Information and Strategic Decision-Making in the Oil and Gas Industry: An Empirical Assessment

Thesis by
Mali Zhang

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2018
Defended April 26, 2018

Mali Zhang
ORCID: 0000-0002-7762-5557

ACKNOWLEDGEMENTS

To the members of my dissertation committee, Matthew Shum, Mike Alvarez, Jonathan Katz, and Lawrence Jin, I am deeply grateful for their guidance. To Matthew Shum, my academic advisor and chair of my committee, I am eternally grateful for his wisdom and advice for my research, as well as his kindness and understanding for my professional pursuit. To Mike Alvarez, I am enormously indebted to his commitment to mentoring and professional guidance, and to Jonathan Katz, my deepest gratitude for his knowledge and encouragement for me to pursue a career outside of academia. To Lawrence Jin, I am grateful for his patience and for introducing me to new fields and topics.

To my colleagues, I am deeply thankful for their friendship and valuable comments for my research. Especially to Lucas Núñez, Marcelo Fernandez, and Alejandro Robinson Cortes, I am indebted to their consistent and honest feedback, and their dedication to helping others improve.

To Laurel Auchampaugh, I thank her for being a wonderful source of support unfailingly for us graduate students.

To my friends at Caltech, who have helped me grow so much, I am extremely honored to call them my friends. To Cecilia Huertas and Leah Sabbeth, I am incredibly fortunate to have known them from the very beginning and to have their unconditional support throughout the years. To Serena Ferraro, I thank her for all the crazy adventures and passion for life. To Matt Orr, I am blessed to have him as a voice of reason and wisdom, who introduced me to more than 30 different kinds of scotch. To Andrei Frimu and Nathan Lawless, I am beyond grateful to have met them on the first day of international orientation, and I cannot imagine a day at Caltech without their calming presence.

To my family and friends outside of Caltech, I would not have made it this far without their unwavering support and guidance. To my mother, who encourages and challenges me every day, I owe her my deepest gratitude for always being my backbone. To Scott Connors and Adam Lewicki, I am incredibly grateful for a decade of friendship from them. To Lisa Wang, I am forever indebted to her for putting her name on the line for me. I look forward to our times ahead in Dallas.

ABSTRACT

This dissertation comprises three essays addressing questions from Industrial Organization Economics concerning the oil and gas industry. The essays offer substantive contributions to the study of joint decision-making (Chapter 2), extrapolative beliefs (Chapter 3), and auctions (Chapter 4).

Chapter 2 investigates the quality of joint operations, where multiple oil and gas companies explore a piece of land together. By developing a discrete-choice model which can be matched to actual drilling data, I show that joint operators consisting of only large companies have the least accurate signals. Further counterfactual analyses show that the best policy governing joint operations depends on government priority: to maximize revenue or to avoid damage to the environment.

Chapter 3, co-authored with Lawrence Jin and Matthew Shum, presents a model of dynamic investment and production in which producers over-extrapolate recent demand conditions into the future. We show theoretically and empirically that, in a volatile industry, these biased beliefs can be beneficial in the long-run by counteracting the general trend in the industry. Calibration of our model to Alaska oil exploration shows that the cushioning effect can be large in reducing price decline and accelerating price recovery.

Chapter 4 examines whether common value or private value auction model best describes the bidding decisions made by oil and gas companies. The common value model suggests that more competition can lead to lower equilibrium bids from bidders and lower revenue. By analyzing tract auction data from Alaska, I find that common value components play a slightly larger role when observable heterogeneity is removed. However, expected revenue still increases with competition and plateaus when competition becomes sufficiently high.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

This dissertation comprises three essays addressing questions from Industrial Organization Economics concerning the oil and gas industry. The essays offer substantive contributions to the study of joint decision-making (Chapter 2), extrapolative beliefs outside of finance (Chapter 3), and auctions (Chapter 4). All three chapters utilize a novel dataset collected from Alaska government agencies, which offers comprehensive and detailed information on oil and gas companies' bidding and drilling decisions.

The dataset used by all chapters covers the leasing of government land and exploration behaviors of oil and gas companies. Figure 1.1 shows a simplified timeline of the exploration process, and data available to the public from each step of the process. When a firm becomes interested in a piece of land, she first decides whether she wants to collaborate with another firm. Then when the land is announced for sale, we observe the land attributes, such as location, date of the sale, and reserve price. The firm then participates in a competitive first-price sealed bid auction either as a solo bidder or joint bidder. We observe the outcome of the auctions, such as the participants, whether the participants are solo or joint entities, the bid amount submitted by each participant, and who won the auction. After the winner starts on her lease, she can conduct further seismic studies of the area and decide whether she wants to drill a well here. We observe her ultimate decisions of whether and how many wells to drill, monthly production of the wells, corresponding oil price of the month, and the cost associated with well drilling.

Each chapter focuses on a different aspect of the exploration process, and examines a different topic. Chapter 2 presents an empirical investigation on the quality of joint operations, where multiple oil and gas companies explore a piece of land together. Despite the prevalence of joint operations in oil and gas exploration, there is little evidence for joint operations making better decisions than solo projects. Additionally, joint operations are generally more likely to drill an initial well on a piece of land than their solo counterparts. To address this lack knowledge and to explain this phenomenon, I develop a discrete-choice model which incorporates priors that companies form through the auction stage, and signals that firms receive

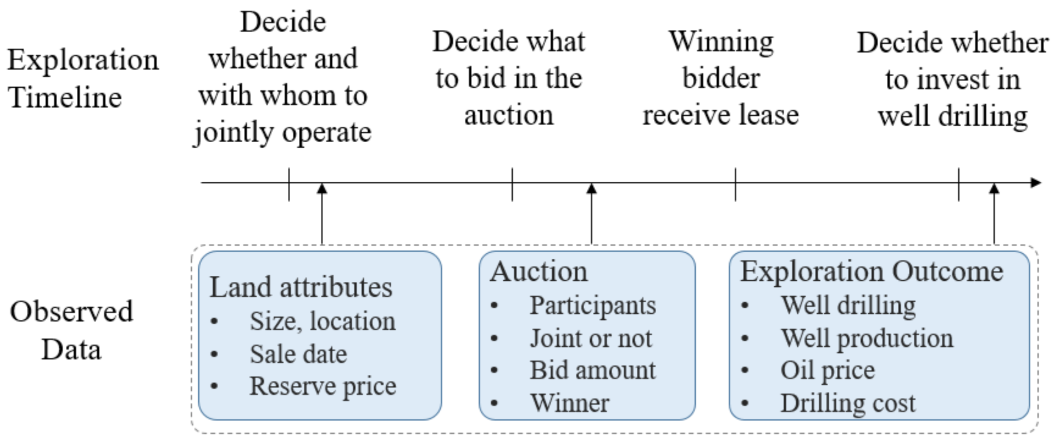


Figure 1.1: **Exploration Timeline and Dataset**

prior to drilling. My structural estimates, based on land clusters operated jointly or solo by at least one large producer, show that joint operators consisting of only large companies have the least accurate signals, followed by solo operators, while partnerships between large and smaller companies have the most accurate signals. The inaccuracy of large-company-only joint operators' signals leads to drilling at places that other types of operators would consider too risky. Furthermore, counterfactual analyses show that if the government forbids joint operations amongst large companies, it can lose up to $5.1 billion in revenue, while saving up to 390 thousand acres of land from unnecessary drilling. Hence, the best policy for a government to implement depends on what its priority is: whether to maximize its revenue or to prevent pollution.

Chapter 3, co-authored with Lawrence Jin and Matthew Shum, presents a model of dynamic investment and production in which producers may have biased beliefs in which they over-extrapolate recent demand conditions into the future. This bias leads producers' beliefs to exhibit *insufficient mean reversion*, as these producers underestimate the degree of mean reversion in the demand process. In a volatile industry, while biased beliefs lead firms to make sub-optimal investment decisions in the short-run, they can be beneficial in the long-run by counteracting the general trend in the industry, "cushioning" the industry against prolonged downturns and aiding faster recovery. As an empirical case study, we consider oil exploration in Alaska. We present evidence that firms in this industry were subject to extrapolation bias, leading to drilling of lower-profit wells after recent price increases. Calibration of our model to Alaska oil exploration shows that the cushioning effect can be large: in a typical episode of oil price decline arising from a sequence of adverse demand shocks, the cushioning effect reduces the decline of the oil price by 8.2% and accelerates the price recovery by 27%. This showcases the potential positive implications that biased beliefs can have on industry dynamics.

Chapter 4 examines which auction model best describes the bidding decisions made by oil and gas companies in Alaska. While the Alaska government encourages the bidding of small companies and individuals to spur competition and increase the state revenue, depending on the actual motivation of the bidder, it may not be the correct policy. A pure common value auction model suggests that more competition can potentially lead to lower equilibrium bids from bidders and thus lower revenue. By analyzing oil and gas tract auction data from Alaska, I find that the private value component plays a much larger role when tracts are assumed to be homogeneous. But common value component starts playing a larger role when we remove all observable heterogeneity. Given a particular tract and signal, companies are likely to bid lower when they face greater competition due to winner's curse effect. However, the winner's curse effect is not enough to offset the greater chance of higher signals. Expected revenue still increases when competition increases and plateaus when competition becomes sufficiently high. I also find that the reserve prices of $5 and $10 do not seem sufficient on the tracts that attracted large oil and gas companies.

Chapter 2

ARE TWO HEADS BETTER THAN ONE?: JOINT VERSUS SOLO OPERATION IN OIL AND GAS EXPLORATION

2.1 Introduction

Overview

Joint operation is a common practice in the oil and gas industry, both in onshore and offshore explorations. They involve two or more companies or individuals entering into a legally binding agreement to share the operational and financial responsibilities in the exploration and development process. Joint bidding, a phenomenon where two or more entities bid on a lease as one joint entity, is usually a result of such operations. On federal land, policymakers are concerned that joint bidding serves the purpose of collusion to stifle competition and lower lease sale prices. As a result, in 1975, both the Department of Interior and Congress passed regulations to prevent eight of the then-largest oil companies worldwide from joint bidding with each other (Hendricks and Porter, 1992).[1] In Alaska, however, joint bidding amongst its largest companies consistently accounted for a big share of winning bids submitted in auctions: 40.6% before 1975, and 30.1% between 1975 and 2000.[2] Joint operations formed by large companies also drilled the highest percentage of their owned land, at around 25% as opposed to just above the 16% from solo large firms or partnerships between large and fringe firms.

While previous industrial organization papers focused exclusively on collusion and the bidding behaviors of joint operations, exploration and development after the auction play a much bigger role for the government revenue and environmental outcomes. Royalty payments from a successful venture can be thousands of times the winning bid in an auction. Meanwhile, the ability to drill more accurately lowers the impact on the environment, avoiding unnecessary alterations to land formation and production of various drilling wastes (O'Rourke and Connolly, 2003). Hence, if joint operations explore more because they are willing to take risks and find oil at places that solo operations would not be able to, then it explains why some governments

[1] The eight companies are Exxon, Gulf, Mobil, Shell, Standard Oil of California, Standard Oil of Indiana, Texaco, and British Petroleum.

[2] The 17 largest companies are ARCO, BP, Unocal, Chevron, Conocophillips, Mobil, Amoco, Exxon, Sinclair, Shell, Texaco, Amerada, Marathon, Anadarko, Union Texas, Sunoco, and Gulf.

do not restrict them despite the collusion concern, and that policymakers focused on maximizing government revenue should continue to sanction such operations. On the other hand, if companies in a joint operation, by pooling their expertise and information, have better signal accuracy and avoid unnecessary drilling, then their operations can lower environmental damage. In this case, policymakers prioritizing environmental protection should encourage such operations.

In this paper, I examine whether joint operations indeed have better information than their solo counterparts. Past organization and finance literatures also raised other hypotheses regarding why companies jointly operate: (1) entering into joint operation (or investment syndication in the venture capital literature) assures participants of the investment quality, when they feel pessimistic about their investment, (2) joint operations allow companies to combine their management expertise and improve project quality, (3) companies enter into joint exploration agreement to collude and avoid competition with each other in the auction stage, and afterwards they divide the land and explore by themselves, where the cost saved in the auction allows them to drill more aggressively, and (4) companies partner with each other for risk sharing purpose, allowing them to invest in more land, and the increased land masses achieved by splitting the cost improve their chances of finding a well location worth drilling. The model explores these potential hypotheses but cannot untangle one from the other.

To tackle these questions, I consider a model in which an operator makes drilling decisions according to a cut-off strategy that incorporates her prior belief on the land value (in terms of oil deposit) formed after the auction, and additional signal received before drilling. This operator, composed of one or multiple oil companies, forms a prior belief of the land value after observing the land characteristics, participating and winning in the auction that sells this land. With this prior in mind, the aforementioned operator, who has won the right to explore and gather additional geophysical data, receives another signal of the true value after additional testing. Updating the prior with the new signal, the operator chooses to drill an initial well on this land if the expected oil production, conditional on her signal, exceeds the break-even amount given her drilling cost. By finding parameters that best match the model prediction with the observed drilling decision for each operation, I can test whether joint operations manage to obtain more accurate signals, and whether companies enter into joint operations due to their pessimism about the land productivity.

Estimation of the structural model uses a sample of land clusters composed of those operated by solo large companies, joint entities composed of only large companies, or partnerships between large and smaller firms. Results of the estimation suggest that solo operations have more accurate signals than joint operations formed only by large companies, but have less accurate signals that joint operations between large and smaller firms. This result suggests that large companies collaborate with one another for reasons outside of expertise or information sharing. Other explanations, such as collusion or risk diversifying may be more applicable. Counterfactual analyses show that this inaccurate signal for large-only joint operators leads to more drilling at places that the other two types of operators are less likely to drill. As a result of their bolder efforts, joint operations between large companies may have contributed additionally up to 140 million barrels of oil and 363 million thousand cubic feet of natural gas in royalty payment to the Alaska government. These quantities equate to approximately 5.1 billion dollars of revenue to the state. The other side of the counterfactual analyses is that banning joint operations between large companies can potentially reduce unsuccessful drilling in up to 18 clusters, with a total of 390,190 acres.

As a whole, this paper sheds light on the benefits and potential harms of joint operations that could be informative to policymakers. For policymakers concerned with revenue generation, joint operation can be a lucrative channel to produce extra royalty revenue for the government, as they tend to explore more boldly and may find oil in unexpected places. However, policymakers concerned with long-term conservation of the ecosystem may need to re-evaluate their policies regarding joint operations, as the extra drilling may cause extra irreversible damage to the areas of exploration, which could have been preserved by reducing partnerships between large companies.

Literature Review

Joint operation is a widely studied subject in the industrial organization (IO) and finance literature. In finance, joint operation takes on the form of investment syndication amongst venture capitalists. Lerner raised three hypotheses on rationales behind syndication and found evidence supporting all of them by looking at the biotechnology industry. The three rationales are (1) assurance of peers on the quality of investment, (2) for later-round investors, avoidance of opportunistic behaviors from the initial investor, and (3) performance boosting to potential investors (Lerner, 1994). Another study by Brander et al. introduced another motivation: (4) syndication to share management expertise. They found evidence supporting the assurance of peers hypothesis, but not the management expertise hypothesis (Brander, Amit, and Antweiler, 2002). My study seems to partially support the management expertise hypothesis, as joint operators consisting purely of large companies have less accurate signal, but partnerships between large and smaller companies have more accurate signals than solo operators.

This paper differs from previous studies in the IO literature substantially by focusing on the drilling decisions after the auction stage, rather than the formation of a joint entity before the auction and its behavior during the auction. An early observational study by Hendricks and Porter examined what tract and company characteristics facilitated joint bidding in federal outer continental shelf (OCS) lease sales (Hendricks and Porter, 1992) . By focusing on three joint bidding types: large companies bidding with each other (L&L), fringe companies bidding with each other (F&F), and large companies joint with fringe companies (L&F), they found that large companies tend to bid solo or jointly with each other on high-valued tracts. Meanwhile, joint ventures between large and fringe companies did not bid substantially higher or lower, which prompted the explanation that this partnership was purely for exchange of knowledge and expertise. My results seem to support this claim. Compared to L&L joint operators, an alliance between large and smaller companies have significantly lower signal noise. Hendricks, Porter, and Tan followed up with a theoretical paper that focused on the formation of two-party bidding rings in common value auctions (Hendricks, Porter, and Tan, 2008). They found that bidding rings do not form when the prior belief of the tract value is pessimistic, when the reserve price is too high, and the competition is low. My results seem to disagree with their prediction, as joint operations seem to be associated with lower expected land values.

In addition to examining hypotheses proposed by previous studies, this paper also introduces a new model and estimation strategy to analyze joint operations. While I apply this model to the oil and gas industry, a key sector in the American economy, this model may also be applicable to other industries as well.

Institutional Background

To better understand the rest of the paper, I provide some background on how joint operation works. Joint operation happens when multiple companies are interested in exploring a certain area together. Once companies set sights on an exploration

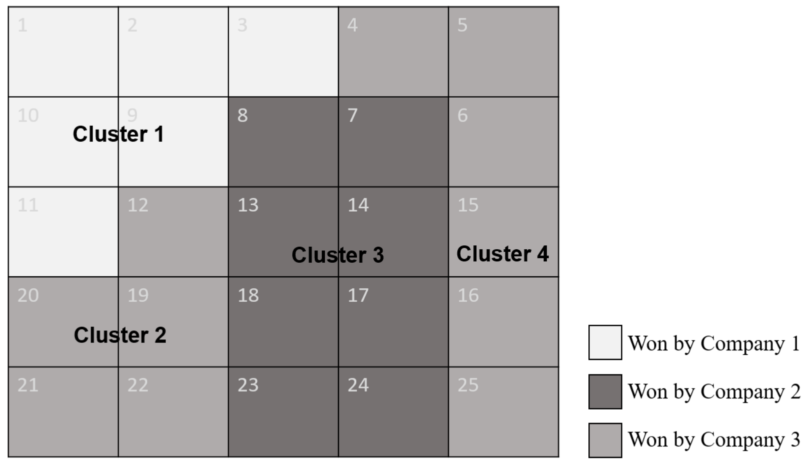


Figure 2.1: **Exhibit of operation cluster**

area, the decision of whether to enter into a legal agreement to collaborate with each other or to explore on their own depends on a variety of factors, such as how familiar the company is with the area and their financial strengths for example. If companies decide to jointly explore, they need to agree on a variety of terms in a participation agreement (also referred to as "exploration agreements", "development agreements", "joint development agreements," etc.). These terms include: who will be in charge of bidding in the auction, which leases in the area they would like to bid on, amount to bid in an auction, how they want to collect and utilize seismic data of the area, how to split financial responsibilities, and so on (Villarreal and Lavoy, 2010). If they win in the auction, they will have the right to explore on the tract (the unit of land sold in government lease auctions) for 5, 7, or 10 years. The winner can further collect geophysical data in the area before deciding whether to drill an initial well, and, if they decide to drill, the exact specifications of the well. The difference in information quality may arise in this process. One possibility is that joint operations may have multiple inputs in collecting and interpreting the geological data, giving them more accurate signals than their solo counterparts. Another possibility is that participants may not interact with each other about technical details or even actively withhold proprietary information from each other, which can lead to lower information quality.

Because companies make their exploration decisions based on prospect areas, the unit of observation in my study should reflect that. Rather than using tracts, which are artificial divisions of land masses that the government utilizes for auction purposes, I use geographically adjacent clusters of tracts sold in the same auction and operated by the same company as a unit of analysis. Figure 2.1 is an exhibition of how this clustering works. Figure 2.1 shows 25 tracts that are adjacent to each other. These 25 tracts are won by 3 different bidders in the same auction, as shown by the three different shades in Figure 1. Because of the geographic proximity of these tracts, it is reasonable to believe that when firms make exploration decisions, they consider them as one prospect area, and make drilling decisions based on the whole area instead of each individual rectangle. However, for company 3, the two blocks of tracts that she won are separated from each other. Because the information from one block may no longer be as useful to the other block due to the separation, company 3 has two clusters in this area. Hence, in this area overall, there are four units of operation. For a particular company, since each cluster is either spatially or temporally separated, I argue that the company's decision to drill in each cluster can be treated as independent. One may argue that for two adjacent clusters sold at different times, finding oil in one cluster may lead to drilling in the other cluster. It is a concern that the model has not been able to address. But to lower the effect of such knowledge on the model estimation, I use an indicator for whether the operator owns adjacent clusters as a control for the information leakage from previously-owned land.

2.2 Model

Companies conduct oil and gas explorations and developments on K independent clusters as defined in Section 1.3. In each operation k, where $k = 1, 2, ..., K$, there is a true value V_k associated with the piece of land. V_k can be thought of as the amount of oil deposit, or the annual outflow of oil if a productive well were drilled, for this cluster of land k. Companies who participate in the auction of this cluster have collected their own information regarding the land, observed the auction outcome, and thus formed prior beliefs on V_k. The winner from the competitive auction then gathers further geophysical data of the land and as a result receives an additional signal S_k of the value of V_k. Based on her prior belief and the additional signal, the winner, and now operator of the cluster, decides whether she want to drill an initial well in this land. Figure 2.2 exhibits the timeline of this process.

An important assumption of the model is that, regardless of whether the operation is solo or joint, only one signal is considered for the drilling decision in each operation. This is because we only truly care about the final outcome of the aggregated signal, not the deliberation process to aggregate individual signals. As the deliberation

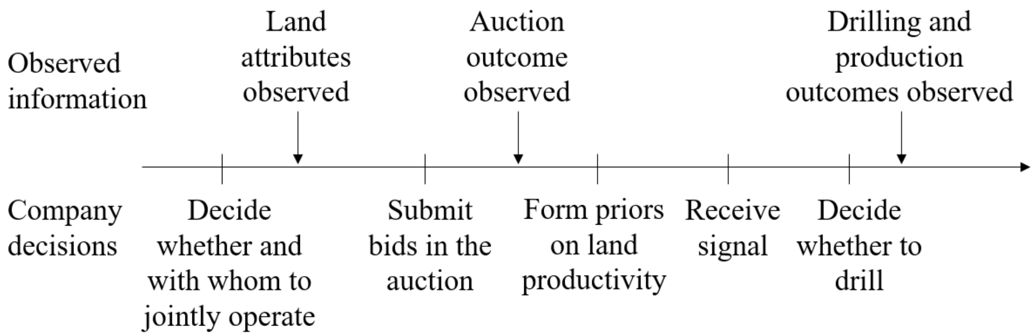


Figure 2.2: **Timeline of Drilling Decision-making in the Model**

procedure and the number of decision-making parties could vary from one joint operation to another, the only factor essential to the final drilling decision is the eventual signal such decision is based upon. We want to compare this final signal between joint and solo operations, and see if accuracy differs between the two.

The specifics of the drilling decision-making are as follows. Each operation is associated with a cost of drilling C_k, which is highly substantial. For such a cost, the well needs to produce a certain amount of oil, Q_k^*, in each period in order for the operator to break even. Suppose the operator drills a well l periods after her lease contract starts and the well is productive for n periods, then Q_k^* is the break-even per period quantity of production that can be solved by the following equation:

$$Q_k^* \cdot \sum_{\tau=l}^{l+n} (1-\delta)^\tau P_{t+\tau} = C_k.$$

Here $P_{t+\tau}$ is the price of oil in period $t+\tau$ and δ is the discount factor.

Knowing the break-even quantity Q_k^*, and after observing her private signal $S_k = s_k$, the operator would choose to drill if the expected per period production, conditional on s_k, exceeds Q_k^*, i.e., $\mathbb{E}[V_k|s_k] \geq Q_k^*$. Under specific assumptions on S_k, which will be discussed in detail in Section 2.4, $\mathbb{E}[V_k|s_k]$ is a monotonic non-decreasing function of s_k. As a result, there exists s_k^* such that $\mathbb{E}[V_k|s_k^*] = Q_k^*$ and for all $S_k \geq s_k^*$, $\mathbb{E}[V_k|s_k^*] \geq Q_k^*$.

Thus, as long as the signal received by the operator exceeds s_k^*, the company would drill. So the probability of drilling for cluster k is $Pr(S_k \geq s_k^*)$. Let $f(\cdot)$ be the density function for the prior distribution of V_k with support $\mathcal{V}$, and let the distribution of S_k follows a conditional density function of $g(\cdot|V_k)$, then the probability of drilling becomes

$$\begin{aligned} Pr(A_k = 1) =& Pr(S_k \geq s_k^*) = \int_{\mathcal{V}} Pr(S_k \geq s_k^* | V_k) f(V_k) dV_k \\ &= \int_{\mathcal{V}} [1 - G(s_k^* | V_k)] f(V_k) dV_k \end{aligned} \tag{2.1}$$

In other words, this model is akin to a discrete choice model where the latent variable is S_k, and we observe the drilling decision of 1 if S_k exceeds the threshold s_k^*, and 0 otherwise.

2.3 Data

Data used in this study come from three main sources: the Alaska Department of Natural Resources (DNR) who governs the sales and contracting of tracts, the Alaska Oil and Gas Conservation Commission (AOGCC) who governs the well permitting and monitoring, and U.S. Energy Information Administration (EIA), who tracks oil prices and costs in major U.S. oil producing regions.

The DNR keeps a comprehensive records on all tracts sold in competitive auctions since 1959. Such records include pre-auction information for the tracts: date of the lease sale, the auction format, and key tract characteristics, such as acreage, royalty requirements, and locations. Post-auction results are also well documented. We can see the value of bid submitted by each bidder, the participant(s) in each bidding party, and whether a winner relinquished her tract after winning. In this study, I limit the scope of tracts to those sold between 1959 and 2000 through cash bids only, and were not relinquished by the owner after they won. I limit the sales year to before 2000 because three of the largest operators in Alaska merged with each other in 2000 and 2001: ARCO first merged with Phillips in 2000, and then Phillips merged with Conoco to form ConocoPhillips (WSJ, 2001; ConocoPhillips, 2017). These mergers substantially reduced the number of large operators in Alaska and the sample of joint bidding between the companies. In addition, the merged firm may behave very differently from the individual firms pre-merger, which could skew results from the pre-merger period. This cutoff also allows for long enough windows to observe the drilling decisions made on tracts sold in 2000. I also remove tracts sold using non-cash payments, such as net profit sharing percentage, because information revealed through these payments is not comparable to cash bids. Finally, relinquished tracts will not have been explored and should not be included in the data. After these eliminations, I retain 4,838 leases for analysis, which can be combined into 1,516 clusters.

The AOGCC, meanwhile, documents all drilling and production activities after the sales and contracting. This data source contains a variety of information on all wells that were permitted to drill in Alaska since 1901, including history for each well such as permit date and drilling date, basic well attributes such as its geo-code, depth and which production unit it belongs to, and monthly productions of oil, gas, and water for each well.

What is special about the Alaskan data, besides the level of details available, is the relatively even mix of joint and solo operations. Joint bidding, especially joint bidding amongst large companies, remains a common phenomenon throughout our observation period. Before 1975, the year joint bidding was outlawed amongst the largest companies in the world on federal land, joint bidding accounted for 40.6% of leases sold to the largest 17 companies in Alaska.[3] After 1975, the percentage remains high, at 30.1%. This balanced mix of joint and solo operations makes the comparison between them possible.

Finally, the EIA provides the U.S. annual crude oil first purchase prices from 1900 to 2015, and the cost of drilling an oil well from 1960 to 2007. Ideally I would like to use the Alaska prices. However, Alaska North Slope prices are not available prior to 1977, which would prevent the use of the large number of joint operations prior to that. By comparing the monthly level data between the U.S. prices and Alaska prices, I find a 99.7% correlation between the prices. Hence, for the oil prices shown in the model ($P_{t+\tau}$), I use the annual U.S. first purchase price per barrel. The crude oil well cost is also an average amount based on key oil producing regions in the U.S. Ideally I would like to have production cost varied by each well in Alaska, since different distances to existing infrastructure, crew relationships and access to pipeline could lead to different production cost. However, these metrics are not available in this dataset. I hence use a constant drilling cost for all wells drilled in the same year.

