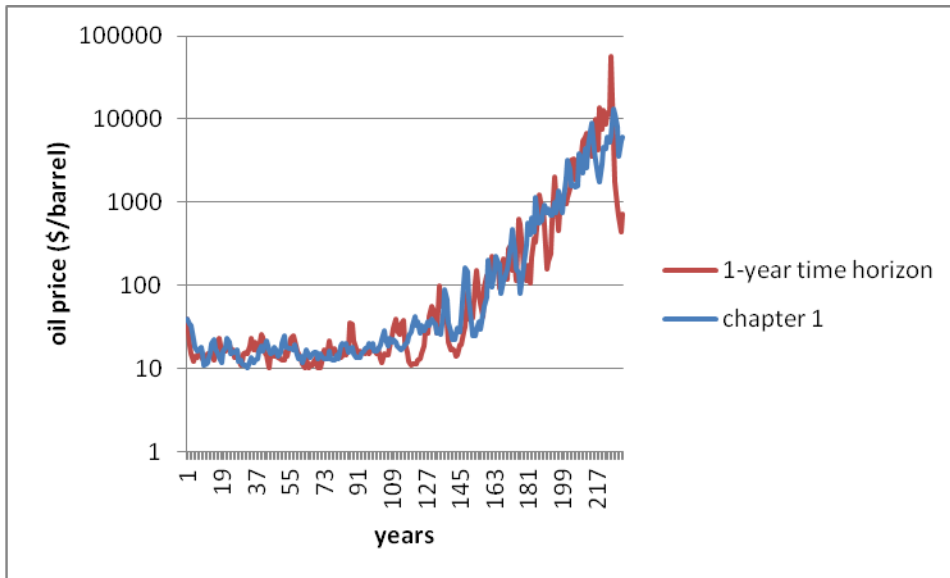
Figure 3.4 shows how the price trajectory compares to the trajectory with long run investors. In addition to the extra fluctuations from year to year, we can observe that prices are lower than in the long-run simulation during the period from years 110 to 170, but get to be higher by the end of the run.

Figure 3.4. Time horizon comparison: two way
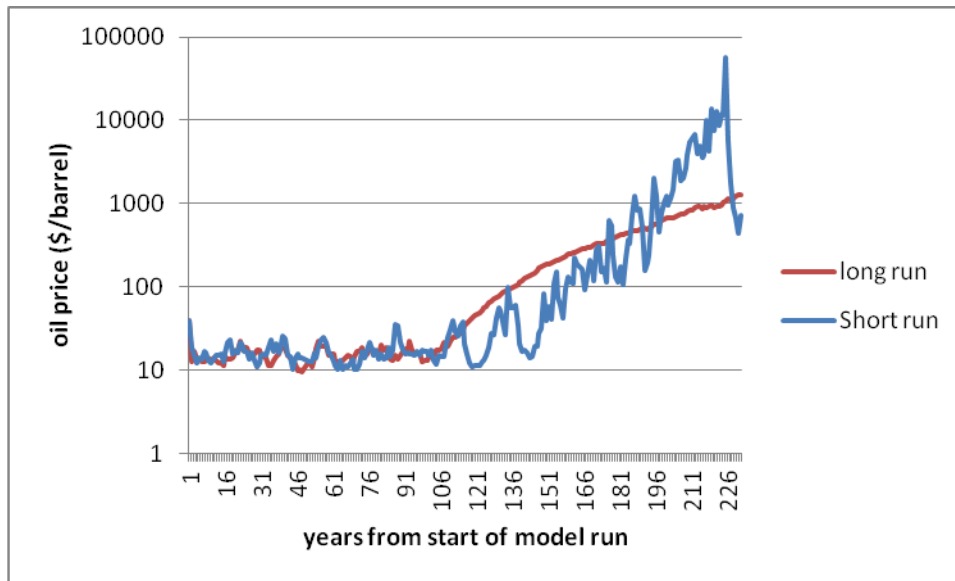
3.2.3.2.2 Seven-year time horizon

If the time horizon is increased, some smoothing is observed. We look at a seven-year time horizon as an example because this is the farthest in the future that oil futures are traded on NYMEX. There are fewer oscillations over the course of the model run than with a one-year time horizon, especially in the second half as the price is increasing. But the long-term trend in prices follows a trajectory that is closer to the one-year horizon trajectory than the long-run trajectory.

Figure 3.5. Price simulation with seven-year investors only
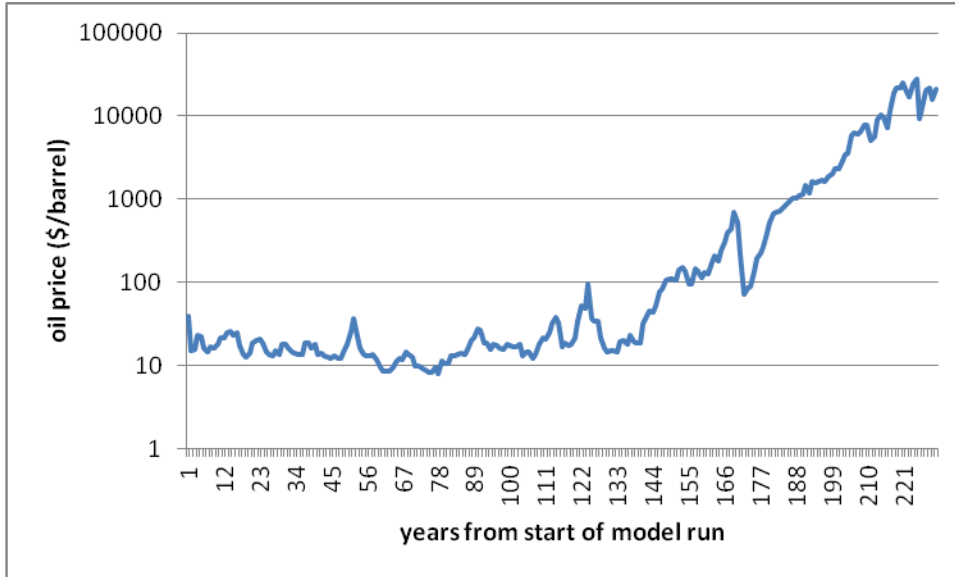
Figure 3.6. Time horizon comparison: three way
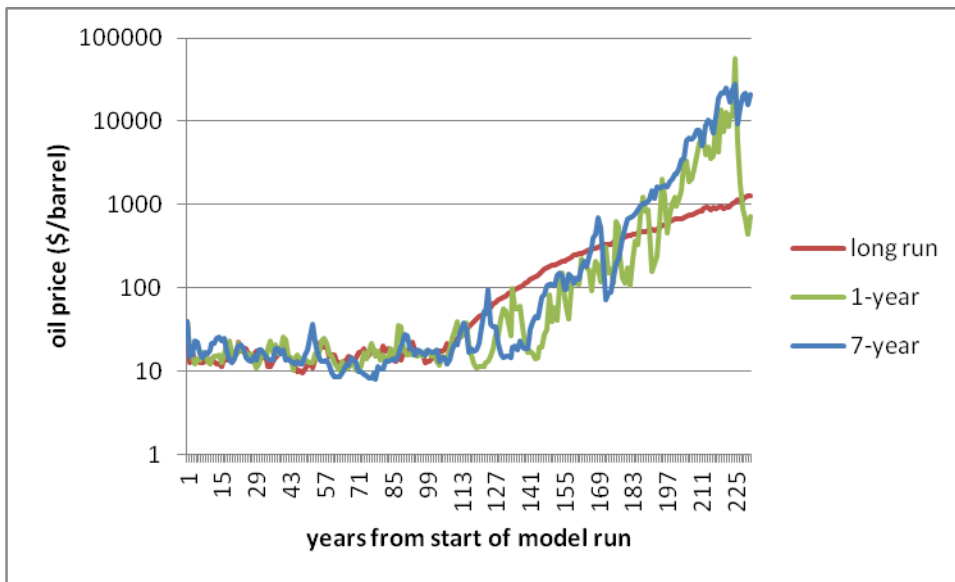
3.2.3.2.3 Twenty-five year time horizon

If investors look 25 years into the future (the longest projections done by the

EIA), the short term variations get smoothed out even further, but the long-run path is

still closer to the short-run model than the long-run model. By the end of 230 years, prices are near $10000/barrel, compared to $1000/barrel in the long run model.

Figure 3.7. Price simulation with 25-year investors only
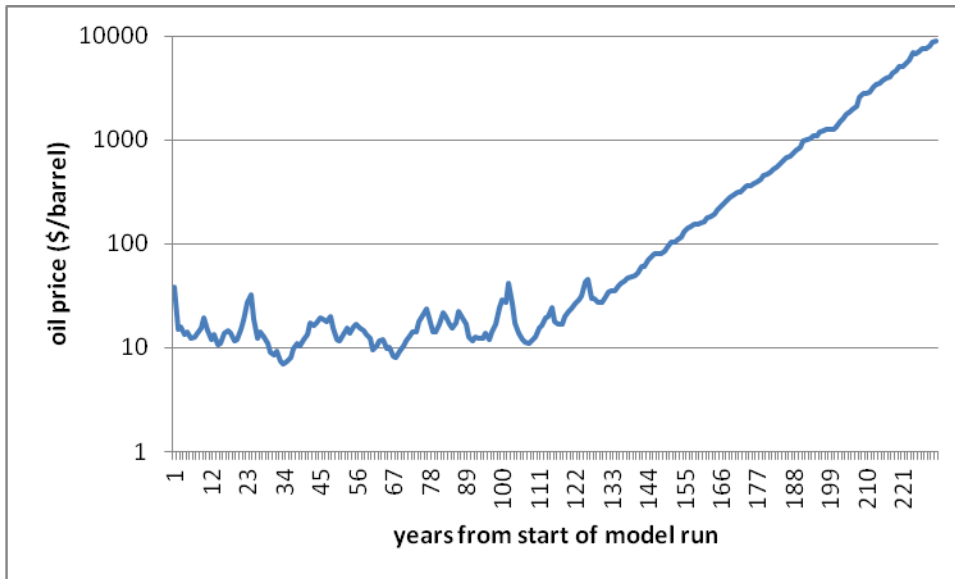
Figure 3.8. Time horizon comparison: four way
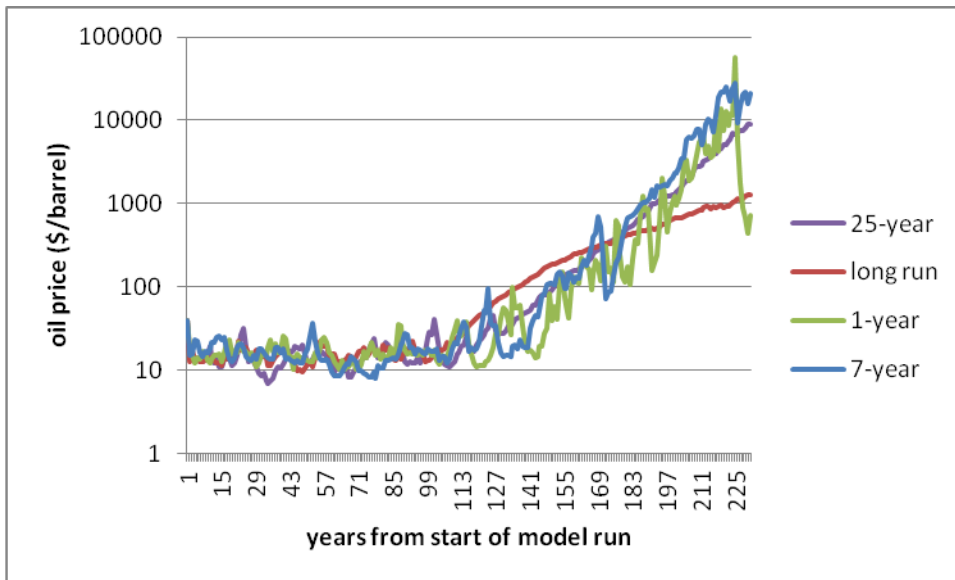
3.2.4 Evolutionary updating

Investors don't know the best strategy for predicting future prices, but they can judge the effectiveness of different strategies by seeing how well they have predicted prices in the past. They will tend to adopt strategies that have proven to be more effective. The goal of the evolutionary algorithm is to have the most effective strategies be the ones that are most commonly used, but with a range of reasonable strategies in use at a time.

There are a number of decisions to make about how the algorithm should work. How fast should investors switch to a new strategy that starts performing well? How much variation should there be in the strategies that are used – should some people keep using strategies that are consistently underperforming?

I use an approach where there is a set of active strategies that evolves over time, and a set of investors that can move between strategies from the active set.

The size of the active set of strategies is important. If it is too large, the running time of the model becomes very slow, because this increases the number of times that the price prediction calculations must be made. However, it has to be large enough to ensure that the space of possible strategies is being explored well, so that everyone doesn't get stuck too long at a sub-optimal strategy.

In order to explore the space of possible strategies efficiently without producing too many unnecessary strategies, it is important to have a good method of targeting strategies that have a good chance of being improvements on the existing strategies.

There is an extensive literature on the use genetic algorithms for optimization purposes. One goal in this literature is to develop efficient ways of exploring a space to

find potential optimums. Since the model uses continuous parameters, we refer to the

literature on continuous genetic algorithms in particular. The approach we came up with

includes a combination of three different types of exploration: experimentation, mutation,

and crossover. Experimentation involves choosing a new random strategy from the

space. Mutation involves making a small change from an existing strategy. Crossover

involves using the information from two existing strategies to produce a third strategy.

3.2.4.1 Updating set of prediction rules

The updating rules use approaches taken from real-coded genetic algorithms,

which we use because two of the factors that vary between rules are continuous variables:

the rate of discovery of new reserves and the number of days of forward cover at the end

of the evaluation period. There is also a binary decision as to the type of rule that is used

(long-run or short-run). There are ways of coding continuous variables using binary

genetic codes that approximate the real numbers being modeled, but there are advantages

of using real coding directly (Herrera *et al.*, 1998). Therefore I use a hybrid approach,

using two real-coded genes and one binary-coded gene.

3.2.4.1.1 Experimentation: new random value

In maximization problems, this is primarily used to generate the starting strategies

that initially populate the strategy space, with no new randomly generated strategies

appearing later in the model. However, in a context where the underlying game is

constantly changing as supply and demand are updated and new information emerges, it

is important for agents to occasionally test the entire strategy space to see if new areas of

high pay-off have appeared. It also makes sense that this would happen in the real world:

while many investors will base their investment strategies on what other people have found to be successful, there will occasionally be investors who make their own independent analysis of investment strategies without taking other people's decisions into account.

New random strategies have a 50% chance of being short run strategies and 50% chance of being long run. The predicted discovery rate is drawn from a log-normal distribution centered at the true expected discovery rate, with scale parameter $\sigma = 0.5$. This creates a slight bias toward picking strategies with the actual rate, but the standard deviation is large enough that a wide variety of them will be chosen.

The expected forward cover is drawn from a log-normal distribution centered at the current number of days of forward cover. This value can change throughout model runs. The scale parameter is set to be $\sigma = 0.2$.

3.2.4.1.2 Mutation: small change from old

Another way that investors may try new strategies is by making small adjustments to existing strategies. The probability of picking a particular strategy for mutation is proportional to the amount of money being invested using that strategy. This way, successful strategies that have been chosen by many investors are more likely to be selected for mutation, so that the region that is close to the successful strategies will be explored most thoroughly.

Each mutation is done to one of the two real-valued genes – discoveries and ending forward cover. Mutations do not switch agents between short-run and long-run rules. The size of the adjustments we use is generated using Muhlenbein's mutation (Herrera *et al.* 1998, Muhlenbein *et al.* 1993):

$$c_i' = c_i \pm rang_i \cdot \gamma_i$$

$$\gamma = \sum_{k=0}^{15} \alpha_k 2^{-k} \tag{3-11}$$

Each α_k is independently chosen to be 1 with probability 1/16 and 0 otherwise. This generates changes that have a high probability of being small (60% chance of being < 1/8), which allows for fine tuning of strategies to come close to the optimum, while still having a non-negligible chance of larger changes (6% chance of being >= 1).

The constant $rang_i$ determines how large these mutations can be. It is set separately for the two terms of the model – the amount of resource discovery and the final stocks – to be proportional to the standard deviation of distribution that random strategies are drawn from, as described in the last section.

3.2.4.1.3 Crossover: combining information from two rules

Crossover occurs when information from two different strategies are combined to form a new strategy. As with mutation, the two strategies are chosen from the set of possible strategies in proportion to the amount of money being invested in the strategy. Also, as with mutation, crossover is used only to adjust the final stocks among short-run investors, or the resource availability among long-run investors. It is not used to switch between short-run and long-run. Therefore, if a long-run and a short-run investor are selected, no crossover occurs.

The standard crossover mechanism in a genetic algorithm is modeled after the process of genetic recombination, where the genome is split at a random point and recombined. In the case of real-coded genetic algorithms with one genome, the genetic recombination approach does not apply; instead another way must be found to use

information from two promising strategies to devise a third possible strategy that has a good probability of being useful.

The approach we use is known as Extended Intermediate Crossover (Muhlenbein *et al.*, 1993) or BLX-0.25 (Eshelman *et al.*, 1993). This rule can be applied to cases with more than one genome, but for our purposes, we only need the simple rule: if the two parents have strategies *c1* and *c2*, the child will have a strategy that is generated randomly from a uniform distribution on the interval $[c_{min} - I \cdot 0.25, c_{max} + I \cdot 0.25]$, where $I = c_{max} - c_{min}$. In words, this means picking a random strategy that falls either between the two parent strategies, or slightly to the outside of them.

The benefit of this approach relative to simple mutation is that agents can explore either a large or a small region depending on how much variation there is between different agents' strategies. When using genetic algorithms for maximization, this allows for more fine-tuning if agents have congregated around a maximum, and more exploration if the agents haven't settled on a clear maximum yet. It is also has a plausible explanation in the context of investors predicting futures prices: if other investors have very different strategies, this is likely to induce more experimentation, whereas if there is a consensus among investors, there is likely to be less exploration of significantly different strategies.

In the usual genetic process, two parent strategies are replaced by two children's strategies. In the world of investors predicting futures prices, there aren't parents producing children in the same way, but it is realistic for one investor to look at another investor's strategy, and use that as a basis for deciding which direction and how far to

explore in choosing a new strategy. To replicate this, I use the crossover operator to replace one of the two parent nodes.

3.2.4.1.4 Elimination of unused strategies

Strategies that are unused by investors for three years in a row are eliminated from the set of possible strategies. If no one adopts them by that time, it is likely because they are not performing very well, so it is not worth continuing to make calculation of the effectiveness of those strategies.

3.2.4.2 Selection of strategies by investors

Selection is the process where investors decide which prediction rules to use when making their investments. The idea is that the most successful strategies become more popular, and the least successful strategies don't get used and eventually disappear.

3.2.4.2.1 Timing of judging fitness

Prediction rules attempt to predict prices for the next period – a year into the future. The effectiveness of a prediction rule can't be evaluated until the period has passed and the actual price has been revealed. If investors adopted a prediction rule before it had been evaluated, there is the potential that they would make wildly erratic price predictions leading them to make investments that would throw off the whole market. To avoid this, we adopt the rule that no prediction rule is used by investors until it has been evaluated. When the set of possible prediction rules is updated, the new strategies are assigned a performance rating of zero – the worst possible. Once its predictions can be tested against actual prices, this performance rating is adjusted based

on the accuracy of the predictions. Investors may then choose to adopt that investment strategy if it has performed well. If no one adopts it after three periods, it is dropped from the set of possible strategies.

3.2.4.2.2 Performance rating

The performance rating of each strategy determines whether or not agents choose to adopt it. This is the equivalent of a fitness function or payoff function used in other evolutionary models. Each new strategy that is tested begins with a valuation of zero. Its performance in each period is judged based on the distance from its prediction to the correct price.

The effectiveness of a strategy in a given period can be judged by how close it comes to correctly predicting prices in the next period, based on the formula:

$$r_{t+1} = Exp\left(d \cdot \left(-|p_{t+1} - f1_t|\right)\right) \tag{3-12}$$

In this formula, $f1_t$ is the predicted future price, predicted at time t of what prices will be one period into the future, and p_{t+1} is what the price turns out to be in the next period. The exponential makes the ratings positive, so that the probability of adopting a strategy can be proportional to its performance rating. The constant d determines how strong the pressure is to adopt the most accurate strategy. If d is large, a small difference in prediction accuracy can greatly increase the probability that a strategy is chosen. This can lead to more precise predictions, but less variation in the strategies available, as strategies that are not making the very best price predictions may be abandoned too quickly, even if their predictions are close to the optimum strategy.

If the same strategy is used over more than one period, its performance rating begins at zero and is updated each period by moving a fixed portion c of the way from its previous rating to the value of its performance in that period:

$$R_0 = 0; \quad R_t = (1-c) \cdot R_{t-1} + c \cdot r_t \tag{3-13}$$

The values of d and c use in the model runs presented in this paper are 2.0 and 0.5 respectively.

3.2.4.2.3 Strategy selection

In each period, each agent compares their current strategy, $S1$ to a randomly chosen strategy, $S2$, from the active strategy set. The probability that they will switch to this new strategy is proportional to its performance rating: $\dfrac{R(S2)}{R(S1) + R(S2)}$

By only evaluating one possible alternative strategy to switch to, this makes it more likely that agents will stick to their old strategy than if they evaluated more possible alternatives in each period. This helps to add some inertia into the system, where investors stick with their old strategies even if they are not the very best out there; however, if a strategy is clearly underperforming relative to most other strategies in the active strategy set it will still be abandoned quickly.

3.2.4.3 Testing evolutionary mechanism

The evolutionary updating rules I use are generally a reasonable representation of how investors might actually behave in the real world – certainly more so than the rational expectations model commonly used in economics – but I do not intend to perfectly replicate the process by which investors actually behave. The goal is that the

evolutionary process will be able to find the most successful strategies and make those strategies more and more commonly used over time, while keeping a range of reasonably successful strategies in use.

To test that the evolutionary algorithm selects strategies properly when there is a clear best strategy, I run a model where the price prediction of a particular strategy is rewarded – specifically the long-run strategy with discovery expectations that match the expected value from the model. Rules that come closest to predicting this price receive the best pay-off. Figure 3.9 shows how closely the average price prediction of investors in the model follows this 'best' price prediction when it is favored by the evolutionary algorithm.

Figure 3.9. Price predictions when selection favors best prediction
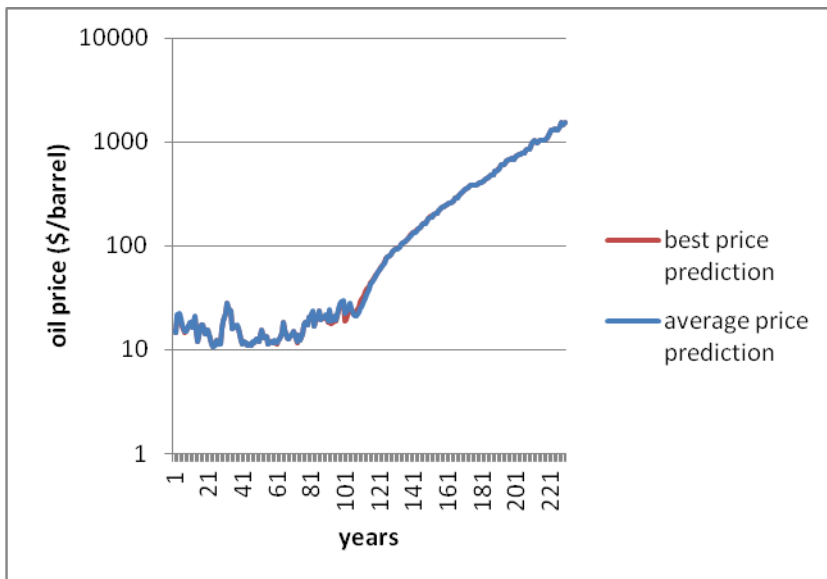
3.3 Model simulations

In this section, we evaluate the outcomes of simulations of the model described in the last section. Simulations are based on runs of the model using Repast for java, with different random number generators and different model assumptions.

3.3.1 The standard model run with one year time horizons

The first simulations we look at include the two extreme types of investors: those using long-run rules and those looking only one year into the future. Investors also can vary in their expectations about the discovery rate of new reserves, and (in the case of short-run investors) in their assumptions about stocks at the end of their evaluation period. The evolutionary algorithm determines which types of investors become the most common and control price-setting.

3.3.1.1 Sample runs

The results from running the full evolutionary model vary greatly from model run to model run, based only on changes in the random number generator used. Three samples of the output are shown below, to get a sense for the different types of outcomes that can occur, and what leads to the differences.[8]

In the first outcome we look at, long run investors dominate, and the price paths aren't too different from the rational equilibrium outcome.

---

[8] Some of the sample runs included here were performed with the number of distinct oil deposits set to 50 instead of 500. Changing this parameter did not substantially affect the results, except by slightly reducing the volatility at the end of the model run as resources approach exhaustion, so I continue to use some examples with the smaller number of deposits for illustrative purposes.