Summary Statistics

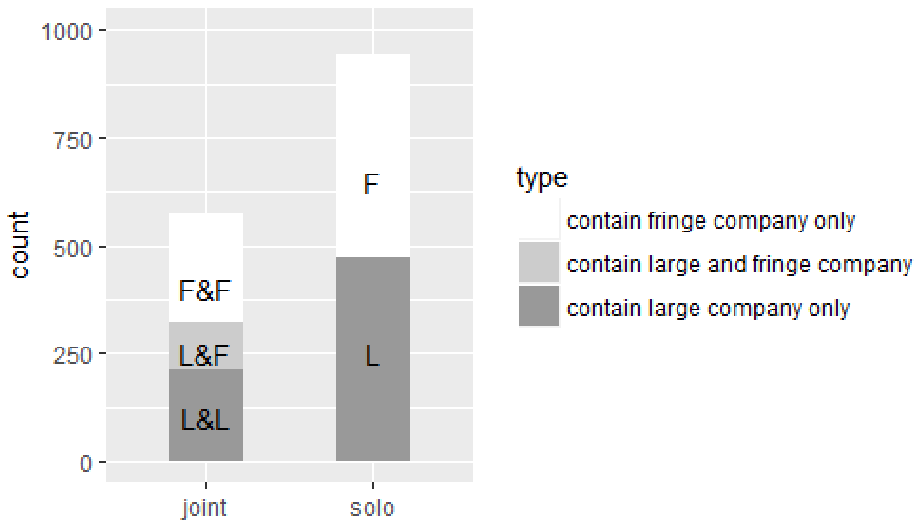
In this section and onward, it is helpful to divide joint operations into three types, similar to those in the 1992 Hendricks and Porter paper: large companies joint operating with each other (L&L), large companies joint with fringe companies (L&F), and fringe companies partnering with each other (F&F) (Hendricks and

[3]These large companies are: ARCO, BP, Unocal, Chevron, Conocophillips, Mobil, Amoco, Exxon, Sinclair, Shell, Texaco, Amerada, Marathon, Anadarko, Union Texas, Sunoco, and Gulf.

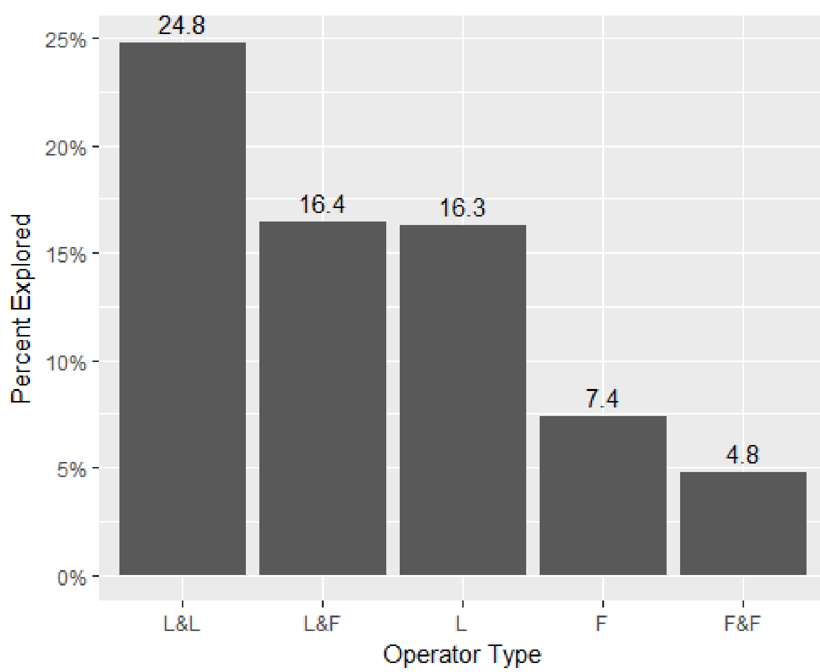
Figure 2.3: **Counts and Exploration by Joint Status and Type**

.

(a) Count of joint and solo ops by type



(b) Proportion of clusters explored by type



Out of 1,516 clusters, the number of clusters operated by L, L&L, L&F, F, and F&F types are: 472, 210, 110, 473, and 251, respectively.

Porter, 1992). This is because each type may have different motivations for obtaining leases. The L&L type is formed by large international companies exclusively, defined as the 17 companies with international presence who bid on at least 150 tracts and won over 90 of them between 1959 and 2000.[4] These companies have the resources and technology to explore and develop oil fields themselves. The L&F type is formed by large companies, as defined above, partnering with smaller local firms. In this case, the large company may be seeking the knowledge of smaller firms in the local formation of the area, as they already have the financial resources and technology. The F&F type is formed solely by smaller firms. These operators may be single investors or smaller local investors. The two solo types are individual large and fringe companies, and I will denote them as L and F, respectively. Figure 2.3a exhibits the total number of clusters for each type under solo and joint operation.

However, the likelihood of exploration varies significantly for different types, as shown in Figure 2.3b. For clusters owned by large companies, joint operators are much more likely to drill than a solo large company. Clusters owned by a joint entity between large and smaller firms, on the other hand, have similar drilling likelihood as clusters owned by solo large operators. Finally, clusters owned by smaller firms, whether solo or joint, are drilled much less often than those owned by their large counterparts. A possible explanation for this is that many of the smaller winning bidders are individual investors. Some of them purchase land not to explore for themselves, but to farm it out to others, or in the hope that their large neighbors find oil and they can be compensated for it. Because of the varying purposes of land ownership for this group, this study will focus on the clusters owned by at least one large company, hence the L, L&L, and L&F types.

Restricting the sample to types consisting of at least one large company, I look at how the clusters vary in attributes for each type. Table 2.1 summarizes how clusters operated by each category of operators differ in some key measures. The clusters won by joint operators are on average larger in area than their solo counterparts. They pay substantially more for these tracts, and are faced with more competition in the auctions. Joint bidders also tend to be less experienced, as they have operated fewer leases in the same region previously. This may suggest that joint operations tend to happen when companies are less familiar with the area. And when companies do jointly bid, they like to bid on tracts adjacent to the ones that they already own, further reducing the risk from a lack of knowledge in the exploration area. Large companies

[4] The exception here is Gulf Oil, who only won 82 tracts in this period. But it is listed as one of the eight largest companies forbidden from joint bidding by the Congress and Department of Interior.

Table 2.1: **Attributes of Clusters by Category**

	Solo-Large (L)		Large w Large (L&L)		Large w fringe (L&F)	
Continuous Variables	Median	Mean	Median	Mean	Median	Mean
Acreage	5,120	9,964	5,243	13,923	5,112	10,626
Win bid ($ per acre)	25.4	255.7	52.8	858.6	47.3	3,313
Number of bids	2	2.4	2	2.8	2.4	3.0
Experience	19.00	27.11	17.50	27.05	11.00	23.34
Discrete Variables	Mean		Mean		Mean	
Knowledge (% owning neighbor)	10.4%		23.3%		20.9%	
Region (% in Cook Inlet)	57.0%		56.2%		49.1%	
Year (% after 1975)	54.2%		39.5%		52.7%	

are also more likely to jointly bid with smaller companies in the North Slope region, than the Cook Inlet region. Finally, joint operations composed solely of large companies happened more before 1975. Because of the underlying heterogeneity across the clusters operated by solo and joint operators, I will incorporate these core measures into the controls for the model.

2.4 Estimation

In this section, I describe the estimation strategy of the model. The goal of estimation is to identify $Pr(A_k = 1)$ for each cluster in the sample. In order to do that, we need to separately identify the conditional distribution of S_k and the distribution of V_k. That is, we need to find the $g(s_k^*|V_k)$ and $f(V_k)$ in Equation 2.1.

However, identifying the signal quality and prior belief distribution is challenging without imposing structural assumptions on these quantities. The difficulty of identifying the continuous distributions of signal accuracy directly from the data arises because signals, private to operators, are not observable to econometricians. Determining the prior belief distribution is also challenging, as we do not directly observe companies' ex-ante estimate of the land productivity. One potential idea is to examine the ex-post production of the land after it is drilled. But this approach is highly biased since companies only choose to drill in areas that they believe to have a good chance of success. The distribution recovered from the ex-post production may look very different from the ex-ante expectation. However, to take advantage of this information, I use the ex-post production amount as guidance the distribution assumption imposed on V_k.

To resolve these aforementioned challenges, I place parametric restrictions on the prior belief distribution of V_k and the distribution of signal S_k. Then I show that by relating the structural setup to the observed drilling data, we can recover the values of these parameters.

Parametric Assumptions and Identification

The formal estimation requires two key assumptions on the true land value V_k and signal S_k. For simplicity, I assume all operations are homogeneous with the exception of whether it is joint or solo. J_k is an indicator here which equals 1 if the operation is joint, and 0 otherwise. In the next section I show that these arguments hold when we allow operations to be heterogeneous.

The first key assumption is that the prior belief on the distribution of V_k follows a log-normal distribution. That is,

$$\tilde{V}_k \equiv \log V_k \sim N(\nu, \sigma_v^2). \tag{2.2}$$

Here σ_v is the uncertainty that operators have around the tract quality after the auction. This assumption is supported by the production amount from all wells drilled. A plot of the production amount can be found in Appendix A.1. In Figure A.1a, productions from all wells is highly skewed to the right, with a mass at zero, where wells produce nothing. Since production is only carried out above a certain threshold, we do not care about the distribution to the left of the threshold. Hence, we can approximate this distribution with a highly right-skewed log-normal distribution. Figure A.1b restricts the wells to productive ones only, whose oil production amounts approximate a log-normal distribution.

The second key assumption is that S_k is a noisy signal of V_k also in the log-normal form, i.e.,

$$\tilde{S}_k \equiv \log S_k = \tilde{V}_k + \sigma_s^k \epsilon, \tag{2.3}$$

where σ_s^k is the level of uncertainty for this signal, and ϵ is distributed standard normal and independent of $\tilde{V}_k$. Thus, we have $\tilde{S}_k|\tilde{V}_k \sim N(\tilde{V}_k,\ (\sigma_s^k)^2)$ and $\tilde{S}_k \sim N(\nu_k,\ (\sigma_s^k)^2 + \sigma_v^2)$. Let $\sigma_s(J_k;\theta_s) = \alpha_0 + \alpha_J \cdot J_k$, where $\theta_s = (\alpha_0,\ \alpha_J)$. Then if $\alpha_J < 0$, there is evidence supporting joint operations having more accurate signals.

By the log-normal assumption (2.2) and (2.3), we have the conditional mean and variance of $\tilde{V}_k$ given $\tilde{s}_k$,

$$\begin{aligned}\mathbb{E}[\tilde{V}_k|\tilde{s}_k] &= \frac{\sigma_v^2 \cdot \tilde{s}_k + (\sigma_s^k)^2 \cdot \nu}{\sigma_v^2 + (\sigma_s^k)^2} \\ var[\tilde{V}_k|\tilde{s}_k] &= \frac{(\sigma_s^k)^2\sigma_v^2}{\sigma_v^2 + (\sigma_s^k)^2}.\end{aligned} \tag{2.4}$$

Then,

$$\begin{aligned}\mathbb{E}[V_k|s_k] &\equiv \exp(\mathbb{E}[\tilde{V}_k|\tilde{s}_k] + \frac{1}{2}var[\tilde{V}_k|\tilde{s}_k]) \\ &= \exp(\frac{\sigma_v^2 \cdot \tilde{s}_k + (\sigma_s^k)^2 \cdot \nu + \frac{1}{2}\sigma_v^2(\sigma_s^k)^2}{\sigma_v^2 + (\sigma_s^k)^2}),\end{aligned} \tag{2.5}$$

which is a monotonically increasing function of $\tilde{s}_k$. Hence, by setting $\exp(\mathbb{E}[\tilde{V}_k|\tilde{s}_k] + \frac{1}{2}var[\tilde{V}_k|\tilde{s}_k]) = Q_k^*$, we can solve for $\tilde{s}_k^*(\theta)$ such that for all $\tilde{S}_k \geq \tilde{s}_k^*$, $\mathbb{E}[V_k|s_k] \geq Q_k^*$. Here Q_k^* is calculated by

$$Q_k^* \cdot \sum_{\tau=l}^{l+9}(1 - 0.05)^\tau P_{t+\tau} = C_k, \tag{2.6}$$

since on average wells produce for 10 years, and the discount rate is at 5%, as suggested by Hendricks, Pinkse, and Porter (2003). Then $\tilde{s}_k^*(\theta)$ is

$$\tilde{s}_k^*(\theta) \equiv \tilde{s}^*(J_k;\theta) = \frac{[\sigma_v^2 + (\sigma_s^k)^2] \cdot \log Q_k^* - (\sigma_s^k)^2\nu - \frac{1}{2}\sigma_v^2(\sigma_s^k)^2}{\sigma_v^2}. \tag{2.7}$$

By applying (2.2), (2.3), and (2.7) to (2.1), we obtain conditional probability of drilling

$$\begin{aligned}Pr(A_k = 1|J_k,\theta) =& Pr(S_k \geq s_k^*|J_k,\theta) = Pr(\exp(\tilde{S}_k) \geq \exp(\tilde{s}_k^*)|J_k,\theta) \\ =& Pr(\tilde{S}_k \geq \tilde{s}_k^*|J_k,\theta) \\ =& 1 - \Phi\Big[\frac{\tilde{s}^*(J_k;\theta) - \nu}{\sqrt{(\sigma_s^k)^2 + \sigma_v^2}}\Big].\end{aligned}$$

Define the log-likelihood for observation k as

$$\begin{aligned}l_k(\theta) &\equiv l(A_k, J_k, \theta) \\ &= A_k \log Pr(A_k = 1|J_k,\theta) + (1 - A_k)\log(1 - Pr(A_k = 1|J_k,\theta)).\end{aligned} \tag{2.8}$$

Then the likelihood function for all observations is

$$\mathscr{L}(\theta) \equiv \frac{1}{K}\sum_{k=1}^{K} l_k(\theta). \tag{2.9}$$

Our goal is to find the $\hat{\theta}$ that maximizes (2.9).

Incorporating Operation Heterogeneity

For simplicity of discussion, the only covariate being considered in the previous subsection is whether the operation is solo or joint. However, this model allows me to incorporate cluster and operator heterogeneities as reduced-form parameters.

Specifically, the prior belief distribution can vary widely based on tract level characteristics, such as cluster acreage, region, and the competitiveness of its auction, as well as operator characteristics, such as their knowledge of the tract, and experience operating in the same region. Let these variables be denoted by $\boldsymbol{W}_k$. To test the hypothesis whether joint operation is associated with a pessimistic belief of the land value, I also include the joint indicator as a parameter associated with prior belief distribution. Hence,

$$\begin{aligned} \nu^k &\equiv \nu(J_k, \boldsymbol{W}_k; \theta_\nu) = \beta_0 + \beta_J \cdot J_k + \boldsymbol{W}_k \beta, \\ (\sigma_\nu^k)^2 &\equiv \sigma_\nu^2(J_k, \boldsymbol{W}_k; \theta_\nu) = \exp(\gamma_0 + \gamma_J \cdot J_k + \boldsymbol{W}_k \gamma), \end{aligned} \tag{2.10}$$

where $\theta_\nu = (\beta_0, \beta_J, \beta)$ and $\theta_\nu = (\gamma_0, \gamma_J, \gamma)$.

In addition, the noise of a signal can also be affected by the operator's experience and knowledge of the land. Familiarity with the area can potentially increase an operator's ability to interpret the signal. Let these variables be denoted by $\tilde{\boldsymbol{W}}_k$, then

$$(\sigma_s^k)^2 \equiv \sigma_s^2(J_k, \tilde{\boldsymbol{W}}_k; \theta_s) = \exp(\alpha_0 + \alpha_J \cdot J_k + \tilde{\boldsymbol{W}}_k \alpha),$$

where θ_s becomes $(\alpha_0, \alpha_J, \alpha)$.

Furthermore, to study how signal qualities vary across different types of operators, I analyze a sample containing clusters that are jointly owned by large and fringe firms, jointly owned by only large firms, and those operated by solo firms. I differentiate the types by having an additional parameter for the L&F type:

$$\begin{aligned} \nu^k &\equiv \nu(J_k, J_k^{lf}, \boldsymbol{W}_k; \theta_\nu) = \beta_0 + \beta_J \cdot J_k + \beta_{lf} \cdot J_k^{lf} + \boldsymbol{W}_k \beta, \\ (\sigma_\nu^k)^2 &\equiv \sigma_\nu^2(J_k, J_k^{lf}, \boldsymbol{W}_k; \theta_\nu) = \exp(\gamma_0 + \gamma_J \cdot J_k + \gamma_{lf} \cdot J_k^{lf} + \boldsymbol{W}_k \gamma), \\ (\sigma_s^k)^2 &\equiv \sigma_s^2(J_k, J_k^{lf}, \tilde{\boldsymbol{W}}_k; \theta_s) = \exp(\alpha_0 + \alpha_J \cdot J_k + \alpha_{lf} \cdot J_k^{lf} + \tilde{\boldsymbol{W}}_k \alpha), \end{aligned} \tag{2.11}$$

where J_k^{lf} is an indicator for L&F operators. Hence, the impact on $(\sigma_s^k)^2$ from the L&F type is $\alpha_J + \alpha_{lf}$, whereas for L&L operators, the influence is only α_J.

Incorporating these heterogeneities, let $\boldsymbol{X}_k$ denote the set of all covariates and $\theta = (\theta_\nu, \theta_\nu, \theta_s)$, then the probability of drilling becomes

$$Pr(A_k = 1 | \boldsymbol{X}_k, \theta) = 1 - \Phi\Big[\frac{\tilde{s}^*(\boldsymbol{X}_k; \theta) - \nu(\boldsymbol{X}_k; \theta_\nu)}{\sqrt{\sigma_s^2(\boldsymbol{X}_k; \theta_s) + \sigma_\nu^2(\boldsymbol{X}_k; \theta_\nu)}}\Big] \tag{2.12}$$

Now the log likelihood for observation k becomes

$$\begin{aligned} l_k(\theta) &\equiv l(A_k, \boldsymbol{X}_k, \theta) \\ &= A_k \log Pr(A_k = 1|\boldsymbol{X}_k, \theta) + (1 - A_k) \log(1 - Pr(A_k = 1|\boldsymbol{X}_k, \theta)) \end{aligned} \tag{2.13}$$

and we again need to find $\hat{\theta}$ that maximizes Equation 2.9 derived from expression 2.13.

Endogeneity of Auction Results Information For prior-related measures, ν^k and σ_ν^k, I use information revealed in the auction, including the winning bid and number of bids submitted. However, the decision to enter into joint operation may be endogenous in determining the winning bid price and number of bids. For example, if the only two bidders interested in the lease decide to collude and submit only one bid, it could lower the winning bid significantly and cut the number of bids by half. Since this model does not account for the selection into joint bidding and auction stage, we do not have a way to control for this possibility. Developing future models which includes the selection process is necessary to eliminate this endogeneity.

2.5 Results

In this section, I first describe the results from estimations using a sample of land clusters operated at least partially by a large company, and discuss the implications of these results. I then conduct two counterfactual analyses where joint operation is not allowed at all, or only allowed in the form of partnership with smaller local firms. I look at the exploration outcomes from the clusters that were won by L&L bidders, and discuss how the state income would be impacted if a solo operator or a L&F joint operator instead were making drilling decisions on these clusters.

Estimates of Structural Parameters

The sample used in the estimation contains 792 clusters that are operated at least partially by a large operator (L, L&L, and L&F types), which include 472 solo-operated clusters, 210 owned by the L&L type, and 110 clusters owned by L&F operators. For this estimation, I use the specifications from Equations 2.10. For a robustness check, I estimate the parameters using specifications from Equations 2.11 and show that the parameters for joint operation in both specifications are not statistically different from each other. The parameter values from using clusters operated solely by large firms are shown in A.3

Table 2.2: **Estimation Results: Clusters At Least Partially Operated by Large Operators**

Description \ Metrics	σ_s Parameter	σ_s Estimate	σ_v Parameter	σ_v Estimate	ν Parameter	ν Estimate
Constant	α	3.197	γ_0	5.630	β_0	-51.699
		(0.441)		*(1.328)*		*(2.911)*
Joint: L&L	α_J	0.413	γ_J	-0.022	β_J	-0.384
		(0.186)		*(0.296)*		*(0.441)*
Joint: L&F	α_{lf}	-0.602	γ_{lf}	0.001	β_{lf}	-0.138
		(0.336)		*(0.561)*		*(0.691)*
Experience	α_e	-0.501	γ_e	0.038	β_e	0.196
		(0.148)		*(0.153)*		*(0.368)*
Knowledge	α_k	-3.989	γ_k	-0.160	β_k	3.181
		(18.294)		*(1.268)*		*(1.247)*
Acreage			γ_a	-0.180	β_a	3.780
				(0.139)		*(0.212)*
Region			γ_r	-0.780	β_r	3.649
(1=Cook Inlet)				*(0.356)*		*(0.387)*
After 1975			γ_{75}	-0.565	β_{75}	1.867
				(0.508)		*(0.413)*
Log winning bid			γ_w	0.125	β_w	1.405
				(0.103)		*(0.167)*
Log number bids			γ_b	-0.052	β_b	0.752
				(0.26)		*(0.318)*

Note: Parameters estimated in this table are based on 792 clusters operated by at least one large company (L, L&L, and L&F types). Numbers in parentheses are standard errors.

In addition to the types of joint operations, I also control for operator and land-specific characteristics in both estimations. For all three measures, σ_s, σ_v and ν, I control for operator attributes: experience, which sums up the number of clusters that the operator had previously owned in the same region, and knowledge, an indicator for the participants having owned a neighboring cluster prior to bidding. This is because an operator's familiarity with the geological formation in the area may influence her belief about the land value and her ability to interpret the information she gathers. For the parameters determining the prior distribution, σ_v and ν, I further control for cluster physical attributes, such as its size in acres and its region (either the North Slope/Beaufort Sea region or Cook Inlet region), as well as its auction characteristics. The auction characteristics I include are whether the auction happened after 1975, the average per acre price from winning bid weighted by the size of each lease in the cluster, and the weighted average of the number of bids submitted for each lease in the cluster.

Table 2.2 presents the results of estimation. The columns below σ_s show how being a joint operator, whether of the L&L or L&F type, having more previous experience in the region, and having greater knowledge of nearby land affect signal uncertainty, respectively. L&L operations have signal noise 0.413 higher than solo operations, whereas L&F operations have signal noise 0.189 lower than solo operations. The lower signal quality for L&L joint operators is possibly due to the fact that large company participants in a joint operation do not pool their expertise together. According to anecdotal evidence from an industry insider, meetings with fellow participants in a joint exploration partnership often focus more on financial planning and accounting, rather than to discuss technical details, such as what each well log means and where to perforate in each well. In the case where one company has a controlling interest in the partnership, drilling decisions are usually left entirely to the controlling company and other participants can only request information through the controlling company and may not have the right to interfere (Derman, 2017). To put it in the words of a proverb: two heads are NOT better than one. The issue here is this: just because there are two heads, it does not mean that more than one head is thinking, or that the two heads interact with each other at all. Estimation results based on the data seem to support this point. A robustness check using the 682 clusters operated solely by large operators has the same sign and similar magnitude for the joint operation indicator.[5] Estimation results using large-only clusters are shown in Table A.1 in Appendix A.3. On the other hand, signal quality is higher for L&F operators, which suggests that, by seeking the expertise of the local fringe companies, joint operators are able to improve signal accuracy, possibly through the local knowledge of the smaller company that large companies do not possess, such as the expertise to better interpret the geological readings of formations specific to Alaska. The parameters of the other variables are intuitive: when companies are more experienced at operating in the same region, or have knowledge of the neighboring land, the uncertainty of their signals decreases.

The columns under σ_v and v examine how different covariates affect the prior distribution. Joint operations are associated with lower means, though neither association is statistically significant. There are three potential explanations for the lower mean: (1) collusion makes bidding cheap, which gives them the financial resource to explore less promising land, and (2) Other companies willing to go into joint operations is a bad sign that the land is not worth as much. Unfortunately, current

[5] I fail to reject the null hypothesis that the parameters on L&L indicators from two samples are different even at 10% level.

model cannot untangle one explanation from another. Companies' experience in the area does not seem to significantly impact the prior distribution. However, having access to adjacent land significantly increases the company's prior mean. This could be due to neighbor firms having better information about the value of this new piece of land or are in a better position to explore (Hendricks and Porter, 1988). Land attributes, such as more acreage and being in the Cook Inlet region, increase the prior belief in the land value, while also decreasing the uncertainty of it. Having a larger piece of land may increase the probability of finding oil and potentially large oil deposits and hence increases the land value and decreases uncertainty. As for the Cook Inlet region, since companies traditionally pay less for the land here, conditional on being willing to pay the same bids as the North Slope region, the companies must be highly optimistic and certain of the productivity of the land. Finally, more competitive auctions, with higher winning bids and a larger number of bids submitted, are associated with more optimism in the land value, as more companies desiring this piece of land suggests that there could be more oil deposit there. However, competitiveness of the auction does not seem to significantly affect the uncertainty of the prior.

To see how well these parameters fit with the actual drilling outcomes, Figure 2.4 plots of the receiver operating characteristic (ROC) curve. The area under the curve is 0.823, which translates to the probability of the model marking a randomly-chosen "explored" cluster as more likely to be explored than not. An AUC value in the range of 0.8 to 0.9 is generally considered good. So this model does a reasonably good job explaining exploration outcomes in the sample.

Counterfactuals

We conduct two counterfactual analyses in this section to show how differences in signal quality across operator types leads to different drilling behaviors. In the first analysis, I look at how drilling decisions would have changed if no joint operation were allowed at all, as the Department of Interior and the Congress outlawed joint bidding in the 1970's. In the model, I force J_k in the data to be 0 for all k, and use parameter estimates from Table 2.2 to calculate the likelihood to drill. In the second analysis, I look at the counterfactual outcome if large companies are only allowed to partner with smaller local firms, an idea consistent with the credits given to smaller explorers by the Alaska government (Department of Revenue, 2012). Here, I force J_k and J_k^{lf} indicators to be 1 for all jointly operated clusters, and then use parameter

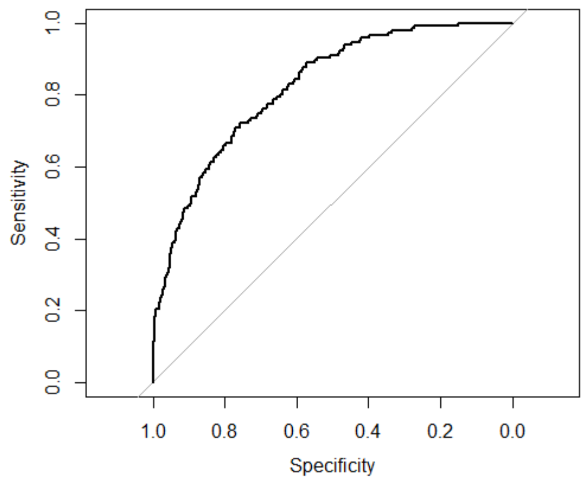


Figure 2.4: **Model ROC curve**
Note: The Area Under Curve (AUC) is 0.823.

estimates from Table 2.2 to predict the drilling outcomes.

The first analysis shows no loss in oil royalty payment and a potential loss of 22.7 million thousand cubic feet (mcf) of natural gas in royalty payment, which approximates to 30.8 million dollars of revenue to the state of Alaska, if no joint operation were permitted. Of the 52 clusters owned by L&L joint operators, solo operators are less likely to drill in 35 of them.[6] A list of these clusters with actual (joint) and counterfactual (solo) probability to drill is presented in Table A.2 in Appendix A.4. Out of these 35 clusters, 10 of them have a counterfactual probability at least 10 percentage points below the actual probability of drilling, as shown by the bolded rows in the table. Assuming that solo operations choose not to drill in these 10 clusters, I sum up the oil and gas production from each lease in these clusters, multiply the production by the royalty requirement from the lease, and arrive at 22.7 million mcf of natural gas paid to the state as royalty. Multiplying the annual production by yearly U.S. natural gas wellhead price (adjusted for inflation to year 2000 dollars), this production would have amounted to around 30.8 million dollars of revenue. To put this value in more perspective, the amount of revenue generated for the state from selling 183 tracts in the year 2000 was around 11.1 million dollars. Thus, 30.8 million dollars in revenue is not insignificant. However,

[6] If the difference in drilling probability is less than 1 percentage point, I consider the two probabilities as equal. There are also 90 clusters not explored by L&L operators, but which have higher probability under solo operation. However, their differences never exceed 5 percentage point. So I consider these clusters as still not explored under solo operation.

considering this 30.8 million comes from a span of over 40 years, the contribution from these clusters to annual state revenue over these 40 years is rather negligible.

The second analysis shows a larger potential loss of 140 million barrels of royalty oil and 363 million mcf of royalty natural gas, which equates to $4.5 billion in oil revenue and $649 million in gas revenue to the state of Alaska, if large companies had been forced to partner with smaller local operators, instead of their large peers. Similar to the analysis above, we look at the 52 clusters owned and explored by L&L operators. 39 of these clusters have lower counterfactual probability to drill for L&F operators, and 22 of them are more than 10 percentage points lower in predicted drilling probability.[7] After summing up the production multiplied by the royalty requirements, I arrive at 140 million barrels of oil and 363 million mcf of natural gas in royalty payment. Multiplying the annual production by yearly U.S. oil first purchase price and natural gas wellhead price (adjusted for inflation to year 2000 dollars) respectively, these quantities predict a royalty revenue of 4.5 billion from oil production, and 649 million from natural gas production. The details of production and revenue from each cluster can be found in Table A.3 in Appendix A.4. While this may seem a rather large sum of revenue lost, the state cannot realistically force large companies to partner with small firms when large firms seek a partner. Rather, it is more practical for the government to outlaw joint operations, forcing large companies either to explore solo or to not participate in the auction at all. Given these considerations, results from the first counterfactual analysis may be more realistic.