Figure 3.10. Run 1: percent of investors with short time horizons
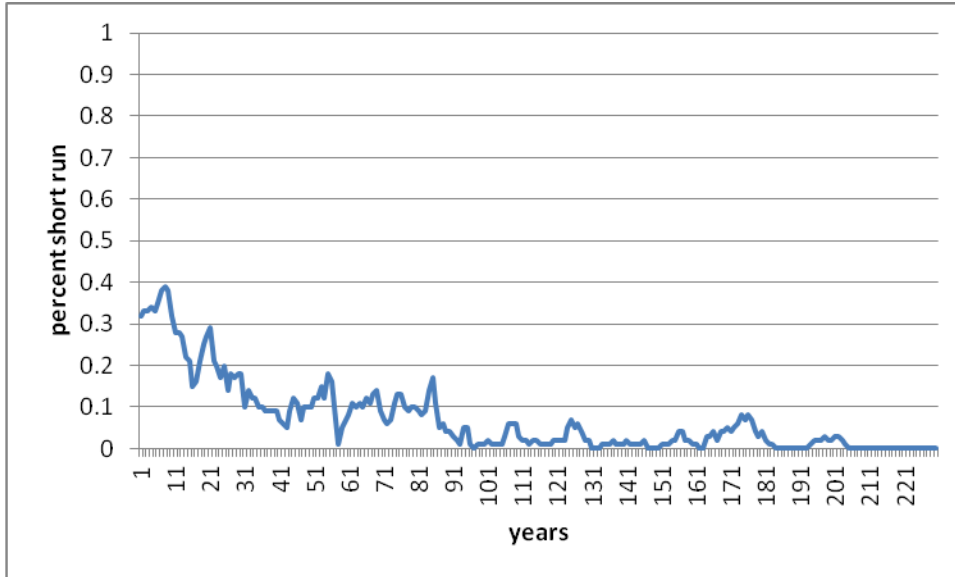
Figure 3.11. Run 1: price trajectory
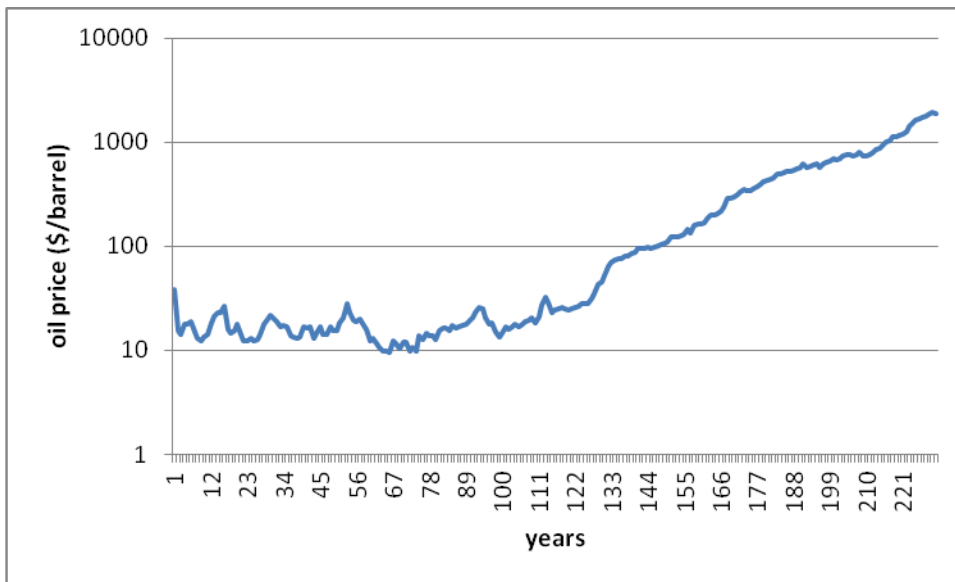
In these runs, the set of investors gradually become more and more dominated by long-run investors, until from year 90 on, almost all investors use long-run rules. The price path looks similar to the path we saw when only long-run investors were allowed.

It is not surprising that an outcome close to the rational expectations outcome is possible. If everyone else is using this strategy, then it leads to optimal price predictions, and so there is no way to improve on it, if this arrangement is reached.

It is, perhaps, more interesting that other outcomes are also possible. In the second model run we look at, short-run investors dominate the market for much of the model run. Prices follow a path closer to that in the scenario with a single type of short run investor, with swings between higher and lower prices. Prices never reach quite as high levels as in the simple short-run scenario, but this appears to be primarily because the amount of oil discovered turns out to be higher than usual in this model run.

Figure 3.12. Run 2: percent of investors with short time horizons
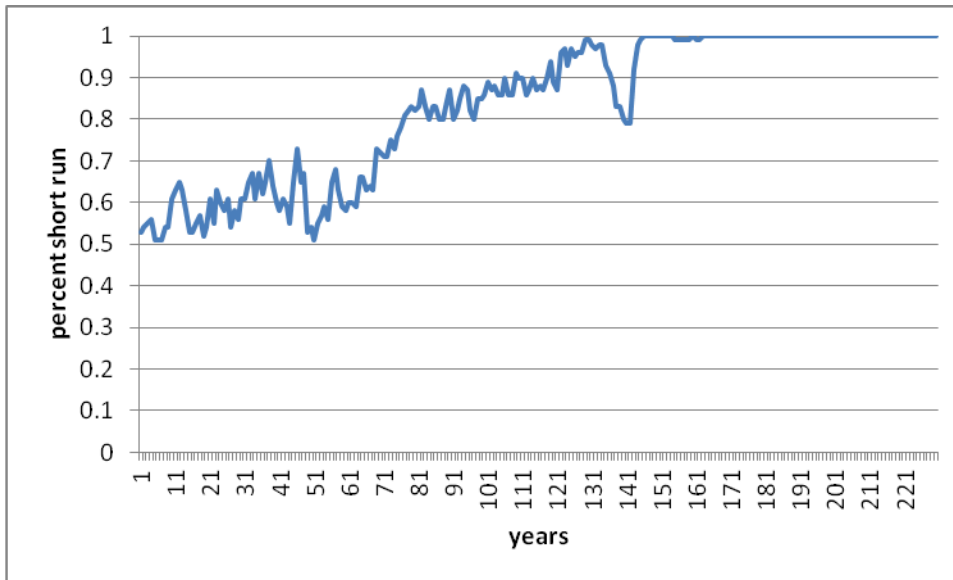
Figure 3.13. Run 2: price trajectory
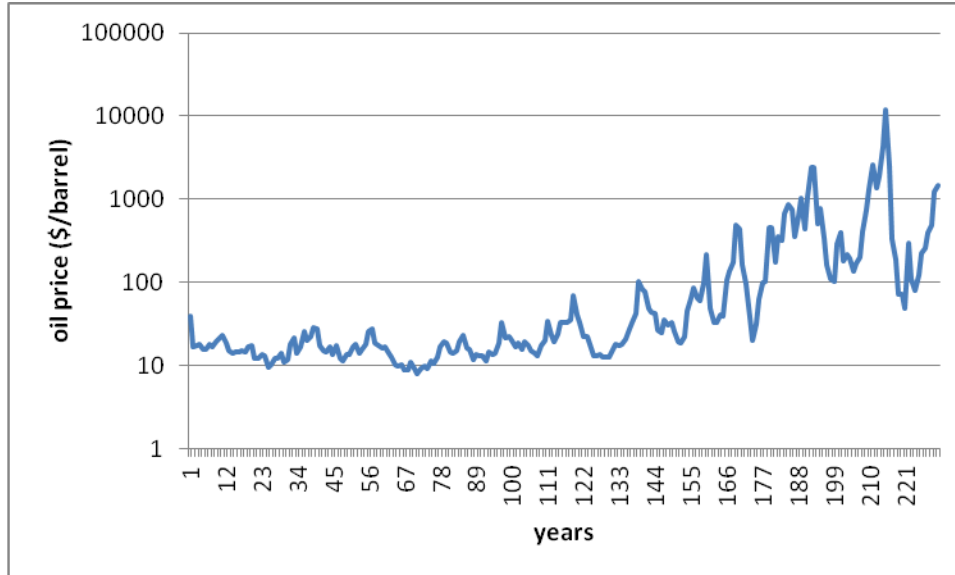
The only difference between this model run and the previous run is the seed used for the random number generator. The fact that random chance can lead to such divergent scenarios suggests that there may be a threshold effect that pushes the model to one equilibrium or another based on small variations in initial conditions. This will be discussed further later on.

A third model run shows a dramatic shift from mostly long-sighted to mostly short-sighted investors part way through the run, after which short-sighted investors dominate. The reason for the shift is explored further in next section.

Figure 3.14. Run 3: percent of investors with short time horizons
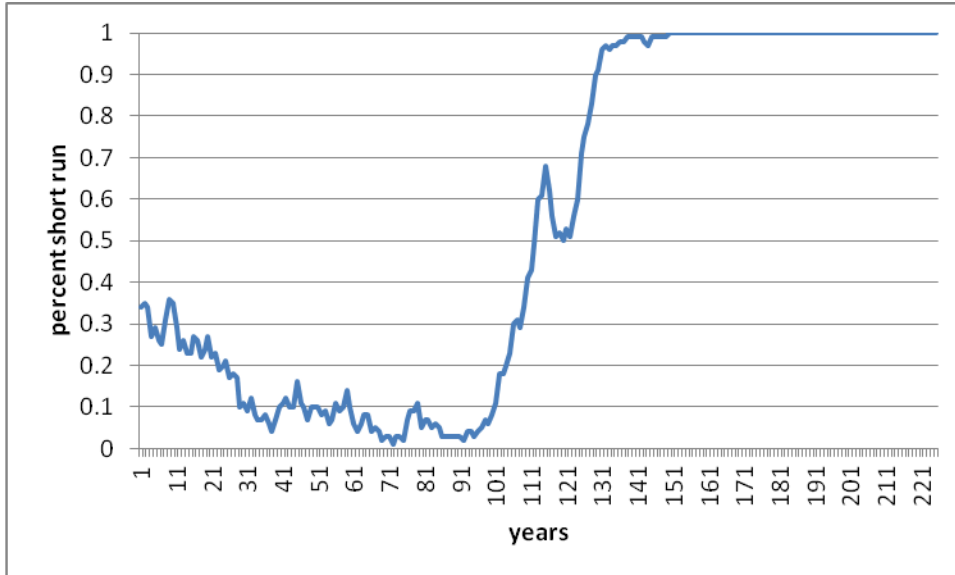
Figure 3.15. Run 3: price trajectory
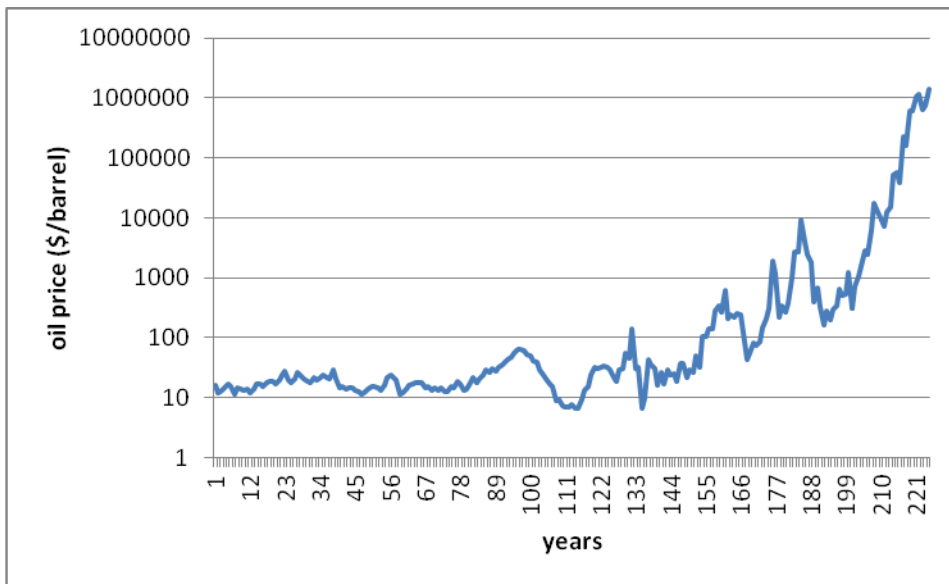
3.3.1.2 Settling on extremes

One striking result from these model runs is that near the end of the model run,

investors seem to either be almost uniformly using short-run rules, or almost uniformly

using long-run rules, with mixes of the two rules being rare. This results from an

evolutionary dynamic that rewards rules that are similar to the rules that everybody else is using. The rational expectations solution produces the correct price predictions as long as everyone else is also using the same rule. However, if everyone else is focusing on short-run concerns, prices may stay low even when long-run shortages should drive up prices. An investor who insists on taking the long run shortages into account when nobody else is may produce a price prediction that is well off from the price that actually occurs. The same is true in reverse – if everyone else is taking long-run shortages into account, people using short-run rules will be less likely to predict the raised prices that are brought on by the long-run investors.

The variation in beliefs about resources and about forward cover of stocks can help to allow investors of the minority type to produce closer price predictions than they would have otherwise, since they can adjust these assumptions to be more consistent with the prices they are seeing.

This can be observed in run 2 mentioned above. In Figure 3.16, we focus on a period where there are mostly short-run investors in the model, but a few long-run investors still there. By the end of this period the long run investors have mostly disappeared. We see that the average price predictions of the long run investors (avg long-run pred) follows roughly the pattern of actual prices, though less closely than short-run investors who are making the calculations that are actually being used in setting prices. This is in contrast with the prices that long-run investors would have predicted if they had stuck with the best assumption about the discovery rate (best long-run pred), which does not follow actual prices at all. By the end of the period, there are so few

long-run investors left around that the evolutionary algorithm stops working as well, and the long-run predictions diverge further from actual prices.

Figure 3.16. Run 2: evolutionary updating of resource expectations
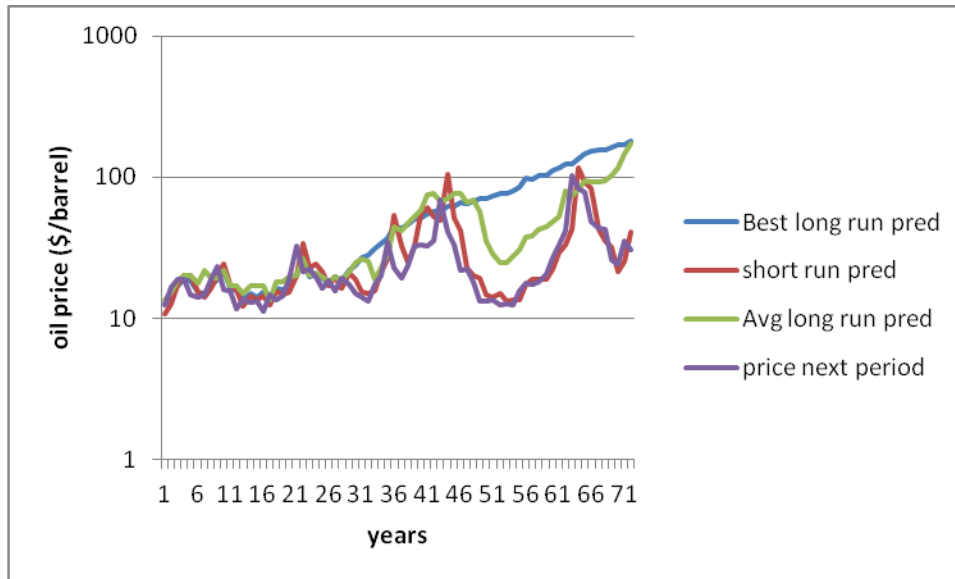
3.3.1.3 Switching between extremes possible

Run 3 from the sample runs is a good illustration of how the model can shift rapidly from having the majority of investors using long-sighted rules to the majority using short-sighted rules.

Some additional graphs from model run 3 help to illuminate what is happening when this shift occurs. The graph below shows actual prices switching from following the average predictions of long run investors to the average predictions of short-run investors. The first thing to note is that the rise in prices starting around year 85 does not follow the pattern predicted by the best long-run rule – instead, it is driven by long-run investors with overly pessimistic views about resource discovery. The shift from long-run to short-run investors is initiated by a downturn in prices that the long-run investors

failed to anticipate. This may have been due to a surprisingly fast discovery of new reserves, it may have been caused by more optimistic investors randomly appearing to drive prices down, or a combination of the two. Once prices start moving down, investors that are making lower price predictions start performing better. This includes both long-run investors with higher expected discovery rates and short-run investors. As these types start becoming more and more common, this makes prices drop further leading to feedback effects that make the short-run predictions, which are the lowest, look more and more attractive, until everyone shifts to using the short-run rules.

It should be noted that the predictions shown here are the fundamental predictions being made, before adjustments based on market information from the actual price of futures (see section 3.2.1.4). The adjustment makes it so fundamental price predictions that are on the right side of the average prediction can be rewarded, even if they are farther away from the true value, which allows the short-run rules to look attractive sooner than one might expect.

Figure 3.17. Run 3: shift from long-run to short-run investors
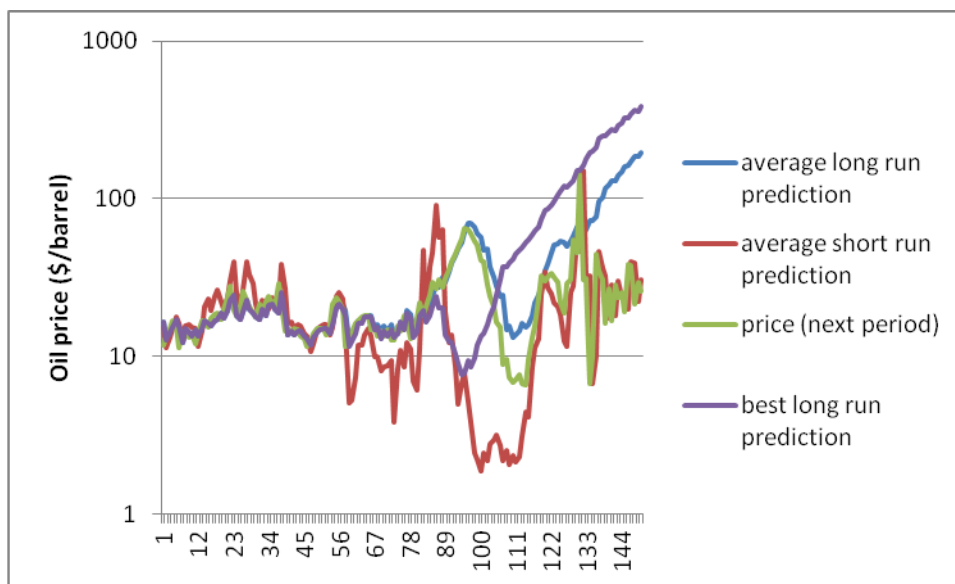
3.3.1.4 Results from more model runs

While it is useful to look at individual model runs to see the different possible behaviors that can occur, it is also important to look at some summary data from many model runs to see which behaviors are most common.

We observed from looking at sample model runs that in the second half of the model runs, it tended to bifurcate into having either all short-run or all long-run investors. We look at the portion of short-run investors at different points in time from 40 model runs to see if this pattern is robust to many model runs, and to identify roughly what portion of the time it settles on the short-run outcome, and what portion it settles on the long-run outcome.