The rough estimations above demonstrate potential revenue losses for the state if joint operations between large companies were not permitted. However, banning joint operations could also come with benefits. Following the argument above, solo operation, if replacing joint operations, could have avoided unsuccessful efforts in 7 clusters, a total of 127,467 acres across them. Similarly, had L&F operators not drilled in such clusters, they could have avoided drilling in 18 unsuccessful ventures, a total of 390,190 acres across all. This would have reduced the deforestation and erosion from building roads, platforms and transportation of heavy equipment, decreased the amount of salty and potentially toxic produced water generated in the drilling process, and prevented chemical contamination of land and water from the drilling waste (O'Rourke and Connolly, 2003). Unfortunately, the exact benefit to the state of Alaska from avoiding these consequences is difficult to quantify. As a

[7]No clusters unexplored by L&L companies have higher probability to drill under L&F joint operation.

result, the upside of allowing joint operations amongst large companies may seem too attractive for the government to pass on, especially if revenue generation is a key consideration for the policymakers. On the other hand, for a government concerned with preserving the environment, 30.8 million dollars of potential revenue loss may be a meager amount in comparison to the long-term benefit from reducing the damage to the ecosystem. For these policymakers, banning joint operation may be the policy to implement.

2.6 Conclusion

This paper is an empirical analysis of oil and gas drilling decisions across three types of operators: solo large companies, partnerships between only large companies (L&L), and collaborations between large and fringe companies (L&F). It seeks to explain why L&L operators drill more often in the land they own than the other two types. In the model, operators decide whether to drill in a piece of land based on the priors held following the auction, and signals received after winning the right to explore the land.

The empirical results by matching the model prediction with the observed drilling outcomes show that L&L joint operators have the least accurate signals, while L&F operators have the most accurate signals, with solo large firms in the middle. This result suggests that large company participants in a joint operation may not receive independent signals, and even if they do, they may not exchange their information and expertise in the exploration process. On the other hand, when large companies collaborate with smaller firms, they may be seeking the local expertise of these firms and thus gain better signals from these partnerships.

Counterfactual analyses based on the empirical results show that, had joint operations between large companies been forbidden, the state of Alaska could have faced a potential loss of $5.1 billion in oil and natural gas royalty revenue, but have avoided exploration on 18 clusters and avoided potential damage to the environment in areas of 390 thousand acres. The implications of this study for policymakers depend on their priorities. If their ultimate goal is to maximize the state revenue, then allowing large companies to jointly operate seems to serve this purpose. However, if the goal is to conserve the environment and minimize damages from excessive drilling, then banning joint operations between large companies can be a policy to consider.

Chapter 3

THE CUSHIONING BENEFITS OF BIASED BELIEFS

3.1 Introduction

Biased belief formation is a bedrock underlying many models in behavioral finance and behavioral economics. In financial settings, many real-world investors, both individual and institutional investors, exhibit extrapolative expectations: they believe that prices will remain high in the future when recent prices have been high. Accordingly, these beliefs can be conducive to price bubbles, and sometimes tend to lengthen the duration of bubbles (Bagehot, 1873, Kindleberger, 1978, Barberis et al., 2017). Conversely, extrapolative expectations can amplify downward price movements and sometimes lead to slow recovery in an industry or economy (Greenwood, Sam Hanson, and L. Jin, 2016, L. J. Jin, 2015). By and large, biased beliefs have negative implications for market dynamics in these settings.

A large literature has focused on consumers having biased beliefs and how their beliefs affect trading in asset markets, In reality, however, biased beliefs arguably also play a role in many "real" investment decisions. Indeed, some recent research explores how biased expectations can impact such decisions. For instance, Gennaioli, Ma, and Shleifer (2015) document that corporate investment plans and actual investments are explained by CFOs' incorrect expectations; Greenwood and Samuel Hanson (2015) study how biased beliefs generate return predictability in the global ship building industry.

In this paper, we study a model of biased beliefs populated by producers who make real investment decisions based on these beliefs. Specifically, the biased beliefs take the form of backward-looking extrapolation: producers' expectations of future consumer demand is formed as a weighted average of these consumers' past demands.[1] These beliefs exhibit *insufficient mean reversion*: producers mistakenly assume that the long-run mean of the demand process is changing and estimate it using recent realizations of demand, thus underestimating the degree of mean reversion in the process. These biased beliefs affect producers' investment behavior, leading to sub-optimal decisions. Obviously, sub-optimal decisions impose welfare

[1] Earlier works of Barberis and Shleifer (2003), Barberis et al. (2015), and Hirshleifer, J. Li, and Yu (2015) propose extrapolative beliefs about stock market returns and GDP growth.

losses on the economy, and much of the existing behavioral literature focuses on these negative effects arising from agents' behavioral biases.

In this paper, however, we go one step further. When firms' investment decisions occur within a dynamic market equilibrium, these "mistakes" can actually translate into long-run gains in the market. Specifically, we show that over time, biased beliefs generate some unexpected effects that counteract the general trend of an industry or economy, "cushioning" the industry or economy against prolonged downturns and therefore leading to faster recovery. During industry upturns, these cushioning effects can shorten the duration of bubbles. Unlike many papers in the existing literature, our focus is on the *positive* implications that biased beliefs may play in market dynamics.

To illustrate these cushioning benefits, consider the oil exploration industry, an industry with pronounced boom-and-bust cycles and volatile prices in which extrapolation can have big effects. In this market, the producers are large oil companies who make important decisions with long-run impact on oil exploration and production. When oil prices spiral downward (as occurred recently in the world oil market), oil companies extrapolate low prices continuing into the future, and therefore cut back on new exploration. Indeed, extrapolative beliefs cause firms to cut exploration *more* than they would in the non-extrapolative benchmark, resulting in large welfare losses in the interim. However, over time, this excessive reduction in oil production will put *upward* pressure on prices, thus reversing the downward trend in prices and aiding the oil industry out of its doldrums. Moreover, this recovery will happen faster when firms have extrapolative beliefs. Conversely, in periods of rising prices, extrapolative producers overinvest in oil exploration, which puts downward pressure on the rising prices. We show that these cushioning effects constitute a generic feature of real investment models with producers having backward-looking extrapolative beliefs.

Our model builds on standard aggregate investment models of Abel (1981) and Abel and Eberly (1994). The price of industry output is positively related to consumer demand and negatively related to total investment from producers. Over time, consumer demand follows a mean-reverting process with a constant long-run mean. Without knowing this long-run mean, however, producers extrapolate past realizations of consumer demands in forming expectations of future demand. Based on these beliefs, producers make investment decisions. For comparison, we also examine a benchmark model in which producers know the long-run mean of consumer demand. To analyze the cushioning effects, we analyze the impulse response of investment, total supply, and product prices with respect to shocks to consumer demands. Our analysis leads to two observations. First, with extrapolative beliefs from producers, a negative demand shock gives rise to persistent underinvestment in subsequent periods, causing total supply to decrease at a faster rate relative to the benchmark model. This rundown in supply lends support to prices, thus "cushioning" the negative impact of the demand shock on the product price. Second, due to the persistence of the cushioning effects, the product price can sometimes even start rising in the midst of a sequence of negative demand shocks.

As an empirical case study, we consider the behavior and experience of oil producers operating in the North Slope of Alaska, one of the most active oil exploration sites in North America. Since oil exploration is not a liquid asset market with ample trading and resale opportunities, the return regressions or survey evidence used in the existing literature to detect extrapolation are not available in this context. Given these challenges, we present several pieces of evidence from Alaska exploration which are consistent with the presence of biased beliefs on the part of producers. First, we find that the number of new wells drilled are positively correlated with past levels of oil prices, with more significant correlation with prices from six to twelve months prior. Second, we find that the five-year production and five-year revenue from oil production for newly drilled wells are both *negatively* correlated with past levels of oil prices. This finding that drilling projects initiated following high prices yield systematically inferior outcomes support our interpretation of these projects as resulting from biased beliefs.

As a further test for firms' biased beliefs, we exploit the availability of data on both oil companies' planned and actual investments to establish that the number of "scrapped" wells — that is, the difference between the number of wells actually drilled and the number of wells planned — is negatively correlated with both past levels of oil prices and the average changes in prices during the subsequent two-year period within which oil companies are approved to carry out the planned drilling. This finding suggests that oil companies over-extrapolate past oil prices when planning for well drilling but subsequently change their mind and forgo these opportunities if oil prices decline during the time after the well was approved but before drilling has commenced.[2]

[2] Similarly, Conlin, O'Donoghue, and Vogelsang (2007) used data on purchased and subsequently returned clothing to identify projection bias.

In addition, we document some path-dependent features of belief formation. When recent oil prices have risen at an increasing rate, oil companies extrapolate to a lesser extent; if the oil price six months ago is 50 dollars per barrel, then all else being equal, an increasing rise of the oil price between six months ago and the time of drilling will result in a 12% reduction of the number of new wells drilled.

Motivated by this empirical evidence of oil companies' suboptimal exploration decisions in response to recent prices, we quantify the magnitude of the cushioning benefit by calibrating our model using parameters appropriate to this industry. In one example, a sequence of adverse demand shocks leads to a price decline which is 8.2% smaller in the extrapolative compared to the benchmark scenario, which quickens the recovery by four months. Another calibration example suggests that the industry downturn during the 2008 financial crisis would have been lengthened by four months if oil industry firms did *not* have biased extrapolative beliefs. These examples show how extrapolative beliefs cushion, or soften, the extremes during the periodic big downturns which punctuate the oil exploration industry. At the same time, the overall welfare calculus is ambiguous; while this cushioning effect shortens the downturn, the biased beliefs lead firms to severely underinvest in drilling activity during the downturn relative to the non-extrapolative benchmark, which involves large costs in lost jobs, underutilized equipment, and so on. Biased beliefs lead to *shorter, albeit direr,* downturns, and the overall welfare effect involves a tradeoff between the short-run costs and the long-run gains.

Our paper adds to both theoretical and empirical research that aims to understand the implications of biased beliefs on asset price movements, consumption and portfolio choices, investment decisions, and individual behavior. On the empirical side, recent papers by Vissing-Jorgensen (2004), Bacchetta, E. Mertens, and Wincoop (2009), Amromin and Sharpe (2013), Greenwood and Shleifer (2014), Koijen, Schmeling, and Vrugt (2015), and Kuchler and Zafar (2016) present survey evidence that real-world investors exhibit extrapolative expectations and they behave according to these beliefs. Cassella and Gulen (2015) show that the extent to which investors extrapolate past returns of the stock market is highly correlated with the degree of predictability of future market returns. And Gennaioli, Ma, and Shleifer (2015) find that CFO expectations about future earnings growth are extrapolative and predictive of planned and actual investments. On the theoretical side, Fuster, Hebert, and Laibson (2011), Choi and T. Mertens (2013), Hirshleifer, J. Li, and Yu (2015), Barberis et al. (2015), L. J. Jin and Sui (2017) show that extrapolative expectations can generate stock market movements that are consistent with the data. Alti and Tetlock (2014) show that over-extrapolation and overconfidence affect investment decisions. Barberis et al. (2017) use extrapolation to explain asset bubbles and trading volume. Gennaioli, Shleifer, and Vishny (2012) and L. J. Jin (2015) connect biased beliefs that arise from risk neglect and availability heuristic with market leverage and financial crashes. Glaeser and Nathanson (2015) tie extrapolation to housing dynamics. And Bordalo, Gennaioli, and Shleifer (2017) and Greenwood, Sam Hanson, and L. Jin (2016) use extrapolative expectations to make sense of facts about credit cycles.

Our paper makes three contributions to this line of research. First, the paper embeds extrapolative expectations into the supply side of a model with real investments and unearths the potential cushioning benefits from these biased beliefs. It is worth pointing out that most research so far has focused on the negative effects and costs associated with biased beliefs; only Dong, Hirshleifer, and Teoh (2017) discuss the potential benefits. Our study adds to the latter strand of literature and counsels greater caution when accessing the overall effect of biased beliefs. Second, we present empirical evidence from the oil exploration industry that supports our model assumptions and predictions. Thus our paper joins a small but growing literature exploring and quantifying the impact of biased beliefs in a specific industrial (i.e., non-financial) setting.[3] Our data allow us to look at the planned and actual investments separately, and studying the difference between the two supports the hypothesis that producers on the supply side exhibit biased beliefs. Finally, we use our data to calibrate model parameters; this allows us to further quantify the cushioning effects highlighted in the paper.

Our paper is related to the works of Greenwood and Samuel Hanson (2015) and Bordalo, Gennaioli, and Shleifer (2017); these works also connect biased beliefs with investment decisions. Different from their studies which analyze the asset price implications of biased beliefs, we focus on the cushioning benefits of these beliefs. Our paper is also related to the work of Glaeser, Gyourko, and Saiz (2008). Their work shows that the elasticity of housing supply can affect the magnitude and duration of housing bubbles, and biased beliefs in their framework come from the demand side. Instead, our study highlights the importance of biased beliefs from the supply side.

[3]The other paper is Greenwood and Samuel Hanson (2015). Kellogg (2014) estimates a structural model of individual oil companies' oil drilling decisions in Texas and also estimates producers' belief process for future oil prices.

The paper proceeds as follows. In Section 3.2, we lay out the model and characterize its solution. We then use impulse responses of the model to illustrate the cushioning benefits. Section 3.3 uses the Alaska exploration data to provide evidence that supports the extrapolative bias amongst oil and gas companies. In Section 3.4, we calibrate model parameters in accordance with the data and further analyze the implications of the model using these parameter values. Section 3.5 concludes. All technical details are in the Appendix.

3.2 The Model

In this section, we first develop a simple aggregate investment model with incorrect beliefs from producers. Then we examine the model implications through impulse response analysis. For congruence with the empirical case study which follows, we will describe the model using terminology from the oil industry. The firms are oil producers who make decisions about the number of wells to drill each period, and obtain revenue from selling the oil extracted from the wells.

Assume that the demand relationship between crude oil prices per barrel H_t and the total number of active wells Q_t is

$$H_t = A_t - BQ_t. \tag{3.1}$$

Here A_t represents a demand factor; this captures outside influences on prices which are exogenous to the firms. Since the oil market is global, these influences can include supply disturbances in other oil-producing areas of the world (such as Texas, Canada, the Middle East, etc.) which will also impact the price that Alaskan producers receive for their oil. Such disturbances evolve randomly and with some serial dependence, so we model the law of motion for A_t as

$$A_{t+1} = \overline{A} + \rho_0(A_t - \overline{A}) + \varepsilon_{t+1}, \tag{3.2}$$

with $\rho_0 \in [0, 1)$ and $\mathbb{V}\text{ar}[\varepsilon_{t+1}] = \sigma_\varepsilon^2$. Q_t, the aggregate number of wells, is an investment decision made by a continuum of risk-neutral firms. At each point in time, each firm chooses its level of investment i_t^G. The relation between the firm's well count q_t and its time-t investment is

$$q_{t+1} = (1 - \delta)q_t + p \cdot i_t^G = q_t + p \cdot i_t, \tag{3.3}$$

where $0 < \delta < 1$ is the depreciation rate, p is the probability of success when producing the industry output, and $i_t = i_t^G - \delta q_t/p$.[4] Effectively, i_t is the choice variable. At the aggregate level, the total number of wells evolves as

$$Q_{t+1} = (1-\delta)Q_t + p \cdot I_t^G = Q_t + p \cdot I_t. \tag{3.4}$$

This law of motion for Q_t implies that there is a one period time-lag in investment I_t before it affects the number of wells Q_{t+1}, and generates cash flow for the firms. In mapping this model to the oil exploration industry, we use a period of a month, which is reasonable given the lag between drilling and well production falls between a few days and a couple of months for most wells in our data.[5]

Similar to Greenwood and Samuel Hanson (2015), we assume that the representative firm earns a net profit of

$$\Pi_t = M \cdot (A_t - BQ_t) - C - \delta P_r \tag{3.5}$$

on each active well, where M is the average number of barrels obtained from each well, C is the operating cost of a well, and P_r is the replacement cost of a well. For an individual firm, given its current well count q_t and its current investment i_t (new drilling), the firm's time-t total profit is

$$V_t = q_t \Pi_t - P_r i_t - k \cdot \frac{i_t^2}{2}, \tag{3.6}$$

where $k \cdot \frac{i_t^2}{2}$ represents the adjustment costs.

Firms' biased beliefs and insufficient mean reversion. A crucial component of our model lies in the specification of firms' expectational errors. Specifically, we assume that, from firms' perspective, the evolution of the demand factor A_t is

$$A_{t+1} = \overline{A}_t^{\alpha} + \rho_f \cdot \left(A_t - \overline{A}_t^{\alpha}\right) + \varepsilon_{f,t+1}, \tag{3.7}$$

where

$$\overline{A}_t^{\alpha} = \alpha \cdot \overline{A}_t + (1-\alpha)\overline{A}, \quad \overline{A}_t = (1-\rho_A)\overline{A}_{t-1} + \rho_A A_t, \tag{3.8}$$

[4]Under risk neutrality and the assumption that success or failure of production is independent across firms, (3.3) is equivalent to leaving the incremental investment stochastic and then taking expectations when deriving the Bellman value function.

[5] In other industries, such as the housing market, it may take much longer (perhaps years rather than months) for new planned housing to be completed, and we conjecture that with such long delays, biased beliefs may actually exacerbate rather than cushion the economy against downturns.

the subscript "f" stands for "firm", and $1 > \rho_f \geq \rho_0 \geq 0$.[6]

Comparing Equation (3.2) to Equations (3.7) and (3.8) shows that such a model of beliefs exhibits *insufficient mean reversion*; $\overline{A}$ is the true long-run mean of the demand process, which is constant over time, but firms mistakenly assume it to be time-varying (denoted $\overline{A}_t^{\alpha}$ in Equation (3.8)), and estimate it using recent realizations of the process. Thus, in each period, firms' beliefs about the long-run mean of the process adjust in the direction of recent realizations: when demand has been slack—that is, when A_t has been low—firms tend to believe that the long-run mean $\overline{A}_t^{\alpha}$ has also fallen, leading to a smaller perceived degree of mean reversion measured by the difference between $\overline{A}_t^{\alpha}$ and A_t; the opposite occurs after periods when demand has been high.

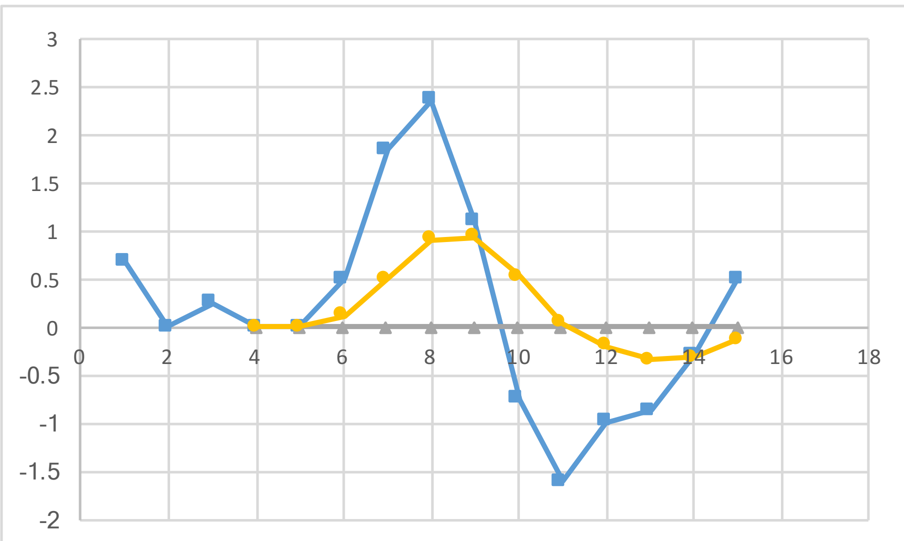


Figure 3.1: **Biased Belief Process and Insufficient Mean Reversion**

This figure contrasts the true process for the demand shocks (Equation (3.2)) from producers' beliefs (Equations (3.7) and (3.8)). The process of demand shocks A_t is plotted using squares. The true long-run mean, $\overline{A}$, is invariant over time and equal to zero, as plotted in triangles. However, producers' beliefs about the long-run mean, $\overline{A}_t^{\alpha}$, change over time, and are plotted in circles. The difference between A_t and $\overline{A}_t^{\alpha}$ measures the degree of mean reversion in A_t perceived by producers. The parameter values used in this example are: $\overline{A} = 0$, $\rho_0 = 0.68$, $\rho_f = 0.68$, and $\alpha = 0.9$

Figure 3.1 contains an illustration of these beliefs. The demand factor process, A_t,

[6] Note that when $\alpha = 0$, $\overline{A}_t^{\alpha}$ equals $\overline{A}$. In this case, the model reduces to the model of Greenwood and Samuel Hanson (2015). In comparison to Greenwood and Samuel Hanson (2015), our way of modeling extrapolative expectations allows us to further make sense of some path-dependent features of belief formation; we discuss this both later in this section and in Section 3.3.

is plotted in squares and there is first an upturn followed by a downturn. However, the true long-run mean, $\overline{A}$, is constant over time and equal to zero, as plotted in triangles. In contrast, firms' beliefs, following Equations (3.7) and (3.8), are characterized by a time-varying long-run mean, $\overline{A}_t^{\alpha}$, which is plotted in circles. Clearly, firms' beliefs about the long-run tendency of the process exhibit insufficient mean reversion: following an upturn in A_t, firms' perceived long-run means also track higher, and as a result, the perceived degree of mean reversion in A_t, measured by $A_t - \overline{A}_t^{\alpha}$, becomes smaller, suggesting "irrational exuberance", while the opposite occurs following the downturn in A_t, suggesting "irrational pessimism".

In addition, conditional on their estimated long-run mean $\overline{A}_t^{\alpha}$, we can also allow firms to perceive less mean reversion of A_t relative to the true process by having $\rho_f \geq \rho_0$. Rearranging (3.7)

$$A_{t+1} - A_t = (1 - \rho_f)\left(\overline{A}_t^{\alpha} - A_t\right) + \varepsilon_{f,t+1} \tag{3.9}$$

illustrates the path-dependent feature of firms' belief formation. With a sequence of steady increase in the demand factor, $\overline{A}_t^{\alpha}$ rises above the true long-run mean of $\overline{A}$. In this case, a high $\overline{A}_t^{\alpha}$ and a high ρ_f both make the perceived evolution of A_{t+1} less mean-reverting than the true data generating process. If, on the other hand, A_t rises at an increasing rate, then $\overline{A}_t^{\alpha}$ increases to a smaller degree compared to A_t, making firms perceive that A_t will mean-revert back to a low level more quickly in the future.

Equations (3.7) and (3.8) are also related to the work of Barsky and De Long (1993). That paper shows how investors learning about the time-varying mean of a dividend process can lead to excess stock market movements; thus investors learning and updating about the time-varying mean of a dividend process can appear to "extrapolate" recent innovations in the dividend process. In contrast to the Barsky-DeLong framework, however, the mean of the demand factor in our framework is *not* time-varying (Equation (3.2)), but firms mistakenly *perceive* it to be (Equation (3.7)). In comparison to Barsky and De Long (1993), then, agents in our model end up "learning too much" from past demand shocks, leading to an excessive degree of extrapolation, and insufficient mean reversion, relative to the full-information benchmark, as pointed out above.

Dynamic investment decision. The model has three state variables at each point in time: A_t, $\overline{A}_t$, and Q_t. For an individual firm, its Bellman equation is

$$J(q_t; A_t, \overline{A}_t, Q_t) =$$
$$\max_{i_t} \left\{ V(q_t, i_t; A_t, \overline{A}_t, Q_t) + \frac{\mathbb{E}_f[J(q_t + p \cdot i_t; A_{t+1}, \overline{A}_{t+1}, Q_{t+1} | A_t, \overline{A}_t, Q_t)]}{1+r} \right\}. \quad (3.10)$$

Here "$\mathbb{E}_f$" means that the expectation is taken under firms' subjective beliefs. The first-order condition gives

$$P_r + k \cdot i_t^* = p \cdot \sum_j^{\infty} \frac{\mathbb{E}_f[\Pi_{t+j} | A_t, \overline{A}_t, Q_t]}{(1+r)^j} \equiv p \cdot P(A_t, \overline{A}_t, Q). \quad (3.11)$$

Here $P(A_t, \overline{A}_t, Q_t)$ is a hypothetical price of discounting future expected per-unit net profit at the required rate of return r under firms' subjective expectations.

We now characterize the optimal level of investment in the proposition below.

Proposition 1 *In the investment model described above, firms' optimal level of investment is*

$$i_t^* = x + y_1 \cdot A_t + y_2 \cdot \overline{A}_t + z \cdot Q_t, \quad (3.12)$$

where

$$z = \frac{BMp^2 + kr}{2kp} - \sqrt{\left(\frac{BMp^2 + kr}{2kp}\right)^2 + \frac{BM}{k}}, \quad (3.13)$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} =$$
$$\begin{pmatrix} (1+r)k + BMp^2 - k\rho_f - kzp & -\rho_A \rho_f k \\ -(1-\rho_f)\alpha k & (1+r)k + BMp^2 - (1-\rho_f)\alpha\rho_A k - k(1-\rho_A) - kzp \end{pmatrix}^{-1}$$
$$\times \begin{pmatrix} pM\rho_f \\ pM(1-\rho_f)\alpha \end{pmatrix}, \quad (3.14)$$

$$x = \frac{(ky_1 + \rho_A k y_2 + pM) \cdot (1-\rho_f)(1-\alpha)\overline{A} - rP_r - p(C + \delta P_r)}{k(r - zp) + BMp^2}. \quad (3.15)$$

Proof: See Appendix. □

Impulse Response Analysis: the Cushioning Benefits

We now examine the model implications through some impulse response analyses. Figure 3.2 plots the net investment (new drilling) I_t, total wells Q_t, and the oil price

H_t from $t = 1$ to $t = 15$ for both the benchmark model ($\alpha = 0$) and the model with biased beliefs ($\alpha = 0.9$ and $\rho_A = 0.25$); from the steady-state, a sequence of half standard deviation negative shocks on A_t are imposed at t = 2, 3, 4, 5, and 6.

The impulse responses presented in Figure 3.2 are computed using model parameters calibrated for the Alaskan oil exploration sector. (These parameter values are presented and discussed in Section 3.4 below.) These impulse responses highlight the cushioning benefits of biased beliefs. Compared to the benchmark case, biased beliefs lead to lower investments (new drilling) in the face of negative demand shocks, which lowers the number of active wells, and persists over many periods after the negative demand shocks are realized. This lower well count “cushions" the negative impact of the adverse demand shocks on the output price, resulting in a smaller price decline and a faster price recovery. To see the mechanism in more details, notice that after a sequence of negative shocks on A_t, over-extrapolation leads firms to lower their estimation of the long-run mean of the demand factor $\overline{A}_t^{\alpha}$, hence becoming pessimistic about future prices and therefore reducing their investment. Relative to the benchmark case, the number of wells and hence total oil production in subsequent periods drops to a larger degree and stays persistently low in the behavioral model. This comes from two reasons: first, firms’ pessimistic beliefs about the future output price are persistent; second, lower past investments cumulatively result in lower total drilling in subsequent periods. This supply effect of biased beliefs partially offsets the negative effect of adverse demand shocks on the output price. In this example, extrapolation reduces the decline of the output price by 8.2% and results in a faster price recovery at month 11 compared to month 15 for the benchmark case, a difference of 27%.

While this example has considered a sequence of negative demand shocks leading to a downturn, the cushioning effects also arise when the market is in an upturn. Indeed, an example with a sequence of positive demand shocks would produce exactly symmetric results: in the upturn, extrapolative firms would overinvest in new drilling projects, leading to excessive production accompanied by downward pressure prices. The cushioning effects here would be of the same magnitude, albeit of the opposite sign: extrapolation would reduce the magnitude of the price by 8.2% and prices would fall back to their pre-shock levels by month 11, four months earlier compared to the non-extrapolative benchmark case.