Figure 3.18. Distribution of investor types by year over 40 model runs

year 50

year 75

year 100

year 125

year 150

year 175
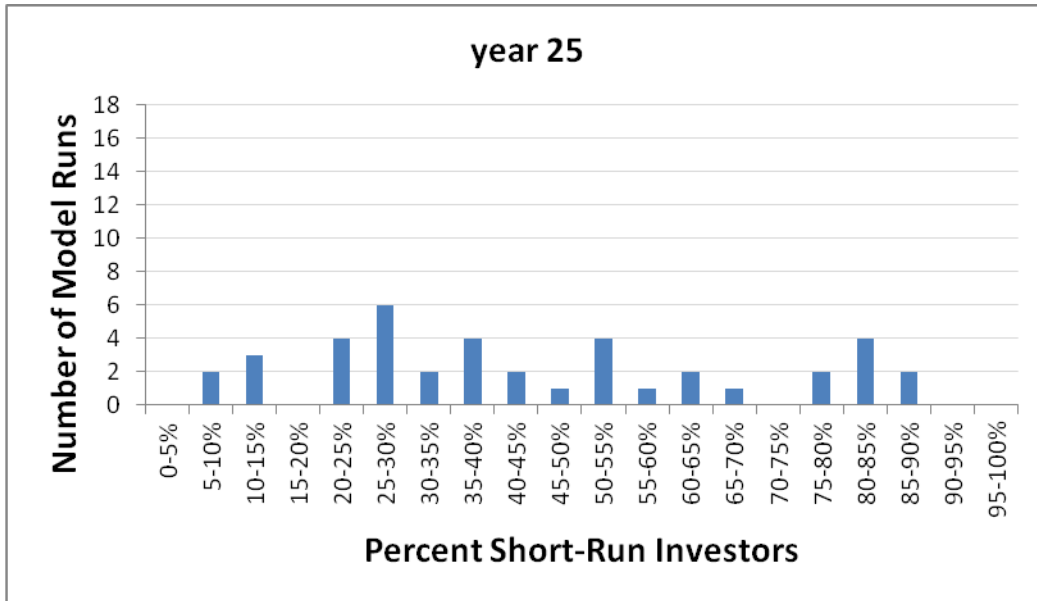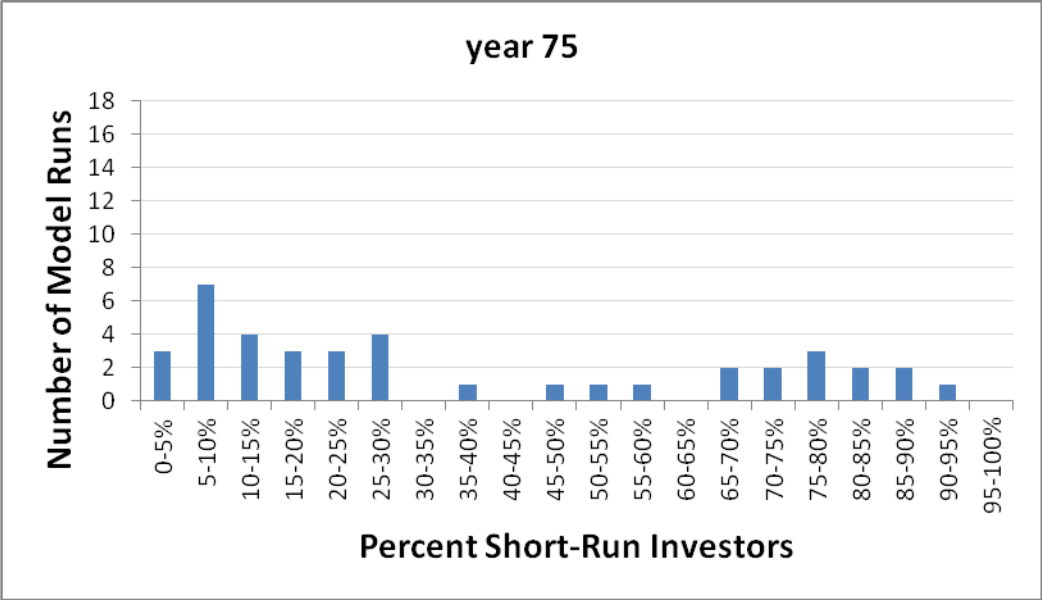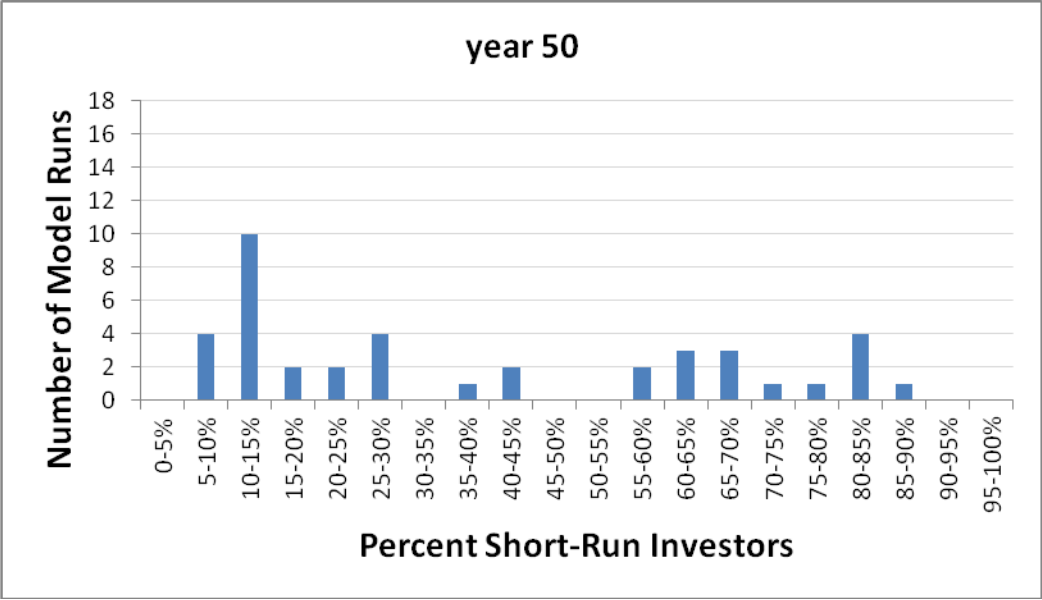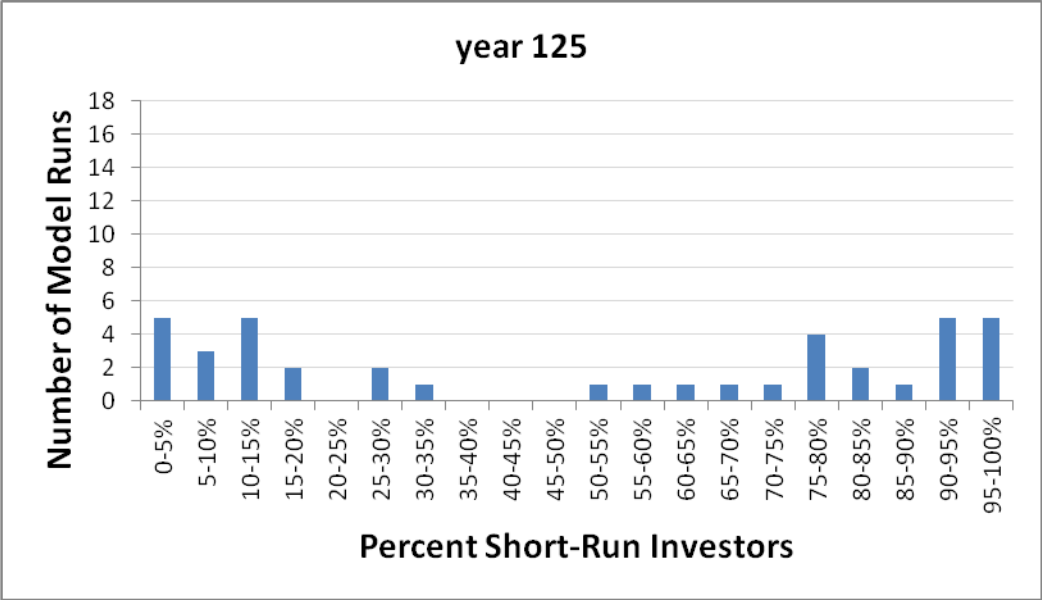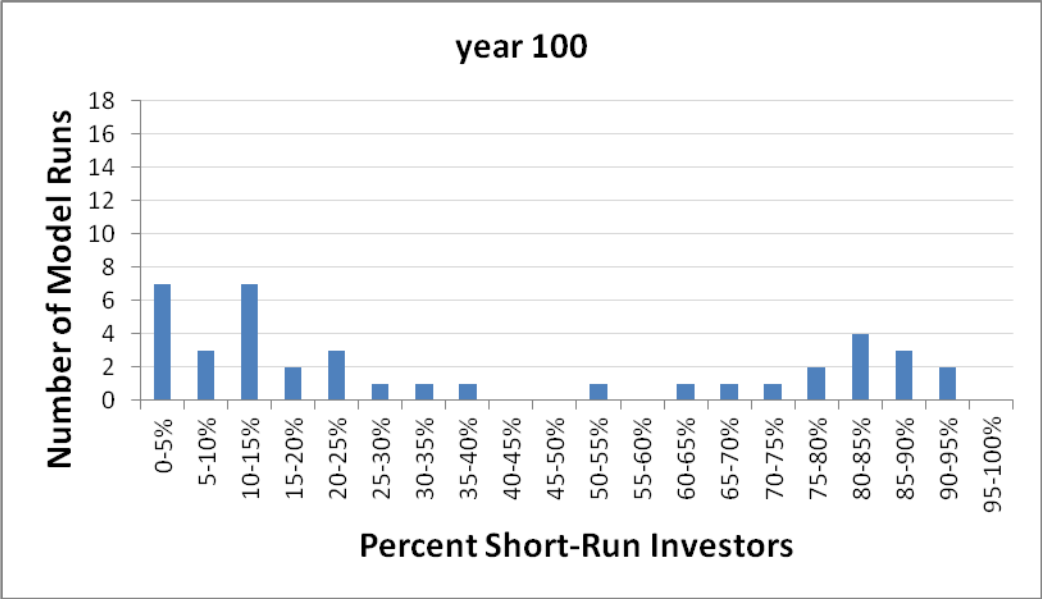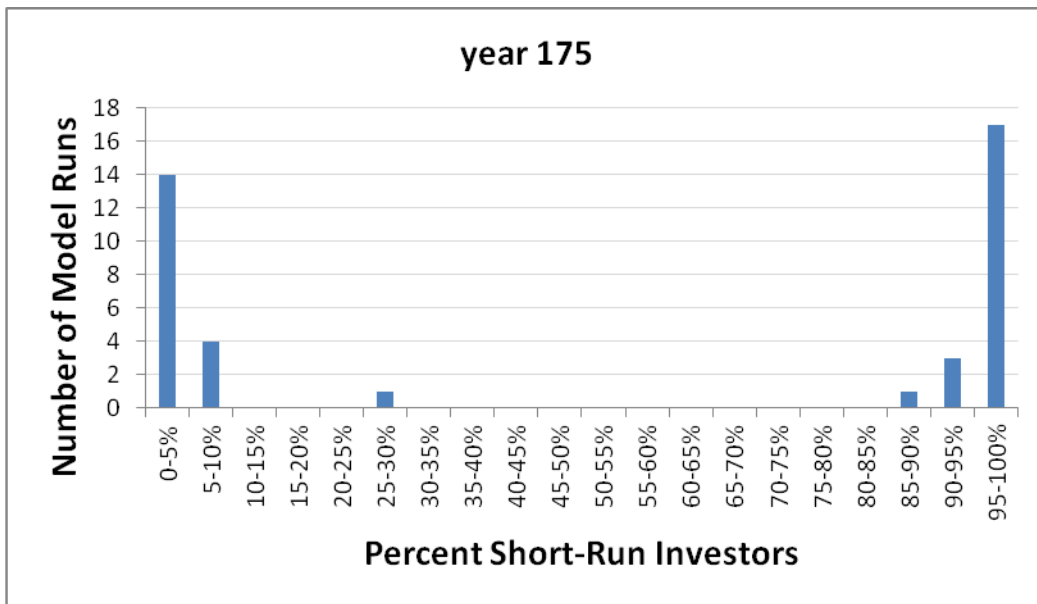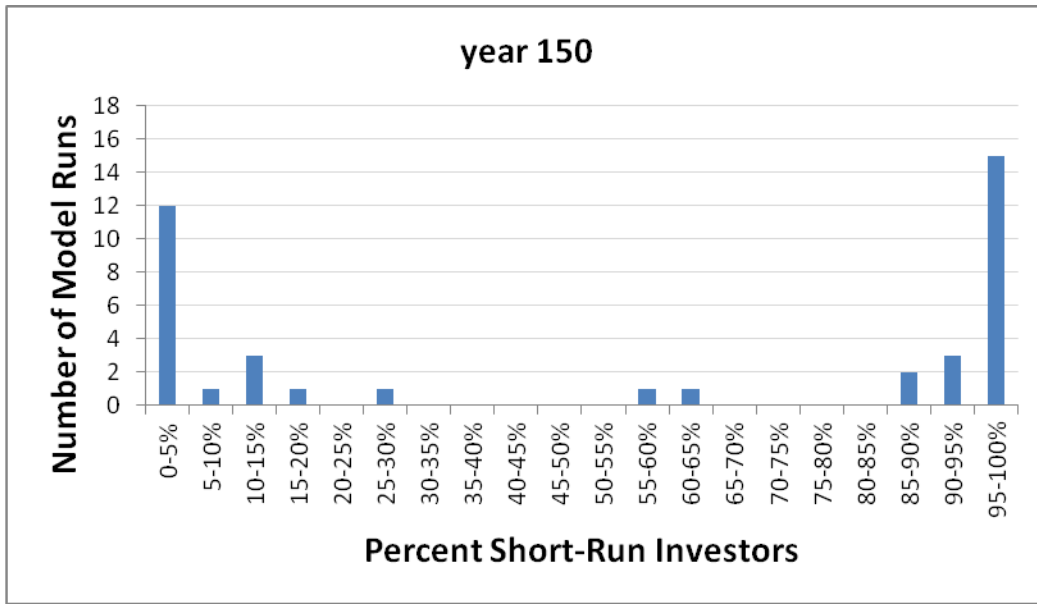
By late in the model run, most runs have all investors at either one extreme or the other. Out of these 40 runs, 21 had most investors with short-run rules, and 19 had most investors with long-run rules. Figure 3.19 shows on average, for the 40 model runs, what percent of investors use short-run rules during the course of the model runs. By the end

investors are slightly more likely to use short-run rules than long-run rules, but it is close to an even split.

Figure 3.19. Average of 40 model runs: percent of investors with short time horizons
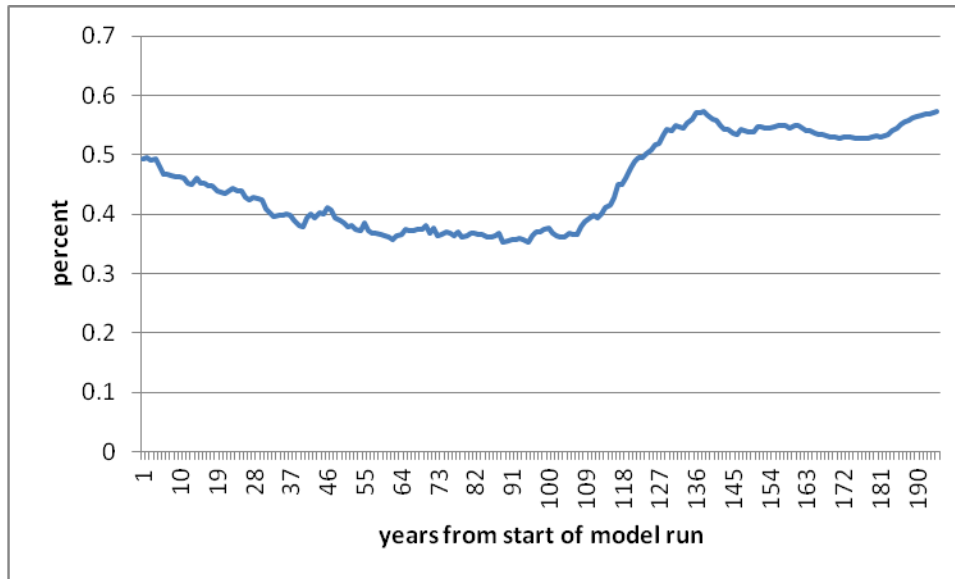
3.3.2 Information cost

The analysis to this point has assumed that investors will gravitate towards whichever policy performs best at predicting oil price movements. This next section considers how the outcomes are affected if we add an additional consideration: the information requirements and calculation difficulties of the long-run approach may discourage investors from taking that approach, all else being equal. If two strategies perform equally well at predicting price movements, the simpler strategy that requires less information should be more appealing. In our model, the 1-year time horizon strategy is much simpler, since it only requires a projection as to what will happen to supply and demand in the next year. The long-run strategy, on the other hand, requires

117

projections for demand, supply and discoveries as many as hundreds of years into the future.

Work by Grossman and Stiglitz (1980) suggests that information cost can be important in determining whether the rational actor model will be an equilibrium. The argument from Grossman and Stiglitz doesn't quite apply in our model, since not investing is not an option, however information costs still are important in determining model outcomes. They play a particularly important role early on in model runs when long-run and short-run rules make similar predictions, and during key transition points when the model could be heading toward one or the other equilibrium.

We add information cost to the model by replacing the payoff function from equation (3-12) with:

$$r_{t+1} = Exp\big(d \cdot \big(-\big|p_{t+1} - f1_t\big| - c\big)\big) \tag{3-14}$$

The size of the information cost measures how much improved accuracy an investor would need to get for it to be worth investing in the additional information and more complicated calculation. An information cost of 0.1, for example, means that forecasts would need to be 10 cents more accurate to justify the information cost.

We run a set of 50 model runs, with information costs ranging from 0 to 0.25. With information costs of 0.05 or higher, the vast majority of model runs converged to an equilibrium with primarily short-run investors.

Figure 3.20. Percent of investors with short time horizons at year 170, with different information costs
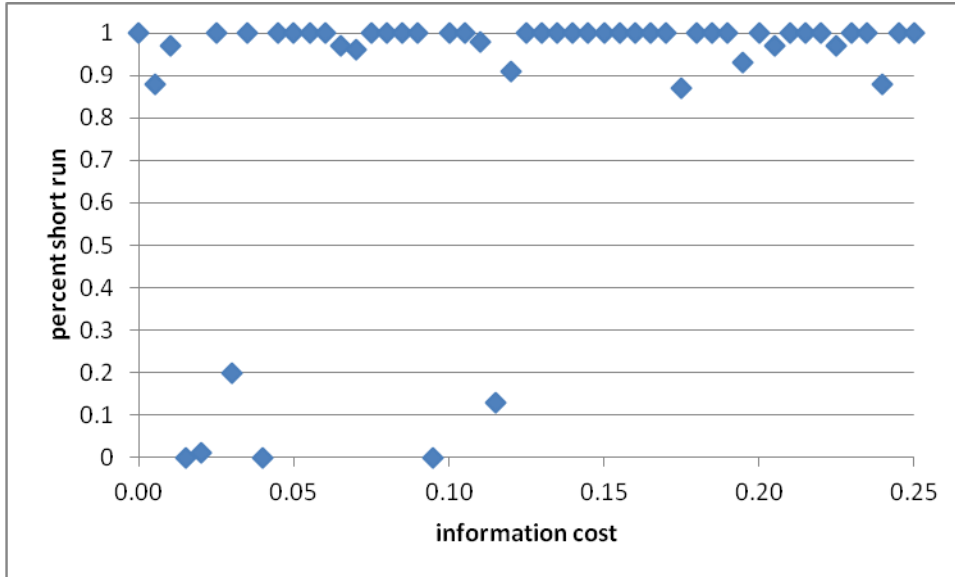
Of 40 model runs with information cost of 0.05, 27 of them, or 67.5%, had

majority short-run investors in year 175.

Figure 3.21. Distribution of investor types in year 175 over 40 model runs with information cost of 0.05
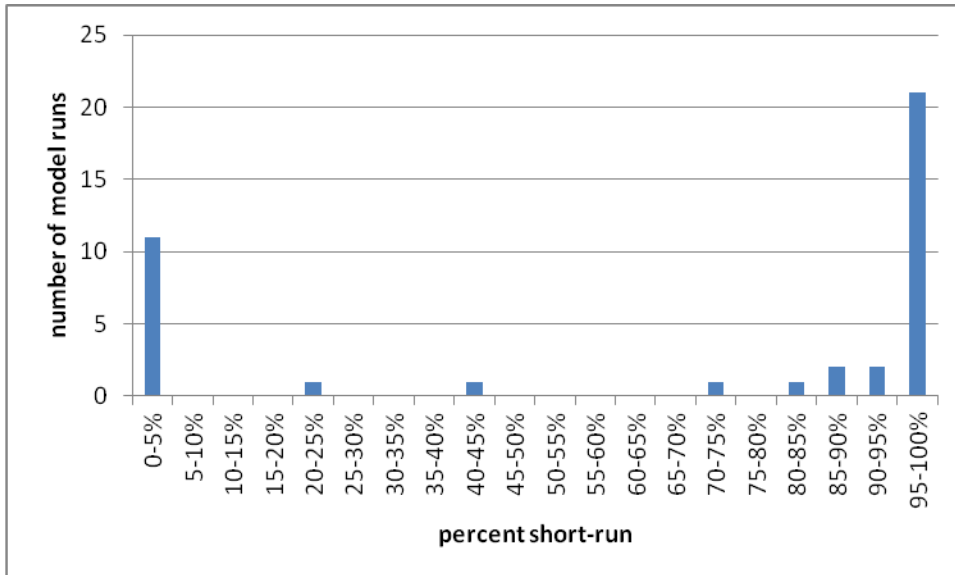
3.3.3 Different time horizons

In future work, it would be interesting to look at how outcomes are affected if a wide range of possible time horizons are possible. However, this would add some complication to the evolutionary model, and would make the results more difficult to interpret. For now, we instead look at models where the prediction rules with 1-year time horizons are replaced by prediction rules with time horizons of 7 and 25 years.

The results from runs with 7-year and 25-tear time horizon investors show similar evolutionary dynamics as with 1-year time horizons. Model runs tend to diverge to outcomes with all investors having the same time horizon, though they are not pushed quite as strongly to the extremes with more chance of switching, especially in the case of 25-year time horizons. This could be because the price predictions made by the short-run approach can be closer to the long-run rule with longer time horizons. Short-run outcomes also become slightly more common.

The model runs dominated by short-run investors have smoother price paths than the corresponding 1-year runs, but are still too low in the middle of the model runs and too high near the end, as was the case when only 7-year or 25-year investors were included. The general conclusion from the 1-year model runs still holds: investors may fail to fully anticipate future shortages leading to sub-optimal price paths. The histograms of the percent of short-run investors at year 175 with 7 and 25 year time horizons are shown below:

Figure 3.22. Distribution of investor types in year 175 over 40 model runs with 7-year time horizons
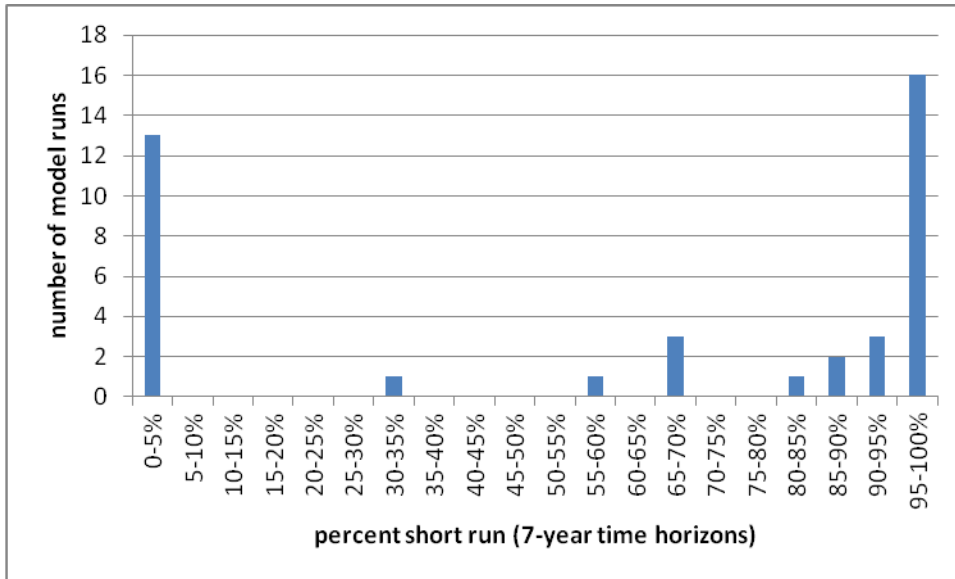
Figure 3.23. Distribution of investor types in year 175 over 40 model runs with 25-year time horizons

121
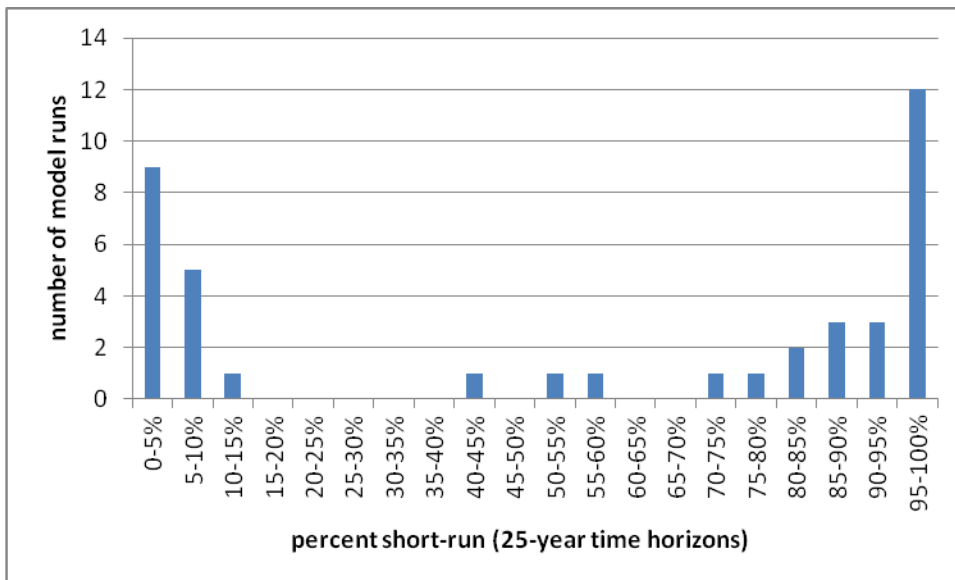
3.4 Conclusions

3.4.1 Goals

The goal of this paper was to develop a theoretical model of oil markets with resource scarcity that would help evaluate whether the conclusions of rational actor models would hold in the presence of smart, knowledgeable but boundedly rational investors. The lack of any behavioral modeling of how oil markets might behave in the face of resource scarcity was a major hole in the literature, given the importance of concerns about resource scarcity.

The model cannot be used to fully resolve questions about whether investors are currently anticipating future scarcity properly, or whether prices in the real world are too low or too high. In some model outcomes, investors do properly anticipate future shortages, and in some they do not, and I have made no attempt to evaluate whether oil market data is more consistent with one or the other of these scenarios. The model outcomes also can be sensitive to the specific assumptions of the model, and changes to these assumptions could produce even more variety of results than have been observed so far.

Still, the fact that this model can produce outcomes that vary substantially from the rational actor predictions is a result that is worth noting. The prevailing view in economics is that anticipating future shortages is something that the market should do well. There are substantial monetary benefits to making an accurate prediction of future price changes, and market structures should make it possible for a diverse set of players to take advantage of any useful information that the rest of the market is not accounting for. It is surprising, therefore, that in a model where rules that correctly predict price

122

movements are rewarded, outcomes can persist in which most market participants are failing to properly anticipate future market developments.

3.4.2 Model contributions

In developing the model, I aimed to choose assumptions that would be simple enough that the model would be manageable and its output possible to interpret, while being realistic enough that the key results from the model runs would have plausible explanations.

The underlying demand and supply models were made to be realistic enough to add credibility to the concerns about long-run scarcity, while also allowing short-run demand and supply constraints to play a role. The negative outcomes associated with the failure to anticipate long-run scarcity are an important part of the conclusions of this study. The model will not resolve all debates as to whether oil shortages will occur before alternatives develop enough to make these unnecessary. However, by grounding the demand and supply models in the literature and (imperfectly) fitting them to data, this chapter does provide evidence that concerns about long-run scarcity are reasonable.

The adaptive model of investor behavior is also designed to provide a reasonable representation of behaviors that occur in the real world. Perhaps the main conclusion from the model runs is that investors can adopt behavioral rules that differ from the rational actor rule, and that this can lead to price paths that are too low in the middle and spike at very high levels as the resource approaches depletion. It is important, then, that the key model assumptions that drive this outcome could reasonably be observed in actual oil markets.

The main reason that sub-optimal outcomes can persist is that when shortages are far enough into the future, there may be little monetary benefit of anticipating those shortages right now. Until the anticipated shortage occurs, the performance of investors calculating the shortage into their pricing will depend on the behavior of other market participants. If other participants also see the shortage coming and start to act on it, prices could go up well before the shortage occurs, leading to immediate benefits. However, if other participants are focused on more short-run concerns, prices may stay low for years or even decades. During that time, the evolutionary algorithm will not favor people who continue to bet on an eventual rise in prices. Investors who factor the future shortages into their price calculation may get discouraged, and are unlikely to be copied by others. Even if it would eventually be a money-making strategy to bet on a long-run increase in prices, the short-term benefits or losses will be what others observe as they decide if it is a good strategy. In this scenario, investors who do not factor the long-run shortage into their pricing decisions will have more immediate success, and therefore will continue to dominate the market. These factors that allow the sub-optimal outcome to persist in the model could plausibly occur in the real world.

The role of information cost in helping determine the model outcome also seems plausible. In runs with no information cost, purely random variation can lead the model to converge on either equilibrium. When a small information cost is added, however, this makes it more likely that it will settle on an outcome with all short-run investors.

The behavioral model can also capture situations where a speculative bubble can emerge. In run 3 from section 3.3.1.1, starting around year 85 prices rose well above where they should have been based on the best long-run rule, driven by long-run

124

investors with overly pessimistic beliefs about new discoveries. The price rise reinforced investors with those beliefs, bringing prices further away from their fundamental value. At some point, prices started to fall and investors with lower price predictions started performing better driving prices down further. The predictions of the investors who initially caused the bubble started looking more and more unreasonable and the market became dominated first by investors with more optimistic views about resource discovery, and eventually by investors who focused only on short-run supply and demand. These dynamics are remarkably similar to what may have happened in the 1970's and 80's. The price rises of the 1970's led to concerns that resource exhaustion was contributing to the rise in prices. When the bottom fell out of the market in the 1980's, these concerns were largely discredited. Through the 1990's, prices stayed low as short run developments tended to dominate discussions of oil prices, with little regard for concerns about future scarcity.

3.4.3 Policy implications

The possible outcome dominated by short-run investors is clearly sub-optimal, suggesting that there may be room for policy interventions to improve the outcome. If the model in this paper is taken at face value, the information is available for policy makers to calculate what prices should be and intervene to be sure that prices followed the desired path. However, this would be risky. In addition to questions about whether policy makers could really make the correct calculation, policy makers would also have more incentive than markets to keep current prices low for political reasons, potentially leading to worse than market outcomes.

While it may not be advisable for the government to intervene directly in setting prices, there may be roles for government agencies in making information and analysis of oil markets readily available to investors to help reduce information costs, making it more likely that an equilibrium with long-run rules is reached. The EIA in the US already plays this role to some extent, but it could contribute further by focusing more on long-run analysis.

The model developed in this paper could also be used to help evaluate policies that have already been proposed, such as a carbon charge or support for technological development to reduce demand. It seems plausible that if we are following a path suggested by the model with short-run investors, then either of these government interventions could help reduce the severity of the shortages late in the model run, providing some additional motivation for implementing these policies. However, more research would be needed to evaluate this claim.

3.4.4 Future work

The model developed in this chapter provides some useful insights into how investor behaviors could influence oil markets. However, it also leaves many questions unanswered. Addressing these questions about how well markets will anticipate scarcity is a huge task, and the model development and analysis done in this paper is just a first step. Much more can be done in future work to build on it.

3.4.4.1 Sensitivity analyses

One big question with this model is how sensitive the results are to some of the specific assumptions made in developing it. There are many parameters that I chose

126

values for along the way with varying levels of justification. Some of the outcomes of the model runs could change if the parameter values are adjusted.

For the purposes of this paper, I have been careful not to draw conclusions that rely heavily on the exact results from the model runs. Even if the particular modeling assumptions are not perfect, the modeling exercise is useful in establishing a possible mechanism that could lead to a failure to anticipate shortages, and in clarifying our thinking about how and under what conditions this could occur. However, other model output may be more sensitive to parameter choices. Some examples include how likely the model is to settle on an extreme equilibrium, when this happens, which equilibrium is reached most frequently, and how common it is for shifts between equilibriums to occur. A thorough sensitivity analysis of how the model output responds to changes in the assumptions would help to establish which results are most robust to changes in the assumptions.

3.4.4.2 More prediction rules

One model assumption that could be relaxed in future iterations as the restriction of price prediction behaviors to the two categories of short-run and long-run rules. This could help to drive the bifurcation of investors into being all of one type or the other. If more variety of rules were allowed, some more interesting dynamics could occur.

3.4.4.2.1 More time horizons

While we experimented with changing the time horizon of short-run investors, we have not tried allowing more than two different time horizons in the model at once. Allowing this could be an interesting next step in the analysis. There is an advantage in

investing in futures markets to predicting market developments one period before everybody else does. This allows you to take a position in futures markets before prices move when others take the same position. For this reason, there could be an advantage of having a slightly longer time horizon than most other players in the market. This could lead to a gradual drift toward longer and longer time horizons. However, the evolutionary pressures are not so clear if there is a mix of people with different time horizons all participating in the market. It is not clear whether or not long enough time horizons will dominate to properly anticipate shortages. This would be an interesting question to address in future analysis.

3.4.4.2.2 More flexible rules

There are many other types of rules that are possible for predicting price movements. Many behavioral models focus on prediction rules that use past price patterns to predict future prices. It would be interesting to see how these types of rules perform relative to rules that use projections of demand and supply to predict prices. Rules that are explicitly trend following could also increase the chances of having speculative bubbles in futures prices.

3.4.4.3 Policy/Demand/Supply manipulations

One interesting use of the model developed in this paper would be to observe how the model output would respond under different imposed scenarios. As discussed in the policy implications section, it would be interesting to see how policies such as a tax on oil or imposed reductions in demand would affect the model output. It could also be interesting to look at how the model responds to a temporary supply disruption in

scenarios with different types of investors. The response to a temporary supply disruption could also be compared to a development that affects supply more permanently such as an unusually large new discovery.

3.4.4.4 Comparisons with data

One of the benefits of developing this model is that it provides a theoretical framework that could ultimately be used to test the results of model runs against historical data. This could provide some validation of the model's ability to explain historical price patterns, as well as some insight into which types of model output have been most consistent with different historical periods.

Matching the model output to data effectively presents some significant challenges. In some cases, the model variables, while sufficient for theoretical modeling purposes, may not match perfectly to the equivalent real-world variables. For example, in matching futures prices from the model with real world futures prices, it would be important to consider whether the hedging role of futures markets leads to any bias in futures prices relative to the best guess of pure speculators. The level of above-ground storage may also be difficult to match to real-world data since this model uses it as a proxy for *in situ* storage by producers when a long time trend of increasing prices is expected. There may therefore be a need to grapple more thoroughly with the question of how suppliers might factor in their future price expectations into their supply decisions.

There also will be challenges in identifying appropriate patterns from the model that can be tested for fit with real world data, and in figuring out how to adjust model assumptions to improve the fit. However, the results could go a long way toward

addressing questions as to how investors have actually behaved up to this point, and

provide more insight into what may be likely to happen moving forward.

CHAPTER 4

AN ANALYSIS OF THE IMPACT OF A CARBON CHARGE AND REVENUE RECYCLING ON US HOUSEHOLDS

4.1. Introduction

This paper addresses the impacts that a carbon charge or cap-and-trade policy would have on households of different income levels in the US. As global warming has gained attention and governments have looked to find ways to address it, some form of carbon charge or cap-and-trade program has emerged as the most promising approach to combating it. Carbon cap-and-trade policies have been adopted in Europe and at the regional level in the US. Cap-and-trade legislation has also gained attention in the US congress, but has yet to be adopted. As a result, studies of the impact of different variations of these policies are currently of great policy relevance.

The effect of a carbon charge on income distribution is an important factor in political discussions. A common objection to cap-and-trade programs is that they would hurt poor and middle class families by raising the price of fuels they rely on. Analyses of the impacts of a carbon charge on households of different income levels are frequently cited in political commentaries and debates (Wall Street Journal, 2009; Yarow, 2009; Abar, 2009).

The distributional effects are also important for normative reasons. The amount of money that could be generated by these programs is enormous: it has been estimated that an auctioned permit program based on the carbon caps being considered in congressional proposals could generate enough revenue to supply a household of four with an average of $1600 to $4900 per year (Paltsev *et al.*, 2007). This means that there

is the potential for a significant redistribution of real incomes, which could be either progressive or regressive depending on how the policy is designed. Any policy that restricts carbon emissions is in effect a reorganization of property rights to emit carbon, moving away from the open-access regime that currently exists, and the design of the policy can be seen as dictating who obtains these property rights. If the rights are given to corporations, this could lead to a strong regressive redistribution of income. A more egalitarian system of distributing the rights could have the opposite effect, increasing the wealth of most lower and middle-income households.

The policy I focus on in this paper, known as cap-and-dividend, would auction permits, and return the revenues to households as equal per-capita dividends. This policy was initially proposed by Peter Barnes (2001), and has been incorporated into bills proposed in the US Congress.[9] The basis for this proposal is that the right to release carbon into the atmosphere should be equally owned by everyone.

Many other variations of cap-and-trade policies have been proposed. Most proposals include giving away a portion of the permits to fossil fuel companies, utilities, or other companies that might be impacted by the policy. To evaluate this type of policy, I also consider the impacts on households if revenues are equally split between payments to households and payments to producers.

While a number of studies have looked at the impacts of these policies before, there are still many unresolved questions. Assumptions about how producers and consumers will respond to the policy vary from study to study, and some simplifying

---

[9] Two cap-and-dividend bills were introduced in the 111th Congress, although neither were adopted. These bills were Van Hollen's H.R. 1826 in the House, and Cantwell and Collins' S. 2877 in the Senate.

assumptions are common to most studies. It is not well understood how these assumptions affect the results.

My goals in this paper are to analyze the distributional impacts of a cap-and-dividend policy, to clarify how different assumptions can affect the results, and to compare the impacts of a cap-and-dividend policy with a policy in which half the permits are given to producers.

In the second section, I look at the current breakdown of carbon consumption in the U.S., and how that consumption can be attributed to households. In the third section, I look at different issues that arise in the design of a carbon pricing policy. In the fourth section, I look at the model of how producers and consumers respond to the policy. In the fifth section, I discuss how the impacts are attributed to households, and how the distributional incidence results are presented. In the sixth section, I present the results of the analyses. In the seventh section, I summarize the main conclusions and discuss areas for further research.

4.2 Description of carbon consumption in US

Carbon dioxide is emitted primarily through the burning of fossil fuels.[10] In the US in 2006, 44% of this was emitted by burning petroleum products, 36% by burning coal, and 20% by burning natural gas. These are consumed by the electricity sector,

---

[10] Carbon dioxide is also emitted when bio-matter is burned or decays, and is removed when plants grow. As a result, land use changes can also contribute to changes in concentrations of carbon dioxide in the atmosphere. Small amounts are also emitted when cement is processed. This paper focuses only on emissions related to the consumption of fossil fuels, as this is the largest contributor. We also do not address emissions of other greenhouse gases that contribute to global warming, though a comprehensive greenhouse gas reduction policy would also put a price on these emissions.

which was responsible for 40% of US carbon emissions, the transportation sector (34%), the industrial sector (17%), the residential sector (6%), and the commercial sector (3%). If electricity-related emissions are attributed to the sectors that consumed the electricity, the percentages become 34% for the transportation sector, 28% for the industrial sector, 20% for the residential sector and 18% for the commercial sector. The full breakdown of consumption by sector and fuel type is shown in Table 4.1.

Table 4.1. 2006 US carbon dioxide emissions by fuel and sector

Sector	Petroleum	Coal	Natural Gas	Other	Total	Percent via electricity
Residential	113.7	717.4	362.4	4.3	1197.9	72.7%
Transportation	1975.6	3.9	33.9	0.0	2013.4	0.2%
Industrial	438.1	722.7	488.3	3.2	1652.4	39.4%
Commercial	68.8	695.7	274.3	4.2	1043.0	80.3%
Total	2596.2	2139.8	1158.9	11.8	5906.7	40.0%
Percent of Total	44.0%	36.2%	19.6%	0.2%	100.0%	

Source: calculations from US EIA, "Energy-Related Carbon Dioxide Emissions from the Residential and Commercial Sectors, by Fuel Type."

4.2.1 Attributing carbon emissions to end users

A more complete picture of the carbon footprint of households requires looking not only at their direct energy consumption, but also at the fossil fuels consumed in making the products they buy. As with electricity, industrial and commercial emissions can be attributed to the end users (households or governments) that consume the products that are produced using fossil fuels. The carbon footprint of a household that buys a car would include the fossil fuels consumed in producing the car, as well as those consumed in extracting and processing the metal used to make the car, and so forth.

Determining a household's carbon footprint therefore requires two steps: determining their consumption levels of different goods and services, and determining the

amount of carbon that was emitted in producing each type of good or service that they consume. For household consumption levels, we use the Consumer Expenditure Survey (CEX) produced by the U.S. Bureau of Labor Statistics (BLS). To calculate the carbon content of each expenditure category, we use input-output tables showing how much each industry consumes of inputs from different industries. These can be used to produce carbon intensities for each industry by tracing indirect carbon consumption levels from when the fuels are burned to when the final products is consumed.

4.2.2 Data and calculations

4.2.2.1 Household expenditures

For expenditure data, I use the quarterly interview survey portion of the Consumer Expenditure Survey. Households participate in the survey for five quarters, with the first quarter discarded from the final data. The quarterly CEX data is combined into extract files (NBER extracts) showing annual expenditures for each participating household by John Sabelhaus and Ed Harris of NBER. I use pooled annual expenditure data for households whose participation in the survey was primarily in 2003 (those that began participating between the 4th quarter of 2002 and the 2nd quarter of 2003.) I use only households that participated in the survey in all four quarters, and use adjusted weights provided by the NBER extracts to account for the households that are lost during this process. This reduces the total number of households in our sample from 7960 to 4470.[11] Expenditures are grouped into 48 categories, including break-downs into fossil-

---

[11] The adjusted weights designed by NBER to make the reduced sample representative assign zero weights to an additional 955 participants, leaving 3515 households with positive weights.

fuel intensive categories such as electricity, natural gas, other household fuels (primarily heating oil) and gasoline.

4.2.2.2 Carbon intensities

To calculate the carbon content of each expenditure categories, we use Input-Output tables from 2003 produced by BLS, combined with EIA data on carbon consumption.[12] The data I use is based on 2003 input output tables. This is the same as used in Boyce and Riddle (2009), but is updated from Metcalf (1999), which uses data from 1992 Benchmark IO tables, and Boyce and Riddle (2007), which use Metcalf's data adjusted for price changes.

Total carbon consumption levels by fossil fuel in 2003 are taken from EIA data[13]. Carbon consumption from coal is attributed evenly to all output from the coal industry. Carbon consumption from oil is attributed primarily to the portion of output of the oil and gas industry that is processed by the petroleum refining industry, and carbon consumption from natural gas is attributed to the output of the oil and natural gas industry that is purchased by other industries[14]. The total carbon intensity of the product of a particular industry i can then be attributed to direct (first level) consumption of fossil fuels (c_{i1}), second level consumption (c_{i2}), which consists of fossil fuels used in

---

[12] The 2003 Input-Output tables do not provide break-downs into sufficiently detailed industry and commodity categories, so I supplement them with 2002 benchmark tables. These are used to determine breakdowns within the 2003 industry and commodity categories.

[13] This data is from EIA's International Energy Annual 2006, available at http://www.eia.doe.gov/iea/carbon.html.

[14] Based on EIA data, 99.4% of oil consumed is processed by refiners, so we attribute this portion of total oil output to consumption by refineries. The rest is divided evenly between all other products of the oil and gas industry, with the exception of that portion that is used by natural gas distributers, which is assumed to be entirely natural gas.

producing the inputs they use, third level consumption (c_{i3}), which consists of those used

in producing the inputs used in producing the inputs they use, and so forth. In matrix

form, we can write $C_{n+1} = A \cdot C_n$, where C_n is a vector of the nth level carbon intensities

for each of the I industries, and A is an $I \times I$ table with entries a_{ij} showing the percent of

industry i's total output value that goes to purchasing the product of industry j. A

vector of total carbon intensity for each industry, C, can then by written as:

$$C = C_1 + C_2 + C_3 + \cdots = \left(I + A + A^2 + \cdots\right) \cdot C_1 \tag{4-1}$$

If A is invertible, this converges to $(I - A)^{-1} \cdot C_1$. This can be translated into

carbon intensities for each commodity using the make table from the I-O accounts, and

adding the direct impact of the charge on the consumption of the commodity.

4.2.2.3 Matching expenditures with carbon intensities

The expenditure categories from the CEX need to be matched with the carbon

intensities from the I-O tables, which are from the National Income and Product

Accounts (NIPA). This is made more difficult by the fact that the data come from

different sources, and are not perfectly consistent with each other. To match the two data

sets, we use a combination of two sources. The first is a bridge matrix provided by the

BEA to match the Input-Output account commodity categories with Personal

Consumption Expenditure (PCE) categories used in the National Income and Product

Accounts. The second is the NBER documentation, which provides relationships

between the PCE categories and the consumption categories in the NBER extracts of the CEX.[15]

An additional difficulty is that the total household expenditure estimates found using these two sources do not match up perfectly, with the data from NIPA being generally larger than the CEX data. These differences can lead to inconsistencies in the estimate of total carbon levels for the economy if they are not reconciled. There are many reasons for the difference, which have been analyzed in some detail (Garner *et al.*, 2006). A number of different approaches have been adopted in studies of carbon control policies to deal with these differences. Dinan and Rogers (2002) adjust CEX expenditures to match with the NIPA estimates. Burtraw *et al.* (2009) use the CEX expenditure data, and adjust the loading factors for indirect expenditures to meet total carbon emissions estimates from EIA data. Boyce and Riddle (2009) adjust the total expenditures from the CEX data to match estimates of household consumption from NIPA, with non-profit consumption removed, but keep the expenditure proportions for different goods from the CEX data. I adopt the approach of Boyce and Riddle (2009), adjusting all expenditures by the same constant ratio so that total household carbon consumption will be consistent with the household share of carbon consumption found using input-output analysis based on NIPA. One difference is that I do not separate household consumption from the consumption of non-profits serving households, since costs borne by these non-profits are likely to ultimately be passed on to households. This

---

[15] In a few cases, the PCE categories provide less detailed break-downs than the IO commodity categories or the CEX consumption categories. Most importantly, natural gas and electricity are grouped into one 'utilities' category in the PCE. As a result, an adjustment is needed to ensure that the natural gas carbon intensity is matched up with the natural gas consumption category, and the electricity carbon intensity is matched up with the electricity consumption category.

leads to an adjustment ratio of 1.62, which is consistent with the ratio of aggregate

expenditures in the CE and PCE found by Garner *et al.* (2006, p. 22).

4.2.3 Differences between poor, rich households

Table 4.2 shows average expenditures patterns for households of different income

levels, based on CEX data. For ease of presentation, the less carbon-intensive of the 48

expenditure categories are aggregated into broader categories, leaving ten broad

expenditure categories: electricity, gasoline, natural gas, heating oil, air transport, other

transport, industrial goods, food, housing, and services/other. Households are broken

down into ten deciles, based on their total expenditure level per household member.[16]

Expenditure on every category of goods increases with total expenditure, but at differing

rates. Poorer households spend a larger share of their expenditures on electricity, natural

gas, heating oil, gasoline, food and housing, while richer households spend a larger share

of their expenditures on air transport, industrial goods and services/other. This suggests a

pattern that will become clearer later – that household fuels are generally necessities that

make up a larger share of the budget of poor households than rich ones. As a result,

policies that raise the price of these goods are likely to be regressive.

---

[16] Per capita household expenditures are used as a proxy for how well-off the household
is. The reasons for this choice are explained in more detail in section 4.5.3.

Table 4.2a. Direct fuel expenditure breakdowns by per capita expenditure decile

Per capita expenditure decile	Per capita expenditure ($)	Electricity expenditure share (%)	Nat. gas expenditure share (%)	Heating oil expenditure share (%)	Gasoline expenditure share (%)
1	4964	6.83%	1.65%	0.49%	6.31%
2	7629	6.37%	2.02%	0.42%	6.39%
3	9925	5.39%	1.49%	0.75%	6.46%
4	12187	4.59%	1.80%	0.43%	5.74%
5	14510	4.27%	1.52%	0.56%	5.57%
6	17290	3.93%	1.64%	0.43%	5.25%
7	20735	3.45%	1.28%	0.44%	4.74%
8	25038	3.17%	1.15%	0.38%	4.57%
9	31460	2.81%	1.09%	0.37%	4.14%
10	53819	2.01%	0.81%	0.29%	2.92%
Average	19756	4.28%	1.45%	0.46%	5.21%

Source: NBER extracts from 2003 Consumer Expenditure Survey

Table 4.2b. Indirect expenditure breakdowns by per capita expenditure decile

Per capita expenditure decile	Per capita expenditure ($)	Air trans. exp. share (%)	Other trans. exp. share (%)	Food exp. share (%)	Indust. good exp. share(%)	Services / oth. exp. share (%)	Housing exp. share (%)
1	4964	0.25%	0.60%	31.37%	11.26%	24.10%	17.14%
2	7629	0.30%	0.20%	28.79%	12.41%	29.74%	13.37%
3	9925	0.41%	0.31%	26.35%	13.97%	32.42%	12.46%
4	12187	0.49%	0.39%	23.98%	16.03%	33.46%	13.10%
5	14510	0.62%	0.30%	22.74%	20.20%	33.26%	10.95%
6	17290	0.68%	0.38%	20.99%	19.27%	36.70%	10.73%
7	20735	0.75%	0.52%	19.63%	22.88%	35.29%	11.02%
8	25038	0.89%	0.43%	18.42%	25.31%	35.46%	10.23%
9	31460	1.00%	0.47%	16.87%	28.75%	34.99%	9.51%
10	53819	1.21%	0.53%	13.78%	36.12%	33.57%	8.77%
Average	19756	0.66%	0.41%	22.29%	20.62%	32.90%	11.73%

Source: NBER extracts from 2003 Consumer Expenditure Survey

4.2.4 Household carbon footprints

Table 4.3 shows the average carbon intensities for each expenditure category.

Not surprisingly, the most carbon intensive categories are direct energy purchases, with

transportation showing intermediate loading factors and services and housing being the least carbon-intensive.

Table 4.3. Carbon intensities by consumption category

Consumption category	tC per $1000 (2003 dollars)
Electricity	2.00
Natural Gas	1.95
Heating Oil	1.92
Car Fuels	1.54
Air Travel	0.37
Other Transport	0.28
Food	0.14
Industrial Goods	0.13
Services and Other	0.12
Housing	0.05

Source: Calculated from 2003 input-output tables; see text for details

Combining these carbon loading factors with the expenditure data from Table 4.2 allows us to calculate the amount of carbon that can be attributed to each household. The results are shown in Table 4.4. For an average household, 61% of the carbon footprint can be attributed to direct fuel expenditures, including electricity, and 39% is due to indirect carbon usage from the consumption of other goods and services. The total carbon footprint of households increases with expenditures. Rich households consume more of every expenditure category, and as a result are responsible for more emissions. However, the carbon intensity of expenditures is higher for poorer households than it is for richer households, reflecting the fact that poor households spend a larger share of their income on more carbon-intensive expenditure categories. This relationship between carbon footprint and total expenditures has important implications for the distributional impacts of a carbon charge, to which we will return later.

141

Table 4.4. Carbon footprints by expenditure

Per capita expenditure decile	Per capita expenditure ($)	Carbon from direct energy expenditures	Carbon from indirect sources	Total carbon footprint per capita	Carbon intensity of expend.
1	4964	1368	524	1892	0.38
2	7629	2086	815	2901	0.38
3	9925	2489	1029	3518	0.35
4	12187	2726	1240	3965	0.33
5	14510	3075	1540	4616	0.32
6	17290	3457	1831	5288	0.31
7	20735	3642	2211	5853	0.28
8	25038	4099	2698	6797	0.27
9	31460	4674	3456	8130	0.26
10	53819	5739	6159	11898	0.22
Average	19756	3336	2150	5486	0.31

Source: author's calculations from Tables 4.2 and 4.3.

4.3 Implementation of charge

To reduce carbon emissions, economists have generally suggested that the most effective methods are policies that put a price on carbon – either through a carbon tax or through a cap-and-trade program. There are many variations on how this can be done, which I will explore in some detail in this section.

4.3.1 Prices vs. Quantities

A price can be placed on carbon emissions either through a carbon tax, or through emission permits, the total number of which is set by a cap. Permits would fix the permissible quantity of emissions, and let the market determine the price, while a tax would fix the price of emissions and let the market determine the quantity. There are advantages to both approaches. Fixing the number of permits would ensure that a scientifically or politically determined target will be met, while fixing the price could help limit the economic costs of the policy if abatement costs turn out to be different than

142

expected. Which policy is likely to be more efficient in the face of uncertainty about abatement costs and climate benefits of a policy depends on the elasticities of the abatement cost and climate benefit curves (Weitzman, 1974). To see why, consider a case where the climate benefits are known and abatement costs are uncertain. If the marginal climate benefits of each unit of abatement are roughly constant, a shift in the abatement cost curve will change the target quantity much more than the target price, so fixing the target price would be more efficient. On the other hand, if the marginal climate benefits decrease dramatically as emissions are reduced, then fixing the quantity can be more efficient. In the case of climate control policy, some researchers have found that a tax is likely to be more efficient (Pizer, 1997).

Some alternative hybrid approaches have also been proposed. A cap with a maximum permit price, or safety valve, is commonly included in policy proposals, intended to limit the economic costs of the policy (Pizer, 1997). However, this approach causes two problems. First, it fails to protect again downward volatility in permit prices, which has proven to be more common in existing policies. Second, it creates insufficient incentive for investment in new technologies (Burtraw *et al.*, 2009). A better approach would be a symmetric safety valve, with both a maximum and a minimum price, so that if abatement costs turned out to be lower than expected, there would be an incentive to reduce emissions beyond the target set by the cap (Burtraw *et al.*, 2009).

While the choice of a price or quantity-based policy may be important in the face of uncertainty, it does not have a significant impact on the distributional impacts of the policy. If a tax leads to the same amount of emissions reduction as an equivalent cap-and-trade program, the impacts of both policies will be the same. In this paper, I use a

143

fixed price, rather than a fixed quantity, in all scenarios I analyze. However, the results can be interpreted either as the result of a carbon tax with that price, or of a cap-and-permit policy that leads to permits being purchased at that price. I refer to both policies interchangeably in the remainder of the paper.

4.3.2 Emissions Covered by Policy

The policies evaluated here cover all emissions of carbon dioxide from the burning of fossil fuels that is associated with U.S. consumption.[17] Covering all emissions from fossil fuel consumption can be done most efficiently through an upstream charge that is imposed on fossil fuel producers and importers as the carbon enters the economy. With this approach, there would be only about 2000 collection points in the U.S., keeping administrative costs to a minimum (Kopp *et al.*, 1999; CBO, 2001).

For the charge to cover emissions associated with all U.S. consumption, but not foreign consumption, import tariffs and export subsidies would need to be imposed on all products based on their carbon content. This would eliminate price differences between domestically and foreign produced goods, reducing the impact that the policy would have on U.S. exporters and U.S. producers that compete with importers. There are some concerns about how this approach would be implemented and whether it could run into problems with international trade agreements (Kang 2010), but a recent WTO report on the subject opened the door for some forms of carbon border tax adjustments to be allowed (WTO 2009). Border adjustments also may not be necessary if our trading partners also have carbon control policies in place. However, it is a useful assumption in

---

[17] As mentioned earlier, other greenhouse gasses and land-use related carbon emissions are not covered. These would have to be addressed separately.

that it simplifies the analysis by avoiding questions about the effect on domestic producers of an increase in domestic production costs that does not affect international production costs. In many studies, this assumption is made implicitly without being stated explicitly by the combination of assumptions that all costs are passed on to final consumers and that all of those costs can be attributed to consumers in the US.

4.3.3 Keeping the government whole

A carbon charge will impose both direct and indirect costs on state, local and national governments. In this paper, I adopt the assumption that the government will set aside enough revenues to offset these costs, so that the policy is revenue-neutral. This is the assumption made by most studies, either explicitly or implicitly by only counting revenues collected from consumers (CBO, 2000; Barnes and Breslow, 2003; Metcalf, 2002; Boyce and Riddle, 2007; Burtraw *et al.*, 2009).

Boyce and Riddle (2008) look at this question in more detail. They find that the government could offset its costs if they distributed their revenues to households as dividends, and taxed those dividends as income. While taxing dividends has advantages, it is more straightforward to assume that dividends will not be taxed and that the government will withhold enough money to offset increases in government payments due to the policy.