Some additional observations are worth making. First, the cushioning benefits come at some cost, as large cutbacks in investment (a 27% reduction from 194 to 142 by

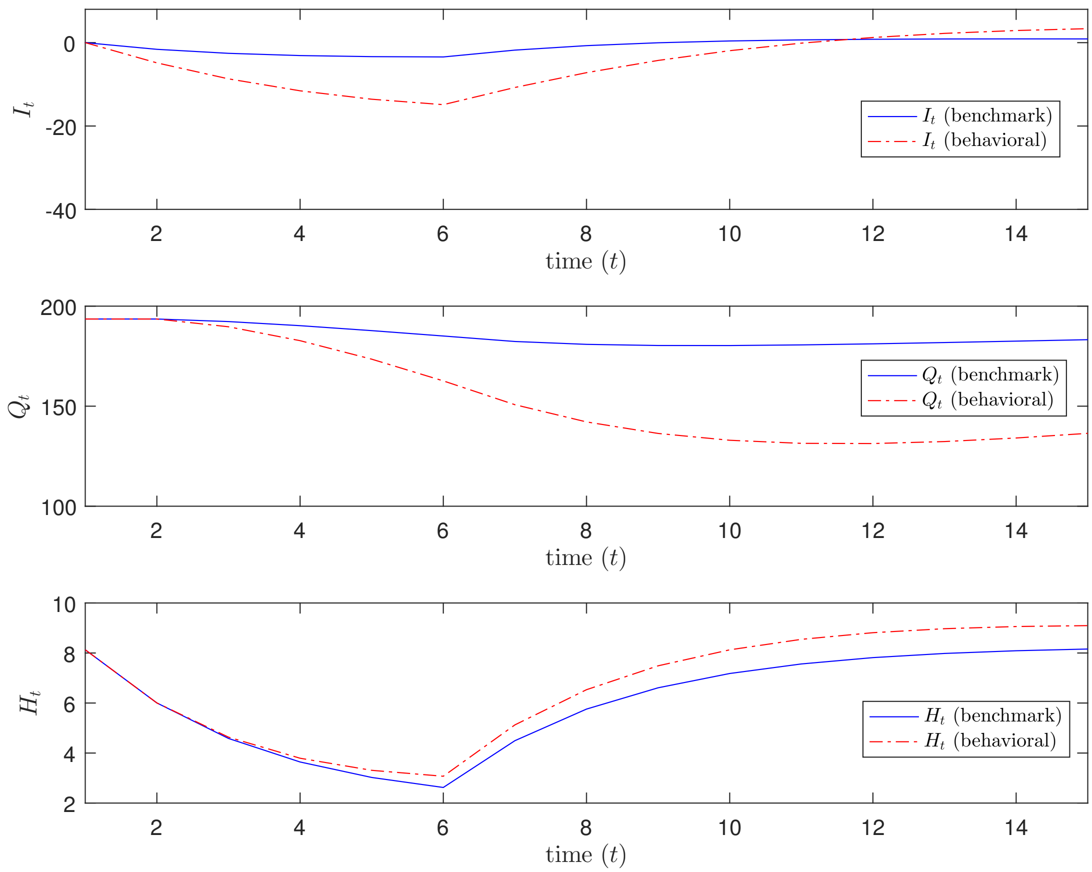


Figure 3.2: **Impulse Responses for the Benchmark Model and the Behavioral Model**

From the steady-state, a sequence of half standard deviation negative shocks on A_t are imposed at $t = 2, 3, 4, 5$, and 6. We plot the net investment I_t, total production Q_t, and the output price H_t from $t = 1$ to $t = 15$ for both the benchmark model ($\alpha = 0$) and the behavioral model ($\alpha = 0.9$ and $\rho_A = 0.25$). The other parameter values are: $B = 0.02$, $\overline{A} = 12$, $\delta = 0.6\text{‰}$, $r = 0.5\text{‰}$, $k = 22.8$, $C = 100$, $P_r = 463$, $\sigma_\varepsilon = 4.25$, $p = 0.8$, $\rho_0 = 0.68$, $\rho_f = 0.68$, and $M = 13$.

period 8 in the middle panel of Figure 3.2) can imply a high level of industrial turmoil; this "shakeout" is more sizable under extrapolative beliefs than in the non-extrapolative benchmark. As such, the overall welfare calculus of the cushioning effects is ambiguous, involving an intertemporal tradeoff between investment and output in the short run vs. faster recovery and higher output in the long run.

Second, the smaller decline of the output price in the behavioral model comes together with a faster recovery. As negative demand shocks continue to arrive at

$t = 4$ and 5, they are basically offset by the lower drilling activity in the behavioral model. As a result, the output price stays relatively flat during these periods. Finally, combining a sequence of small negative demand shocks into a big shock tends to limit the cushioning effect. Figure 3.3 shows that if we clump all the half standard deviation negative demand shocks from Figure 3.2 into a large negative shock at $t = 2$, the decline of the output price is of the same magnitude in both the benchmark model and the behavioral model, although over-extrapolation still leads to a faster recovery.

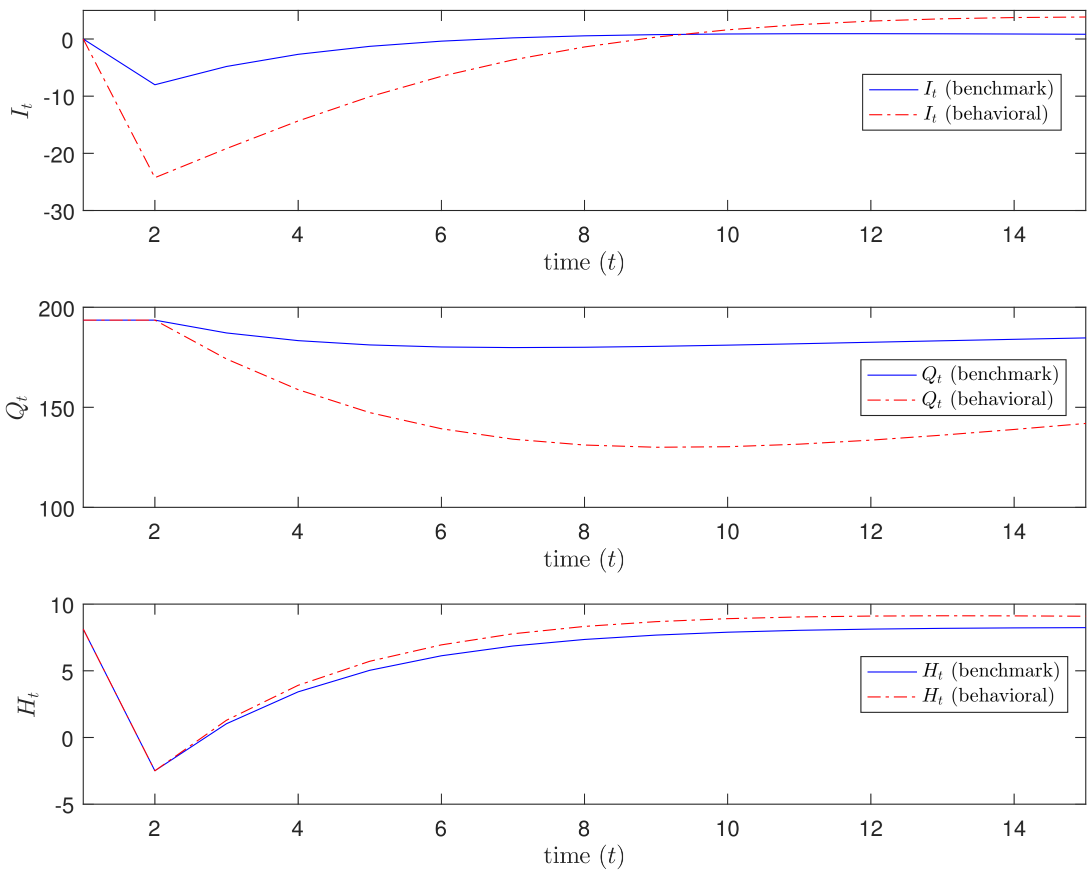


Figure 3.3: **Impulse Responses for the Benchmark Model and the Behavioral Model**

From the steady-state, a two and a half standard deviation negative shock on A_t is imposed at $t = 2$. We plot the net investment I_t, total production Q_t, and the output price H_t from $t = 1$ to $t = 15$ for both the benchmark model ($\alpha = 0$) and the behavioral model ($\alpha = 0.9$ and $\rho_A = 0.25$). The other parameter values are: $B = 0.02$, $\overline{A} = 12$, $\delta = 0.6\%$, $r = 0.5\%$, $k = 22.8$, $C = 100$, $P_r = 463$, $\sigma_\varepsilon = 4.25$, $p = 0.8$, $\rho_0 = 0.68$, $\rho_f = 0.68$, and $M = 13$.

Empirical Implications of Extrapolative Producers

In what follows, we study oil exploration in Alaska as an empirical application of the model. Since oil exploration is not a liquid asset market with ample trading and resale opportunities, the return regressions or survey evidence used in the existing literature to detect extrapolation are not available in this context. For that reason, before moving on to the empirical section of the paper, we derive some theoretical results to guide our empirical strategy of detecting extrapolative beliefs. The following corollary shows how Proposition 1 pins down the coefficients of regressing current and future investments on the current output price.

Corollary 1 *The regression coefficient for regressing I_t on H_t, both conditional and unconditional on A_{t-1}, $\overline{A}_{t-1}$, and Q_{t-1}, is $\beta_0 = y_1 + \rho_A y_2$. The regression coefficient for regressing I_{t+1} on H_t, both conditional and unconditional on A_{t-1}, $\overline{A}_{t-1}$, and Q_{t-1}, is $\beta_1 = (y_1 + \rho_A y_2)\rho_0 + y_2(1 - \rho_A)\rho_A + zp(y_1 + \rho_A y_2)$.*

Proof: See Appendix. □

We plot in Figure 3.4 the coefficients β_0 and β_1 as functions of α and ρ_A. Compared to the benchmark case of $\alpha = 0$ or $\rho_A = 0$, higher values of α and ρ_A make firms tend to overestimate the long-run mean of the demand factor after a sequence of positive demand shocks, hence reducing firms' perceived degree of mean reversion about future prices and causing them to overinvest; overinvestment after a sequence of high prices — high prices are caused by positive demand shocks — therefore leads to higher values of β_0 and β_1. Note that β_1 tends to be lower than β_0 as firms still anticipate some degree of mean reversion about future prices so they tend to scale back future investment. Also note that as ρ_A further increases, β_1 increases at a lower rate or decreases in some other cases: a higher ρ_A makes the estimated long-run mean of the demand factor $\overline{A}_t^{\alpha}$ less persistent, so adverse demand shocks in the future tend to affect investment decisions to a larger extent.

Finally, we consider some implications of our model for firms' investment after different price pattern scenarios. In Figure 3.5 we plot the impulse responses for our model under two different price pattern scenarios. From the steady-state, two different sequences of shocks on A_t are imposed from $t = 2$ to $t = 6$, resulting in one case a steady rise in oil price and in another case an increasing (or accelerating) rise in price. The top panel of the figure then suggests that firms overinvest less (by 8.9%) in the accelerating price scenario compared to the steady price increase

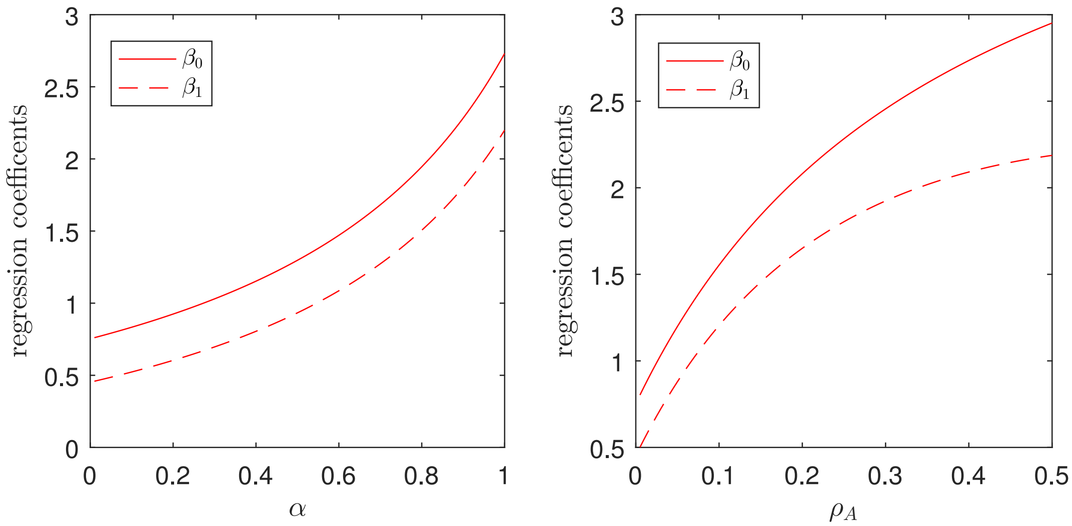


Figure 3.4: **Coefficients of Regressing Current and Future Investments on Current Price Level**

The figure plots the coefficients of regressing current and future investments on the current price level, β_0 and β_1, as functions of the belief-based parameters α and ρ_A. The default values for α and ρ_A are 0.9 and 0.25, respectively. The other parameter values are: $B = 0.02$, $\overline{A} = 12$, $\delta = 0.6\%$, $r = 0.5\%$, $k = 22.8$, $C = 100$, $P_r = 463$, $\sigma_\varepsilon = 4.25$, $p = 0.8$, $\rho_0 = 0.68$, $\rho_f = 0.68$, and $M = 13$.

scenario. Indeed, our model suggests that, in comparison with a steady increase in price, an accelerating price increase leads to smaller revisions in the firms' estimated long-run mean $\overline{A}_t^{\alpha}$, giving rise to stronger perceived mean reversion in future prices and less overinvestment.

It is worth noting that leading asset pricing models of extrapolation such as Barberis, Shleifer, and Vishny (1998) and Barberis et al. (2015) cannot explain this finding: in these models, an increasing rise will result in stronger extrapolative beliefs and therefore more investment. On the other hand, our finding is consistent with the empirical results of Barber, Odean, and Zhu (2009) and Greenwood, Shleifer, and You (2017) in the context of the stock market. These findings suggest some reasonable caution on the part of extrapolators in the face of accelerating price increase, and hence a belief formation process that is more sophisticated than simple extrapolation studied in the literature. To the best of our knowledge, the belief dynamic we proposed in equations (3.7) and (3.8) is among the first that can, at least in part, capture this path-dependent feature of belief formation.[7]

[7] Additional computations also showed that the biased belief specification for ship producers in

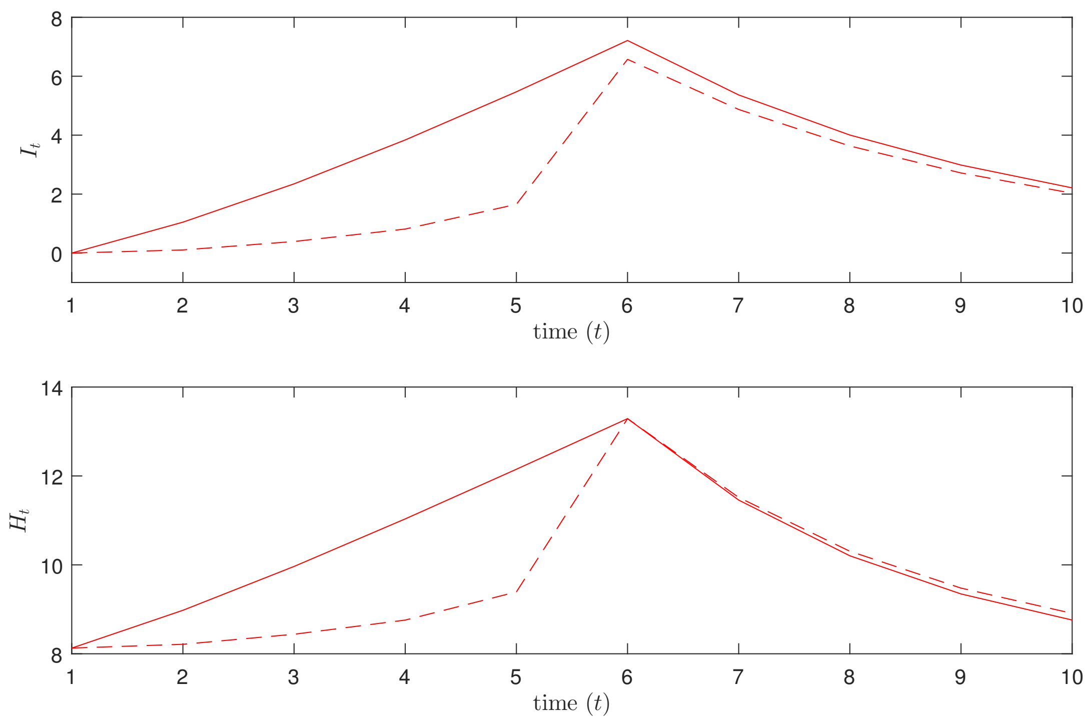


Figure 3.5: **Impulse Responses for the Behavioral Model with Different Price Patterns**

From the steady-state, two different sequences of shocks on A_t are imposed from $t = 2$ to $t = 6$, resulting in a steady rise in price (solid) and an increasing rise in price (dashed), respectively. We then plot, for these two cases, the net investment I_t and the output price H_t from $t = 1$ to $t = 15$ for both the benchmark model ($\alpha = 0$) and the behavioral model ($\alpha = 0.95$ and $\rho_A = 0.05$). The other parameter values are: $B = 0.02$, $\overline{A} = 12$, $\delta = 0.6\%$, $r = 0.5\%$, $k = 22.8$, $C = 100$, $P_r = 463$, $\sigma_\varepsilon = 4.25$, $p = 0.8$, $\rho_0 = 0.68$, $\rho_f = 0.68$, and $M = 13$.

3.3 Oil Exploration: An Empirical Application

Having established the existence of cushioning benefits from extrapolative beliefs in a theoretical model, we now proceed to quantify them in a real world setting. In choosing a suitable industry for a case study, we seek an industry with pronounced boom and bust periods, where cushioning can play a more important role. For that reason, we focus on oil exploration and production in Alaska. There are several reasons which make Alaska a suitable stage for our analysis: the size of its

Greenwood and Samuel Hanson (2015) cannot generate this pattern of decreasing overinvestment with an accelerating price pattern.

market, the purity of the North Slope oil price index, and the ability for Alaskan operators to respond to price changes. Alaska has long been one of the biggest oil producing states in the U.S, consistently ranked amongst the top five across oil producing regions in the country since the 1970s. Its production peaked at around two million barrels a day in 1988, behind only Texas. In addition to the sheer size of the production in Alaska, petroleum activities in Alaska are centered in the remote North Slope region (along the northern Arctic coast of Alaska). North Slope field production has accounted for over 95% of all field production in Alaska each year since the latter half of 1970s (EIA, 2017). As a result, the Alaska North Slope crude oil price is a pure indicator of the supply and demand for Alaska-produced oil, as almost no supply from regions outside of Alaska is in the mix. Furthermore, due to its Arctic location, the Alaska North Slope region has a topography and geology which differ from other main oil-producing regions in the continental United States, such as California and Texas, and which pose unique challenges for oil drilling and production. Successful drilling in this region requires a fair amount of local know-how and specialization. Due to such reasons and also different state regulations, international large operators, such as British Petroleum and ConocoPhillips, all have independent operations in Alaska that are separate from their other U.S. operations. BP, for instance, has an independently incorporated subsidiary in Alaska, BP Exploration (Alaska) Inc (BP, 2016). This independence allows these subsidiaries to make drilling decisions themselves, and it is thus reasonable to model their decisions in Alaska as independent from the decisions in other regions. These independent operations are also arguably more responsive to the boom and bust cycle in the Alaska oil industry. For that reason, the cushioning effects that we pointed out in the previous section, which soften out the boom and the busts, may be especially important in Alaska compared to other oil-producing regions.

Before proceeding to the calibration exercises which quantify the magnitude of the cushioning benefits for several historical episodes in Alaskan oil exploration, we present some suggestive evidence from regressions which indicate that oil producers are affected by extrapolative beliefs in their oil production decisions. In these regressions, we are guided by the empirical implications from Corollary 1 above.

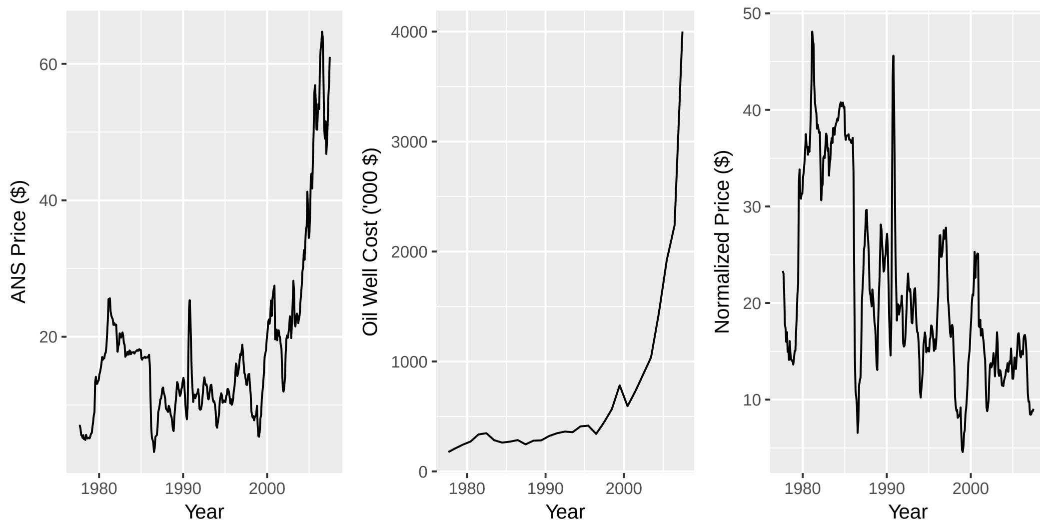


Figure 3.6: **Alaska North Slope Oil Price and U.S. Well Cost Trend**

The left panel of the figure plots monthly Alaska North Slope first purchase price; the middle panel plots the interpolated monthly U.S. well drilling cost (actual data is on an annual basis); the right panel plots the monthly price normalized by drilling cost (with the drilling cost of June 2000 as the basis). The sample period for all three plots is August, 1977 to June, 2007.

Data Description

The data used in our case study is drawn from multiple sources. For oil prices, we use the monthly Alaska North Slope (ANS) first purchase price per barrel from the U.S. Energy Information Administration (EIA) from 1977 to 2016. As depicted by the first panel of Figure 3.6, the ANS price remained relatively low prior to 2000, but increased drastically after 2000. In the meantime, cost to drill an oil well, using the nominal average cost per crude oil well drilled each year in the US provided by the EIA, followed a similar trend. The drilling cost remained relatively stable but skyrocketed after 2000, as shown in the middle panel of Figure 3.6.[8] To more accurately capture firms' revenue and cost considerations, we introduce a *normalized oil price*, defined by ANS oil price divided by the oil well cost for the corresponding month, and then multiplied by the drilling cost of June, 2000. Hence the normalized price in June 2000 is the same as the nominal ANS price, but the normalized oil price is higher in the 1980s and much lower in the 2000s, as shown in the right panel in Figure 3.6.

To examine how firm investments respond to past levels of oil prices, we focus on the

[8]The cost data is only available on an annual basis from 1960 to 2007. We use linear interpolation to obtain monthly drilling cost data.

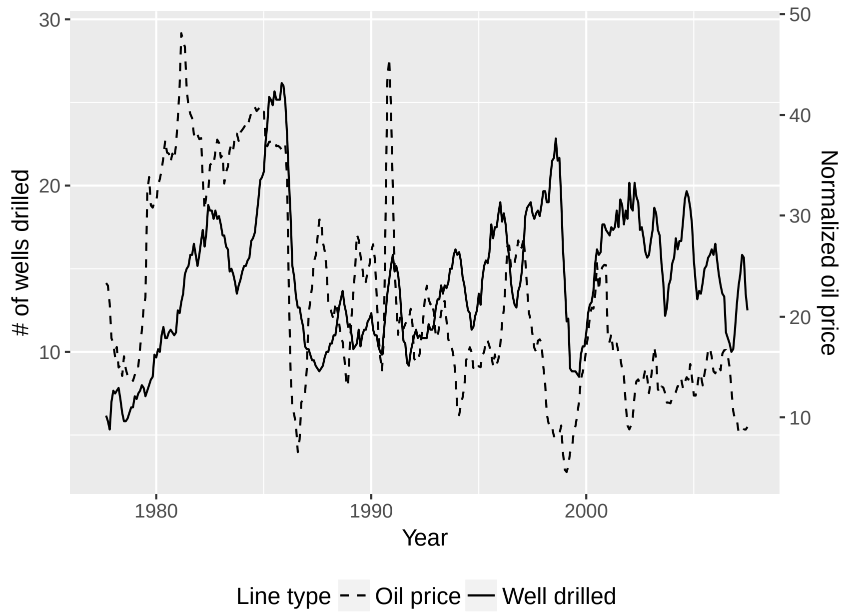


Figure 3.7: **Well Investment Trend**

The solid line shows the trend for the number of oil-related wells drilled by operators in Alaska in each month. The dashed line plots the time series of normalized oil price. The sample period is August, 1977 to June, 2007.

number of wells drilled for oil producing or servicing purposes, using historical well activity data from the Alaska Oil and Gas Conservation Commission (AOGCC). The drilled wells are those with records showing the dates of actual construction activities, such as well spudding or well completion, or those with positive well depth. A firm with extrapolative beliefs is more likely to drill a well when it observes high oil prices in the recent past because the firm perceives high prices and hence high revenue moving forward. In Figure 3.7, we plot the time series of the number of wells drilled in each month, together with the normalized oil price. It suggests that indeed a larger number of wells are drilled following high price periods.

While Figure 3.7 suggests that recent price levels seem to be associated with firms' well drilling decisions, these decisions may be fully rational. Rather than over-extrapolating, the firm may correctly foresee high oil prices in the future after observing high oil prices in the recent past, and therefore wells drilled at these high price level times create more profit due to the rising price. Additional evidence is needed before concluding that firms over-responded to the recent high prices, in line with the over-extrapolation hypothesis. In Figure 3.8, we plot, on a monthly basis, the time series of the average per-well profit and production in the first 60 months of production for wells that are drilled, along with the normalized oil prices.[9] Both plots do *not* seem to support that the firms are rationally foreseeing future oil prices. On the contrary, both plots in Figure 3.8 suggest that *lower-revenue and less productive* wells are drilled following periods of high oil prices, more consistent with the interpretation of these wells as mistakes arising from over-extrapolative beliefs.

To further examine firms' extrapolation bias, we utilize the availability of well approval data to study the percentage of scrapped wells in each month. Within the well history data, we observe all wells that received permits to drill. However, not all permitted wells ended up being drilled. Scrapped wells are those that received permits but the constructions of which never took place within 24 months of the permit approval.[10] As stipulated by the Alaska State Legislature (Title 20 Chapter 25), if a well is not drilled within this period, a new permit needs to be applied for. If a firm has extrapolative beliefs, then it might become overly exuberant and apply for more well permits after observing high oil prices, but subsequently reverse its investment decision when price drops after the permits are issued. To capture this reversal of investment decisions, we look at the proportion of wells scrapped each month, calculated as the number of wells ending up scrapped over the total number of wells approved in each month. Figure 3.9 plots the time series of wells scrapped percentage, as well as the normalized oil price. It suggests that the percentage of wells scrapped increases following periods of recent low prices.

Regression Results

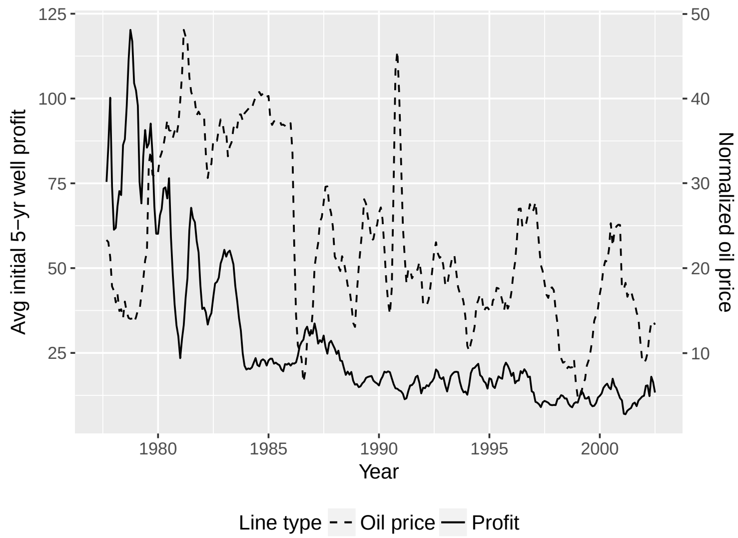
Based on what we observe in Figures 3.7, 3.8, and 3.9, we now run a series of regressions to formally test whether firms extrapolate recent oil prices when making investment decisions, whether this extrapolative belief is unbiased in terms of well profit and production, and whether firms take actions to correct their investment decisions after receiving drilling permits.

First we look at the relationship between the number of wells drilled each month

[9] Profit is calculated using normalized price multiplied by the production. As the normalized price is a measure of how profitable oil production is, we multiply it by production volume for total profit.