4.3.4 Permit price

In all the scenarios analyzed in this paper, I use a carbon charge or permit price of $100 per metric ton of carbon, or $27.3 per metric ton of carbon dioxide. This is within the range of current legislative proposals. For example, CBO (2009) estimates that the

145

Waxman-Markey bill that was passed by the House in 2009 would lead to a permit price of $28 per metric ton of carbon dioxide in 2020.

The permit price could vary significantly depending on the details of the policy. Changes in the permit price would change the magnitude of the effect of the charge on households, the amount that carbon emissions are reduced, and the extent of the adjustment costs that producers and consumers would face. However, the relative impact on households of different income levels should not change significantly if the carbon price changes, so the distributional results found in this paper should be valid for a range of possible permit prices.

4.3.5 Distributing Revenues (or permit value)

The most important aspect of the policy design for studies of distributional impacts is how the government distributes the value of the permits, or the revenue from a carbon charge. The government can give some or all of the permits for free, or they can auction the permits (or collect tax revenue) and use the money they collect for a wide range of purposes. Giving away free permits is essentially equivalent to auctioning the permits and giving away the permit value, so it is simpler to assume that the permits are sold initially and consider different methods for distributing the revenue.

The uses of the revenue that are most often discussed fall into three categories: returning revenues to households, compensating other parties that might be harmed by the policy, or investing in technologies that could aid in the clean energy transition. Each of these could be done in different ways. Revenues could be returned to households through lump-sum payments, reductions in taxes, or distributions to electricity rate payers by way of utility companies. Other parties that may be compensated include fossil fuel

companies and exporters of carbon-intensive products, workers in coal mining and other industries that may face job losses, and state and local governments that face higher fossil fuels prices. Finally, a wide variety of technologies could be targeted for investments, ranging from renewable energy technologies such as wind and solar, to 'clean' coal and carbon sequestration. The impacts of a policy will depend greatly on which of these uses of the revenue are chosen.

I will consider the distributional impacts of two of the most commonly analyzed proposals: a cap-and-dividend policy where all revenues not needed to offset government expenditures are returned to households as equal per capita dividends; and a hybrid policy where some of the revenues are given to producers and some of the revenues returned to households as dividends.

4.3.6 Time frame

The analysis conducted here is a short-run, static analysis. It looks at the effects that might be expected in the first few years after the implementation of the policy. Also, I do not try to project changes in baseline expenditures or emissions in the absence of a policy; instead, I assume that the expenditures and emissions with no policy would be the same as they were in 2003, the year from which the data for the analysis are taken.

4.4 Modeling Response to Carbon Charge

As a carbon charge is applied, producers and consumers will respond to this by adjusting their prices and consumption levels. Producers, faced with higher prices for fossil fuels and other inputs, are likely to raise output prices, and may also adjust their production levels and the ratio of inputs that they consume. Consumers will respond to

changes in prices by adjusting their consumption levels of different goods. It is through this process that a carbon charge works to reduce carbon emissions. The effect of the carbon charge on companies and households depends on how they respond to the new price structure.

4.4.1 Supply model

In analyses of the incidence of carbon charges, the most common approach to determining how a carbon charge would affect commodity prices has been to use input-output accounts to trace price changes through the economy, based on work by Leontief (1986). The methodology commonly used with this approach is based on the assumption that suppliers do not adjust their input ratios, and the entire charge is passed on to consumers through higher prices (Hassett *et al.*, 2007, p. 21).

The assumption of fixed input ratios eliminates an important mechanism through which emissions could be reduced in response to a carbon charge. With higher fuel prices, companies may be able to engage in fuel substitution or other process shifting in the short to medium run, and may invest in more efficient technologies in the long run, in order to reduce their use of the higher priced inputs without reducing output (EAAC, 2010, pp. 23-25). A number of models have been built to model supplier response to carbon pricing policies (see Stern, 2006 for a review). However, most of these models are designed to determine the overall costs of climate control policies, and not the incidence of the policies on different groups.

Our model does not attempt to model supplier efforts to adjust input ratios in response to the new price structure; instead we adopt the usual assumption that input ratios are fixed. As a result, we will most likely underestimate how much total carbon

emissions will be reduced in response to a rise in prices. However, the incidence calculations that are central to this paper should not be much affected by this assumption. The extent to which it could affect the incidence calculations will be discussed later.

The other simplifying assumption used in standard input-output analysis is that price increases are fully passed through to consumers. This plays a more important role in determining the incidence of the charge. The assumption of full pass-through is commonly used, and has been supported by some general equilibrium models that have found that consumption taxes are fully passed forward to consumers (Holak *et al.*, 2008, p. 33). It is consistent with a model of perfect competition with constant returns to scale if there are no fixed costs, and it simplifies the analysis considerably.

However, there are several reasons to question the validity of this assumption in models of supplier response to carbon charges. In any short-run analysis, fixed cost can be expected to lead to supplier adjustment costs and incomplete pass-through. In the electricity industry, fixed costs can be particularly long-lived, as power plants can remain in use for many decades (Burtraw *et al.*, 2009a, p. 20). Extractors of fossil fuels may also have inelastic supply curves due to the limited availability of the resource (Holak *et al.*, 2008, p. 33). Also, industries may not be perfectly competitive, and firms may change the amount that they mark up their final price above their marginal costs in response to changes in input prices.

A few studies have looked at how the incidence of a carbon charge would change if a portion of the costs of the charge is borne by producers rather than consumers (Shah and Larsen, 1992; Boyce and Riddle, 2007). These studies do not explicitly model how prices change in the presence of partial pass-through; instead, they simply combine

incidence results for households from a full pass-through scenario with information on the distribution of stock ownership by income level to determine the overall incidence with different assumptions about the share of costs that are borne by households.

A few studies have used explicit models of supplier behavior in key industries to look at how the costs of a carbon charge might be broken down between producers and consumers. Bovenberg and Goulder (2000) focus on the role of fixed capital, and look at the loss of producer surplus that would occur in the short run in response to cost increase resulting from a carbon charge. Holak *et al.* (2008) uses the EPPA model developed at MIT to determine how much suppliers will raise prices in response to a system of carbon permits. Burtraw *et al.* (2009a) assume full pass-through in all industries except electricity generation, but use a complicated model of the electricity market developed by Resources For the Future, known as Haiku, to analyze the effects of fuel price increases on the electricity industry.

No studies that I am aware of have evaluated how changing the pass-through assumption could influence the commodity price increases and supplier impacts that would result from a carbon charge, or used this to determine the distributional incidence. This study aims to fill this gap by evaluating several alternative pass-through assumptions using a variation of the commonly used approach to input-output analysis. In addition to a full pass-through scenario, two alternative scenarios with different pass-through rates assigned uniformly to all industries are considered, as well as one scenario with variable pass-through rates by industry.

4.4.1.1 Full pass-through

The baseline analysis in this paper takes the commonly used assumption that costs faced by suppliers will be fully passed on to consumers. With this assumption, the full impact of the carbon charge will be borne by consumers, in proportion with their share of carbon consumption. The price increase for each good can be calculated simply by taking the carbon intensity calculations from section 4.2.2.2, and multiplying by the carbon charge or permit price. Here we use an alternative way of making this calculation that produces the same result, but that is easier to adjust in subsequent alternative scenarios. First, we set up equations that show how the price of the product of each industry depends on the prices of the inputs, as well as on the direct charge that they face:

$$x_{11}(p_1 + t_{11}) + x_{21}(p_2 + t_{21}) + \cdots + x_{N1}(p_N + t_{N1}) + v_1 = x_1 p_1$$

$$x_{12}(p_1 + t_{12}) + x_{22}(p_2 + t_{22}) + \cdots + x_{N2}(p_N + t_{N2}) + v_2 = x_2 p_2 \qquad (3\text{-}2)$$

$$\ldots$$

$$x_{1N}(p_1 + t_{1N}) + x_{2N}(p_2 + t_{2N}) + \cdots + x_{NN}(p_N + t_{NN}) + v_N = x_N p_N$$

In these equations, x_{ij} is the consumption of the product of industry i by industry j, x_i is the total production of industry i, p_i is the price of good i, and t_{ij} is the tax on industry j's consumption of industry i's product, expressed as dollars per unit of industry i's goods.

We can break prices down into their initial prices from the input-output tables, $\bar{p}_i$, plus the change in their price, Δp_i:

$$p_i = \bar{p}_i + \Delta p_i \qquad (3\text{-}3)$$

The initial prices satisfy the input-output table identities for each industry:

$$x_{1i}\bar{p}_1 + x_{2i}\bar{p}_2 + \cdots + x_{Ni}\bar{p}_N + v_i = x_i\bar{p}_i \tag{3-4}$$

By breaking down the prices and subtracting the input-output table identities from each line, we get:

$$x_{11}(\Delta p_1 + t_{11}) + x_{21}(\Delta p_2 + t_{21}) + \cdots + x_{N1}(\Delta p_N + t_{N1}) = x_1\Delta p_1$$

$$x_{12}(\Delta p_1 + t_{12}) + x_{22}(\Delta p_2 + t_{22}) + \cdots + x_{N2}(\Delta p_N + t_{N2}) = x_2\Delta p_2 \tag{3-5}$$

$$\ldots$$

$$x_{1N}(\Delta p_1 + t_{1N}) + x_{2N}(\Delta p_2 + t_{2N}) + \cdots + x_{NN}(\Delta p_N + t_{NN}) = x_N\Delta p_N$$

After dividing both sides of each equation by x_i, this can be expressed in matrix form as:

$$A'\Delta P + D_t = \Delta P \tag{3-6}$$

A is a matrix with entries $a_{ij} = \dfrac{x_{ij}}{x_j}$ as before, and D_t is a vector with the direct impacts of the carbon charge on input prices, $d_i = a_{1i}t_{1i} + a_{2i}t_{2i} + \cdots + a_{Ni}t_{Ni}$. Solving for ΔP, we get

$$\Delta P = (I - A')^{-1} \cdot D_t. \tag{3-7}$$

This gives us the increases in prices paid to each industry for their product. As before, to translate this into commodity prices, I multiply by a make table matrix, and add the direct tax placed on final consumption of each commodity. The household share of carbon consumption is calculated from the commodity price increases using lines from the input-output tables that give the amount of final consumption and private investment of each commodity by households. The government share comes from eight lines of the

input-output accounts representing federal, state and local government final consumption and investment.

This calculation is similar to that presented in Metcalf (1999) and in Hassett *et al.* (2007), but with one important difference: Metcalf expresses the charge levied on each fossil fuel as a percentage increase in the price of the fuel (an ad valorem tax), while we express it as a charge that is in proportion to the quantity of the fuel consumed (a quantity tax). This makes a difference, even in a static analysis where prices aren't changing independently over time, because the tax will increase the cost of inputs used in extracting the fuels, and therefore the pre-tax fuel prices. This is a slight problem with the widely used loading factors calculated by Metcalf, as the tax rates they use are designed to collect a specified amount of revenue from each industry at the initial fuel prices, but the revenues collected will actually be a bit higher due to the rise in pre-tax fuel prices. The size of the difference may also vary between fuels, so the relative charges for different fuels may also be slightly off. Carbon intensities based on ad valorem tax calculations using Metcalf's methodology are used in virtually all distributional analyses of carbon charges in the U.S., with the exception of Boyce and Riddle (2009). The approach used in this paper corrects this problem.

4.4.1.2 Partial pass-through

The first alternative supply model I examine is one in which all industries pass the same fixed portion of their costs onto consumers. In this model, all costs faced by producers are paid either directly to the carbon charge or indirectly via higher input prices – there are no additional adjustment costs. This is consistent with a model with no fixed costs, constant marginal costs and fixed input ratios, but with imperfect competition, so

153

that companies change the amount that they mark up their product in response to a change in costs.

To model this, I modify equations (4-5), so that only a portion r of the increase in costs will be passed on as higher prices:

$$r \cdot \left[x_{11}(\Delta p_1 + t_{11}) + x_{21}(\Delta p_2 + t_{21}) + \cdots + x_{N1}(\Delta p_N + t_{N1}) \right] = x_1 \Delta p_1$$

$$r \cdot \left[x_{12}(\Delta p_1 + t_{12}) + x_{22}(\Delta p_2 + t_{22}) + \cdots + x_{N2}(\Delta p_N + t_{N2}) \right] = x_2 \Delta p_2 \qquad (4\text{-}8)$$

$$\ldots$$

$$r \cdot \left[x_{1N}(\Delta p_1 + t_{1N}) + x_{2N}(\Delta p_2 + t_{2N}) + \cdots + x_{NN}(\Delta p_N + t_{NN}) \right] = x_N \Delta p_N$$

After dividing both sides of each equation by x_i, this can be expressed in matrix form as:

$$r \cdot A' \Delta P + r \cdot D_t = \Delta P \qquad (4\text{-}9)$$

As before, A is a matrix with entries $a_{ij} = \dfrac{x_{ij}}{x_j}$ and D_t is a vector with entries

$d_i = a_{1i} t_{1i} + a_{2i} t_{2i} + \cdots + a_{Ni} t_{Ni}$. Solving for ΔP, we get:

$$\Delta P = (I - r \cdot A')^{-1} \cdot r \cdot D_t. \qquad (4\text{-}10)$$

As before, to translate this into commodity prices, I multiply by a make table matrix, and add the direct tax placed on final consumption of each commodity. The household share of policy costs is calculated from the commodity price increases using lines from the input-output tables that give the amount of final consumption and private investment of each commodity by households. The government share of costs is calculated from lines that give consumption and investment levels for each commodity by state, local and national governments. The share of costs borne by each industry is found

154

by multiplying the cost increase they face by $1 - r$, and the total industry share is the sum of the individual industry shares.[18]

A portion of the change in industry profits will be passed on to governments via reduced payments into corporate income taxes. We follow CBO (2000, p. 20) in assuming that 45% of a change in corporate profits – whether an increase or a decrease – will be passed on to governments.[19]

It should be noted that we use the initial output and consumption numbers from the input-output accounts in these calculations, and do not adjust for changes in demand. If we used adjusted demand quantities, it would not affect the price increases, since they depend only on the input shares for each industry which would not change. It would have a small effect on the shares of the carbon cost borne by consumers, producers and government. However, adjusting for this would require a complicated iterative process, since quantities consumed would affect the share of corporate burden and the size of dividends, which would affect consumer incomes and therefore the quantities consumed, and so on. The minor changes that would result from this process do not justify this added effort.

Changes in the quantity demanded could have a more substantial effect on individual supplier profits. However, the effects of these changes in demand on the profits of different industries should roughly offset each other, because with revenue

---

[18] The import and export lines of final consumption from the input-output tables are not included in the calculation of total permit value, since the policy being analyzed only covers emissions associated with domestic consumption. Private inventory adjustments are also not included.

[19] This assumed rate seems unrealistically high, but we adopt it since this CBO study appears to be the most reliable source that has made this calculation in a comparable setting.

recycling and no supplier adjustment costs, total demand will not change. Since we don't

have data on which households own shares in which industries, we are more interested in

changes in total industry profits than in identifying the effect on specific industries, so we

do not calculate the effects of changing demand on the profits of particular industries.

4.4.1.2.1 Pass-through rate

The rate of pass-through depends on several factors, including the degree of

market power of the firms, the shape of the demand and marginal cost curves, and even

consumers attitudes about the legitimacy of the price increases. In some situations, more

than 100% pass-through may be possible if companies can use the new policy as an

excuse to raise prices more than necessary.

There is surprisingly little empirical work that addresses this question. The few

studies that do look at rates of pass-through of price increases are based on estimates of

supply and demand curve elasticities, but do not consider whether there will be changes

in the mark-up rate. Holak *et al.* (2008, p. 20), in their 287 bmt emissions scenario, find

that coal producers will pass on 98% of their costs to consumers in 2015, and 94% in

2030; oil producers will pass on 89% of their costs to consumers in 2015 and 84% in

2030; and natural gas producers will pass on 73% of their costs to consumers in 2015 and

252% in 2030. The rise in natural gas prices in 2030 is due to the model's finding that

there will be a large increase in demand for natural gas as electricity producers shift from

coal to natural gas production. Burtraw *et al.* (2009, p. 20), report that in the electricity

industry, consumer impacts are eight times as great as producer impacts. Bovenberg and

Goulder (2000) do not explicitly provide the pass-through rates that result from their

analysis, but suggest that the costs to producers are relatively small compared to the total

value of the permits: at a carbon tax rate of $25 per ton, producer equity values could be maintained by giving 4.3% of permits to the coal industry and 15% to the oil and gas industry.

Theoretical models of price setting in a monopoly do not provide much guidance either. If the monopoly profit-maximizing price and quantity are p^* and q^* and the inverse demand function is $p = DI(q)$, and the marginal cost function is $MC(q)$, then the change in p^* in response to a marginal change in costs is given by:

$$\frac{\partial p^*}{\partial c} = \frac{DI'(q^*)}{DI''(q^*) - MC'(q^*) + 2DI'(q^*)}. \tag{4-11}$$

This can be any number depending on the shape of the demand curve. Even knowing the elasticity of the demand curve does not help, since it depends crucially not only on the derivative but also the second derivative of demand. In the simple case with a linear demand curve and constant marginal costs, this simplifies to 0.5, or 50% pass-through.

We analyze two alternative pass-through scenarios: one with 90% pass-through, and one with 50% pass through. The 50% pass-through scenario is intended to illustrate an extreme case, based on the monopoly model with linear demand, which will demonstrate the effect of large changes in this assumption. The 90% pass-through scenario is intended to be a more realistic illustration of the impact of a carbon charge with partial pass-through, and is chosen as a round number that is roughly in line with an average of the estimated pass-through rates found in the literature.

It should be noted that the rate of pass-through is not the same as the portion of the charge that is ultimately borne by consumers. When 90% of the cost increases are

passed through several intermediate industries before reaching consumers, each industry

passes only a portion the cost increases it faces on to the next industry in line, so the

amount of the charge that is ultimately passed on to consumers will be less than 90%.

4.4.1.3 Demand-dependent pass-through rates

The second alternative supply model I analyze is based on a model of perfect

competition with fixed costs and increasing marginal costs. In this model, the change in

the price of each product depends not only on the shift in the marginal cost curve, but

also on the change in demand for each product.

With revenue recycling, nominal incomes will rise, shifting the demand curve up

at the same time as costs are increasing. Some products will see increases in the quantity

demanded and others will see decreases, depending on whether the demand increase or

the marginal cost increase is larger. Products that are more carbon-intensive, and

therefore face a greater increase in marginal costs, are more likely to see a decrease in

quantity demanded, while less carbon-intensive products are more likely to see an

increase in quantity demanded. If the quantity demanded increases, this will lead to a

movement up the marginal cost curve, so prices will rise by more than the shift in

marginal costs. If the quantity demanded decreases, prices will rise by less than the

increase in marginal costs. In other words, pass-through rates will be less that 100% for

carbon-intensive goods, and greater than 100% for goods that don't use as much carbon.

To illustrate the effects of pass-through rates that vary based on the carbon

content of the product, I set the pass-through rates using a simple linear demand and

supply model. First, I assume that the increase in demand from revenue recycling will be

the same for all products, and will be the same as the increase in marginal costs in an

158

industry with average carbon content. This way, industries with average carbon content will have 100% pass-through, industries with above average carbon content will have less than 100% pass-through, and industries with below average carbon content will have more than 100% pass-through. Since the average pass-through rate is about 100%, the average impact on consumers should be similar to the impact in the full pass-through scenario; the difference will be in the relative prices of the different products.

Second, I use linear demand and supply curves, with identical price elasticities at the point of consumption of ε_d=-0.14 and ε_s=1.0. While these assumptions are not perfectly consistent with the assumptions of the full demand model discussed in the next section, they are sufficient to provide a set of pass-through rates that vary by industry based on the carbon content of the industry's product. [20]

With these assumptions, the change in price that results from a shift in the marginal cost curve by $\Delta MC_i(q_i)$ can be calculated with the set of equations:

$$\Delta p_i = r \cdot \Delta MC_i(q_i) + s$$

$$r = \frac{\varepsilon_s}{\varepsilon_s - \varepsilon_d} \qquad\qquad (4\text{-}12)$$

$$s = (1 - r) \cdot \overline{\Delta MC}$$

---

[20] The demand elasticity is chosen to replicate the amount of demand shifting that occurs in the full demand model, as measured by the reduction in carbon emissions. It is less than the final demand elasticities used in the demand model for several reasons; the most significant is that it reflects the elasticities of demand from all users, including intermediate users as well as end users, and intermediate demand elasticities are small because input ratios are sticky or fixed.

Here, r is the rate of pass-through of cost increases, s is the increase in prices caused by the increase in demand, and $\overline{\Delta MC}$ is the average of the marginal cost curve increases in a 100% pass-through scenario, weighted by industry output.

This leads to the system of equations:

$$r \cdot \left[\frac{x_{11}}{x_1} (\Delta p_1 + t_{11}) + \frac{x_{21}}{x_1} (\Delta p_2 + t_{21}) + \cdots + \frac{x_{N1}}{x_1} (\Delta p_N + t_{N1}) \right] + s = \Delta p_1$$

$$r \cdot \left[\frac{x_{12}}{x_2} (\Delta p_1 + t_{12}) + \frac{x_{22}}{x_2} (\Delta p_2 + t_{22}) + \cdots + \frac{x_{N2}}{x_2} (\Delta p_N + t_{N2}) \right] + s = \Delta p_2 \qquad \text{(4-13)}$$

$$\ldots$$

$$r \cdot \left[\frac{x_{1N}}{x_N} (\Delta p_1 + t_{1N}) + \frac{x_{2N}}{x_N} (\Delta p_2 + t_{2N}) + \cdots + \frac{x_{NN}}{x_N} (\Delta p_N + t_{NN}) \right] + s = \Delta p_N$$

This can be solved for ΔP to give

$$\Delta P = (I - r \cdot A')^{-1} \cdot (r \cdot D_t + S) \qquad \text{(4-14)}$$

As before, A is a matrix with entries $a_{ij} = \dfrac{x_{ij}}{x_j}$, and D_t is a vector with entries

$d_i = a_{1i} t_{1i} + a_{2i} t_{2i} + \cdots + a_{Ni} t_{Ni}$. S is a vector with all entries equal to s.

The price increases by commodity and the impact on consumers in this scenario can be calculated from ΔP as before. However, the impact on producers is more complicated. The amount that producers pay into the carbon charge, either directly or indirectly through changes in input prices, can be calculate as before, by subtracting the increase in price from the increase in the marginal cost curve. This amount could be positive or negative depending on the rate of pass-through.

In addition to this cost, firms may also face transition costs caused by the change in quantity demanded, which represent a true economic cost of the policy rather than a

transfer. In a standard short-run market supply model under perfect competition with fixed costs, average costs are at a minimum near the equilibrium point. If the quantity demanded shifts either right or left, average costs will increase. The size of these transition costs can be calculated by finding the area of the triangle between the original marginal cost curve and the original price level, running from the old to the new quantity level, using the formula:

$$tc_i = \frac{\Delta q_i \cdot (\Delta p_i - \Delta MC_i(q_i))}{2} \tag{4-15}$$

The changes in quantities for each industry are estimated based on the same supply and demand elasticities used to estimate pass-through rates for each industry. These transition costs are added to the charge costs faced by producers. This makes the total costs faced by industries and consumers add to more than the total revenue available to be returned to households.

Since we are assuming that input ratios are fixed, these cost calculations do not include the effects of supplier efforts to reduce emissions on the costs faced by suppliers. Since abatement efforts are done because they become profitable with the new price structure, these efforts should reduce the costs faced by suppliers in response to a given increase in input prices. For example, if the carbon charge makes natural gas cheaper than oil, the cost increase faced by producers will be less if they do switch to natural gas than if they don't. This on its own should not have much effect on whether producers or consumers would bear more of the cost of the charge, however. A smaller cost increase would mean lower costs for both producers and consumers, assuming that the same portion of the cost increase is passed on to consumers in both cases.

The distributional impacts could be affected slightly by the decrease in the amount paid into the charge due to the abatement efforts, which would leave less money available to distribute back to households. Payments into the charge would decrease by more than the decrease in supplier costs, so this would be another source of economic costs of the charge. A full discussion of the effects of this assumption on the economic costs of the charge is put off until after discussing the demand model, since the abatement costs faced by consumers would also be affected by the smaller price increases that would result from this abatement activity.

4.4.2 Demand model

As with supply, the most common demand model used in studies of the incidence of carbon charges is the simplest: that demand does not change in response to changes in product prices (Burtraw *et al.*, 2009, p. 4). This approach is taken by several U.S. studies, including Metcalf (1999), CBO (2000) and Barnes and Breslow (2003).

On the other extreme are studies that estimate a complete demand model directly from a panel of expenditure survey data, combined with commodity price data. This approach was taken in studies of the UK (Symons et al. 1994) and Spain (Labandeira *et al.* 2004), but has not been done in the US. These studies combine data from a number of sources to estimate parameters representing own price elsticities and income elasticities, as well as cross price elasticities for every pair of products. They use this fitted demand model to predict demand responses to price changes caused by a carbon charge policy.

Other studies take intermediate approaches. Cornwell and Creedy (1996, pp 25-26) estimate income (or total expenditure) elasticities from expenditure survey data, and derive price elasticities from them using theoretical restrictions based on the assumption

of an additive utility function. Burtraw *et al.* (2009a) use a partial equilibrium analysis where they adjust consumption levels of each good independently based on elasticity estimates from the literature. However, they do not adjust expenditures in response to income changes caused by revenue recycling. Boyce and Riddle (2007; 2008; 2009) use price elasticity estimates from the literature to adjust expenditures in response to changes in relative prices, and also increase or decrease expenditures on all products in response to income changes.