[10] However, if a different well was drilled at the exact same longitude and latitude within two years of the initial approval, the initial well does not count as a scrapped well, since a better well may be planned to replace the current one.

(a) Average initial five-year profit per well ($MM)



(b) Average initial five-year production per well (MM bbl)

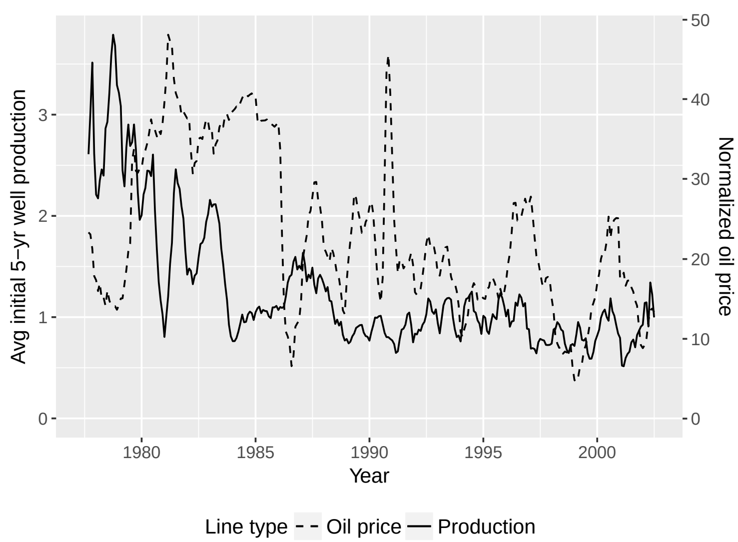


Figure 3.8: **Well Production and Profit Trend**

Figure 3.8a plots on a monthly basis the average profit (in million dollars) in the first 60 months of production (hence the initial 5 years) for wells that are drilled. Figure 3.8b plots on a monthly basis the average production (in million barrels) in the first 60 months of production for wells that are drilled. The dashed lines in both figures plot the time series of normalized oil price. The sample period for both plots is August, 1977 to June, 2002.

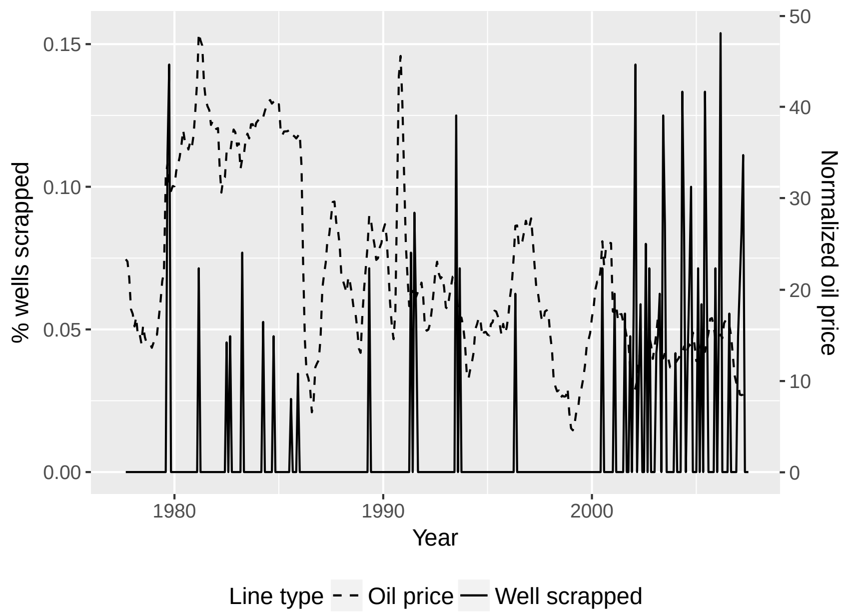


Figure 3.9: **Well Scrappage Trend**

The solid line plots the percentage of wells scrapped in each month among the wells that are given permits to drill. The dashed line plots the time series of normalized oil price. The sample period is August, 1977 to June, 2007.

and normalized oil prices lagged by various time periods: one, three, six, nine, and twelve months,

$$Y_t = \alpha + \beta P_{t-\tau} + \gamma S_t + \varepsilon_t, \quad \text{where } \tau = 1, 3, 6, 9 \text{ or } 12 \text{ months.} \tag{3.16}$$

Here, Y_t is the number of wells drilled in month t; $P_{t-\tau}$ is the normalized oil price from τ periods before; and ε_t is the error term for each period. In this equation, we also include price pattern variable S_t, an indicator variable for whether prices have been changing at an increasing or decreasing rate in the previous months leading up to month t. We categorize price paths into five patterns: no clear pattern in the rate of change as the reference level; decreasing rate of decline; increasing rate of decline; decreasing rate of rise; and increasing rate of rise. These categories are calculated as follows. Let n be the maximum number of consecutive months prior to t that prices have gone up or down. For n that is greater than or equal to three, we calculate the rate of change as the difference of price changes between the second and first half of these periods.[11] If the rate of change has a different sign from

[11] If n is even, the difference is $(P_t - P_{t-n/2}) - (P_{t-n/2} - P_{t-n})$. If n is odd, the difference is $[P_t - (P_{t-(n-1)/2} + P_{t-(n+1)/2})/2] - [(P_{t-(n-1)/2} + P_{t-(n+1)/2})/2 - P_{t-n}]$, which can be simplified to $(P_t - P_{t-(n-1)/2}) - (P_{t-(n+1)/2} - P_{t-n})$.

the price trend, then this period experiences decreasing rate of change; otherwise it experiences increasing rate of change. For instance, if the rate of change is less than zero but the price has been going up, then this period is going through a decreasing rise in price.

With extrapolative beliefs, firms perceive high oil prices moving forward after observing high recent prices, resulting in high investments in the current period. In other words, extrapolation implies a positive relation between the number of wells drilled and past oil prices; coefficient β in (3.16) is expected to be positive. As for the price patterns, some empirical results from previous papers for financial markets suggest that γ should be negative when S_t corresponds to an increasing rate of rise; that is, accelerations in the rate of price increase tend to reduce extrapolative effect of past prices on current behavior (see Barber, Odean, and Zhu (2009) and Greenwood, Shleifer, and You (2017)). It is an empirical question, however, whether such a result will obtain outside financial markets. Also note that, as we described earlier, traditional extrapolation models cannot capture this fact, whereas the belief dynamic proposed in equations (3.7) and (3.8) tends to generate it.

Table 3.1 summarizes the regression results based on equation (3.16). In this table, we see a significant positive relation between the monthly number of wells drilled and the lagged normalized oil prices from six months or more prior to drilling; the relation becomes insignificant when the lag of past oil price is less than six months. This reduced significance for the more recent months' coefficients is perhaps due to two reasons. First, it takes time to move equipment and set up rigs before drilling can actually start. Second, it takes time for firms to attend to past oil prices; this interpretation with limited attention is consistent with models of extrapolation in the behavioral finance literature (see, for instance, Barberis et al. (2017), Barberis and Shleifer (2003), and Harrison Hong and Stein (1999)). Overall, Table 3.1 suggests that oil companies in Alaska increase their investment in well drilling when the oil price levels are high in recent months.

In addition, we also find that the extrapolation is tempered by the pattern of increasing rise in price, as the coefficient on the "increasing rise" price patten is negative and statistically significant. That is, when price has been going up at an increasing rate, all else equal firms become less extrapolative and therefore reduce investments. This confirms the predictions of the extrapolation models illustrated in Figure 3.5. For instance, if the normalized oil price is \$50 six months ago with no observable price pattern, then a company with extrapolative beliefs would drill 18.1 wells in the

	Dependent variable:				
	Number of wells drilled				
	(1)	(2)	(3)	(4)	(5)
Normalized oil price					
1-month lag	0.070				
	(0.054)				
3-month lag		0.083			
		(0.055)			
6-month lag			0.109**		
			(0.054)		
9-month lag				0.121**	
				(0.052)	
12-month lag					0.127**
					(0.051)
Rate of change					
decreasing decline	−0.867	−1.241	−1.602	−1.384	−1.661
	(1.022)	(1.034)	(1.108)	(1.082)	(1.065)
increasing decline	−1.094	−1.240	−0.434	−0.428	−0.927
	(1.108)	(1.110)	(1.002)	(0.965)	(0.938)
decreasing rise	−1.914*	−1.741	−1.675	−1.903*	−2.174**
	(1.102)	(1.149)	(1.107)	(1.056)	(1.044)
increasing rise	−2.630***	−2.398**	−2.174**	−2.205***	−2.477***
	(0.952)	(0.943)	(0.903)	(0.854)	(0.844)
Constant	13.528***	13.302***	12.690***	12.477***	12.573***
	(1.290)	(1.311)	(1.303)	(1.210)	(1.121)
Observations	358	356	353	350	347
R-squared	0.063	0.073	0.093	0.108	0.123
Adjusted *R*-squared	0.050	0.059	0.079	0.095	0.110

Table 3.1: **Total Wells Permitted on Normalized Oil Prices and Price Trends**

The table is based on a sample of 5,121 oil-related wells drilled in Alaska between August 1, 1977 and July 1, 2007. The lagged prices are the normalized oil prices 1, 3, 6, 9, and 12 months prior to the month of drilling. The rate of change categories are defined as periods with continuous increase or decrease for at least 3 periods and the rate of change follows a convex or concave shape. Numbers in the parentheses are Newey-West standard errors allowing for 6-month maximum lag in autocorrelation. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

current month, 1.1 more than if the price was \$40 six months ago. Furthermore, if the price has been rising at an increasing rate in the months leading up to the current month, then the number of wells drilled would become 16, which is 12% lower than the original level of 18.1. In other words, the impact of observing increasingly rising oil prices leading up to the month of drilling is similar to a decrease of \$20 in

	Dependent variable:				
	Initial five-year profit				
	(1)	(2)	(3)	(4)	(5)
Normalized oil price					
1-month lag	−0.106**				
	(0.050)				
3-month lag		−0.108**			
		(0.053)			
6-month lag			−0.258***		
			(0.083)		
9-month lag				−0.095	
				(0.110)	
12-month lag					−0.425**
					(0.192)
Rate of change					
decreasing decline	1.806	2.483**	3.432***	1.857*	2.240**
	(1.164)	(1.019)	(1.044)	(1.057)	(0.916)
increasing decline	1.350	2.910***	2.399***	1.965**	3.004***
	(1.009)	(1.033)	(0.916)	(0.863)	(0.993)
decreasing rise	3.151	2.900	2.494	3.116	3.209
	(2.775)	(2.885)	(2.757)	(2.869)	(2.801)
increasing rise	1.995***	1.764***	1.527***	2.075***	2.142***
	(0.572)	(0.683)	(0.560)	(0.723)	(0.682)
Constant	74.669***	60.102***	104.908***	116.458***	130.062***
	(19.740)	(14.882)	(28.849)	(27.434)	(32.291)
Unit FE	Y	Y	Y	Y	Y
Drilling Year FE	Y	Y	Y	Y	Y
Observations	4,055	4,044	4,027	4,005	3,991
R-squared	0.368	0.368	0.361	0.364	0.357
Adjusted *R*-squared	0.345	0.345	0.338	0.342	0.335

Table 3.2: **Five-year Well Profit on Normalized Prices**

This table is based on a sample of 4,055 oil-related wells with tract number drilled in Alaska between August 1, 1977 and July 1, 2002. Well profit is calculated by summing up the monthly production times monthly normalized price for each well in the first 60 months of production. The lagged prices are the normalized oil prices 1, 3, 6, 9, and 12 months prior to the month of drilling. The rate of change categories are defined as periods with continuous increase or decrease for at least 3 periods and the rate of change follows a convex or concave shape. Numbers in the parentheses are cluster-robust standard errors clustered by units. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

oil prices from six months prior.

The previous regression provides some evidence on firms extrapolating high levels of past prices that they observe. However, firms' extrapolative beliefs can be fully

	Dependent variable:				
	Initial five-year production				
	(1)	(2)	(3)	(4)	(5)
Normalized oil price					
1-month lag	−0.004				
	(0.002)				
3-month lag		−0.004**			
		(0.002)			
6-month lag			−0.009*		
			(0.005)		
9-month lag				−0.002	
				(0.007)	
12-month lag					−0.010
					(0.008)
Rate of change					
decreasing decline	0.096	0.119**	0.153**	0.100*	0.109**
	(0.065)	(0.060)	(0.069)	(0.057)	(0.055)
increasing decline	0.079	0.129**	0.104*	0.092*	0.116**
	(0.060)	(0.058)	(0.059)	(0.051)	(0.055)
decreasing rise	0.125	0.113	0.099	0.125	0.126
	(0.138)	(0.142)	(0.140)	(0.142)	(0.139)
increasing rise	0.070**	0.058	0.052	0.074*	0.075**
	(0.034)	(0.039)	(0.036)	(0.041)	(0.037)
Constant	2.363***	1.900***	3.071***	3.329***	3.651***
	(0.665)	(0.491)	(0.809)	(0.737)	(0.988)
Unit FE	Y	Y	Y	Y	Y
Drilling Year FE	Y	Y	Y	Y	Y
Observations	4,055	4,044	4,027	4,005	3,991
R-squared	0.258	0.257	0.251	0.250	0.244
Adjusted R-squared	0.231	0.229	0.224	0.224	0.218

Table 3.3: **Five-year Well Production on Normalized Prices**

This table is based on a sample of 4,055 oil-related wells with tract number drilled in Alaska between August 1, 1977 and July 1, 2002. Well production is calculated by summing up the monthly production for each well in the first 60 months of production. The lagged prices are the normalized oil prices 1, 3, 6, 9, and 12 months prior to the month of drilling. The rate of change categories are defined as periods with continuous increase or decrease for at least 3 periods and the rate of change follows a convex or concave shape. Numbers in the parentheses are cluster-robust standard errors clustered by units. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

rational; empirically, this regression is not able to disentangle over-extrapolative drilling behavior from "rational exuberance" when recent oil prices have been high. To address this issue, we run regressions of initial five-year profit and production of wells on lagged normalized prices one, three, six, nine, and twelve months prior to the drilling date, controlling for price trend patterns. The initial five-year production is calculated as the amount of oil produced in the first 60 months of the well production. The initial five-year profit is the monthly production multiplied by monthly normalized oil price in the first 60 months of well production. Tables 3.2 and 3.3 summarize the results for these two regressions, which confirms the graphical evidence presented in Figure 3.8 which we discussed earlier. Table 3.2 shows that well profits are actually *lower* for wells drilled following high price levels, suggesting that firms over-extrapolate and fail to foresee the price reversal when they observe periods of high prices. Similarly, Table 3.3 shows that wells drilled following high price levels are not more productive, hence ruling out the possibility that firms save the most productive wells for high price periods. Taken together, these results support the interpretation that firms drill *excessively* after a period of high prices, leading to lower profit and production, and symptomatic of biased beliefs.

Just as how firms can over-extrapolate when they drill a well, firms can also over-extrapolate when they apply for the permit to drill. When they observe oil price drops after the approval, they can retract their initial plan and decide not to drill. Hence, observing an increased likelihood of scrapped wells when prices drop after initial approval can be evidence of firms being overly exuberant when they make drilling plans in the first place. To test this hypothesis, we regress the indicator of whether a well is scrapped on the change in oil prices 24 months after the permit issuance, controlling for the current normalized price in the approval month. Table 3.4 summarizes the results. The changes in oil prices in both columns are calculated as the percentage difference between average oil prices in the next 24 months and the price in the approval month. In Column 2, we also control for additional well heterogeneity, such as the region of the wells, the operator and the unit that the wells belong to.

Qualitatively, the results in both columns agree, and we will focus on the results in Column 2. Starting from the top, the negative and significant coefficient on the normalized oil price (–0.001) shows that when the oil price levels are high at the time of approval, firms are less likely to scrap these wells. This observation confirms the earlier results from Table 3.1 that when oil prices are high, firms are more likely to actually drill the wells, and are less likely to scrap these wells.

The negative and significant coefficient (–0.009) on the 2-year post-approval price change shows that an increase in the price after a well has been approved lowers the

	Dependent variable:	
	Whether a well is scrapped	
	(1)	(2)
Normalized oil price	-0.0004^{*}	-0.001^{***}
	(0.0002)	(0.0001)
2-yr avg perc. change in price	-0.009^{***}	-0.009^{***}
	(0.002)	(0.001)
Region = Other		0.0004
		(0.020)
Region = Cook Inlet		0.988^{***}
		(0.019)
Constant	0.018^{***}	0.022^{***}
	(0.003)	(0.004)
Operator FE	N	Y
Unit FE	N	Y
Observations	4,757	4,757
R-squared	0.002	0.272
Adjusted R-squared	0.001	0.248

Table 3.4: **Likelihood to Scrap on Normalized Oil Price and Price Change after Approval**

This table is based on a sample of 4,757 oil-related wells with tract number that are approved between August 1, 1977 and July 1, 2007. The two-year average percent change in price is calculated as the 2-year average normalized oil price after approval minus the normalized oil price at the time of approval, and then divided by the price at the time of approval. For region, the reference level is the North Slope and Beaufort Sea region. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Parameter	Value	Justification
$\overline{A}$	12 + Year fixed effect	Coefficients from regressing monthly price over
B	0.02	productive well count
ρ_0	0.68	Coefficient from regressing residual prices on its lag
σ_ε	4.25	Standard deviation of residual from the price residual regression
δ	0.6‰	Inverse of median length of life of development wells
p	0.8	Proportion of wells producing $\geq$ 1000 barrels of oil
P_r	463	Coefficients from regressing well cost over # new wells
k	22.8	
C	100	Umbrella term for all operating costs
M	13	Average of the median well production for each month

Table 3.5: **Calibrated Parameter Values**

This is a list of parameters in the model that can be directly or indirectly inferred from the Alaska data. When point estimate of the parameters is available, we use the point estimate. When the estimated values fall within a range, we round to a reasonable number within the range. The operating costs here include lease operating expense, gathering, processing and transport expense, water disposal costs, and any general and administrative (G&A) costs.

probability of scrapping a well. This result is consistent with extrapolative producers, as extrapolative firms, which may be eager to initiate new drilling projects when prices have been high (as demonstrated in earlier results) end up scrapping these projects if oil prices decline after obtaining drilling permits (but before drilling has begun).

3.4 Model Calibration

Having presented empirical evidence consistent with the hypothesis of extrapolation in the oil exploration sector, we now turn to the calibration of the theoretical model presented in Section 3.2 using data from Alaska to set model parameters.

To more realistically measure the effect of firms' extrapolation bias, we calibrate the key parameters in the model using the available Alaska data. Table 3.5 summarizes the list of parameters, their values and our justification for choosing each value.

Demand-related parameters are associated with equations (3.1) and (3.2). A_t in equation (3.1) can be viewed as the long-term mean $\overline{A}$ plus a stochastic term specific to period t. Hence, to obtain $\overline{A}$ and B, we regress the monthly per-barrel North Slope price, H_t, on the number of productive wells in Alaska in each month, Q_t.[12] We

[12]Being "productive" here is defined as a monthly production of at least 2000 barrels, around the

	Dependent variable:	
	Alaska North Slope crude oil first purchase price	
	(1) OLS	(2) IV
Number of productive wells	−0.021***	−0.017**
	(0.007)	(0.007)
Constant	12.200***	11.400***
	(3.270)	(3.620)
Year FE	Y	Y
Observations	465	464
R-squared	0.963	0.963
Adjusted *R*-squared	0.959	0.959

Table 3.6: **Demand Regression**

These regressions can be expressed as $H_t = \overline{A} - B \cdot Q_t + \epsilon_t$, where H_t is the per-barrel Alaska North Slope First Purchase price, and Q_t is the number of wells producing at least 2000 barrels in that month. The first column uses OLS regression. The second column uses IV regression where the instrument is the number of productive wells in the previous month.

will refer to this regression as the demand regression. The resulting intercept from the demand regression is then $\overline{A}$ and the coefficient associated with well count is then $-B$. The result of the demand regression can be found in Table 3.6.[13] Next, in equation (3.2), notice that $A_{t+1} - \overline{A}$ is simply the residual from the demand regression for period $t + 1$. Thus, to obtain ρ_0, we regress the residual on its one-month lag, and to obtain σ_ε, we calculate the standard deviation of the residuals from this lag regression. The results of these regressions can be found in Table 3.7.[14]

median of what the bottom 10% of wells produce monthly in our sample.

[13]For robustness, we ran simple linear regression as well as IV regressions using lagged well counts as the instruments. All regressions give us comparable range of parameter values.

[14] The estimate of $\rho_0 = 0.68$ indicate a substantial degree of mean reversion in the process of A_t, the demand factor. Note that this does not rule out the possibility that oil *prices* may not be mean reverting.

	Dependent variable:	
	Residual from OLS	Residual from IV
	(1)	(2)
Lagged residual (from OLS or IV)	0.673***	0.684***
	(0.034)	(0.034)
Constant	0.004	0.004
	(0.198)	(0.196)
Observations	464	463
R-squared	0.452	0.467
Adjusted R-squared	0.451	0.466
σ_ε	4.25	4.22

Table 3.7: **Residual Demand Regression**

The residuals are calculated as $H_t - \hat{A}_t - \hat{B}Q_t$ where $\hat{A}_t + \hat{B}Q_t$ comes from predictions from regressions in Table 3.6. Column 1 uses residuals from the OLS regression and column 2 uses residuals from the IV regression. In both columns, we regress the current period residual on the previous period's residual. σ_ε is the standard deviation of the residuals of each lagged residual regression.

To determine the representative firm's investment level, as shown in equation (3.3), we need to estimate the depreciation rate δ and the probability of successful drilling p. For the depreciation rate, we look at all the development wells that were drilled within our time frame, and see for how long they produce. The median length of production life for our wells is around 180 months, and hence the depreciation rate is on average 1/180. The success rate amongst wells for production purposes, exploratory and development wells, varies by the definition of success. If the definition of success is producing over 1000 barrels, then the success rate is around 80%.[15]

To calculate a representative firm's profit, we need well production in a month for an representative well, M, and operating cost C, as well as drilling-related cost parameters P_r and k. We look at the median well for each month since July, 1978, and find that the average of all the months come to around 13,000 barrels. The operating costs include lease operating expenses, gathering, processing and

[15] If we make a more stringent requirement that the development well production needs to be at least 1 million barrels, then $p = 0.55$. For a threshold of 0.5 million, $p = 0.65$; for a threshold of 100,000 barrels, $p = 0.75$.

	Dependent variable:
	Oil well drilling cost
Number of newly-drilled wells	11.399*
	(6.912)
Constant	462.581***
	(104.387)
Observations	360
R-squared	0.008
Adjusted R-squared	0.005

Table 3.8: **Demand Regression Calibration**

We regress the per-well drilling cost associated with oil wells on the number of newly-drilled wells in the current month. The intercept is the P_r and the coefficient associated with the number of wells is $k/2$. $p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

transporting costs, as well as water disposal and General and Administrative costs. These costs vary widely across well locations, performances or the amount of production (EIA, 2016). We use C as an umbrella term for all of these costs, and set C to be around 100,000 a month as an estimate for the monthly level of all costs mentioned above. Finally, drilling cost-related parameters can be extracted from the well cost data. In equation (3.6), the cost per well is modeled as $P_r + k \cdot i_t/2$, which increases as the number of wells drilled increases. Knowing the total cost of drilling per well and the number of newly-drilled well in each month, i_t, we regress the monthly well drilling cost on the number of wells drilled each month, and the intercept from the regression is P_r and the coefficient for the number of wells is $k/2$. The results of the cost regression can be found in Table 3.8.

The impulse responses emerging from the model simulations utilizing these parameters are in Figures 3.2 and 3.3, and we already discussed them earlier.

Two Historical Episodes: 2008-2011 and 1986-1987

Taking the calibration exercise one step further, we next present two additional model-fitting exercise to quantify the cushioning benefits for specific historical episodes of price downturns for Alaskan crude oil. Since oil produced in Alaska is mostly supplied to other parts of the U.S. and can be easily substituted with oil produced from other regions in the U.S. and areas outside of the U.S., Alaska oil prices are sensitive to domestic and international events that affect oil demand and supply. We focus on episodes following two such events, the U.S. financial crisis in 2008 (the Great Recession) and Saudi Arabia's dramatic increase in production in 1986. These are illustrated in Figures 3.10 and 3.11.

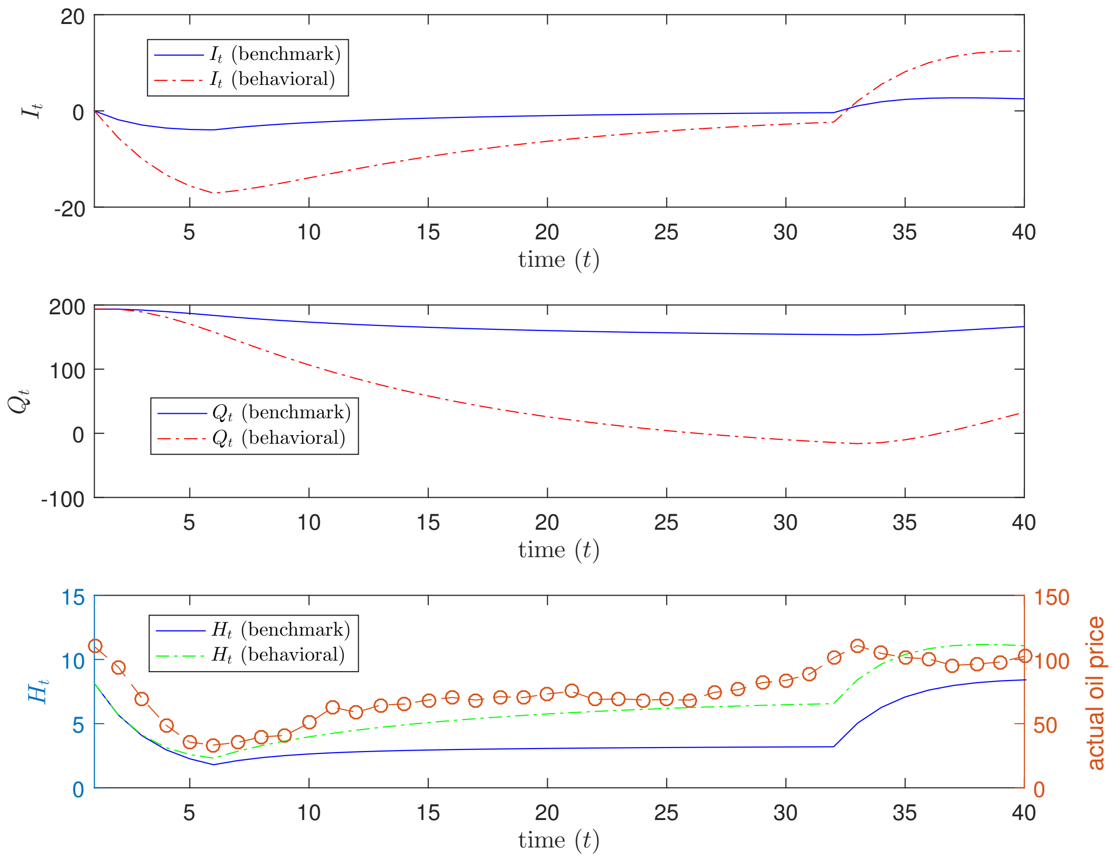


Figure 3.10: **Calibrated Impulse Responses for Historical Episode 1: the 2008 Great Recession**

We plot the net investment I_t, total production Q_t, and the output price H_t from $t = 1$ to $t = 15$ for both the benchmark model ($\alpha = 0$) and the behavioral model ($\alpha = 0.9$ and $\rho_A = 0.25$). In the bottom panel, we also plot (on the right) the actual oil price from August 2008 to November 2011. From the steady-state, we choose a sequence of demand shocks on A_t so that the price pattern implied by the behavioral model roughly matches the actual oil price movements. The other parameter values are: $B = 0.02$, $\overline{A} = 12$, $\delta = 0.6\%$, $r = 0.5\%$, $k = 22.8$, $C = 100$, $P_r = 463$, $\sigma_\varepsilon = 4.25$, $p = 0.8$, $\rho_0 = 0.68$, $\rho_f = 0.68$, and $M = 13$.

In Figure 3.10, we consider a forty-month period from August 2008 to November 2011, coinciding with the most recent financial crisis in the US. As illustrated in the bottom panel of this figure (in the bubble-dashed line), oil prices fell sharply during the first six months (08/2008-01/2009) by over 60% from peak to trough, and then recovered very slowly and gradually, regaining the initial price level at month 33 (04/2011). Based on this, we calibrated a shock process for the extrapolative model (dashed green line) to match the relative magnitudes and shape of the actual price process. The recovery for this calibrated extrapolative price process takes 33 months, as in the actual price process. Using this calibrated shock process, we also simulated prices in the benchmark non-extrapolative industry (graphed in solid blue). For this benchmark economy, we see that the recovery takes longer; only at month 37 (08/2011) do prices reach the initial level. Thus, in this way, we find that the cushioning benefits shortened the recovery process by roughly 4 months, or 11%.

The two top panels in the figure show striking differences in the extrapolative and benchmark industry during the price drop and recovery process, which illustrates the complex welfare effects of biased beliefs. With extrapolation, investment falls sharply as prices drop, leading to large differences in the well-count. Since the oil exploration sector is composed of many small firms, such a large drop in drilling activity will entail a sizable “shakeout” as firms become inactive and are forced to leave the market. We can get a sense of the size of this “shakeout" by looking at a period with similar drastic decline of oil prices, though to a smaller extent and without the woes of the global financial crisis. Following the oil price collapsing by 40% in the second half of 2014, 128 oil and gas companies filed for bankruptcy between 2015 and 2016, up from around 30 between 2013 and 2014 (Haynes and Boone, 2016, Egan, 2016). At 30 months, indeed, the number of active wells in the extrapolative industry is only a fraction of the well-count in the benchmark industry. Clearly, the accompanying decrease in output allows prices to recover more quickly. Extrapolative beliefs lead to a direr but shorter duration, and results in ambiguous welfare effects.

Figure 3.11 illustrates a similar exercise for an earlier and less pronounced price drop episode, in the mid-1980’s. Between 1981 and 1985, Saudi Arabia reduced its oil production by three quarters in order to combat the price collapse caused by the world consumption decline. However, beginning in 1986, Saudi Arabia decided to abandon its effort and ramp up its production, causing oil prices to fall further in 1986 (Hamilton, 2011). We again compare the lengths of recovery period following this event under the behavioral and benchmark models. By design, the

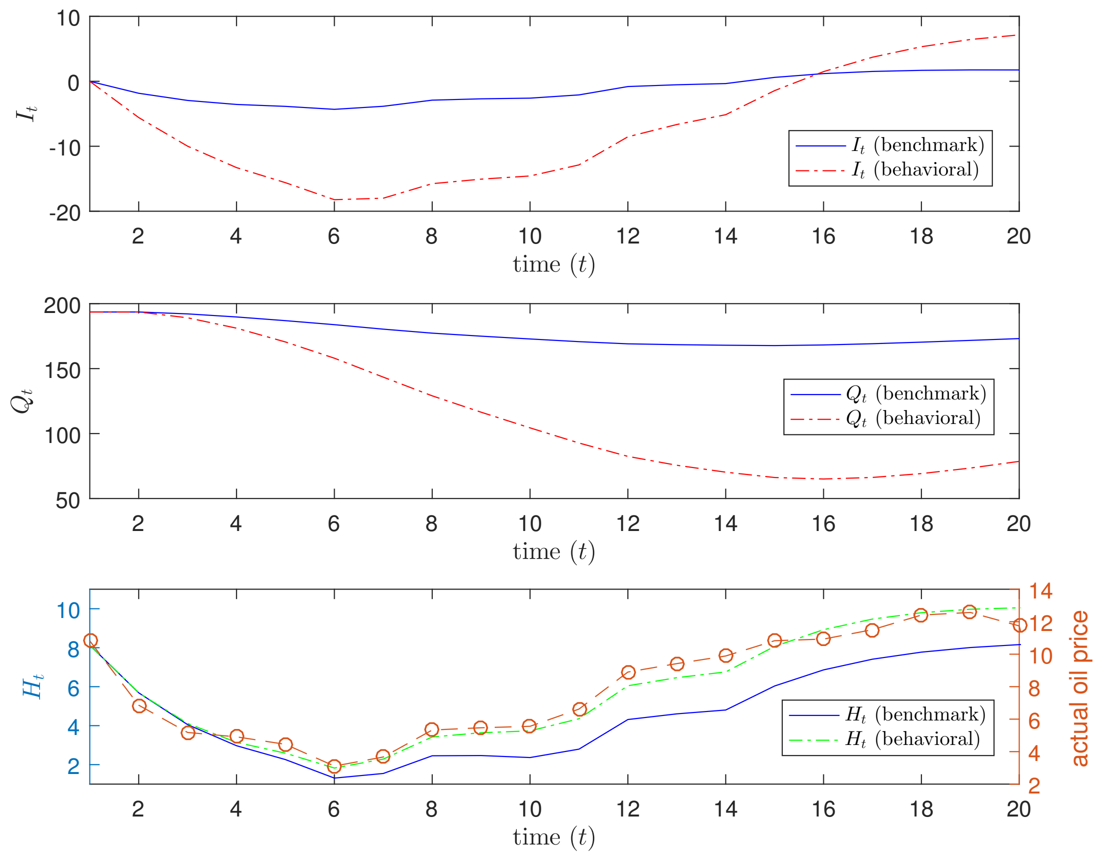


Figure 3.11: **Calibrated Impulse Responses for Historical Episode 2: the 1986 Saudi Oil Glut**

We plot the net investment I_t, total production Q_t, and the output price H_t from $t = 1$ to $t = 15$ for both the benchmark model ($\alpha = 0$) and the behavioral model ($\alpha = 0.9$ and $\rho_A = 0.25$). In the bottom panel, we also plot (on the right) the actual oil price from February 1986 to September 1987. From the steady-state, we choose a sequence of demand shocks on A_t so that the price pattern implied by the behavioral model roughly matches the actual oil price movements. The other parameter values are: $B = 0.02$, $\overline{A} = 12$, $\delta = 0.6\%$, $r = 0.5\%$, $k = 22.8$, $C = 100$, $P_r = 463$, $\sigma_\varepsilon = 4.25$, $p = 0.8$, $\rho_0 = 0.68$, $\rho_f = 0.68$, and $M = 13$.

recovery occurred in the fifteenth month (April 1987) in the actual data, and also for the behavioral model. For the benchmark non-extrapolative model, however, the recovery did not occur until September 1987; in this example, then, the cushioning effects shorten the recovery process by five months.

Naturally, these are stylized examples, but they illustrate the real benefits that "cushioning" can have in a real-world setting, in an industry notorious for its boom and busts sequences. Indeed, a lesson from these examples is that the cushioning benefits of extrapolation can soften the extremes of the cycles.

3.5 Conclusion

Much of the existing literature in behavioral economics and finance has focused on the negative and undesirable effects of behavioral biases and biased beliefs. In contrast, we point out in this paper that in certain settings, such as industries prone to periodic boom and bust cycles, biased beliefs can have benefits in terms of softening the up and downs of the economic cycle. In these industries, biased beliefs cause firms making investment decisions to respond more quickly to recent information in market prices. Thus, for instance, a price downturn will trigger a more immediate decrease in investment; in turn, this leads to lower supply which "cushions" and prevents prices from falling too quickly and leads to a quicker recovery. Modeling and quantifying these positive implications of biased beliefs on industry dynamics are important contributions of this paper.

We develop a theoretical framework, based on a standard aggregate investment model to illustrate these cushioning benefits. We then apply this model to the oil exploration industry in Alaska, a highly volatile industry characterized by sharp price fluctuations. One striking calibration example shows that the industry downturn during the 2008 financial crisis would have been lengthened by four months if oil industry firms did *not* have biased extrapolative beliefs. This suggests that the cushioning benefits can be sizable in a real-world setting. In ongoing work, we are exploring other sectors in which extrapolative beliefs may be important and hence in which the cushioning benefits can play an important role.

Chapter 4

OIL AND GAS TRACT AUCTION IN ALASKA: PRIVATE VERSUS COMMON VALUE

4.1 Introduction

Traditionally, state and federal governments in the U.S. lease out their land through competitive auctions for oil and gas explorations. Oil and gas companies submit their bids in a sealed-bid first-price auction, and the highest bidder wins the lease if his/her bid exceeds the reserve price. The bid submitted by each company reflects how much they value this piece of land based on a variety a factors: (1) common value factors shared by all companies, such as how much oil deposit is in this land, and (2) private value factors such as company's experience, strategies and financial constraints. But which of these two types of components play a more substantial role? Can the variations in bidding prices be explained mainly by one of them? By looking at competitive lease auctions in Alaska between 1993 and 2003, this paper seeks to identify which component plays a bigger role in determining bid prices in Alaska.

Previous studies have come to different conclusions regarding which component matters more. Li, Perrigne, and Vuong, modeling bidders' private information as the product of common and private value components and using a two-step nonparametric estimation procedure, found that private components explain bids variation better than common components (T. Li, Perrigne, and Vuong, 2000). On the other hand, Hendricks, Pinkse, and Porter analyzed bid markups and rents using *ex post* production amount and found that these values are more consistent with a common value model (Hendricks, Pinkse, and Porter, 2003). Outside of the oil and gas industry, studies also look into the role of common versus private model in procurement auctions. Hong and Shum found that the average procurement cost in some auctions strictly increases with the number of bidders, evidence for a larger role played by the common value. Their results suggest that more competition leads to a worse outcome for the auctioneer (Han Hong and Shum, 2002). Another study focusing on Michigan highway procurement auctions by Paulo Somaini allows for asymmetric bidders with nonindependent private information and common values. Using a semi-parametric approach and distance to bidders' plants as a cost shifter, he also finds evidence against a private value model for 6 of the 10 large firms, which, however, can be rationalized by an affiliated private value model with common cost component (Somaini, 2011). In this paper, I adopt Li, Perrigne, and Vuong's and Hong and Shum's model of decomposing private information into two components and use the parametric approach used by Hong and Shum to determine which component plays a key role.

A better understanding of the auction model has two implications. The first is that the correct model allows us to better evaluate the rationality of corporate bidders. If the winning bidder has a private signal that prompts them to bid higher, then using the conditional expected common valuation could overstate the valuations of other bidders. This would thus make them look more rational than they truly are. The second implication is to see whether policies implemented by the government make economic sense. For example, Alaska shifted from Area-specific sale to Areawide sale in 1998 (specific differences will be discussed in Section 2), partially to encourage bidding from smaller companies and individuals (Alaska DOR, 2012). However, if the appropriate model is pure common value and the firms foresee the winner's curse, then increased competition could potentially lead to lower revenue for the State.

In this study, I use the competitive lease auction data in Alaska to look at how an increase in the number of potential bidders impacts bidding prices submitted by companies. Given a participant's private signal, a pure common value model predicts that he would bid lower as the number of fellow bidders increases, whereas an independent private value (IPV) model would show no difference. To determine which model is more suitable for the Alaska oil and gas lease auctions, I first conduct a correlational analysis to look for preliminary evidence supporting one model over another. Then I use a parametric structural model developed by Hong and Shum to draw further inference on which component, common or private, plays a more important role in the bidding strategy. The paper is organized as follows. In Section 2, I briefly describe the background for Alaska oil and gas tract auctions. In Section 3, I present the equilibrium bidding strategy for first-price sealed-bid auction with a reserve price. This will serve as the theoretical prediction of how bidding strategies shift with more competition. Section 4 introduces the data and different variables, and summarizes the results of correlational analyses. Section 5 discusses the specifics of the structural model, its results, and counterfactual analyses. Section 6 concludes this paper.

4.2 Alaska Oil and Gas Tract Auction

To help readers better understand the study, I describe how the Alaska oil and gas tract auctions work. The State of Alaska oil and gas tract auctions involve the sale of leases on state-owned land. Starting on December 10, 1959, the Alaska Department of Natural Resources (DNR) has conducted lease sales through competitive bidding. A parcel of land is identified for leasing, and is divided into mostly rectangular pieces of land known as tracts, and companies can bid on one or more such tracts. The way the DNR identifies a parcel of land for lease changed in 1998. Prior to 1998, the DNR used a format called Area-Specific sale, which involves requesting industry input on preferred locations for the next sale. The DNR would then conduct the title review to determine the ownership of the land and establish legal descriptions prior to the bidding. However, this process took a long time and sometimes no bids were received. In 1998, the DNR transitioned into Areawide Sale, which announces tracts for sale in advance each year in the five known oil producing regions. Title and legal reviews are only conducted after the sale is final. This change in identifying land for sale needs to be considered when looking at factors affecting companies' valuations.

During the auction, there are different types of offers that the government seeks. A large majority of tracts fall under this category called "fixed royalty, cash offering," meaning that the government specifies the royalty percentage paid to the government and the percentage stays fixed for the duration of the lease, while companies compete with each other based on the cash bids submitted during the auction. Many auctions also have reserve prices set by the DNR which govern the minimum bids that the government finds acceptable. The reserve prices are mostly \$5 and \$10, but a few tracts have a reserve price of \$100. The company with the highest bid that exceeds the reserve price wins the lease. The State of Alaska generally allows for a variety of bidders in the auction, including large international firms, smaller or local companies, and individuals. These bidders can either bid solo or jointly with each other.

The terms of a tract relevant to this study are the location, size and fixed royalty amount. The location of the tract is important as it gives us information on the tract's region and proximity to previously explored land. For example, North Slope tracts were highly valuable in the 1980s after the discovery of Prudhoe Bay field. Depending on the region, the size of the tract can vary widely, ranging from below 1,000 acres to 5,760 acres. The fixed royalty requirement is 12.5% for a large majority of the tracts, with the rest being 16.67% and a few at 20%.

4.3 Equilibrium Bidding with Reserve Price

With the auction format in mind, I describe the equilibrium bidding strategy using notations of this paper. For a given tract t, denote the number of potential bidders as $N = n$, and each bidder is indexed by $i = 1, \ldots, n$. Assuming that each bidder is risk neutral, symmetric, and affiliated, we can derive the equilibrium bidding function for an arbitrary bidder i. Let the valuation of each bidder be denoted by U_i, which can be described as a function of the private signal, denoted by X_i, and the unknown common value, denoted by V, i.e., $U_i = u(V, X_i)$. The private signal X_i is assumed to be one-dimensional and distributed on the interval $[\underline{x}, \bar{x}]$. In an independent private value model, we have $U_i = X_i$, whereas in a pure common value model, $U_i = V$. Let Y_i denote the highest bid submitted by the rivals of bidder i, and let the individual with the highest bid be indexed by j and his bidding strategy be b_j. So in an first-price sealed-bid auction, after observing its signal X_i, company i chooses a bid b_i to maximize its expected payoff U_i. Hence,

$$b_i = \arg\max_b \mathbb{E}_{U_i, Y_i}[(U_i - b)\mathbf{1}(b_j(Y_i) < b)|X_i = x, N = n]. \tag{4.1}$$

Here I make two assumptions also made by the previous literature. First I assume the oil and gas companies are symmetric, in the sense that the joint distribution of U_i's and X_i's is exchangeable when you interchange index i and j. I also assume that the random variables $(U_1, \ldots, U_n, X_1, X_2, \ldots, X_n)$ are affiliated, where large values for some variables makes other variables more likely to have large values (Milgrom and Weber, 1982). These assumptions give us a unique pure-strategy Bayesian-Nash Equilibrium where each company bids according to a monotonically increasing strategy identical across all firms, i.e., $b_i(\cdot) = b^*(\cdot)$ for $i = 1, \ldots, n$. Equation 4.1 thus becomes

$$b_i = \arg\max_b \mathbb{E}_{U_i, Y_i}[(U_i - b)\mathbf{1}(b^*(Y_i) < b)|X_i = x, N = n]. \tag{4.2}$$

Solving for the first-order condition of (4.2) gives us

$$b^*(x) = v_n(x, x) - \frac{F_{Y_i|X_i,N}(x|x, n)}{f_{Y_i|X_i,N}(x|x, n)} b^{*\prime}(x), \tag{4.3}$$

where $f_{Y_i|X_i,N}(\cdot)$ denotes the density of Y_i conditional on bidder i's signal X_i and the number of potential bidders N.

Equation 4.3 gives a good summary of the conflicting effects that increased competition has on the equilibrium bid. The first term $v_n(x,x)$ is the expected utility for bidder i conditional on his signal being x and that no one else observes a higher signal. For a fixed x, as n increases, the fact that his bid is still the highest becomes “bad news" for bidder i, since it increasingly suggests that his observed signal is too high. As a result, bidder i would lower his bid. This effect is called the “winner’s curse effect" (Han Hong and Shum, 2002). On the other hand, as n goes to infinity, the probability of all signals being less than x_i goes to 0. The second term then decreases to 0. Hence, if the oil and gas auction follows a common value model, then the effect of the first term dominates, and we see a decrease in bid prices when n increases. However, if the private value component dominates, then the effect is more ambiguous when n is finite.

In the case that a binding reserve price r exists, then given the monotonicity of the bidding strategy, bidder i must observe an x_i above a certain threshold, denoted by $x^*(r)$, to be willing to submit a bid of at least r, i.e., $b^*(x^*(r)) = r$. $x^*(r)$ can be formally expressed as

$$x^*(r) \equiv inf\{x : \mathbb{E}[U_i|X_i = x, Y_i < x, N = n] \geq r\}.$$

Note that in the private value case, $x^*(r) = r$ since $inf\{x : x \geq r\} = r$. In the common value case, $x^*(r)$ is less straightforward and requires numerical approximation. The equilibrium bidding strategy, given a reserve price r, can then be obtained by solving the differential equation in Equation 4.3 using the initial condition $b^*(x^*(r)) = r$. Therefore,

$$b^*(x) = \begin{cases} L(x^*(r)|x,n)[r - v_n(x^*(r), x^*(r))] + v_n(x,x) - \int_{x^*(r)}^{x} L(\alpha|x,n)\frac{d}{d\alpha}v_n(\alpha,\alpha)d\alpha \\ \qquad \text{for } x \geq x^*(r) \\ < r \qquad \text{for } x < x^*(r), \end{cases} \tag{4.4}$$

where

$$L(\alpha|x,n) = exp(-\int_{\alpha}^{x} \frac{f_{Y_i|X_i,N}(s|s,n)}{F_{Y_i|X_i,N}(s|s,n)}ds).$$

Equation 4.4 will serve as the equilibrium bidding strategy that I estimate in the structural model. The goal of the analysis is to match the predicted bids using (4.4) with the observed bids as closely as possible.

4.4 Data and Descriptive Analysis

Before moving to the structural estimation, I first describe the datasets and variables, and then conduct correlational analyses to see the relationship between bid prices and competition. The three datasets relevant to this study are lease sales and status, drilled wells and their productions, and oil price and volatility. The main dataset for this study is the lease sales and status data from the Alaska Department of Natural Resources (DNR). This data includes all tracts offered for competitive bidding starting December 10th, 1959. It provides information on the acreage, location, fixed royalty requirement, and reserve price of each tract. We also observe bids submitted by all bidders in each auction, including the losing bids. In the case of joint bidding, we observe who the participants are, but we only observe the split of interests only amongst winning bidders.

Lease data can be linked to well drilling and production data from the Alaska Oil and Gas Conservation Committee (AOGCC). The AOGCC data shows which wells are drilled in what leases, well locations, well operators, and oil, gas, and water production for each well on a monthly level. This dataset allows me to determine which nearby wells are drilled before the sale of a tract and how productive these wells are.

In addition to tract sales data, I also obtained data on expected oil prices and price volatility from the replication materials of Kellogg (2014). The expected oil prices are from futures contracts 18 months to maturity. 18 months is the farthest-out horizon for regularly-traded NYMEX futures contracts. Since oil and gas gas exploration is a lengthy process, companies may focus more on the expected prices in the more distant future, when they would actually be producing oil. Hence, we use the 18-month futures price as oil price instead of current oil price at the time of bidding. The expected oil volatility is the implied volatility according to NYMEX futures options prices, and is calculated by Kellogg for his paper. Due to the limitation of the price and volatility data, I focus on tracts sold between 1993 and 2003.

Variables and descriptive statistics

Between 1993 and 2003, the Alaska Department of Natural Resources conducted 21 sales involving more than 2,600 tracts, which attracted over 1,900 bids. In this study,

Table 4.1: **Tract characteristics by number of bids received**

	After 1998		Size	Reserve		
# bids	0	1	Total Acreage	5	10	100
0	1,056	98	4,958	964	182	8
1	235	1,000	4,042	983	246	6
2	105	100	2,985	130	75	0
3	40	19	2,200	18	41	0
4	11	1	1,934	1	11	0
5	7	1	1,481	2	6	0
Total	1,454	1,219	4,298	2,098	561	14

I focus solely on tracts with "Fixed Royalty, Cash Offer" terms.[1] A summary of these tracts is shown in Table 4.1. Though the total number of tracts offered is similar before and after 1998, areawide sales implemented in 1998 seem to substantially increase the number of tracts auctioned off. The size of the tract seems to decrease with the number of bidders. This could be due to the sizes of tracts near known pools of oil deposit being smaller but attracting more bidders. Furthermore, all tracts sold in this period have a reserve price of at least $5.

In this dataset, 72 bidders participated in at least 1 auction. Of the 72 bidders, exactly half are companies and the other half are individuals. However, company bidders account for the majority of the bids received, more than three times the number of bids submitted by individuals. Table 4.2 presents a summary of the number bids submitted by top 9 company bidders, ranked by the total number of bids submitted between 1993 and 2003. It also shows whether these bids were submitted before 1998, whether they are solo or joint with another firm, and how often these top firms win in the auctions that they compete in.[2]

Next I describe the dependent and covariates to be included in the regression. For the dependent variable, I define the bid price as the offer price submitted by each bidder divided by the size of the tract. If, for instance, BP and Chevron submit a joint bid of one million dollars, then this counts as one bid and the joint entity of BP and Chevron counts as one participant. In the regression, I use log of bid per acre

[1] This is because these tracts account for the vast majority of tracts, and also the "Cash Offer" part in the term refers to the cash bid amount from the auction, whereas some other tracts offered for sale may not involve a cash portion at all.

[2] In the case of mergers and acquisitions, if both companies existed individually before acquisition, I count them as two different companies and pick one as the main company after the merger. If only one company operated in Alaska before the merger, then the post-merger company is considered the same as the pre-merger company.

Table 4.2: **Bids submitted by top bidders on leases sold between 1993 and 2003**

Name	Total	Before 1998	Joint bid	Win percentage
ARCO	322	297	168	72%
Anadarko Production	259	82	163	85%
BP Exploration	212	167	94	88%
Chevron	165	63	120	84%
Encana Oil and Gas	139	0	130	95%
Phillips Petroleum	105	34	105	94%
Unocal	95	39	29	84%
Union Texas	92	92	46	77%
Conoco/ConocoPhillips Alaska	44	7	15	91%

as the dependent variable to eliminate the wide range of per acre bid prices. The first row of Table 4.4 contains the summary statistic of the raw bid per acre for the 1,519 tracts that received bids between 1993 and 2003.

The variable of interest for the regression analysis is the number of potential bidders. I define a potential bidder as one who has shown interest in nearby tracts and who is still active at the time of auction. A previous study by Hendricks and Porter showed that companies in adjacent tracts tend to have better information regarding the value of the target tract (Hendricks and Porter, 1988). Based on this finding, Hendricks, Pinkse, and Porter (2003) then proposed that the number of potential bidders can be calculated as the sum of all bidders who submitted bids for this tract and bidders who competed for adjacent contracts in the past. Adapting this methodology, I identify adjacent tracts using latitude and longitude information of each tract, and identify bidders who have submitted bids for adjacent tracts previously. The summary statistic for the number of potential bidders is presented in Table 4.4 as well.

Given the aforementioned definitions of bid prices and potential bidder count, a simple tabulation, shown in Table 4.3, suggests that the median bid per acre increases with the number of potential bidders. This seems to be evidence against a common value model. However, in order to better discern the relationship between bid prices and competition, we need to control for tract and firm-specific characteristics.

To address tract and firm heterogeneity, I run an ordinary least squares regression of log bid price over the log of the number of potential bidders, tract/auction attributes, and firm fixed effects. The reduced form of this regression can be summarized by the following equation

$$b_{it} = \beta_0 + \beta_1 n_t + \beta_1 \mathbf{Z}_t + \mu_i + \varepsilon_{it}, \tag{4.5}$$

Table 4.3: **Median bids submitted for different numbers of potential bidders**

	Count	Median Bid
1	857	8.58
2	331	11.14
3	197	15.65
4	77	18.31
5+	57	18.71

Bid prices are in 2000 dollars

where i indexes the firm and t indexes the tract. Here, n_t is the number of potential bidders competing for tract t, $\mathbf{Z}_t$ is the set of variables corresponding to tract/auction attributes and μ_i denotes firm fixed effects.

Heterogeneity in tract sales, such as timing, contract terms and location, can impact tract value and should be controlled for. Timing-related variables that I include are: whether the sale happened before or after 1998, 18-month oil futures price, and oil price volatility implied by the 18-month futures contract. Recall that Table 4.1 suggests that the policy change in 1998 may be associated with higher percentage of tracts being sold. We could also expect this policy change to impact bid prices. At the time of auction, higher expected oil prices in the future may signal higher revenue and higher profit at the time of production, which may consequently encourage oil companies to bid more aggressively. In addition, Kellogg found that oil companies are less willing to invest in well drilling when price volatility is high (Kellogg, 2014). Inferring from Kellogg's finding, higher volatility may also lower companies' willingness to bid on tracts. Based on these considerations, I control for expected oil prices and volatility 18 months from the time of sale. A summary of these variables can be found in Table 4.4.

Terms specified in each tract sale may also impact bids submitted by firms. Some of the key terms are the reserve price and the fixed royalty amount. The reserve price determines the lowest bid allowed to be submitted in an auction. Royalty amount specifies the percentage of oil production belonging to the government. Usually a higher reserve price and a greater royalty amount are associated with more valuable land, and they are often highly correlated with each other. Of the 1,519 tracts which received bids in the data, 1,056 tracts have a reserve price of $5 and a fixed royalty rate of 12.5%. Another 318 have a $10 reserve price and a 16.67% fixed royalty rate. In other words, more than 90% of tracts have perfectly correlated reserve price and royalty amount. Hence I focus solely on reserve price in my analysis. In addition,

Table 4.4: **Summary statistics for numeric variables**

Variable	Type	Min.	Median	Mean	Max.	Std. dev
Bid per acre	dependent variable	0.65	11.80	41.07	6464.22	206.86
# potential bidder	variable of interest	1.00	2.00	2.32	10.00	1.74
Distance to	location					
-productive well		0.00	0.39	0.63	3.40	0.66
-unproductive well		0.00	0.12	0.28	2.68	0.43
Nearby well count	location					
-# productive wells		0.00	0.00	8.16	195.00	23.78
-# unproductive wells		0.00	0.00	2.80	41.00	5.69
-avg production of productive wells ('000 barrels)		0.00	0.00	302	9,278	1,100
Oil price	timing	15.65	22.74	23.24	26.84	2.56
Oil price volatility	timing	2.31	3.04	2.99	3.37	0.24
Experience in area	firm	0.00	36.00	112.92	857.00	176.48
Whether joint bid	firm	0.00	0.00	0.36	1.00	0.48

Revenue and price in 2000 dollars

since only two tracts in my sample have a reserve price of $100, I exclude these tracts from the analysis.

Tract location factors I consider include: the area that the tract is located in, productivity of nearby wells, and the tract's distance to nearby productive and unproductive wells. Alaska's North Slope and Beaufort Sea areas historically produced more oil than the other areas, and one can expect companies to value the tracts in these areas more. The productivity of nearby wells that were drilled before a tract goes on sale could also impact a company's valuation of the tract. Here I include a number of variables to control for the effect of learning from nearby wells. I first look at the distance to the closest productive and unproductive well. Close distance to a productive well could mean that the tract is located in a oil producing pool, whereas being close to unproductive wells may signal low-productivity for the target tract. I also include the number of productive wells within 0.1 degree in longitude or latitude distance, or approximately 11 kilometers, of the target tract. The rationale is that when a large number of productive wells have been drilled in the area already, it may signal that the area has been fully explored and the company is unlikely to discover more oil. However, if a large number of unproductive wells have been drilled, it could mean that these other firms are still optimistic about this area, and that it is "due" for oil discovery. Finally, I include the average production of the nearby wells: nearby wells being highly productive could raise the value of the land around these wells.

In addition to tract heterogeneity, firm behavior can also influence bid prices. In this category, I control for firm experience, whether a bid is submitted jointly or solo, and firm fixed effects. Firm experience is determined by the number of tracts that this firm has previously won in the region. If a firm has operated in a region for a long time, like BP in North Slope, it might have better local relationships and easier access to equipment. These advantages in the exploration or production process allow the firm to bid higher in the auction stage on tracts that it is interested in. Whether a bid is submitted jointly may influence the bid price but the direction of the relationship is not clear. While joint bidding may increase competition by allowing smaller companies to participate through partnerships, joint bidding amongst large companies may be a sign of collusion to reduce the price that they have to pay on the leases (Hendricks and Porter, 1992). Additionally, controlling for firm fixed effects eliminates how factors, such as firms' financial strength, could impact their bidding behaviors. One would expect that a powerful company like BP should be able to afford higher bids for valuable tracts than its individual investor competitors. If a bid is submitted by one firm, then the fixed effect is just the mean bid price from bids submitted by this firm. However, if a bid is submitted jointly by multiple firms, the fixed effect is a weighted average of these companies. Summary statistics for firm experience and joint bid indicator can be found in Table 4.4.