There are problems with each of these approaches. Failing to account for demand changes does not allow an analysis to address one of the goals of a climate policy – to stimulate behavioral responses that will reduce emissions – and may lead to an overstatement of the costs faced by consumers. On the other hand, directly estimating demand model parameters, as in Symons *et al.* (1994) and Labandeira *et al.* (2004), requires very extensive data, and even with good data produces imprecise parameter estimates. For some parameters this is acceptable, but for the key parameters, particularly the own-price elasticities of demand for fuels, it may be preferable to use elasticity estimates that are based on results from the large body of work that has focused specifically on estimating these parameters. Cornwell and Creedy's (1996) approach reduces the data requirements, but their price elasticity estimates are even less reliable, as they are derived from questionable theoretical assumptions, rather than from data.

Burtraw *et al.* (2009) use more reliable price elasticity estimates from the literature, but by failing to account for demand responses to changes in income they are missing what could be an important part of the picture of how demand will respond to a carbon charge. Their approach would predict that households that receive more than

163

enough in dividends to compensate for the higher prices they face will still reduce their expenditure levels. This will lead to estimated emission reductions that are due as much to decreases in overall expenditure as to shifts from more carbon-intensive to less carbon-intensive products. Boyce and Riddle (2007; 2008; 2009) produce more reasonable estimates of demand response by accounting for responses to income changes as well as price changes, but there is little theoretical basis for the demand response rules they use.

The approach taken in this paper is to model how each individual in the expenditure survey adjusts their expenditures in response to the new price levels, and construct decile averages based on these individual responses. The demand model used to adjust expenditures satisfies theoretical restrictions on demand such as homogeneity and additivity, and is consistent with the own-price demand elasticities used by Boyce and Riddle, as well as initial consumption shares from the consumer expenditure survey.

The demand model I use to implement this is a simplified version of the Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980). The AIDS is among the most commonly used demand systems in studies of how demand responds to price changes, including studies that focus on demand responses to changes in energy prices (Labandeira *et al.*, 2004, Symons *et al.*, 1994, West and Williams, 2004). It has a number of advantages: it is flexible, allowing price elasticities of demand for each good to be specified; it satisfies the homogeneity and additivity properties of a demand system as long as the parameters used follow simple rules; and it is consistent with an indirect utility function that can be used to estimate changes in utility in response to demand changes.

The full AIDS demand equations are given by:

$$w_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log\left(\frac{x}{P}\right) \tag{4-16}$$

In this equation, w_i is the budget share $\dfrac{p_i q_i}{x}$ for good i, p_j is the price of good j,

x is total expenditures on all goods, P is a price index term[21], α_i are parameters that

determine the relative demand for each product if prices were equal, γ_{ij} are parameters

that determine both the own-price and cross-price elasticities for each product, and β_i are

parameters that determine the income elasticity of demand for each product. In order for

the model to satisfy homogeneity and additivity, the parameters must satisfy the

following set of restrictions:

$$\sum_i \alpha_i = 1$$

$$\sum_i \gamma_{ij} = 0 \tag{4-17}$$

$$\sum_j \gamma_{ij} = 0$$

$$\sum_i \beta_i = 0$$

To simplify the model, I assume, as in Boyce and Riddle (2009), that if prices are

fixed, a change in total expenditure will lead to a proportionate change in expenditure on

each good. This is equivalent to assuming $\beta_i = 0$ for all i, so it simplifies the demand

equations to:

$$(18) \quad w_i = \alpha_i + \sum_j \gamma_{ij} \log p_j$$

---

[21] See Deaton and Muellbauer (1980, p. 314), for a definition of the price index term. I
do not discuss it here, because with the simplifications I make, it is no longer needed.

This simplifies the analysis considerably, as it allows me to avoid the use of the price index parameter, P, and ensures that carbon consumption will be a linear function of income, making equilibrium calculations that will be discussed later much more straightforward.

If total expenditures change by the same amount as income, this would be equivalent to assuming income elasticities of 1.0 for all goods. However, total expenditures may not change by as much as income changes. I assume that 90% of income changes from revenue recycling will be translated into increased expenditures in the short run. This is based on estimates in Carroll (2001), of the effect of an unexpected permanent change in income on consumption, using a theoretical model with parameters based in real data. A review by Jappelli and Pistaferri (2009), suggests that this is the best estimate of its kind, and that empirical studies that try to identify the effect of this type of change in income have found mixed results, some higher than 0.9 (0.91-1.02, p. 37) and some lower (0.65, p. 41).

The γ_{ij} parameters are chosen using a two-part approach, with the own-price elasticity parameters γ_{ii} specified first, followed by the cross-price elasticity parameters γ_{ij}, $i \neq j$. The own-price elasticity parameters are chosen to be consistent with estimated elasticities from the literature. The most important elasticities in determining how demand will respond to a carbon charge are the own-price elasticities for energy products. These have been extensively studied, and these studies have been summarized in several reviews (Bohi, 1981; Dahl, 1993). The price elasticities for heating oil, gasoline, natural gas and electricity are short-run elasticities taken from Dahl (1993), which is the most comprehensive review of this literature. The price elasiticities for other

products are chosen based on Williamson's (2006) "stylized facts of demand". They are

presented in Table 4.5.

Table 4.5. Price elasticities of demand

Consumption Category	Own price elasticity of demand (short run)
Food	0.60
Industrial Goods	1.30
Services/Other	1.00
Electricity	0.20
Natural Gas	0.20
Heating Oil	0.27
Car Fuels	0.26
Air Transport	0.25
Other Transport	0.25

sources: Dahl (1993); Willamson(2006)

The own-price elasticity parameters γ_{ii} are related to actual own-price elasticities

based on the formula:

$$\gamma_{ii} = (\varepsilon_{ii} + 1) \cdot w_i \tag{4-19}$$

Where ε_{ii} is the own price elasticity from the literature, and w_i is the share of

expenditure on good i. I assume that all individuals have the same own price elasticity

parameters γ_{ii}, and calculate these common parameter values using aggregate CEX data

on expenditure shares, so that the aggregate elasticity estimate will match the elasticities

from the literature. Individual elasticities may be different from the aggregate elasticity if

the expenditure shares for the individual are different from the aggregate.

Cross price elasticities for each pair of products are not easily available, so instead

of using estimates from the literature, I use restrictions on the parameters in the AIDS

demand models to generate cross-price elasticity parameters that will satisfy these

conditions. This requires making the simplifying assumption that the cross-price

elasticities between two products depend only on the own price elasticities of the two

products. Products that have high (negative) own price elasticities, and therefore are

easily substitutable will also have high cross price elasticities. We do not try to identify

pairs of products that might substitute for each other particularly well for other reasons.

Based on this assumption, the cross-price elasticity parameters needed to satisfy

the parameter restrictions $\sum_i \gamma_{ij} = 0$, and $\sum_j \gamma_{ij} = 0$ are found to be:

$$\gamma_{ij} = \frac{(\bar{\gamma} - \gamma_{ii}) + (\bar{\gamma} - \gamma_{jj})}{n-2} - \frac{\bar{\gamma}}{n-1} \tag{4-20}$$

$$\bar{\gamma} = \frac{\sum_i \gamma_{ii}}{n}$$

One shortcoming of the AIDS demand model is that it can predict negative

expenditures for some sets of input prices (Rothman *et al.*, 1994).[22] This occurs

primarily when the initial expenditures on a particular category of goods is zero. To limit

the extent of this problem, we make the assumption that households whose initial

expenditures are zero on a good will continue not to buy that good if prices change. So,

for each individual, we construct the AIDS parameters based only on those goods for

which they have positive expenditures. This results in slightly different cross-elasticity

parameters for each household, as the *n* in the denominator of equations (4-20) can be

---

[22] This is true of most of the flexible demand systems that are used, including the
Rotterdam model. One model that does not allow this is the generalized logit model, but
this model has the disadvantage that to calculate demand shares for a given set of prices
and expenditures, it requires iteration of a nonlinear equation that is not guaranteed to
converge (Rothman, Hong and Mount, 1994).

different for each household depending on the number of goods for which they have positive expenditures[23].

The last parameters that need to be calculated are α_i, which would determine the share of expenditures on each good if prices were equal. They will be different for each individual in the expenditure survey, and are calculated using the formula:

$$\alpha_i = w_i - \sum_j \gamma_{ij} \log p_j \tag{4-21}$$

I define units so that prices are all initially equal to 1, which simplifies this expression to $\alpha_i = w_i$, where w_i are shares of expenditure from the CEX data for the household.

Once the demand system parameters have been calculated, we can estimate demand responses to the commodity price changes generated by the supply model, and the income changes generated by recycling the carbon charge revenue. Since the amount of revenue generated by the carbon charge depends on the amount of carbon consumption, which in turn depends on the amount of revenue recycling, it is necessary to determine equilibrium consumption levels, carbon charge revenue, and revenue recycling levels that are consistent with each other. The linear relationship between changes in income and carbon consumption can be estimated simply by running the

---

[23] Cross price elasticities are undefined in the case of two or fewer goods, which reflects the fact that it is impossible to specify independent own-price elasticities for both goods in a two good demand system. In our study there are two households with zero expenditure, and 17 with expenditures on only two goods – these households are dropped from the analysis. There remain a few cases where the model predicts that an individual will have negative expenditures on a category of goods, but this is rare, and should not have much effect on decile averages.

model twice with arbitrarily chosen income changes, which makes this equilibrium calculation straightforward.

4.5 Attributing impacts to households

Households can be impacted by a carbon charge in several different ways. Increases in the prices of products they use will decrease the amount that they can consume, while dividend payments will increase their well-being. Changes in producer profits, either because of changes in input prices and demand or because of free permit receipts, will also affect households that own shares in those companies.

Reductions in carbon emissions will also provide environmental benefits to households. A full accounting of the impacts of the policy should include these benefits as well as the costs. However, it is difficult to know the magnitudes of the environmental benefits, or how they will be distributed among households. Because of this, I adopt the common approach of allocating only the costs (including abatement and permit costs) and the monetary benefits to households, and presenting the reduction in carbon emissions separately.

4.5.1 Impacts on consumers

Impacts on consumers are divided into permit costs, which are monetary payments that go directly or indirectly to the carbon charge, and abatement costs caused by reductions in consumer utility from shifting to less carbon-intensive consumption patterns. These consumer abatement costs are the primary economic costs in our model since producers do not change their input ratios. Boyce and Riddle (2007; 2009) focus only on permit costs because including abatement costs but not environmental benefits

can lead to misleading conclusions that total costs exceed total benefits. I sympathize

with this view, but since it is more common to include abatement costs (US CBO, 2000;

Dinan and Rogers, 2002; Burtraw *et al.*, 2009a, etc.), and since the distributional

incidence of these cost can be analyzed more easily than the climate benefits, it is useful

to present the incidence results with these costs included.

Monetary payments are found based on final consumption levels, subtracting the

amount they would have paid for their final consumption bundle at old prices from the

amount they pay at the new prices. Additional utility costs are calculated using the

indirect utility function associated with the cost function from which the AIDS demand

model is derived (Deaton and Muellbauer, 1980, p. 313). The cost function associated

with the demand model we use is given by:

$$\log(c(u, p)) = \alpha_0 + \sum_k \alpha_k \log(p_k) + \tfrac{1}{2} \sum_k \sum_j \gamma_{kj} \log(p_k) \log(p_j) + u\beta_0 \qquad (4\text{-}22)$$

The new parameters in this equation, α_0 and β_0, determine the size of the utility

numbers, but are eliminated once the utility numbers are converted to dollar equivalents,

so for simplicity I set $\alpha_0 = 0$ and $\beta_0 = 1$.

The indirect utility function associated with this is:

$$u(x, p) = \log(x) - \left(\sum_k \alpha_k \log(p_k) + \tfrac{1}{2} \sum_k \sum_j \gamma_{kj} \log(p_k) \log(p_j) \right) \qquad (4\text{-}23)$$

From this, the utility associated with both the initial and final prices and

expenditure levels can be calculated. This can be converted to dollar equivalents by

calculating the amount of expenditure that would be needed to generate each of these

utilities at the original prices using the cost function. Finally, extra unspent income is

added to the final utility numbers, since households are assumed to spend only 90% of

new income they receive from dividend payments and changes in profits. The change in

the dollar equivalent of utility as a result of the policy can then be broken down into

monetary payments resulting from the charge and abatement costs.

4.5.2 Allocating producer profit impacts to households

　　　　Changes in producer profits are allocated to households based on the distribution

of stock ownership by income level. The data for this are the same as those used in

Boyce and Riddle (2007), and are taken from the 2004 Survey of Consumer Finances.

They include both direct ownership of stocks and indirect ownership through mutual

funds and other sources. The distribution is shown in Table 4.6.[24]

Table 4.6. Stock ownership by income decile

Per capita income decile	Stock ownership	Share of total stock ownership
1	7437	0.8%
2	4564	0.5%
3	8697	0.9%
4	16069	1.7%
5	23066	2.4%
6	40296	4.2%
7	54571	5.7%
8	67427	7.0%
9	116542	12.1%
10	626335	64.9%

Source: 2004 Survey of Consumer Finances

---

[24] I would like to have data on the distribution of stock ownership by per capita expenditure, rather than per capita income, as this is how we present the rest of our results. However, there do not appear to be good data available on the distribution of stock ownership by expenditure. Holak et al. (2008) use capital holdings from the consumer expenditure survey for this purpose, but it is not clear exactly what is included in this variable, and it is likely to be less reliable than data from the Survey of Consumer Finances, which is designed for this purpose. I therefore use the more reliable Survey of Consumer Finances data and assume that the distribution of holdings by expenditure will be the same as the distribution by income.

4.5.3 Grouping households

The goal in presenting results is to show how the impacts of the policy vary by household, depending on how well-off the household is. There is no consensus on the best way to show this. A common approach is to break households into income groups, and show the impacts for each income group. However, many researchers have called into question whether annual income is the best measure of how well off a household is. Stratifying households by annual income produce results that look more regressive than they would if a better measure of lifetime income were used (Poterba, 1992; Metcalf, 1999; Hasset *et al.*, 2007).

For this study, I use total expenditures, rather than income, to group households, and present results as a portion of total expenditures. Expenditures are often seen as a better proxy for lifetime incomes than annual income (Poterba, 1992). This is also the approach taken by Boyce and Riddle (2007; 2008; 2009).

Another question that has received less attention is whether to group households by total income/expenditure or per capita income/expenditure. This study uses per capita expenditures to group households, and presents all impacts in per capita terms. Deciles are also constructed so that there are the same number of people in each decile, rather than the same number of households.

While using total household incomes or expenditures is more common, per capita income is often seen as a better measure of household well-being than total household income, since larger households have more people to support, and therefore will be less well off with the same total income (Datta and Meerman, 1980). Neither total income nor per capita income is a perfect measure of household well being, but more

complicated measures are less transparent and require more detailed data. Using per

capita numbers is also consistent with the approach of the cap-and-dividend policy,

which makes payments proportional to the number of people in the household.

4.6 Incidence results

In this section, the results of the analysis are presented. They are divided into

several different scenarios, which represent different assumptions about how the policy is

implemented and how suppliers respond to price changes.

4.6.1 Baseline Scenario

The baseline analysis looks at the impact of a $100/ton carbon charge if all the

revenues not needed to offset increased government expenditures are returned to

households, and suppliers pass-through all of their increased costs to consumers.

The results are presented in Tables 4.7 through 4.10. Table 4.7 shows the

distributional impacts of a carbon charge on its own, both in per person terms and as a

percentage of expenditures. Table 4.8 shows the distributional impacts of the policy after

the dividends are distributed. Tables 4.9 and 4.10 show in more detail how the demand

model assumptions influenced the distributional impacts.

4.6.1.1 Emission reductions and abatement costs

The baseline policy leads to a reduction in carbon emissions by 3.6%. This

relatively small reduction reflects the short-run nature of the analysis, and the fact that it

only represents carbon reductions caused by changes in final demand. In all scenarios

evaluated in this paper, suppliers do not change their input ratios in response to a change

in prices, which means they can not improve the energy efficiency of their production, or substitute low-carbon fuels for high-carbon ones. As a result, the carbon reductions we find should be interpreted as a lower bound estimate on the total emission reduction that is likely to occur in response to the policy.

The abatement costs calculated are also relatively small. In this model, the only abatement costs are those faced by consumers due to their shifting consumption patterns. These per person abatement costs are $8.8 per year, or only 1.63% of the average cost faced by households. The abatement costs of a $100/ton carbon charge could be higher if producers were able to adjust their input ratios to reduce their emissions. Allowing more abatement options would lower the marginal abatement cost curve, so more abatement would take place at the same permit price, leading to a higher total abatement cost. However, if the amount of emission reduction were fixed, allowing more abatement options would reduce the abatement cost. Our calculation of $8.8 per person therefore represents an upper bound on the cost of achieving a 3.6% reduction in carbon emissions. Higher emission reductions would require higher abatement costs, but would produce greater climate benefits as well.

To check that our average abatement costs are reasonable, we compare them to a simple estimate based on calculating the area under the marginal abatement cost curve between the business as usual emission levels and the reduced levels. If the marginal abatement cost curve is linear, the total abatement costs will equal the size of reduction in carbon emissions times the marginal cost of abatement, or the permit price, divided by two. This calculation produces an estimated total abatement cost of $9.8 per person, which is close to the $8.8 per person cost we found with our full calculations. The

175

difference reflects the fact that the demand function used in the full calculations does not produce a linear abatement cost curve.

Several other studies have looked at the abatement costs associated with a carbon charge, and found them to be higher. For example, CBO (2000) finds substitution costs to be 7.9% of the total costs faced by households, and Burtraw *et al.* (2009a) find them to be 9.3% of the total costs[25]. These higher abatement costs can be partly explained by different assumptions about the permit price and the reduction in carbon emissions that are being analyzed. For example, the higher adjustment costs found by CBO (2000) are consistent with the abatement cost curve calculation presented above, based on a larger assumed reduction in carbon emissions (15%) and a higher permit price ($100 per ton CO_2).[26]

4.6.1.2 Distributional incidence of charge

The results in Table 4.7 show that although rich households pay more into the charge than poor households, the charge on its own would be regressive in that its impact on poor households is greater as a percentage of their total expenditure. This is consistent with the carbon footprint patterns shown in section 4.2.4: low income households have lower carbon footprints than high income households, but the carbon intensity of their

---

[25] While Burtraw *et al.* (2009a) do not explicitly report abatement costs, these costs should be reflected in the average net impacts on household of the policies they analyze if revenues are returned to households. The abatement cost ratio we report is found by dividing the average net impact on households from Figure 7 by the average cost of the charge to households from Table 3.

[26] The numbers in Burtraw *et al.* (2009a), on the other hand, do not appear to be consistent with this calculation. The average net impacts of revenue-neutral policies on households appear to be about twice what they should be based on the estimated abatement costs calculated from their permit price of $20.87 per ton CO_2 and 7.4% reduction in carbon emissions.

consumption is higher. It is also consistent with the results of other studies done in the

US (Poterba, 1989; Metcalf, 1999; CBO, 2000; Dinan and Rogers, 2002; Barnes and

Breslow, 2003; Parry, 2003b; Boyce & Riddle, 2007; Hassett *et al.*, 2007; Holak *et al.*,

2008; Burtraw *et al.*, 2009a).

Table 4.7. Distributional incidence of carbon charge

Per capita expenditure decile	Initial per capita expenditure ($)	Cost per person	Cost as % of expend.
1	4964	193	3.9%
2	7629	288	3.8%
3	9925	346	3.5%
4	12187	388	3.2%
5	14510	451	3.1%
6	17290	515	3.0%
7	20735	571	2.8%
8	25038	663	2.6%
9	31460	793	2.5%
10	53819	1169	2.2%

Based on a carbon charge of $100 per tC
Source: Author's calculations

4.6.1.3 Incidence with dividends

In this scenario, 86% of the revenue from the carbon charge is devoted to

dividend payments, which are divided among households on a per capita basis. The

remaining 14% is retained to offset higher costs paid by local, state and national

governments.

Table 4.8 presents the full incidence of a cap-and-dividend policy as a percentage

of total expenditures. The distribution of equal per capita dividends more than

compensates for the regressive impact of the charge. The efficiency costs are too small to

have much effect on the overall incidence of the policy.

The net impacts range from a benefit of 6.8% of initial expenditures for the bottom decile to a cost of 1.2% of expenditures for the top decile. The average net impacts of the policy on households in the bottom six deciles are positive, while the averages for the top four deciles are negative. This is also consistent with the results of other studies that have looked at this, although the number of deciles that benefit and the number that are hurt vary somewhat among the studies (CBO, 2000; Dinan and Rogers, 2002; Barnes and Breslow, 2003; Parry, 2003b; Boyce & Riddle, 2007; Holak *et al.*, 2008; Burtraw *et al.*, 2009a).

Table 4.8. Distributional incidence of cap-and-dividend

Per capita expenditure decile	Per capita expenditure ($)	% of expenditures		
		Costs	Dividend	Net Impacts
1	4964	-3.9%	10.7%	6.8%
2	7629	-3.8%	6.9%	3.2%
3	9925	-3.5%	5.3%	1.8%
4	12187	-3.2%	4.3%	1.2%
5	14510	-3.1%	3.6%	0.5%
6	17290	-3.0%	3.1%	0.1%
7	20735	-2.8%	2.6%	-0.2%
8	25038	-2.6%	2.1%	-0.5%
9	31460	-2.5%	1.7%	-0.8%
10	53819	-2.2%	1.0%	-1.2%

Based on a carbon charge of $100 per tC
Source: Author's calculations

4.6.1.4 Role of demand model

Since we devoted some attention to improving the model of how consumer demand responds to changes in prices, it is useful to evaluate how the demand response affects our results. The most significant affect of the demand model is on our conclusions about how much emissions are reduced, and the size of the abatement costs. The model used in this paper produces a more accurate estimate of expected quantity-

based emission reductions than could be found through a simple elasticity calculation. While this is only part of the total carbon emission reductions that would be likely to occur, since producer responses are not included, it is useful to have a good estimate of the size of the demand response.

The demand response has less impact on the relative incidence of a cap-and-dividend policy on different deciles. Table 4.9 compares the incidence results from the full demand model with the incidence results with no demand response to price changes. The demand response leads to only small changes in the impact of the charge on all deciles, with the middle income deciles receiving slightly more benefit from their shifts in demand than the top and bottom deciles. There is very little change in the regressivity of the charge payments on their own. There is also little impact on the conclusions about the net impact of a cap-and-dividend policy. There is slightly less redistribution from rich households to poor households because dividend payments are reduced slightly, but the changes are very small.

Table 4.9. Distributional incidence of cap-and-dividend with different demand responses

Per capita expenditure decile	Initial per capita expenditure ($)	Charge		Net Impact	
		Fixed Demand	Full demand response	Fixed Demand	Full demand response
1	4964	-3.8%	-3.9%	7.2%	6.8%
2	7629	-3.8%	-3.8%	3.4%	3.2%
3	9925	-3.5%	-3.5%	2.0%	1.8%
4	12187	-3.3%	-3.2%	1.2%	1.2%
5	14510	-3.2%	-3.1%	0.6%	0.5%
6	17290	-3.1%	-3.0%	0.1%	0.1%
7	20735	-2.8%	-2.8%	-0.2%	-0.2%
8	25038	-2.7%	-2.6%	-0.5%	-0.5%
9	31460	-2.6%	-2.5%	-0.8%	-0.8%
10	53819	-2.2%	-2.2%	-1.2%	-1.2%

Based on a carbon charge of $100 per tC
Source: Author's calculations

This neutral impact of the demand response to a cap-and-dividend policy masks

two offsetting features of the demand response. Demand for carbon-intensive products

decreases in response to the new price structure, but increases because of the dividend

payments. The positive demand response to dividend payments is greater for lower

income deciles, because the payments are a higher percentage of their income. However,

the negative demand response to the price changes is also greater for lower income

deciles. The size of the demand responses to the income and price changes, as well as the

total demand response, are presented in Table 4.10.

Table 4.10. Demand responses to cap-and-dividend

Per capita expenditure decile	Initial per capita expenditure ($)	% change in carbon emissions due to dividend payment	% change in carbon emissions due to price changes	Total % change in carbon emissions
1	4964	9.7%	-9.4%	-0.6%
2	7629	6.2%	-9.2%	-3.5%
3	9925	4.8%	-8.6%	-4.2%
4	12187	3.9%	-7.9%	-4.3%
5	14510	3.3%	-7.5%	-4.5%
6	17290	2.8%	-7.2%	-4.6%
7	20735	2.3%	-6.3%	-4.2%
8	25038	1.9%	-5.8%	-4.0%
9	31460	1.5%	-5.2%	-3.7%
10	53819	0.9%	-3.1%	-2.2%

Based on a carbon charge of $100 per tC
Source: Author's calculations
Note that the last column is the compounded effect of the previous two columns rather than the sum, so (1+a)(1+b)=(1+c) rather than a+b=c

It is an interesting conclusion of our model that lower income deciles decrease

their carbon emissions more in response to higher carbon prices. This is a question that

has attracted some attention, though there has been little empirical work that has

addressed this issue. I am not aware of any studies that address this question in the

context of a US carbon charge. Studies conducted in other countries or for particular

fuels such as gasoline have found mixed results, with some reporting higher price elasticities for rich households, and others reporting higher elasticities for poor households (West and Williams, 2004; Kayser, 2000; Labandeira *et al.*, 2004).

This differential demand responses by income group found in this paper are not based on any empirical evidence of how different households have responded to price changes in the past. Instead, they are based on our theoretical demand model, combined with information from the expenditure survey on the expenditure shares of different households on different products. The logic behind the demand model helps to build an argument why greater adjustments by lower income deciles are likely, but this conclusion is tentative due to the lack of empirical evidence.