Regression Results

The results of the OLS regressions are presented in Table 4.5. In the first column, I run a simple univariate regression between the log of bid price per acre and the log of number of potential bidders. Similar to Table 4.3, Column 1 shows that bid prices and the number of potential bidders are positively correlated.

Column 2 and 3 incorporate controls for tract heterogeneity and firm characteristics with Column 3 also controlling for firm fixed effects. Compared to Column 1, tract value still increases with the number of potential bidders and this correlation is still significant at 10% and 1% level, respectively, though on a much smaller magnitude. Without controlling for any heterogeneity, a 10% increase in the number of potential bidders is associated with a 4.4% increase in bid per acre. After including tract and firm heterogeneity into the regression, a 10% increase is now associated with around 2% increase in tract value. Based on results from the three regressions, private value

Table 4.5: Regression of bid price on number of potential bidders

	Dependent variable:		
	Log bid per acre		
	(1)	(2)	(3)
Log # of potential bidder	0.444***	0.167*	0.220***
	(0.155)	(0.096)	(0.084)
Reserve = $10		0.046	0.256
		(0.276)	(0.279)
Area⊛			
Cook Inlet		−1.154***	−0.616**
		(0.360)	(0.298)
North Slope Foothills		−0.687**	−0.545
		(0.310)	(0.355)
North Slope		−0.568***	−0.361*
		(0.193)	(0.195)
After 1998		0.096	0.187
		(0.350)	(0.332)
Oil price		−0.031	−0.042
		(0.042)	(0.037)
Oil price volatility		−0.982	−0.672
		(0.769)	(0.663)
Log distance to productive wells		−0.085	−0.029
		(0.065)	(0.065)
Log distance to unproductive wells		−0.035	−0.051
		(0.039)	(0.039)
Log # productive wells nearby		−0.118*	−0.071
		(0.064)	(0.047)
Log # unproductive wells nearby		0.082	0.007
		(0.097)	(0.068)
Log average production of wells nearby		−0.007	−0.008
		(0.011)	(0.012)
Log experience		0.078***	−0.048*
		(0.020)	(0.029)
Joint bidding		−0.261*	−0.388***
		(0.135)	(0.083)
Constant	2.429***	6.465**	5.621**
	(0.170)	(3.028)	(2.531)
Observations	1,904	1,904	1,904
R^2	0.077	0.300	0.415
Adjusted R^2	0.076	0.294	0.407

Note: *p<0.1; **p<0.05; ***p<0.01
⊛: reference region is "Beaufort Sea"
⊛⊛: "nearby" is defined as 0.1 degree in longitude/latitude, approximately 11 km
Bid prices are in 2000 dollars
Numbers in parentheses are clustered standard errors clustered on sale date

model seems to explain Alaska oil and gas tract auctions better.

Column 2 and 3 regressions also show what other factors do or do not impact tract value. First, timing related variables seem to matter little. Switching to Areawide sales in 1998 do not seem to generate significantly higher bids. Meanwhile, similar to Kellogg (2014)'s finding that companies are less likely to invest in well drilling when oil price volatility is high, I also find that an increase in volatility lowers the bid amount. However, neither the impact of oil price nor price volatility is at all significant. When it comes to contract terms, both Column 2 and 3 show that a reserve price of $10 increases bids, but not significantly. This could potentially be due to the reserve prices set by the Alaska government being too low to be binding in these auctions (McAfee and Vincent, 1992). Location, however, seems to matter for tract value. As expected, tracts in the North Slope and Beaufort Sea regions are worth more than those in the Cook Inlet or the North Slope Foothills region. Both regressions also show that the productivity of wells nearby seems to influence tract value, especially the number of productive wells. Both regressions show that having a large number of productive wells nearby is associated with a drop in tract values. But we fail to find a consistent significant relationship between tract value and the number of unproductive wells. In terms of firm characteristics, I find that after controlling for firm fixed effects, owning more tracts in an area actually decreases the firm's evaluation of the target tract. This could be explained by firms owning enough tracts to know that the target tract is no longer valuable. Finally, both regressions consistently show that joint bidding is associated with lower bid prices. In summary, evidence from these correlational analyses seems to suggest that private value components play a larger role in companies' bidding strategy.

4.5 Structural Model

While the OLS regressions offer some insights into what factors correspond to higher tract prices, they do not account for the strategic interactions between firms in the auction process. In an auction equilibrium, firms not only look at tract terms or their own attributes, but also potential bids submitted by other firms. This section uses a parametric structural model to incorporate such strategic behaviors accounted for in auction theory. This model, developed by Hong and Shum, includes both common and private values, and allows for the differentiation of roles played by each factor (Han Hong and Shum, 2002). The parametric assumptions are essential to such identification, as Laffont and Vuong have famously shown that nonparametric distinction between common and private models is not possible without parametric assumptions on the common value component and bidder signals (Laffont and Vuong, 1996).

Model Setup

Let the joint distribution of the valuations and signals be parameterized by θ, i.e., $F(U_1, ..., U_N,\ X_1, ..., X_N; \theta)$. The valuation of bidder i is assumed to take a multiplicative form of common value V and private value A_i:

$$U_i = V \times A_i.$$

Further assume that V and A_i's are independently log-normally distributed. The independence assumption is reasonable since private value components, such as the strategic planning, the crew coordination or the equipment movement specific to a firm, are unlikely to be correlated with the oil and gas deposit. Then by the log-normality, we have

$$\tilde{V} \equiv \log V = \nu + \varepsilon_v \sim N(\nu, \sigma_v^2)$$
$$\tilde{A}_i \equiv \log A_i = \bar{a} + \varepsilon_{a_i} \sim N(\bar{a}, \sigma_a^2).$$

Define that $\tilde{U}_i \equiv \log U_i$, then

$$\tilde{U}_i = \tilde{V} + \tilde{A}_i \sim N(\nu + \bar{a}, \sigma_v^2 + \sigma_a^2).$$

Each bidder observes a noisy signal X_i of their true valuation U_i, which can be modeled as

$$X_i = U_i \times e_i, \text{ where } e_i = e^{\sigma_e \xi_i} \text{ and } \xi_i \sim N(0, 1).$$

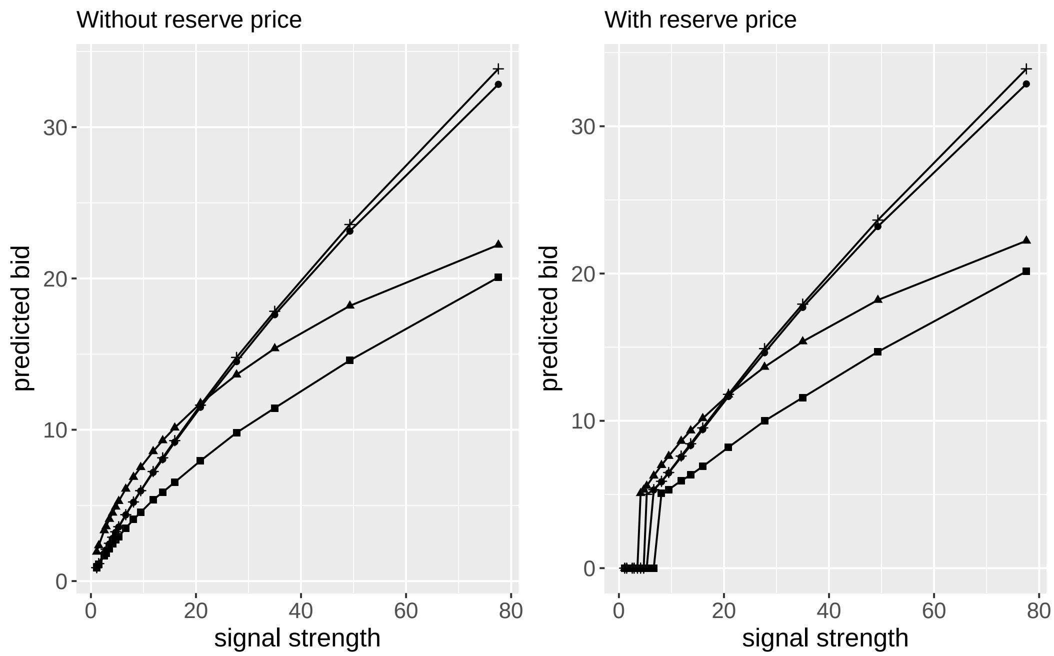
Further define that $\tilde{X}_i \equiv \log X_i$, then conditional on $\tilde{U}_i$,

$$\tilde{X}_i | \tilde{U}_i = \tilde{U}_i + \sigma_e \xi_i \sim N(\tilde{U}_i, \sigma_e^2).$$

Since we cannot effectively differentiate common value from private value components in the data, we cannot separately identify parameters ν and $\bar{a}$. Hence define $\mu = \nu + \bar{a}$. The set of parameters to estimate is then $\theta = (\mu, \sigma_v, \sigma_a, \sigma_e)$.

Using this model, we can infer the importance of the common and private value components through the magnitude of their corresponding parameters. As σ_e goes

Figure 4.1: **Simulated Equilibrium Bids**



● : $\mu = 2, \sigma_v = 1, \sigma_a = 1, \sigma_e = 1$; ▲ : $\mu = 2, \sigma_v = 0.5, \sigma_a = 1, \sigma_e = 1$;
■ : $\mu = 2, \sigma_v = 1, \sigma_a = 0.5, \sigma_e = 1$; + : $\mu = 2, \sigma_v = 1, \sigma_a = 1, \sigma_e = 0.5$.

to zero, there is no uncertainty in the individual's private signal and information asymmetry disappears. This model then resembles a pure affiliated private value model. Meanwhile, uncertainty about the common value disappears as σ_v goes to zero, and in this case we have an independent private value model with noisy signals. If σ_a goes to zero, then private value variation goes away and we have a pure common value model.

Figure 4.1 shows simulated equilibrium bids given different signal values with no reserve price and with a reserve price of $5. Here the number of potential bidders is set arbitrarily to 6. Simulation 1 involves parameter values of $\mu = 2, \sigma_v = 1, \sigma_a = 1$ and $\sigma_e = 1$. Simulation 2 differs from Simulation 1 by having $\sigma_v = 0.5$. Simulation 3 differs from Simulation 1 in σ_a, such that $\sigma_a = 0.5$. And finally, in Simulation 4, $\sigma_e = 0.5$ but all other parameters remain identical to Simulation 1. A comparison of Simulation 1 and 2 shows that when the variance of common value goes down, bidders bid more aggressively at lower signals but less so at higher signals. Holding common value variance constant, comparing Simulation 1 and Simulation 3 shows that when private value components become less important, the equilibrium bidding price decreases. Finally, when uncertainty of one's signal decreases, the equilibrium bid price goes up, as shown by Simulation 4. Comparing the left and right figures also demonstrates that when there is a reserve price, bidders bid more aggressively at signals closer to $x^*(r)$. But this aggressiveness slowly disappears as signal value increases and bids with or without reserve price become identical when the signals become sufficiently high.

Estimation and Results

Under the parametric setup described in the previous sub-section, for each parameter value θ, I estimate individual parts of Equation 4.4 to determine the equilibrium bids at different values of individual signal x. I then identify parameter values that best match the predicted bids with the observed bids. This process can be achieved using the quantile estimator. As Hong and Shum pointed out, given the monotonicity assumption of the bidding function $b^*(\cdot)$, the quantiles of the signal x are equivalent to the quantiles of $b^*(x)$ (Han Hong and Shum, 2002). Let $x_\tau, \tau = 1, ..., \mathcal{T}$ denote the τ^{th} quantile of the marginal distribution of signal X_i. Then we can obtain the parameters by minimizing the sum of difference between equilibrium bidding function evaluated at each quantile, $b^*(x_\tau)$, and the actual bid b_{it} for bidder $i = 1, ..., n'$ and auction $t = 1, ..., T$. Note here that n' denotes the number of observed bids, instead of the number of potential bidders. Due to the reserve price requirement in each auction, not all potential bidders end up submitting a bid. Furthermore, given different values of θ and the reserve prices from different auctions, the x_τ's need to be drawn from the truncated distribution above $x^*(r)$. Hence, the estimated parameter $\hat{\theta}$ minimizes the quantile objective function defined as:

$$Q(\theta) = \sum_{t=1}^{T} \sum_{i=1}^{n'} \sum_{\tau=1}^{\mathcal{T}} \rho_k [b_{it} - b_i^*(x_\tau; \theta, r_t)],$$

where r_t is the reserve price for auction t, and $\rho_k(\cdot)$ is defined as

$$\rho_\tau(x) = [\tau - \mathbf{1}(x \leq 0)]x.$$

The details of the estimation procedure for $b^*(\cdot)$ are discussed in Appendix C.1.

I estimate two sets of parameters under the assumptions of homogeneous and heterogeneous tracts. Under the homogeneity assumption, the tracts are all similar to each other in tract-related attributes mentioned in section 4.2, i.e., contract terms, location and timing of the sale. Under this assumption, we directly estimate the parameter of θ using quantile estimator. However, this assumption is unrealistic: we can see from Table 4.4 that tracts differ widely from one another in various attributes. Hence, under the heterogeneity assumption, I incorporate tract and firm attributes when estimating the parameters. One alternative is to include these attributes into μ, such that $\mu = \mathbf{X}\varphi$ where $\mathbf{X}$ are the covariates, and estimate φ separately. However, this approach requires a large dataset when the covariate list becomes large. Considering the limited size of my dataset, this approach would introduce large errors in the estimation process. Another alternative, proposed by Haile, Han Hong, and Shum (2003) is more suitable, if we are willing to assume

$$v(x, x, n, \mathbf{y}) = v(x, x, n) + \Gamma(\mathbf{y})$$

with covariates $\mathbf{Y}$ independent of signals $X_1, ..., X_n$. Under this assumption, the equilibrium bid also has a separable form $s(x; n, \mathbf{y}) = s(x, n) + \Gamma(\mathbf{y})$ by Lemma 4 in the paper. Hence, we only need to regress the observed bids on the covariates and a set of dummy variables for each value of potential bidder count n. Then new observed bids are derived by summing up the residual and the intercept estimate corresponding to the value of potential bidder count n. The derived bids are also equivalent to subtracting variations associated with tract and firm attributes from initial observed bids, while only keeping the variation due to different number of potential bidders. Using this approach, we can incorporate a large number of covariates even when the data set size is moderate. For this estimation, I use the set of covariates listed in Table 4.5, and use the derived bids after removing the tract and firm heterogeneity as the target for predicted bids to match using the quantile estimator.

Table 4.6 shows the estimated parameter values under both specifications. Without removing observable tract and firm heterogeneity, private value plays a much larger role in a firm's valuation. However, the role of private value components reversed when observable heterogeneity is removed. Common value now accounts for more variation in submitted bids and is much more significant than private value. However, both common value and private value variance becomes statistically insignificant even at the 10% level.[3] Overall, under the homogeneity assumption, my results agree with Li, Perrigne, and Vuong's results, while under the heterogeneity assumption, my results agree more with Hendricks, Pinkse, and Porter's arguments. Since Li, Perrigne, and Vuong did not remove tract heterogeneity in their study, their finding of private value dominating could potentially be due to tract or firm heterogeneity.

[3]The lack of significance could be due to the lack of variation in the number of potential bidders in the data. Further examination using more dynamic range of potential bidders can potentially improve the accuracy of the parameter estimates.

Table 4.6: **Estimated Parameter Values**

Parameter	Homogeneous	Heterogeneous
μ	1.599	5.741
	[1.332, 1.862]	*[5.415, 6.063]*
σ_v	0.226	0.300
	[0.209, 0.242]	*[0.063, 1.394]*
σ_a	2.902	0.209
	[2.206, 3.808]	*[0.017, 2.470]*
σ_e	1.863	0.906
	[1.355, 2.550]	*[0.618, 1.319]*
Simulation draws	100	
Quantiles	0.25, 0.4, 0.5, 0.6 & 0.75	

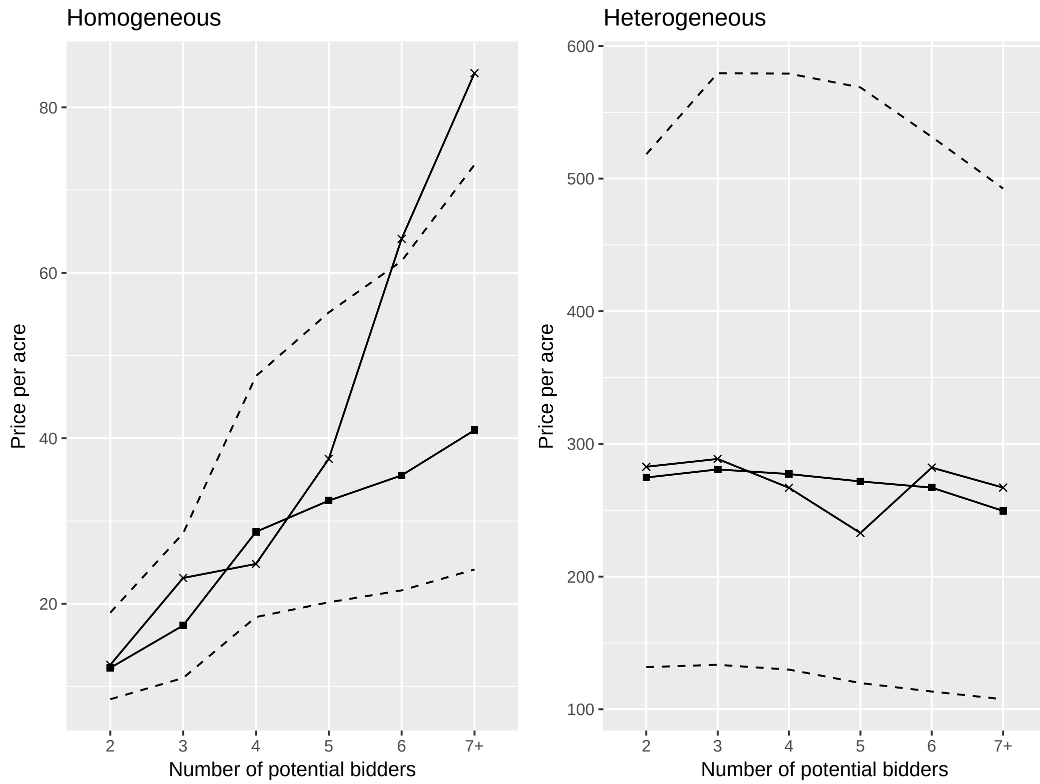
Note: The numbers in brackets are 95% confidence intervals.
Simulation draws are the number of simulations used to calculate $v_n(x, x)$ and to determine signal values in each quantile given each θ.
Quantiles are the set of quantile values used to fit the predicted bids with the observed bids.

As a robustness check, Figure 4.2 plots a comparison between the observed median bids for each tract and the predicted bids at median signal level above $x^*(r)$ at each competition level. The predicted bids trace the observed values relatively well, except when the number of potential bidders become large and the number of observations becomes small.

Knowing the relative roles of common and private value to each bidder, we can see how equilibrium bids differ at different competition levels and signal values. Figure 4.3 plots the relationship between equilibrium bids and competition under the homogeneity and heterogeneity assumptions. Figure 4.3a shows that equilibrium bids, for any given signal value, increase as the number of potential bidders increases, since private value dominates under the homogeneity assumption. This trend exists whether we have a reserve price of $5 or not. Figure 4.3b shows the opposite, where equilibrium bids decrease as the number of potential bidders increases, because common value is more prevalent under the heterogeneity assumption. Here we use a reserve price of $100, since when removing the heterogeneity all bids are artificially inflated. $100 corresponds to the lowest reserve price across tracts after removing tract heterogeneity.

While Figure 4.1 shows how bidders react to changes in competition, it does not address how increased competition impacts the revenue for the State. To offset the effect that each bidder bids lower with greater competition, i.e. the winner's curse effect, the increased number of participants also makes it more likely that a high

Figure 4.2: **Comparing Observed and Predicted Bids**



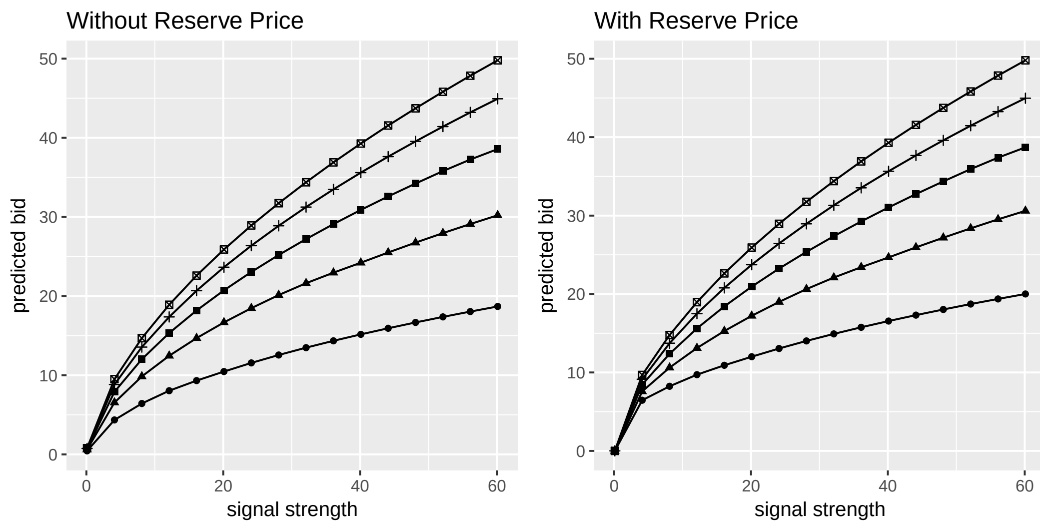
× : Median observed bids
■ : Median predicted bids
Notes: The dashed lines are 95% confidence intervals.

signal occurs, which I will refer to as the "likelihood effect." When the "likelihood effect" is sufficiently large, the revenue could in fact increase with more competition, even when common value components play a larger role. To determine which effect prevails, I simulate the State's expected revenue with the estimated parameters.[4] Figure 4.4 exhibits the shift in expected revenue as competition increases.[5] Under the homogeneity assumption where private value dominates, the expected revenue increases sharply as the number of potential bidder increases. However, under the heterogeneity assumption where common value plays a larger role, the expected revenue remains relatively flat as competition increases. This simulation shows that while the revenue may still increase when common value plays a larger role,

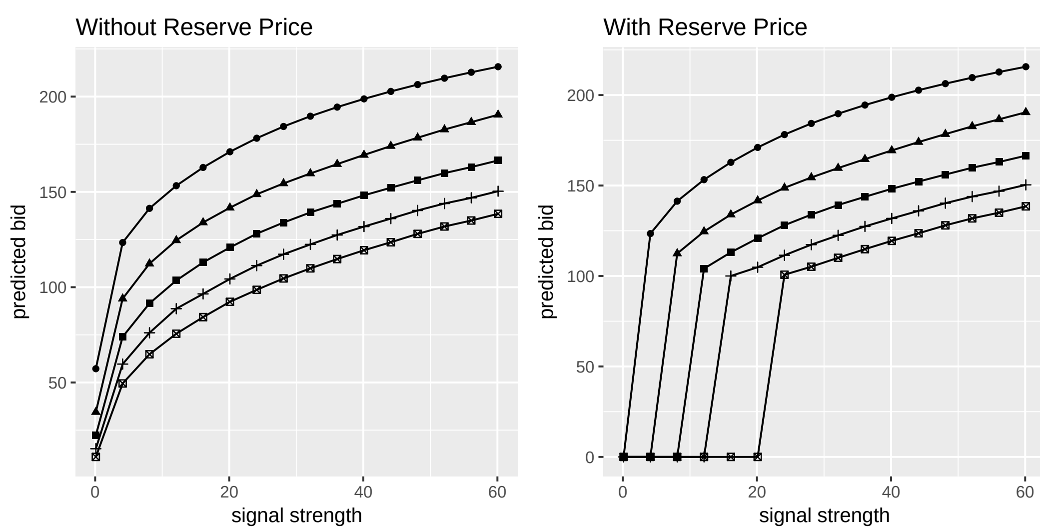
[4]The simulation process is as follows: (1) Simulate signal values for each potential bidder using the estimated parameters 500 times (2) Calculate the bid price for the highest signal amongst the n potential bidders in each iteration (3) Take the average of all the iterations. This serves as the expected revenue.

[5]Due to computational constraints, confidence intervals for expected revenues are being calculated but are not included in this thesis.

Figure 4.3: **Predicted bids varying by number of potential bidders**



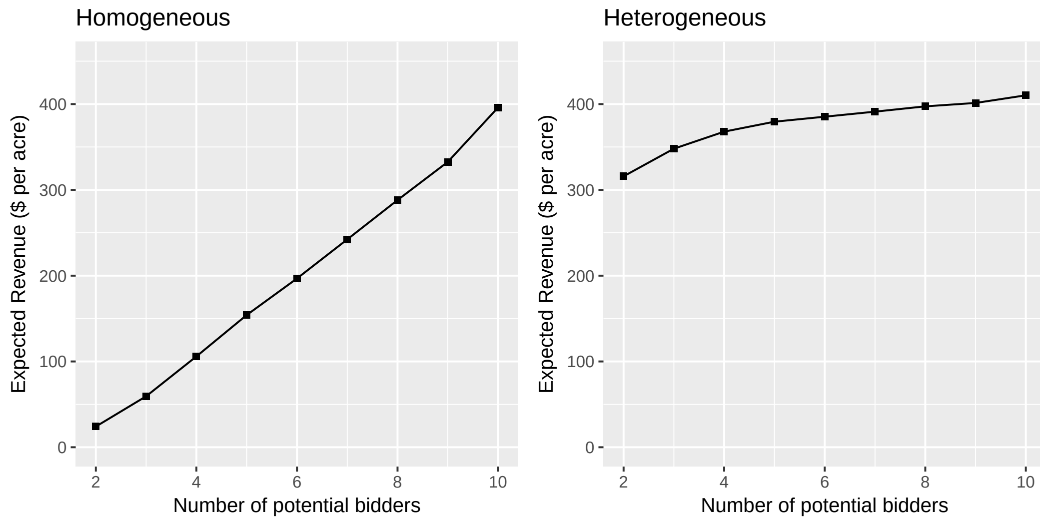
(a) Using parameters from homogeneity assumptions



(b) Using parameters from heterogeneity assumptions

● : n = 2; ▲ : n = 3; ■ : n = 4; + : n = 5; ⊠ : n = 6.

Figure 4.4: **Expected Revenue By Number of Potential Bidders**



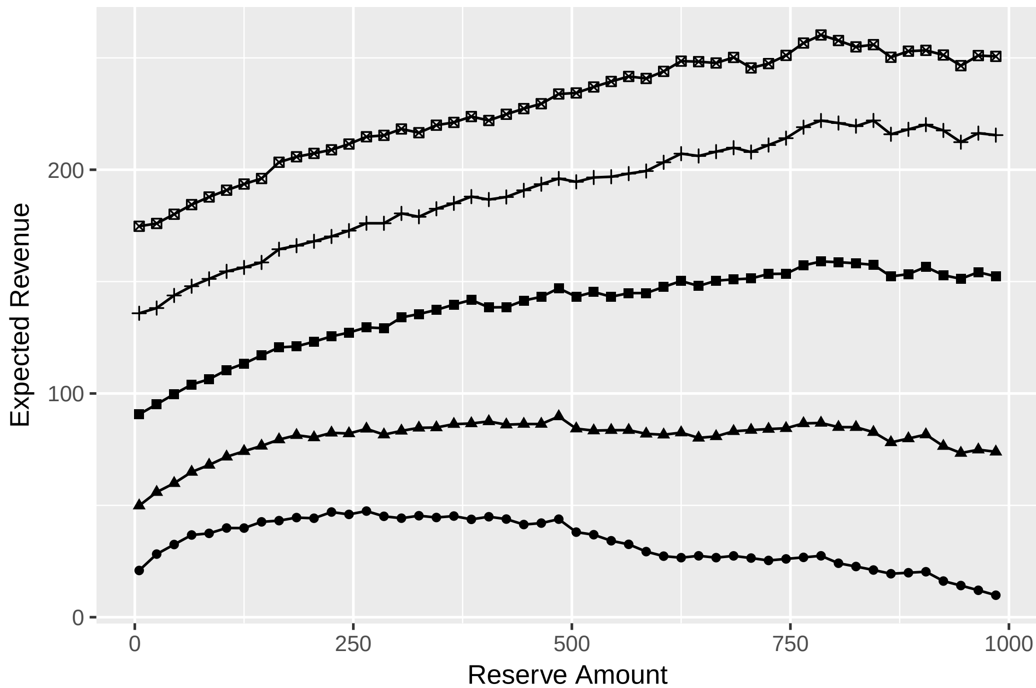
encouraging greater competition may not be nearly as effective as the government would hope for.[6] In other words, the government should evaluate the revenue impact carefully before blindingly giving out tax credit to encourage greater competition.