The relative demand responsiveness of different income groups is generated by our assumption that all households have the same set of elasticity parameters γ_{ii}, and that own price elasticities ε_{ii} are related to the elasticity parameters γ_{ii} based on the formula:

$$\varepsilon_{ii} = \frac{\gamma_{ii}}{w_i} - 1 \qquad\qquad (4\text{-}24)$$

This means that as a household spends a higher share of their budget on a particular good, the own price elasticity for that good will get closer to -1.0. This is consistent with a theoretical restriction that any demand model must satisfy: as the budget share approaches one, the own-price elasticity must also approach one in order to keep total expenditures from changing in response to a price change.

In the case of energy goods, which are generally inelastic (that is, they have own-price elasticities that are less than one in magnitude), households with higher budget shares spent on energy goods will have higher own-price elasticities for those goods. The

household expenditure data shows that poorer households spend a larger share of their budgets on energy goods, which leads to the conclusion that poorer households will on average reduce their carbon emissions more that rich household in response to a change in carbon prices.

4.6.2 Partial pass-through scenarios

Two different partial pass-through scenarios are analyzed, one with 90% pass through and one with 50% pass through. In the 90% pass-through scenario, carbon emissions are reduced by 3.3%. In the 50% pass-through scenario, emissions are reduced by just 1.7%. Reducing the pass-through rate lowers the reduction in carbon emissions because prices do not increase by as much, so consumers do not adjust their consumption patterns as much. The adjustment costs faced by consumers are also reduced: with 90% pass-through, the average adjustment cost is 1.34% of the total costs faced by consumers, and with 50% pass-through it is 0.41% of total costs.

The incidence results are presented in Tables 4.11 and 4.12. The costs are increased for the top decile, as they bear the brunt of the producer costs by virtue of their larger share of corporate stock ownership. The remaining deciles benefit or stay even. In the 90% pass-through scenario the charge on its own is less regressive than with full pass-through. In the 50% pass-through scenario, it becomes progressive, with the costs being a higher percentage of expenditures for the top decile than all other deciles. The net impact with dividend payments also becomes positive for more households. In the 90% pass-through scenario, the net impact for the seventh decile becomes barely positive on average, and in the 50% pass-through scenario, the bottom eight deciles all see net positive impacts.

182

Table 4.11. Distributional incidence of cap-and-dividend with 90% pass-through

Per capita expenditure decile	Per capita expenditure ($)	% of expenditures				
		Consumer costs	Producer costs	Total costs	Dividend	Net Impacts
1	4964	-3.3%	-0.1%	-3.4%	10.2%	6.8%
2	7629	-3.2%	0.0%	-3.2%	6.6%	3.4%
3	9925	-2.9%	-0.1%	-3.0%	5.1%	2.1%
4	12187	-2.7%	-0.1%	-2.8%	4.2%	1.4%
5	14510	-2.6%	-0.1%	-2.7%	3.5%	0.8%
6	17290	-2.5%	-0.2%	-2.7%	2.9%	0.3%
7	20735	-2.3%	-0.2%	-2.5%	2.4%	0.0%
8	25038	-2.2%	-0.2%	-2.4%	2.0%	-0.4%
9	31460	-2.1%	-0.3%	-2.3%	1.6%	-0.7%
10	53819	-1.7%	-0.8%	-2.6%	0.9%	-1.6%

Based on a carbon charge of $100 per tC
Source: Author's calculations

Table 4.12. Distributional incidence of cap-and-dividend with 50% pass-through

Per capita expenditure decile	Per capita expenditure ($)	% of expenditures				
		Consumer costs	Producer costs	Total costs	Dividend	Net Impacts
1	4964	-1.5%	-0.4%	-2.0%	9.3%	7.4%
2	7629	-1.5%	-0.2%	-1.7%	6.1%	4.4%
3	9925	-1.3%	-0.3%	-1.6%	4.7%	3.1%
4	12187	-1.2%	-0.4%	-1.6%	3.8%	2.2%
5	14510	-1.2%	-0.5%	-1.6%	3.2%	1.6%
6	17290	-1.1%	-0.7%	-1.8%	2.7%	0.9%
7	20735	-1.0%	-0.8%	-1.7%	2.2%	0.5%
8	25038	-0.9%	-0.8%	-1.7%	1.9%	0.2%
9	31460	-0.9%	-1.1%	-1.9%	1.5%	-0.4%
10	53819	-0.7%	-3.3%	-4.0%	0.9%	-3.1%

Based on a carbon charge of $100 per tC
Source: Author's calculations

There are two reasons that incomplete pass-through might affect the distributional incidence. First, and most importantly, part of the impact of the charge is now borne by producers rather than consumers. In the 90% pass-through scenario, 86.4% of the costs

are borne by consumers, and 13.6% by producers. In the 50% pass-through scenario, 40.5% of the costs are borne by consumers, and 59.5% by producers. The difference between the pass-through rate and the ultimate share of costs reflects the fact that when the costs of the charge are passed through several industries before reaching consumers, 90% pass-through in each intermediate industry can lead to less than 90% of the charge ultimately being passed on to end users. The government share of carbon consumption and the tax rate on corporate profits also play a role in the calculation of final impacts.

The effect of a shift in costs from consumers to producers has been studied before. Since a large share of producer profits is attributed to households in the top decile, it is not surprising that these studies have found that shifting costs to producers will lead to more progressive incidence results (Boyce and Riddle, 2007; Holak *et al.*, 2008).

An additional consequence of assuming incomplete pass-through is that the relative prices of different products are different than with full pass-through. For all products, prices do not increase by as much as they did with full pass-through, but the difference is greater for products that consume fossil fuels indirectly than for direct fuel purchases. This is because each firm along the production line will not pass on the full impact of the increase in their input costs. The changes in relative prices with incomplete pass-through actually contribute to making the carbon charge slightly more regressive because lower income households consume a higher share of their carbon directly through fuel purchases, while rich households use more carbon indirectly. This is consistent with findings from other studies that charges on the direct consumption of fossil fuels are more regressive than charges on indirect consumption (Bull *et al.*, 1994;

184

Hassett *et al.*, 2007). This is incorporated into the incidence results presented above, but the effect of relative prices is overshadowed by the transfer of costs between consumers and producers. Table 4.13 isolates the effects of relative prices on consumers by looking at the share of the total consumer costs borne by each decile in each of the three pass-through scenarios. The impacts on the bottom deciles are highest relative to impacts on other deciles when the pass-through rates are lowest. However, the differences are small, especially in the more reasonable 90% pass-through case.

Table 4.13. Comparison of consumer impacts with different pass-through rates

er capita expenditure decile	decile share of all consumer costs of charge		
	100% pass-through	90% pass-through	50% pass-through
1	3.6%	3.7%	4.0%
2	5.3%	5.5%	6.0%
3	6.4%	6.6%	7.1%
4	7.2%	7.3%	7.8%
5	8.4%	8.5%	8.9%
6	9.6%	9.7%	10.1%
7	10.6%	10.7%	10.8%
8	12.3%	12.3%	12.3%
9	14.7%	14.6%	14.3%
10	21.7%	21.1%	18.8%

Based on a carbon charge of $100 per tC
Source: Author's calculations

4.6.3 Fixed-cost supply model

This section presents the results using the supply model based on fixed costs and an increasing supply curve. In this model, pass-through rates vary by industry, with more carbon-intensive goods having lower pass-through rates. Suppliers also face adjustment costs as they shift their production to meet changes in the quantity demanded of their product. The exact assumptions are explained in more detail in section 4.4.1.3.

In this scenario, carbon emissions are reduced by 2.8%. This is less than in the

full pass-through scenario because price increases in all industries are closer to the

average prices increase, which sends a less strong signal to consumers as to how they

should adjust their demand to reduce emissions. Consumer abatement costs are 1.27% of

total costs, and producer abatement costs are 0.21% of total costs. This makes the total

efficiency costs 1.48% of total costs, which is less than in the full pass-through scenario,

and consistent with the smaller reduction in carbon emissions.

Table 4.14. Distributional incidence of cap-and-dividend with fixed cost supply model

Per capita expenditure decile	Per capita expenditure ($)	% of expenditures				
		Consumer costs	Producer change in profits	Total costs	Dividend	Net Impacts
1	4964	-3.6%	0.0%	-3.6%	10.1%	6.4%
2	7629	-3.5%	0.0%	-3.5%	6.5%	3.0%
3	9925	-3.2%	0.0%	-3.2%	5.0%	1.8%
4	12187	-3.0%	0.0%	-3.0%	4.1%	1.1%
5	14510	-2.9%	0.0%	-2.9%	3.4%	0.5%
6	17290	-2.8%	0.0%	-2.8%	2.9%	0.1%
7	20735	-2.6%	0.0%	-2.6%	2.4%	-0.2%
8	25038	-2.5%	0.0%	-2.5%	2.0%	-0.5%
9	31460	-2.4%	0.0%	-2.4%	1.6%	-0.8%
10	53819	-2.1%	0.0%	-2.1%	0.9%	-1.1%

Based on a carbon charge of $100 per tC
Source: Author's calculations

Table 4.14 presents the incidence results. Overall, there is very little difference

from the 100% pass-through scenario. Consumers bear all of the costs, while the effect

on producer profits is negligible. This is because the model was set up so that the

industries with more than 100% pass-through would roughly offset the industries with

less than 100% pass-through. The costs are slightly lower for all deciles, as are the

dividend payments. There is a slightly greater reduction in costs for low-income

households than high-income households, reflecting their higher share of consumption of

high carbon-intensity products, which see smaller price increases than in the full pass-through scenario. However, this effect is barely noticeable.

4.6.4 Partial grandfathering of permits

This section looks at how the incidence of the policy would change if 50% of the permit value that is available for recycling were given away to producers, with the other 50% given to consumers as before (this amounts to 43% of the total permits for each purpose, with the remaining 14% retained to offset government expenses). The full pass-through supply model is used. Carbon emissions are reduced by 3.8%, slightly more than when all permit value is recycled to households. This is because the transfer of income from poor to rich households leads to a shift toward the less carbon-intensive expenditures of rich households. Consumer adjustment costs are also slightly higher at 1.64%.

Table 4.15 shows the distributional incidence of the policy. It is dramatically different from the cap-and-divided results. Instead of having a progressive impact with the bottom six deciles benefitting, the only deciles that benefit are the ones at the extremes, the 1st and 10th deciles, while there is no effect on the 2nd decile. The lowest decile benefits because their carbon consumption is low enough that 50% of the permit revenue is enough to compensate them. The top decile benefits because they get the largest share of the increase in profits resulting from the free permits. The remaining deciles of middle-class households do not receive enough in dividends to offset the higher costs they face, but also don't own enough stock to benefit much from the higher producer profits.

Table 4.15. Distributional incidence of cap with revenues split between dividends and payments to producers

Per capita expenditure decile	Per capita expenditure ($)	% of expenditures				
		Consumer costs	Change in producer profits	Total costs	Dividend	Net Impacts
1	4964	-3.7%	0.4%	-3.3%	5.3%	2.0%
2	7629	-3.7%	0.2%	-3.5%	3.5%	0.0%
3	9925	-3.4%	0.2%	-3.2%	2.7%	-0.5%
4	12187	-3.1%	0.4%	-2.8%	2.2%	-0.6%
5	14510	-3.1%	0.4%	-2.6%	1.8%	-0.8%
6	17290	-3.0%	0.6%	-2.3%	1.5%	-0.8%
7	20735	-2.7%	0.7%	-2.0%	1.3%	-0.7%
8	25038	-2.6%	0.7%	-1.9%	1.1%	-0.9%
9	31460	-2.5%	1.0%	-1.5%	0.8%	-0.7%
10	53819	-2.2%	3.2%	1.0%	0.5%	1.4%

Based on a carbon charge of $100 per tC
Source: Author's calculations

4.7 Conclusions

This study confirms what has been found elsewhere in the literature: that a carbon charge on its own is likely to have a regressive impact, and that equal per capita redistribution of revenues to households will more than offset these regressive impacts.

In addition to confirming these basic results, this paper contributes to the literature by developing new modeling techniques to evaluate supply and demand responses to the charge. It improves on the methods used by Metcalf (1999), Hassett *et al.* (2007) and others to calculate price increases for different products from input-output tables, by modeling the policy as a quantity tax, rather than an ad valorem tax. It also develops new techniques for using input-output tables to model price increases based on supply models with incomplete pass-through. Specifically, this paper examines two ways of relaxing the full pass-through assumption – either by assuming a fixed rate of pass-through that is less than one, or by using different pass-through for different products,

based on a model of perfect competition with fixed costs. It then evaluates how changing these assumptions affects the distributional impacts.

It seems that the most important parameter in determining the distributional incidence is how much of the costs are ultimately borne by producers and how much by consumers. The impact is seen most clearly in the extreme scenario with only 50% pass-through, which shifts the impacts of the charge very heavily toward the top income decile. This effect is also apparent, though less dramatically, in the 90% pass-through scenario.

In addition to affecting how much of the charge is borne by producers vs. consumers, these assumptions can also have an impact on which consumers bear the burden of the charge. With the partial pass-through assumption, products that use carbon directly show more of a price increase than those that only use carbon indirectly, leading to more regressive incidence results on the consumption side. With the fixed costs assumption, the price increases for the most carbon-intensive products are reduced, while the price increase for less carbon-intensive products are increased. The magnitude of the distributional impacts on consumers is relatively small, however, suggesting that the simple full pass-through assumption may be reasonable if the only consumer impacts are being considered.

The paper also improves on the demand models used in prior studies, and evaluates how the demand model assumptions affect the distributional incidence of the charge. It finds that under reasonable assumptions about demand behavior, low-income consumers will be more responsive to changes in energy prices than high-income consumers, and that the difference in demand response to price changes by income level

can be fairly substantial. However, this effect can be offset by demand responses to income changes if revenues are returned to households on an equal per capita basis, increasing incomes in the bottom deciles by a larger percent than in the top deciles.

Finally, this paper looks at the abatement costs faced by suppliers and consumers under different model assumptions, and finds that they are small, and do not have much impact on the results. The size of the abatement costs could increase for any given carbon charge if suppliers were allowed to adjust their input ratios in response to a change in input prices, but so would the environmental benefits from further reductions in emissions.

4.7.1 Areas for further study

There are several limitations of this analysis that could be improved on with further study. One limitation is the static nature of this study. Few attempts have been made to evaluate how the distributional incidence of a policy could change over time, and this is a rich area for future research. It would also be instructive to look at how allowing firms to change the inputs they use in response to input price changes could affect the results. This could lead to a more complete picture of the emissions reductions that would result from a given carbon price, and of the adjustment costs that firms would face. It would also be interesting to see how the role of adjustment costs changes as the price set for carbon changes.

The demand model could be improved by allowing income elasticities to vary by product. If income elasticities for energy products are lower than average, this could dampen the increase in carbon consumption by low-income households in response to the dividend payments they receive. It would also be interesting to look at how much the

190

results would change if price elasticity parameters were allowed to vary depending on the total expenditures of the household.

4.7.2 Policy conclusions

The cap-and-dividend policy evaluated here has many attractive features. By putting a price on carbon emissions it would spur changes throughout the economy to reduce carbon emissions. Unlike many other cap-and-trade proposals, it would do this without hurting the pocketbooks of lower and middle income consumers. In fact, most low-to-middle income households would come out ahead financially as a result of the policy.

This stands in stark contrast to the mixed policy we evaluate, in which only half of the permits are auctioned and the other half are given away for free to firms. If revenues are thereby split between producers and consumers, most households in middle-income deciles are not compensated sufficiently to offset their costs. This difference could have important implications for the reception each of these policies would receive. While there may be short-run political advantages to giving some of the permit revenues to producers in key industries to maintain their support, a cap-and-dividend policy is more likely to generate sustained support from the general public.

APPENDIX A

RELATIONSHIP BETWEEN SUPPLY MODEL AND HUBBERT'S PEAK

The relationship between this paper's model and Hubbert's peak can be seen by looking at the model if prices are constant at $\overline{p}$, there is no quadratic term in the regression $(b_{2S} = 0)$ and there is no stochastic variation. With these assumptions, the shape of the curve will satisfy:

$$QS_t = (1 - e^{-A\overline{p}}) \cdot \frac{c}{(1 + e^{-(b_0 + b_1 t + LRE \cdot \overline{p})_t})} \cdot R_t$$

$$R_0 = \overline{R_0}$$

$$R_t = R_{t-1} - QS_{t-1}$$

Letting $\hat{b}_0 = b_0 + LRE \cdot \overline{p}$, and substituting for R_t, we get:

$$QS_t = (1 - e^{-A\overline{p}}) \cdot \frac{c}{(1 + e^{-(\hat{b}_0 + b_1 t)})} \cdot (R_0 - \sum_{\tau=0}^{t-1} QS_\tau)$$

The continuous time equivalent of this equation is:

$$Q_S(t) = (1 - e^{-A\overline{p}}) \cdot \frac{c}{(1 + e^{-(\hat{b}_0 + b_1 t)})} \cdot (R_0 - \int_0^t Q_S(\tau) d\tau) \tag{A-1}$$

The solution is a flexible variation of Hubbert's curve, where the rate of increase going up the curve need not be equal to the rate of decrease going down. The variable b_1 determines the rate of increase on the way up, and the rate of decrease on the way down is set by the quantity $c \cdot (1 - e^{-A\overline{p}})$. This can be shown rigorously by demonstrating that if the solution to this equation is extended to $-\infty$ on the left, it will approach an

exponential curve $D_1 \cdot e^{b_1 t}$, with rate of increase b_1 as $t \to -\infty$, and that as $t \to \infty$, the

solution will approach $D_2 \cdot e^{-rt}$ where the rate of decrease is $r = c \cdot (1 - e^{-A\bar{p}})$.[27]

To look at the limit as $t \to -\infty$, it is useful to notice that the solution $Q^*s(t)$ to

equation (A-1) will also satisfy:

$$Q_S(t) = (1 - e^{-A\bar{p}}) \cdot \frac{c}{(1 + e^{-(\hat{b}_0 + b_1 t)})} \cdot \left(R_{-\infty} - \int_{-\infty}^{t} Q_S(\tau) d\tau \right)$$

if $R_{-\infty} = R_0 + \int_{-\infty}^{0} Q^*s(\tau) d\tau$.

As $t \to -\infty$, the term $1 + e^{-(\hat{b}_0 + b_1 t)}$ approaches $e^{-(\hat{b}_0 + b_1 t)}$, because $e^{-(\hat{b}_0 + b_1 t)}$ becomes

arbitrarily large and dominates the 1. Similarly, the term $R_{-\infty} - \int_{-\infty}^{t} Q_S(\tau) d\tau$ approaches $R_{-\infty}$

because the integral approaches 0 and is dominated by $R_{-\infty}$. So as $t \to -\infty$, the integral

equation approaches:

$$Q_S(t) = (1 - e^{-A\bar{p}}) \cdot \frac{c}{e^{-(\hat{b}_0 + b_1 t)}} \cdot R_{-\infty}$$

which simplifies further as follows:

$$Q_S(t) = R_{-\infty} \cdot (1 - e^{-A\bar{p}}) \cdot c \cdot e^{\hat{b}_0} \cdot e^{b_1 t},$$

$$Q_S(t) = D_1 \cdot e^{b_1 t}$$

$$D_1 = R_{-\infty} \cdot (1 - e^{-A\bar{p}}) \cdot c \cdot e^{\hat{b}_0}$$

---

[27] The two functions are equivalent as t approaches $-\infty$ in the sense that the ratio
between the actual solution and this exponential curve approaches one.

Therefore, for small values of t, the curve approximates an exponential curve that is increasing at rate b_1. The constant term D_1 determines the level of the curve at any given time, but not the rate of increase.

As $t \to \infty$, the solution to equation (2) approximates $D_2 \cdot e^{-rt}$, where $r = c \cdot (1 - e^{-A\bar{p}})$. This can be shown by trying out an equation of this form, and showing that the integral equation approaches being satisfied as $t \to \infty$. First note that for an exhaustible resource, over an infinite time horizon, the entire resource stock will be exhausted, so the initial stock of the resource will equal the total cumulative production over time:

$$R_0 = \int_0^\infty Q_S(\tau)d\tau.$$

The final term in the equation simplifies:

$$R_0 - \int_0^t Q_S(\tau)d\tau = \int_0^\infty Q_S(\tau)d\tau - \int_0^t Q_S(\tau)d\tau = \int_t^\infty Q_S(\tau)d\tau.$$

Plugging in a solution of the form $D_2 \cdot e^{-rt}$, the initial equation becomes:

$$D_2 \cdot e^{-rt} = (1 - e^{-A\bar{p}}) \cdot \frac{c}{(1 + e^{-(\hat{b}_0 + b_1 t)})} \cdot \int_t^\infty D_2 \cdot e^{-r\tau} d\tau$$

$$= (1 - e^{-A\bar{p}}) \cdot \frac{c}{(1 + e^{-(\hat{b}_0 + b_1 t)})} \cdot D_2 \cdot \frac{1}{-r}\left[e^{-r\tau}\right]_t^\infty.$$

$$= (1 - e^{-A\bar{p}}) \cdot \frac{c}{(1 + e^{-(\hat{b}_0 + b_1 t)})} \cdot D_2 \cdot \frac{1}{r}e^{-rt}$$

As $t \to \infty$, the middle term $\dfrac{c}{(1 + e^{-(\hat{b}_0 + b_1 t)})}$ approaches c, since $e^{-(\hat{b}_0 + b_1 t)}$ approaches

0 and is dominated by the 1. So we get: $D_2 \cdot e^{-rt} = (1 - e^{-A\bar{p}}) \cdot c \cdot D_2 \cdot \frac{1}{r}e^{-rt}$

194

Simplifying, this becomes: $r = c \cdot (1 - e^{-A\bar{p}})$, so for that value of r, the equation

will be satisfied for all t. Therefore, as $t \to \infty$, the solution approaches a decreasing

exponential function that is decreasing at rate $c \cdot (1 - e^{-A\bar{p}})$. (Any value for D_2 would

satisfy this equation – the exact value for D_2 will depend on the level of R_0 and making

sure the equation $R_0 = \int_{-\infty}^{\infty} Q_S(\tau) d\tau$ is satisfied.)

If the initial rate of increase and the final rate of decrease are equal – that is

$b_1 = c(1 - e^{-A\bar{p}})$ – then the solution to this equation is exactly Hubbert's Curve, with

$w = \frac{1}{b_1}$, $h = R_0 \cdot b_1$ and $t_{max} = -\frac{\hat{b}_0}{b_1}$. Plugging these in, the equation for Hubbert's curve

becomes:

$$Q_S(t) = R_0 \cdot b_1 \cdot \frac{e^{-\hat{b}_0 + b_1 t}}{\left(1 + e^{-\hat{b}_0 + b_1 t}\right)^2} .$$

Plugging this into the integral equation (A-1), we get:

$$R_0 \cdot b_1 \cdot \frac{e^{-\hat{b}_0 + b_1 t}}{\left(1 + e^{-\hat{b}_0 + b_1 t}\right)^2} = (1 - e^{-A\bar{p}}) \cdot \frac{c}{(1 + e^{-(\hat{b}_0 + b_1 t)})} \cdot \left(R_0 - \int_{-\infty}^{t} R_0 \cdot b_1 \cdot \frac{e^{-\hat{b}_0 + b_1 \tau}}{\left(1 + e^{-\hat{b}_0 + b_1 \tau}\right)^2} d\tau \right)$$

Evaluating the integral, this becomes:

$$R_0 \cdot b_1 \cdot \frac{e^{-\hat{b}_0 + b_1 t}}{\left(1 + e^{-\hat{b}_0 + b_1 t}\right)^2} = (1 - e^{-A\bar{p}}) \cdot \frac{c}{(1 + e^{-(\hat{b}_0 + b_1 t)})} \cdot \left(R_0 - R_0 \frac{1}{1 + e^{-(\hat{b}_0 + b_1 t)}} \right)$$

Simplifying further, this reduces to:

$$b_1 = c(1 - e^{-A\bar{p}})$$

So the equation is satisfied by Hubbert's curve as long as the rate of increase, b_1,

and the rate of decrease, $c(1 - e^{-A\bar{p}})$, are equal.

APPENDIX B

HEDGING AND SPECULATION ROLES OF FUTURES MARKETS

I consider a problem where an oil market participant (the agent) is simultaneously making a decision that affects how much oil they will buy or sell in the next period, while at the same time deciding how many oil futures to buy or sell. The agent chooses an amount x that represents the amount of oil it will sell in the next period minus the amount they will buy. Their overall profit of the agent will be $f(x) + x \cdot p_{t+1}$, where f(x) catches everything else that affects profits and is assumed to be independent of the price of oil in the next period (is this assumption necessary? There may be an alternative formulation without this assumption, but risk offsetting in futures markets would be more complicated.) In addition, they choose to take a long or short position in futures markets. They choose qf, which can be positive if they take a long position or negative if they take a short position. Their net profit from this decision is $qf \cdot (p_{t+1} - p_f)$.

I assume that they are risk averse, with constant absolute risk aversion (CARA). This can be represented by a utility function u of the form $u(c) = -Exp(-\alpha c)$.

The producer's problem is to choose x and qf to maximize their expected utility:

$$E\left[u\left(f(x) + x \cdot p_{t+1} + qf \cdot (p_{t+1} - p_f)\right)\right].$$

A change of variables to $qf = qf_2 - x$ helps separate the expression into two independent terms. The problem becomes choosing qf_2 and x to maximize

$$E\left[u\left(f(x) + x \cdot p_f + qf_2 \cdot (p_{t+1} - p_f)\right)\right]$$

Using CARA utility function, this becomes

$$E\left[-u\left(f(x) + x \cdot p_f\right) \cdot u\left(qf_2 \cdot (p_{t+1} - p_f)\right)\right]$$

With independence assumption this becomes

$$- E\big[u\big(f(x) + x \cdot p_f\big)\big] \cdot E\big[u\big(qf_2 \cdot (p_{t+1} - p_f)\big)\big]$$

Since only x is in the first term, and only qf_2 is in the second term, the values of x and qf_2 that maximize each term can be calculated separately, independent of the value of the other term.

The choice of qf_2 is the speculators' role, and the choice of x is the producers/storage holders' role. In total, the producer will 'buy' $qf = qf_2 - x$ futures. The $-x$ is done to offset risk – the hedging function of futures markets. In principle, by this logic, everyone who expects to sell oil next year should take a short position in futures markets to offset the risk from price volatility, and everyone who expects to buy oil should take a long position. These effects would cancel each other out, leaving the net effect on futures markets to be driven by the speculative part of participation of futures markets, which is governed by participants' predictions of what prices will be.

BIBLIOGRAPHY

Abar, George. 2009. "Options for Protecting Households in Climate Policy: A State-By-State Analysis." Unpublished White Paper.

Amin-Nazeri, Mohammad Reza and Ehsan Ahmadi Gharacheh, (2007) "A Hybrid Artificial Intelligence Approach to Monthly Forecasting of Crude Oil Price Time Series" Conference proceedings for 10th International Conference on Engineering Applications of Neural Networks. Working paper available at http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-284/page160.pdf.

Arthur, W. B., J. H. Holland, et al. (1996). "Asset Pricing Under Endogenous Expectation in an Artificial Stock Market," Santa Fe Institute Working Paper.

Bacon, R. (1991), "Modeling the price of oil." *Oxford Review of Economic Policy* 7 (2), 17-34.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. *Journal of Financial Economics* 49, pp. 307-343.

Barnes, P. 2001. *Who Owns the Sky?: Our Common Assets and the Future of Capitalism*. Washington, D.C.: Island Press.

Barnes, P. and M. Breslow. 2003. "The Sky Trust: The Battle for Atmospheric Scarcity Rent." In *Natural Assets: Democratizing Environmental Ownership*, ed. J. K. Boyce, 135-149. Washington: Island Press.

Barnett, H.J., and C. Morse (1963), *Scarcity and Growth*, Baltimore, MD: John Hopkins University Press and Resources for the Future.

Berck, Peter and Michael Roberts (1996), "Natural Resource Prices: Will They Ever Turn Up?" *Journal of Environmental Economics and Management*, 31: pp. 65-78.

Bohi, Douglas R. 1981. *Analyzing Demand Behavior: A Study of Energy Elasticities*. Washington DC: Resources for the Future, Inc.

Bovenberg, A. L. and L. H. Goulder. 2000. "Neutralizing the Adverse Industry Impacts of CO2 Abatement Policies: What Does it Cost?" National Bureau of Economic Research Working Paper 7654.

Boyce, James K. and Matthew Riddle. 2007. "Cap-and-Dividend: How to Curb Global Warming while Protecting the Incomes of American Families." Political Economy Research Institute Working Paper 150.

Boyce, James K. and Matthew Riddle. 2008. "Keeping the Government Whole: The Impact of a Cap-and-Dividend Policy for Curbing Global Warming on Government Revenue and Expenditure." Political Economy Research Institute Working Paper 188.

Boyce, James K. and Matthew Riddle. 2009. "Cap-and-Dividend: A State-by-State Analysis." Political Economy Research Institute and Economics for Equity and the Environment Network.

BP. 2011. *BP Statistical Review of World Energy*.

Bucks, Brian K., Arthur B. Kennickell, and Kevin B. Moore. 2006. "Recent Changes in U.S. Family Finances: Evidence from the 2001 and 2004 Survey of Consumer Finances." *Federal Reserve Bulletin,* 92: A1-A38.

Bull N., K. Hassett and G. Metcalf. 1994. "Who Pays Broad-based energy taxes? Computing Lifetime and Regional Incidence." *Energy Journal*, 15:145-64.

Burtraw, Dallas, Karen Palmer, Ranjit Bharvirkar, and Anthony Paul. 2001. "The Effect of Allowance Allocation on the Cost of Carbon Emission Trading." Resources for the Future Discussion Paper 01-30.

Burtraw, Dallas, Richard Sweeney and Margaret Walls. 2009a. "The Incidence of U.S. Climate Policy: Alternative Uses of Revenues from Cap-and-Trade Auctions." Resources for the Future Discussion Paper 09-17.

Burtraw, Dallas, Karen L. Palmer and Danny Kahn. 2009b. "A Symmetric Safety Valve." Resources for the Future Discussion Paper 09-06.

Campbell, Colin J. and Jean H. Laherrere (1998), "The End of Cheap Oil," *Scientific American*, March 1998, pp. 78-83.

Campbell, Harry F. (1980), "The Effect of Capital Intensity on the Optimal Rate of Extraction of a Mineral Deposit," *Canadian Journal of Economics*, 13:2, pp. 349-56.

Carroll, CD. 2001. "Precautionary Saving and the Marginal Propensity to Consume out of Permanent Income." National Bureau of Economic Research Working Paper 8233.

Chevillon, G., Rifflart, C., 2009. Physical market determinants of the price of crude oil and the market premium. Energy Economics 31, 537-549.

Cochrane and Orcutt (1949), "Application of least squares regression to relationships containing autocorrelated error terms," *Journal of the American Statistical Association* 44, pp 32-61.

Considine, T. and T.D.Mount. 1984. "The Use of Linear Logit Models for Dynamic Input Demand Systems." *The Review of Economics and Statistics,* 66(3): 434-44.

Cornwell, Antonia and John Creedy. 1996. "Carbon Taxation, Prices, and Inequality in Australia." *Fiscal Studies*, 17(3): 21-38.

Cremer, Jacques and Djavad Salehi-Isfahani (1991), Models of the Oil Market (book in 'Fundamentals of Pure and Applied Economics' series) Amsterdam: Hardwood Academic Publishers.

Dahl, Carol. 1993. "A Survey of Energy Elasticities in Support of the Development of the NEMS," Prepared for: United States Department of Energy.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor Psychology and security market under- and over-reactions. *The Journal of Finance*. Vol 53, No 6, pp. 1839-1885

Deaton, Angus, and John Muellbauer. 1980. "An Almost Ideal Demand System." *The American Economic Review*, 70(3): 312-326.

DeCanio, Stephen J. 2007. "Distribution of Emissions Allowances as an Opportunity." *Climate Policy,* 7(2) 91-103.

Dees, Stephane, Pavlos Karadeloglou, Robert Kaufmann and Marcelo Sanchez (2007), "Modeling the World oil Market: Assessment of a quarterly econometric model," *Energy Policy* 35 178-191.

DeLong, J.B., A. Shleifer, L Summers and R Waldmann (1990a), "Noise trader risk in financial markets," *Journal of Political Economy*, 98, pp. 703-38.

Dinan, Terry M. and Diane Lim Rogers. 2002. "Distributional Effects of Carbon Allowance Trading: How Government Decisions Determine Winners and Losers." *National Tax Journal,* 55(2): 199-222.

Economic and Allocation Advisory Committee (EAAC). 2010. "Allocating Emission Allowances under a California Cap-and-Trade Program" available at http://www.climatechange.ca.gov/eaac/

Ederington, Louis H., Chitru S. Fernando, Thomas K. Lee, Scott C. Linn, and Anthony D. May, 2011, "Factors Influencing Oil Prices: A Survey of the Current State of Knowledge in the Context of the 2007-08 Oil Price Volatility", accessed at http://38.96.246.204/finance/markets/reports_presentations/factors_influencing_oil_prices.pdf on 10/25/2011

Ellen, S. Ter and R.C.J. Zwinkels. (2010) "Oil price dynamics: A behavioral finance approach with heterogeneous agents," Energy Economics, 32, 1427-1434. 2010.

Energy Information Administration (EIA). 2004. "Long Term World Oil Supply Scenarios," written by John H. Wood, Gary R. Long and David F. Morehouse.

Energy Information Administration (EIA). 2007a. *NEMS International Energy Module (IEM), Model Documentation Report*, DOE/EIA-M071, available at http://tonto.eia.doe.gov/FTPROOT/modeldoc/m071(2007).pdf.

Energy Information Administration (EIA). 2007b. *Annual Energy Review 2006*. Washington, DC: US Department of Energy, DOE/EIA-0384(2006).

Energy Information Administration (EIA). 2008. "Domestic Crude Oil First Purchase Prices by Area," available at http://tonto.eia.doe.gov/dnav/pet/pet_pri_dfp1_k_m.htm.

Energy Information Administration (EIA). 2011. *Annual Energy Outlook 2011*, available at http://www.eia.gov/forecasts/aeo/

Eshelman, L.J. and J.D. Schaffer (1993). "Real-Coded Genetic Algorithms and Interval-Schemata." *Foundation of Genetic Algorithms 2*, L. Darrell Whitley (Ed.) (Morgan Kaufmann Publishers, San Mateo), 187-202.

Etemad, B., J. Luciani, P. Bairoch, and J.-C. Toutain (1991), *World Energy Production 1800-1985*, Librarie DROZ, Switzerland.

Fama, Eugene (1965), "The Behavior of Stock Market Prices," *Journal of Business*, 38, pp. 34-106.

Fama, Eugene (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25:2, pp. 383-417.

Fama, Eugene (1998), "Market Efficiency, Long-term Returns and Behavioral Finance," *Journal of Financial Economics*, 49, pp. 283-306.

Fattouh, Bassam (2007), "The Drivers of Oil Prices: The Usefulness and Limitations of Non-Structural model, the Demand-Supply Framework and Informal Approaches," Oxford Institute for Energy Studies, WPM 32.

French, Mark W. 2005. "Why and When do Spot Prices of Crude Oil Revert to Futures Price Levels." *Finance and Economics Discussion Series*, 2005-30.

Garner, Thesia I., George Janini, William Passero, Laura Paszkiewicz and Mark Vendemia. 2006. "The CE and the PCE: A Comparison." *Monthly Labor Review* 129(9): 20-46.

Gately, Dermot (1993), "The Imperfect Price-Reversibility of World Oil Demand," *The Energy Journal*, 14(4), pp. 163-182.

Gately, Dermot (2001), "How Plausible is the Consensus Projection of Oil below $25 and Persian Gulf Oil Capacity and Output Doubling by 2020?" *The Energy Journal*, 22(4), pp. 1-27.

Gately, D. and Huntington, H. (2002), "The Asymmetric effects of changes in prie and income on energy and oil demand," *The Energy Jounal*, 23(1), pp. 19-58.

Gilbert, Christopher L. (2010), "Speculative Influences on Commodity Futures Prices 2006-2008," Discussion Paper 197, United Nations Conference on Trade and Development 2010.

Graham, Daniel J. and Stephen Glaister (2002) "The Demand for Automobile Fuel: A Survey of Elasticities," *Journal of Transport Economics and Policy*, 36(1): 1-26.

Grossman, S. J. and J. E. Stiglitz (1980). "On the Impossibility of Informationally Efficient Markets." *The American Economic Review* **70**(3): 393-408.

Hamilton, James (2003). "What is an Oil Shock?" *Journal of Econometrics*, 113(2): 363-398.

Hamilton, James (2005a). "Oil and the Macroeconomy." *Palgrave Dictionary of Economics.*

Hamilton, James (2005b), "How to talk to an economist about peak oil." Posted on Econbrowser blog: July 11, 2005. Available at http://www.econbrowser.com/archives/2005/07/how_to_talk_to.html.

Hamilton, James D., 2009b. "Understanding Crude Oil Prices," The Energy Journal, International Association for Energy Economics, vol. 30(2), pages 179-206.

Hamilton, James, 2009a, "Causes and Consequences of the Oil Shock of 2007-08", NBER working paper 15002. http://www.nber.org/papers/w15002

Hamilton, James (2011), "Fundamentals, speculation, and oil prices." Posted on Econbrowser blog: August 28, 2011. Available at http://www.econbrowser.com/archives/2011/08/fundamentals_sp.html

Hassett, Kevin A., Aparna Mathur and Gilbert E. Metcalf. 2007. "The Incidence of a U.S. Carbon Tax: A Lifetime and Regional Analysis." National Bureau of Economic Research Working Paper 13554.

Heinberg, R. (2003). *The Party's Over: Oil, War and the Fate of Industrial Societies*. Gabriola Island, Canada, New Society Publishers.

Herrera F., M. Lozano and J.L. Verdegay (1998). "Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioral Analysis." *Artificial Intelligence Review* 12: 265-319.

Hirsch, Robert L., Roger Bezdek, Robert Wendling, (2005). "Peaking World Oil Production: Impacts, Mitigation and Risk Management."

Holak, J.F., H. Jacoby, G. Metcalf, S. Paltsev and J. Reilly. 2008. "Analysis of U.S. Greenhouse Gas Tax Proposals." National Bureau of Economic Research Working Paper 13980.

Horn, M.K. (2007). "Relationship Between Giant Field Data and Ultimate Recoverable Oil." Search and Discovery article # 40236, available at http://www.searchanddiscovery.com/documents/2007/07032horn/images/horn.pdf

Hotelling, H. (1931). "The Economics of Exhaustible Resources." *The Journal of Political Economy* **39**(2): 137-175.

Hubbert, M. King (1962), *Energy Resources: A Report to the Committee on Natural resources*, National Academy of Sciences.

Huntington, H.G. (1994), "Oil price forecasting in the 1980s: what went wrong?" *The Energy Journal* 15 (2), 1-22.

International Energy Agency (2007), "World Energy Model, Methodology and Assumptions." http://www.worldenergyoutlook.org/docs/weo2007/WEM_Methodology_07.pdf

International Energy Agency. 2011, *Oil Market Report*, October 12, available at www.oilmarketreport.org

Irwin, Scott H., Dwight R. Sanders and Robert P. Merrin (2009) "Devil or Angel? The Role of Speculation in the Recent Commodity Price Boom (and Bust)." Journal of Agricultural and Applied Economics, 41(2) 377-391.

Jappelli, Tullio and Luigi Pistaferri. 2009. "The Consumption Response to Income Changes." Working paper online at http://www.stanford.edu/~pista/ann_rev.pdf. Written for Annual Reviews in Economics, 2010.

Jegadeesh, N and S. Titman (1993), "Returns to buying winners and selling losers: implications for stock market efficiency," *Journal of Finance*, 48, pp. 65-91.

Kang, Sungjin. 2010. "Carbon Border Tax Adjustments from WTO Point of View." Society for International and Economic Law Online Proceedings, Working Paper # 2010/08.

Kaufmann, Robert K., Andrew Bradford, Laura H. Belanger, John P. Mclaughlin and Yosuke Miki1 (2008), "Determinants of OPEC production: Implications for OPEC behavior," *Energy Economics*, 30(2), pp. 333-351.

Kayser, Hilke A. 2000. "Gasoline Demand and Car Choice: Estimating Gasoline Demand Using Household Information." *Energy Economics,* 22: 331-348.

Keynes, J. M. (1936), The General Theory of Employment, Interest and Money, London: Macmillan.

Killian, Lutz, 2009, "Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market," American Economic Review, 99, 1053-1069

Kilian, L., Murphy, D., 2010, The Role of Inventories and Speculative Trading in the Global Market for Crude Oil, Working Paper, Center for Economic Policy Research and University of Michigan at Ann Arbor.

Knetsch, Thomas A. (2007), "Forecasting the price of Crude Oil via Convenience Yield Predictions," Journal of Forecasting? Or Deutsche Bundesbank Discussion paper, series 1: Economic Studies No 12/2006.

Kopp, Raymond, Richard Morgenstern, William Pizer and Michael Toman. 1999. "A Proposal for Credible Early Action in U.S. Climate Policy." Washington, DC: Resources for the Future.

Krautkraemer, J. A. (1998). "Nonrenewable Resource Scarcity." *Journal of Economic Literature* **36**(4): 2065-2107.

Krugman, P., 2008, Speculation and Signatures, NYTimes.com, June 24. http://krugman.blogs.nytimes.com/2008/06/24/speculation-and-signatures/.

Labandeira, Xavier, Jose M. Labeaga, and Miguel Rodríguez. 2004. "Microsimulating the Effects of Household Energy Price Changes in Spain." EconWPA Public Economics Series WorkingPaper 0412001.

Lagi, M. Yavni Bar-Yam, K.Z. Bertrand, Yaneer Bar-Yam, "The Food Crises: A Quantitative Model of Food Prices Including Speculators and Ethanol Conversion." arXiv:1109.4859, September 21, 2011.

Lasserre, Pierre. 1984. "Reserves and Land Prices with Exploration Under Uncertainty," *Journal of Environmental Economics and Management*, 11:3, 191-201.

LeBaron, Blake (2006), "Agent-based Computational Finance," from: *Handbook of Computational Economics*, edited by H.M. Amman, D. A. Kendrick and J, Rust, Elsevier, Vol. 2, pp. 1187-1233.

Leontief, W. 1986. *Input-Output Economics, 2nd Edition*. New York: Oxford University Press.

Lynch, M.C. (2002), "Forecasting oil supply: Theory and Practice." *The Quarterly Review of Economics and Finance*, 42, pp. 373-389.

MacKenzie, James J. (2000), "Oil as a finite resource: When is global production likely to peak?" Report for World Resources Institute, available at http://www.wri.org/publication/oil-finite-resource-when-global-production-likely-peak.

Manthy, Robert S. (1978), *Natural Resource Commodities – A Century of Statistics*, Baltimore: John Hopkins University Press.

Marland, G., T.A. Boden, and R.J. Andres (2007), "Global, Regional and National Fossil Fuel CO2 Emissions." In *Trends: A Compendium of Data on Global Change*. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., USA.

Masters, Michael, 2008, Testimony to US Senate Committee on Homeland Security and Governmental Affairs, available at http://hsgac.senate.gov/public/_files/052008Masters.pdf

Masters, M. W., White, A., 2009, The 2008 Commodities Bubble: Assessing the Damage to the United States and its Citizens, Masters Capital Management and White Knight Research and Trading. http://media.mcclatchydc.com/smedia/2009/05/20/16/Hall-2008-Commodities.source.prod_affiliate.91.pdf

Meadows, Donella H., Dennis L. Meadows, Jorgen Randers and William W. Behrens III (1971), *The Limits to Growth*, New York: Universe Books.

Metcalf, G. E. 1999. "A Distributional Analysis of Green Tax Reforms." *National Tax Journal* 52(4): 655-681.

Muhlenbein, H. and D. Schlierkamp-Voosen (1993). "Predictive Models for the Breeder Genetic Algorithm I. Continuous Parameter Optimization." *Evolutionary Computation* 1, 25-49.

Okushima, Shinichiro. 2003. "Global Warming Policy and Distributional Effects: A General Equilibrium Analysis." Society for Computational Economics's Computing in Economics and Finance Series 2003, 58.

Paltsev, Sergey *et al.* 2007. "Assessment of U.S. Cap-and-Trade Proposals," Cambridge, MA: Massachusetts Institute of Technology Joint Program on the Science and Policy of Global Change, Report No. 146, April. Available at http://web.mit.edu/globalchange/www/MITJPSPGC_Rpt146.pdf.

Parry, Ian. 2003a. "Fiscal Interactions and the Case for Carbon Taxes over Grandfathered Carbon Permits." *Oxford Review of Economic Policy*, 19(3): 385-399.

Parry, Ian. 2003b. "Are Emissions Permits Regressive?" Resources for the Future Discussion Paper 03-21.

Phillips, Peter C.B. and Jun Yu (2010), "Dating the Timeline of Financial Bubbles during the Subprime Crisis," Cowles Foundation Discussion Paper 1770, Cowles Foundation for Research in Economics 2010.

Pindyck, Robert S. (1978), "The Optimal Extraction and Production of Nonrenewable Resources," *The Journal of Political Economy*, 86(5), pp. 841-861.

Pindyck, Robert S. (1999), "The Long-Run Evolution of Energy Prices," *The Energy Journal,* Vol 20, No 2.

Pindyck, Robert S. (2001), "The Dynamics of Commodity Spot and Futures Markets: A Primer," *The Energy Journal*, 22(3) pp. 1-29.

Pizer, William A. 1998. "Prices vs. Quantities Revisited: The Case of Climate Change." Resources for the Future Discussion Paper 98-02.

Poterba, J. M. 1989. "Lifetime Incidence and the Distributional Burden of Excise Taxes." American Economic Review, 79(2): 325-30.

Potter, Neal and Francis T. Christy, Jr. (1962), *Trends in Natural Resource Commodities*, Baltimore: John Hopkins University Press.

Ramcharran, Harri (2002), "Oil production responses to price changes: an empirical application of the competitive model to OPEC and non-OPEC countries," *Energy Economics* 24(2), pp. 97-106.

Rothman, Dale S., J. Ho Hong, and Timothy D. Mount. 1994. "Estimating Consumer Energy Demand Using International Data: Theoretical and Policy Implications." *The Energy Journal*, 15(2): 67-88.

Rubinstein, Mark (2001), "Rational Markets: Yes or No? The Affirmative Case," *Financial Analysts Journal*.

Saul, Jonathan and Christopher Johnston, 2010, "Crude oil storage play at sea over for now" Reuters news story March 11, 2010, http://www.reuters.com/article/2010/03/11/us-oil-seastorage-analysis-idUSTRE62A29U20100311.

Shah, A. and B. Larsen. 1992. "Carbon Taxes, the Greenhouse Effect, and Developing Countries", Office of the Vice President, Development Economics, World Bank.

Shambora, William E. and Rosemary Rossiter (2007), "Are there Exploitable Inefficiencies in the Futures Market for Oil?" *Energy Economics* 29:1, pp. 18-27.

Shiller, Robert (1981), "Do Stock Prices Move Too Much to be Justified by Subsequent Movements in Dividends?" *American Economic Review*, 71:3, pp. 421-36.

Shiller, Robert (1990), "Speculative Prices and Popular Models," *Journal of Economic Perpectives*, Spring 1990, pp 55-65.

Shiller, Robert (2005), *Irrational Exuberance*, New York: Doubleday.

Shleifer, Andrei (2000), "Inefficient Markets: An Introduction to Behavioral Finance," New York: Oxford University Press.

Slade, Margaret E. (1982), "Trends in Natural Resource Commodity Prices: An Analysis of the Time Domain," *Journal of Environmental Economics and Management,* 15: 189-205.

Slade, Margaret E. (1992), "Do Markets Underprice Natural-Resource Commodities?" Working Paper No. 962. Washington, DC: The World Bank.

Spyrou, Spyros (2006), "Unobservable Information and Behavioral Patterns in Futures Markets: The case for Brent Crude Oil, Gold and Robusta Coffee Contracts," Derivatives Use, Trading Regulation, 12:1, pp. 48-59.

Stern, Nicholas. 2006. *The Stern Review: The Economics of Climate Change.* London: HM Treasury.

Symons, Elizabeth, John Proops, and Philip Gay. 1994. "Carbon Taxes Consumer Demand and Carbon Dioxide Emissions: A Simulation Analysis for the UK." *Fiscal Studies,* 15(2): 18-43.

Tilton, J. E. (2003). *On Borrowed Time?: Assessing the Threat of Mineral Depletion*, RFF Press.

U.S. Bureau of Labor Statistics, Division of Consumer Expenditure Surveys. 2004. "2003 Consumer Expenditure Interview Survey Public Use Microdata Documentation" available at http://www.bls.gov/cex/2003/cex/csxintvw.pdf.

US Commodity Futures Trading Commission (CFTC). 2011. "The Economic Purpose of Futures Markets and How they Work," available at: http://www.cftc.gov/ConsumerProtection/EducationCenter/economicpurpose

U.S. Congressional Budget Office. 2000. "Who Gains and Who Pays Under Carbon-Allowance Trading? The Distributional Effects of Alternative Policy Designs." Washington, DC: CBO.

U.S. Congressional Budget Office. 2001. "An Evaluation of Cap-and-Trade Programs for Reducing U.S. Carbon Emissions." Washington, DC: CBO.

USGS World Energy Assessment Team (2000), "US Geological Survey World Petroleum Assessment 2000 – Description and Results," available at http://pubs.usgs.gov/dds/dds-060/index.html#TOP

US Senate (2006). The Role of Market Speculation in Rising Oil and Gas Prices: A Need to Put the Cop Back on the Beat. Staff Report prepared by the Permanent Subcommittee on Investigations on the Committee on Homeland Security and Governmental Affairs United States Senate.

Wall Street Journal. 2009. "Who Pays for Cap and Trade?" Editorial. March 9, 2009. http://online.wsj.com/article/SB123655590609066021.html.

Weitzman, Martin L. 1974. "Prices vs. Quantities." *Review of Economic Studies*, 41(4): 477-91.

West, Sarah E. and Robertson C. Williams III. 2004. "Estimates from a Consumer Demand System: Implications for the Incidence of Environmental Taxes," *Journal of Environmental Economics and Management* 47: 535-558.

Williamson, Jeffery G. 2006. *Did British Capitalism Breed Inequality?* London: Routledge.

World Trade Organization (WTO). 2009. "Trade and Climate Change," WTO-UNEP Report, available at http://www.wto.org/english/res_e/booksp_e/trade_climate_change_e.pdf

Yarow, Jay. 2009. "Cap and Trade will cost Households just $175 Annually" The Business Insider Green Sheet, available at: http://www.businessinsider.com/cap-and-trade-will-cost-households-just-175-annually-2009-6.