Another question that Figure 4.1 does not address is whether the $5 reserve price is ideal. Could the government generate higher revenues had they chosen a different reserve price? Figure 4.5 shows that *ex ante*, at various level of competition, the Alaska government is setting their reserve price too low, at least for the land that generated interests from large international firms. In this graph, I evaluate the expected revenue with different numbers of potential bidders at reserve prices ranging from 5 to 1000 dollars. Here I use the parameters from homogeneity assumption, since a $5 reserve price can vary widely from $110 to $275 after removing heterogeneity.[7] Figure 4.5 shows that as the number of potential bidders increases, the ideal reserve amount increases. However, while the reserve price for each lease is either $5 or $10, the optimal reserve price for the tracts in scope is much higher. The best reserve price appears to be around $250 when there are only two potential bidders. For three potential bidder, the ideal reserve price rises to $500, and it further rises to $800 with four or more potential bidders. This result is consistent with findings from previous literature that a $15 reserve price is too

[6] A dataset with greater variation in the number of potential bidders is needed in order to better pin down the parameters in the heterogeneity case in order to have tighter confidence interval around the expected revenue to draw better conclusion on the revenue impact.

[7] This means that for different tracts, the same reserve price could mean very differently for each bidder. This makes setting the ideal optimal reserve price uniformly for all bidders highly challenging.

Figure 4.5: **Expected Revenue By Different Reserve Price**



● : $n = 2$; ▲ : $n = 3$; ■ : $n = 4$; + : $n = 5$; ⊠ : $n = 6$.

low for the Federal OCS auctions (McAfee and Vincent, 1992). Nevertheless, this exercise has its drawbacks, as the leases in the estimation received bids from one of the top bidders, who have much stronger financial strength than an average bidder. Ex ante, we may not know if these firms were interested. However, this exercise has some implications for the state: if large international firms are interested or if many firms have shown interests in an area, it could increase the state revenue by raising the reserve prices. In popular areas with lots of interest, \$5 or \$10 reserve prices may be too conservative.

4.6 Conclusion

In this study, I set out to identify the roles that common and private values play in the Alaska oil and gas tract auctions. By analyzing the auction data from 1993 to 2003, the results suggest that private value dominates in the equilibrium bidding strategy when we treat each tract as homogeneous. However, common value plays a larger role when we remove tract heterogeneity. My results under the more realistic heterogeneity assumption are more consistent with Hendricks, Pinkse, and Porter's results. Foreseeing winner's curse, the predicted equilibrium bid decreases as the number of potential bidders increases. While the expected revenue may still increase when the common value plays a larger role, encouraging greater competition may not have the level of positive impact on revenue as the government would hope for. Furthermore, I find that the existing reserve price of $5 or $10 is too low for the areas that receive interests from large international firms, a potential avenue for the state to explore to increase state revenue in the auction process.

However, the results from the current study are by no means conclusive. Further analyses need to be conducted to better understand auction strategies in Alaska. Joint bidding has long been causing difficulties in the study of auctions, as collusions amongst companies could lower the State revenue, but reduced uncertainty thanks to information sharing could increase the bids. Empirical studies accounting for joint bidding have not been possible due to theoretical complications. But to better understand firms' bidding behaviors and their effects on the State revenue, these factors need to be accounted for. Another extension involves removing the symmetry assumption. Firms that participate in Alaska auctions come in a variety of sizes and capital capacity: individuals, medium-sized private firms, and international conglomerates. Focusing the analysis only on tracts that received bids from top bidders prevents us from extending the analysis to more general issues. For instance, are the $5 and $10 reserve prices still too low if we take all potential bidders, large and small, into consideration? However, asymmetry no longer allows for the the existing estimation procedures. Hong and Shum proposed a two-step structural estimation method, using a non-parametric first step to determine the joint signal distribution of bidders, and then a parametric second step to identify equilibrium bidding strategy (Han Hong and Shum, 1999). This proposed methodology needs more observations than the current dataset provides. Other states, such as Texas, with more frequent and larger number of tracts could be a good candidate for this exercise.

Chapter 5

EPILOGUE

Chapters 2 through 4 study the behaviors of oil and gas companies while focusing on different aspects of the exploration process. Their findings have various implications for the understanding and governance of oil and gas exploration. Overall, we found that large companies jointly operating with each other, and higher past oil prices tend to increase the drilling activities. Such increase in drilling activities benefits the government by increasing the government income and benefits consumers by combating rising oil prices through a faster recovery of the previous lower oil prices. On the other hand, large companies jointly operating with smaller, local firms and declining oil prices lead to less drilling. As a result, the government prevents the environment on this land from being damaged in the exploration process by sacrificing its revenue potential, and the oil industry shortens its period of "suffering" through a quicker recovery from the oil price downturns. These findings provide the government an incentive to evaluate its policies regarding joint operations, and contribute to the knowledge that seemingly irrational behaviors could actually benefit the overall industry and the economy.

We also add to the existing understanding of companies' evaluations of land prospects. While private value components, such as exploration strategy, local relationships, and financial constraints, still play a role in a company's evaluation, common value plays a larger role. As a result, increased competition may not lead to nearly as much increase in revenue as the government would expect. Furthermore, joint operation seems to be associated with lower evaluation of the land, evidenced by lower mean in their prior for land productivity (Table 2.2), and the lower bids submitted by joint bidders (Table 4.5). There are multiple possible explanations for this observation, such as collusion in the auction process, or the distributed risk from joint operating, but existing data would not be able to disentangle these hypotheses.

Despite the best effort of this dissertation, many questions still remain unanswered. These questions include (1) to what extent do large companies jointly operate to collude with each other, (2) if we take selection into joint operation into account, how does banning joint operation influence government revenue, (3) what is the cushioning benefit of biased beliefs in other cyclical industries outside of oil and gas, and (4) how much common value matters in the case of asymmetric auctions instead of symmetric auctions. These questions and many more others will need to be answered by future studies on these topics.

BIBLIOGRAPHY

Abel, Andrew B. (1981). “A Dynamic Model of Investment and Capacity Utilization”. In: *Quarterly Journal of Economics* 96, pp. 379–403.

Abel, Andrew B. and Janice C. Eberly (1994). “A Unified Model of Investment under Uncertainty”. In: *American Economic Review* 84, pp. 1369–1384.

Alti, Aydogan and Paul Tetlock (2014). “Biased Beliefs, Asset Prices, and Investment: A Structural Approach”. In: *Journal of Finance* 69, pp. 325–361.

Amromin, Gene and Steven A. Sharpe (2013). *Expectation of Risk and Return Among Household Investors: Are their Sharpe Ratios Countercyclical?* Working Paper. Federal Reserve Board, Washington DC.

Bacchetta, Philippe, Elmar Mertens, and Eric van Wincoop (2009). “Predictability in Financial Markets: What Does Survey Expectations Tell Us?” In: *Journal of International Money and Finance* 28, pp. 406–426.

Bagehot, Walter (1873). *Lombard Street: A Description of the Money Market*. Wiley, New York.

Barber, Brad M., Terrance Odean, and Ning Zhu (2009). “Systematic Noise”. In: *Journal of Financial Markets* 12, pp. 547–569.

Barberis, Nicholas and Andrei Shleifer (2003). “Style Investing”. In: *Journal of Financial Economics* 68, pp. 161–199.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny (1998). “A Model of Investor Sentiment”. In: *Journal of Financial Economics* 49, pp. 307–343.

Barberis, Nicholas et al. (2015). “X-CAPM: An Extrapolative Capital Asset Pricing Model”. In: *Journal of Financial Economics* 115, pp. 1–24.

– (2017). “Extrapolation and Bubbles”. In: *Journal of Financial Economics*. forthcoming.

Barsky, Robert B and J Bradford De Long (1993). “Why does the stock market fluctuate?” In: *Quarterly Journal of Economics* 108.2, pp. 291–311.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2017). “Diagnostic Expectations and Credit Cycles”. In: *Journal of Finance*. forthcoming.

BP (2016). “BP Annual Report and Form 20-F 2016”. In: `https://www.bp.com/content/dam/bp/en/corporate/pdf/investors/bp-annual-report-and-form-20f-2016.pdf`.

Brander, James A, Raphael Amit, and Werner Antweiler (2002). “Venture-Capital Syndication: Improved Venture Selection vs the Value-Added Hypothesis”. In: *Journal of Eonomics & Management Strategy* 11.3, pp. 423–452.

Cassella, Stefano and Huseyin Gulen (2015). *Extrapolation Bias and the Predictability of Stock Returns by Price-Scaled Variables*. Working Paper. Purdue University.

Choi, James and Thomas Mertens (2013). *Extrapolative Expectations and the Equity Premium*. Working Paper. New York University and Yale University.

Conlin, Michael, Ted O'Donoghue, and Timothy J Vogelsang (2007). "Projection bias in catalog orders". In: *American Economic Review* 97.4, pp. 1217–1249.

Derman, Andrew B. (2017). "1989 Joint Operating Agreement with Commentary". In: URL: http://www.tklaw.com/files/Publication/ef72d02f-06e2-490f-8806-747ea801df05/Presentation/PublicationAttachment/7719816f-42b1-47ff-9a57-afa788bb4aef/New%20and%20Improved%20-%20Derman.pdf.

Dong, Ming, David Hirshleifer, and Siew Hong Teoh (2017). *Stock Market Overvaluation, Moon Shots, and Corporate Innovation*. Working Paper. University of California, Irvine.

Egan, Matt (2016). "U.S. Oil Bankruptcies spike 379%". In: http://money.cnn.com/2016/02/11/investing/oil-prices-bankruptcies-spike/index.html.

EIA (2016). *Trends in U.S. Oil and Natural Gas Upstream Costs*. https://www.eia.gov/analysis/studies/drilling/pdf/upstream.pdf.

– (2017). *Crude Oil Production*. https://www.eia.gov/dnav/pet/pet_crd_crpdn_adc_mbblpd_a.htm.

Fuster, Andreas, Benjamin Hebert, and David Laibson (2011). "Natural Expectations, Macroeconomic Dynamics, and Asset Pricing". In: *NBER Macroeconomics Annual* 26, pp. 1–48.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer (2015). "Expectations and Investment". In: *NBER Macroeconomics Annual* 30, pp. 379–442.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny (2012). "Neglected Risks, Financial Innovation, and Financial Fragility". In: *Journal of Financial Economics* 104, pp. 452–468.

Glaeser, Edward L., Joseph Gyourko, and Albert Saiz (2008). "Housing Supply and Housing Bubbles". In: *Journal of Urbban Economics* 64, pp. 198–217.

Glaeser, Edward L. and Charles G. Nathanson (2015). *An Extrapolative Model of House Price Dynamics*. Working Paper. Harvard University and Northwestern University.

Greenwood, Robin, Sam Hanson, and Lawrence Jin (2016). *A Model of Credit Market Sentiment*. Working Paper. California Institute of Technology and Harvard University.

Greenwood, Robin and Samuel Hanson (2015). "Waves in Ship Prices and Investment". In: *Quarterly Journal of Economics* 130, pp. 55–109.

Greenwood, Robin and Andrei Shleifer (2014). "Expectations of Returns and Expected Returns". In: *Review of Financial Studies* 27, pp. 714–746.

Greenwood, Robin, Andrei Shleifer, and Yang You (2017). *Bubbles for Fama*. Working Paper. Harvard University.

Haile, Philip, Han Hong, and Matthew Shum (2003). "Nonparametric Tests for Common Values in First-Price Sealed-Bid Auctions". In: URL: http://www.nber.org/papers/w10105.

Hamilton, James D. (2011). "Historical Oil Shocks". In: *NBER Working Paper Series* 16790.

Haynes and LLP Boone (2016). "Oil Patch Bankruptcy Monitor". In: http://www.haynesboone.com/~/media/files/attorney\%20publications/2016/energy_bankruptcy_monitor/oil_patch_bankruptcy_20160106.ashx.

Hendricks, Kenneth, Joris Pinkse, and Robert Porter (2003). "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions". In: *The Review of Economic Studies* 70.1, pp. 115–145.

Hendricks, Kenneth and Robert Porter (1988). "An Epirical Study of an Auction with Asymmetric Information". In: *American Economic Review* 78.5, pp. 865–883.

– (1992). "Joint Bidding in Federal OCS Auctions". In: *The American Economic Review, Papers and Proceedings* 82.2, pp. 506–511.

Hendricks, Kenneth, Robert Porter, and Guofu Tan (2008). "Bidding Rings and the Winner's Curse". In: *The RAND Journal of Economics* 39.4, pp. 1018–1041.

Hirshleifer, David, Jun Li, and Jianfeng Yu (2015). "Asset Pricing in Production Economies with Extrapolative Expectations". In: *Journal of Monetary Economics* 76, pp. 87–106.

Hong, Han and Matthew Shum (1999). *Two-Step Structural Estimation of Parametric Asymmetric Auction Models*. Proposal.

– (2002). "Increasing Competition and the Winner's Curse: Evidence from Procurement". In: *Review of Economic Studies* 29, pp. 871–898.

Hong, Harrison and Jeremy C. Stein (1999). "A unified theory of underreaction, momentum trading and overreaction in asset markets". In: *Journal of Finance* 55, pp. 2143–2184.

Jin, Lawrence J. (2015). *A Speculative Asset Pricing Model of Financial Instability*. Working Paper. California Institute of Technology.

Jin, Lawrence J. and Pengfei Sui (2017). *Asset Pricing with Return Extrapolation*. Working Paper. California Institute of Technology.

Kellogg, Ryan (2014). "The effect of uncertainty on investment: evidence from Texas oil drilling". In: *American Economic Review* 104.6, pp. 1698–1734.

Kindleberger, Charles P. (1978). *Manias, Panics, and Crashes: A History of Financial Crises*. Basic Book, New York.

Koijen, Ralph S. J., Maik Schmeling, and Evert B. Vrugt (2015). *Survey Expectations of Returns and Asset Pricing Puzzles*. Working Paper. London Business School.

Kuchler, Theresa and Basit Zafar (2016). *Personal Experiences and Expectations about Aggregate Outcomes*. Working Paper. New York University and Federal Reserve Bank of New York.

Laffont, Jean-Jacques and Quang Vuong (1996). "Structural Analysis of Auction Data". In: *The American Economic Review, Papers and Proceedings* 86.2, pp. 414–420.

Lerner, Josh (1994). "The Syndication of Venture-Capital Investments". In: *Financial Management* 23.3, pp. 16–27.

Li, Tong, Isabelle Perrigne, and Quang Vuong (2000). "Conditionally Independent Private Information in OCS Wildcat Auctions". In: *Journal of Econometrics* 98, pp. 129–161.

McAfee, Randolph and Daniel Vincent (1992). "Updating the Reserve Price in Common-Value Auctions". In: *American Economic Review* 82.2, pp. 512–18. URL: `https://EconPapers.repec.org/RePEc:aea:aecrev:v:82:y:1992:i:2:p:512-18`.

Milgrom, Paul R. and Robert J. Weber (1982). "A Theory of Auctions and Competitive Bidding". In: *Econometrica* 50.5, pp. 1089–1122.

O'Rourke, Dara and Sarah Connolly (2003). "Just Oil? The Distribution of Environmental and Social Impacts of Oil Production and Consumption". In: *Annual Review of Environmental Resources* 28, pp. 587–617.

Revenue, Alaska Department of (2012). "Alaskas Oil and Gas Fiscal Regime: A Closer Look from a Global Perspective". In:

Somaini, Paulo (2011). "Competition and Interdependent Costs in Michigan Highway Procurement Auctions". In: URL: `https://economics.yale.edu/sites/default/files/somaini-120906.pdf`.

Villarreal, Debra J. and Lucas Lavoy (2010). "Oil and Gas Exploration and Development Agreements". In: *Energy & Mineral Law Institute*. Chap. 10, pp. 323–375.

Vissing-Jorgensen, Annette (2004). "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions". In: *NBER Macroeconomics Annual 2003* 18, pp. 139–194.

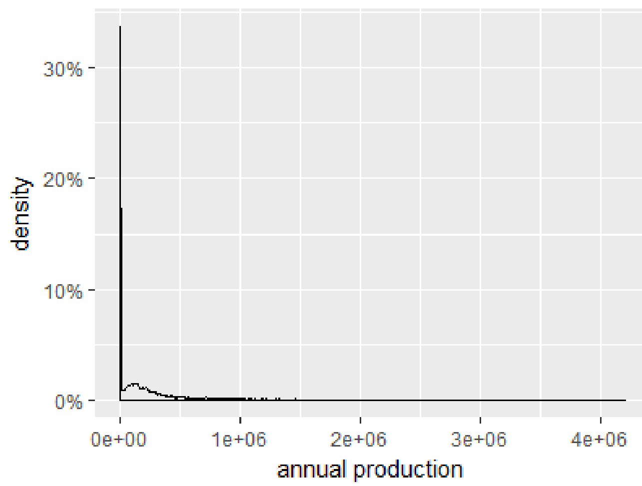
Wooldridge, Jeffrey M. (2010). *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, Massachusetts: The MIT Press. ISBN: 9780262232586.

Appendix A

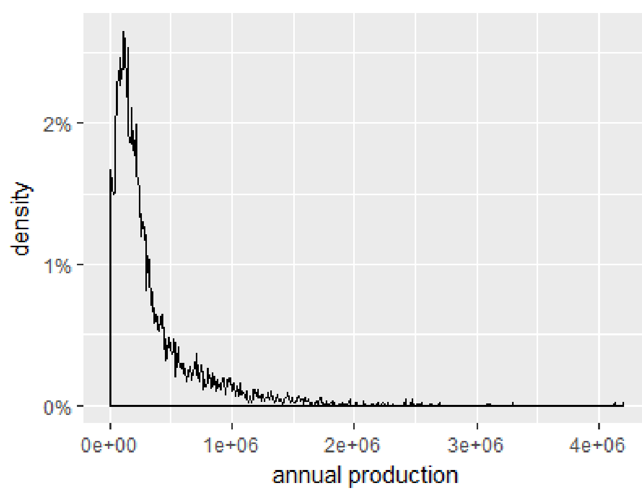
CHAPTER 2 APPENDIX

A.1 Ex-Post Well Production Distribution

(a) Production distribution of all wells



(b) Production distribution of wells with positive oil production



A.2 Standard Error Estimation

The finite sample variance of θ can be estimated using

$$V\hat{a}r(\hat{\theta}) = \frac{1}{K}[\sum_{k=1}^{K} s_(\theta)s_k^T(\theta)]^{-1},$$

where $s_k(\theta)$ is the score of the likelihood function, $l_k(\theta)$ (Wooldridge, 2010).

Since

$$\begin{aligned} l_k(\theta) &\equiv l(A_k, \boldsymbol{X}_k, \theta) \\ &= A_k \log Pr(A_k = 1|\boldsymbol{X}_k, \theta) + (1 - A_k)\log(1 - Pr(A_k = 1|\boldsymbol{X}_k, \theta)) \\ &= A_k \log p_1^k(\theta) + (1 - A_k)\log(1 - p_1^k(\theta)), \end{aligned}$$

where $p_1^k(\theta) = Pr(A_k = 1|\boldsymbol{X}_k, \theta)$.

To get the score, we calculate

$$\begin{aligned} \frac{\partial l_k(\theta)}{\partial \theta_j} &= \big[\frac{A_k}{p_1^k(\theta)} - \frac{1 - A_k}{1 - p_1^k(\theta)}\big]\frac{\partial p_1^k(\theta)}{\partial \theta_j} \\ &= \big[\frac{A_k}{p_1^k(\theta)} - \frac{1 - A_k}{1 - p_1^k(\theta)}\big]\Big(- \phi(\frac{\tilde{s}^*(\theta) - \nu^k}{\sqrt{(\sigma_v^k)^2 + (\sigma_s^k)^2}})\Big)\frac{\partial \frac{\tilde{s}^*(\theta) - \nu^k}{\sqrt{(\sigma_v^k)^2 + (\sigma_s^k)^2}}}{\partial \theta_j}. \end{aligned} \quad (3)$$

Let $\frac{\partial \frac{\tilde{s}^*(\theta) - \nu^k}{\sqrt{(\sigma_v^k)^2 + (\sigma_s^k)^2}}}{\partial \theta_j}$ be represented by $\circledast(\theta_j)$.

Since $\tilde{s}_k^*(\theta)$ contains all parameters, The first step is to derive $\partial \tilde{s}_k^*(\theta)/\partial \theta_j$.

Deriving $\partial \tilde{s}_k^*(\theta)/\partial \theta_j$ for all θ_j

- For σ_s^k related variables
  - Constant

    $$\frac{\partial \tilde{s}_k^*(\theta)}{\partial \alpha_0} = \frac{(\sigma_s^k)^2}{(\sigma_v^k)^2}\big(\log Q^* - \nu^k - \frac{1}{2}(\sigma_v^k)^2\big)$$

  - Non-Constant (i.e. J_k and $\tilde{W}_k$)

    Here I use J_k as an example. But the same goes for $\tilde{W}_k$.

    $$\frac{\partial \tilde{s}_k^*(\theta)}{\partial \alpha_J} = \frac{(\sigma_s^k)^2}{(\sigma_v^k)^2}\big(\log Q^* - \nu^k - \frac{1}{2}(\sigma_v^k)^2\big)J_k = \frac{\partial \tilde{s}_k(\theta)}{\partial \alpha_0}J_k$$

- For σ_v^k related variables
  - Constant

$$\frac{\partial \tilde{s}_k^*(\theta)}{\partial \gamma_0} = -\frac{(\sigma_s^k)^2}{(\sigma_v^k)^2}\big(\log Q^* - \nu^k\big)$$

  - Non-Constant (i.e. W_ks and J_k)

$$\frac{\partial \tilde{s}_k^*(\theta)}{\partial \gamma_j} = -\frac{(\sigma_s^k)^2}{(\sigma_v^k)^2}\big(\log Q^* - \nu^k\big)W_k = \frac{\partial \tilde{s}_k(\theta)}{\partial \gamma_0}W_k$$

- For ν^k related variables
  - Constant

$$\frac{\partial \tilde{s}_k^*(\theta)}{\partial \beta_0} = -\frac{(\sigma_s^k)^2}{(\sigma_v^k)^2}$$

  - Non-Constant (i.e. W_k and J_k)

$$\frac{\partial \tilde{s}_k^*(\theta)}{\partial \beta_j} = -\frac{(\sigma_s^k)^2}{(\sigma_v^k)^2}W_k = \frac{\partial \tilde{s}_k(\theta)}{\partial \beta_0}W_k$$

Deriving $\circledast$ for all θ_j

- For σ_s^k related variables
  - Constant

$$\circledast(\alpha_0) = \frac{\frac{\partial \tilde{s}_k^*(\theta)}{\partial \alpha_0}((\sigma_v^k)^2 + (\sigma_s^k)^2) - \frac{\partial (\sigma_s^k)^2}{\partial \alpha_0}}{\big((\sigma_v^k)^2 + (\sigma_s^k)^2)\big)^{3/2}}$$

  - Non-Constant (i.e. J_k and $\tilde{W}_k$)

$$\circledast(\alpha_J) = \circledast(\alpha_0) \cdot J_k$$

- For σ_v^k related variables
  - Constant

$$\circledast(\gamma_0) = \frac{\frac{\partial \tilde{s}_k^*(\theta)}{\partial \gamma_0}((\sigma_v^k)^2 + (\sigma_s^k)^2) - \frac{\partial (\sigma_v^k)^2}{\partial \gamma_0}}{\big((\sigma_v^k)^2 + (\sigma_s^k)^2)\big)^{3/2}}$$

  - Non-Constant (i.e. W_ks and J_k)

$$\circledast(\gamma_j) = \circledast(\gamma_0) \cdot W_k$$

- For v^k related variables
  - Constant

$$\circledast(\beta_0) = \frac{\frac{\partial \tilde{s}_k^*(\theta)}{\partial \beta_0} - \frac{\partial v^k}{\partial \beta_0}}{\big((\sigma_v^k)^2 + (\sigma_s^k)^2)\big)^{1/2}}$$

  - Non-Constant (i.e. W_k)

$$\circledast(\beta_j) = \circledast(\beta_0) \cdot W_k$$

Calculating $s_k(\theta)$

Suppose there are m variables in $\tilde{\boldsymbol{W}}_{\boldsymbol{k}}$ and n variables in W_k for each k, then by the results from previous sections,

$$\begin{aligned} s_k(\theta) =& \nabla_\theta l_k(\theta)^T \\ =& \Big(\frac{A_k}{p_1^k(\theta)} - \frac{1 - A_k}{1 - p_1^k(\theta)}\Big) \cdot \begin{bmatrix} \partial p_1^k(\theta)/\partial\alpha_0 \\ (\partial p_1^k(\theta)/\partial\alpha_0)J_k \\ (\partial p_1^k(\theta)/\partial\alpha_0)\tilde{W}_k^1 \\ \dots \\ (\partial p_1^k(\theta)/\partial\alpha_0)\tilde{W}_k^m \\ \partial p_1^k(\theta)/\partial\gamma_0 \\ (\partial p_1^k(\theta)/\partial\gamma_0)J_k \\ (\partial p_1^k(\theta)/\partial\gamma_0)W_k^1 \\ \dots \\ (\partial p_1^k(\theta)/\partial\gamma_0)W_k^n \\ \partial p_1^k(\theta)/\partial\beta_0 \\ (\partial p_1^k(\theta)/\partial\beta_0)J_k \\ (\partial p_1^k(\theta)/\partial\beta_0)W_k^1 \\ \dots \\ (\partial p_1^k(\theta)/\partial\beta_0)W_k^n \end{bmatrix} \end{aligned}$$

In the case with both L&L and L&F operators, we just need three additional rows for J_k^{ls} for the score.

A.3 Parameters for Clusters Operated Solely by Large Operators

Table A.1: **Estimation Results: Clusters Operated Solely by Large Operators**

Description \ Metrics	σ_s Parameter	σ_s Estimate	σ_v Parameter	σ_v Estimate	ν Parameter	ν Estimate
Constant	α	3.099	γ_0	5.512	β_0	-49.999
		(0.264)		*(1.468)*		*(2.361)*
Indicator - L&L	α_J	0.506	γ_J	-0.125	β_J	-0.418
		(0.198)		*(0.427)*		*(0.323)*
Experience	α_e	-0.825	γ_e	0.073	β_e	0.174
		(0.175)		*(0.294)*		*(0.27)*
Knowledge	α_k	-4.998	γ_k	-0.261	β_k	2.339
		(2.146)		(1.403)		*(0.973)*
Acreage			γ_a	-0.210	β_a	3.889
				(0.157)		*(0.172)*
Region			γ_r	-0.557	β_r	2.889
(1=Cook Inlet)				*(0.454)*		*(0.393)*
After 1975			γ_{75}	-0.654	β_{75}	1.473
				(0.891)		*(0.479)*
Log winning bid			γ_w	0.171	β_w	1.420
				(0.151)		*(0.137)*
Log number bids			γ_b	-0.108	β_b	0.182
				(0.393)		*(0.342)*

Note: Parameters estimated in this table is based on 682 clusters operated by large operators exclusively (only the L and L&L types). Numbers in parentheses are standard errors. Details of the standard error calculation can be found in Appendix A.2.

A.4 Counterfactuals

Table A.2: **Clusters explored by joint operators but may not have by solo operators**

cluster	sale date	drilling probability		royalty production		royalty revenue	
		actual (joint)	counter (solo)	oil (MM bbl)	gas (MM mcf)	oil (MM $)	gas (MM $)
1	1965-07-14	0.95	0.87	140.3	340.4	4,479.1	618.4
2	1965-07-14	0.13	0.09	87.6	336.9	2,580.1	579.7
3	**1961-12-19**	**0.66**	**0.56**	**0**	**17.3**	**0**	**25.1**
4	**1959-12-10**	**0.86**	**0.71**	**0**	**5.4**	**0**	**5.7**
5	**1959-12-10**	**0.69**	**0.53**	**0**	**0.002**	**0**	**0.003**
6	**1959-12-10**	**0.61**	**0.45**	**0**	**0**	**0**	**0**
7	**1959-12-10**	**0.85**	**0.69**	**0**	**0**	**0**	**0**
8	1960-07-13	0.12	0.11	0	0	0	0
9	1961-12-19	0.82	0.73	0	0	0	0
10	1962-07-11	0.38	0.37	0	0	0	0
11	1962-07-11	0.39	0.38	0	0	0	0
12	1963-12-11	0.53	0.50	0	0	0	0
13	**1965-07-14**	**0.83**	**0.68**	**0**	**0**	**0**	**0**
14	1965-07-14	0.25	0.18	0	0	0	0
15	**1965-07-14**	**0.73**	**0.61**	**0**	**0**	**0**	**0**
16	**1965-07-14**	**0.84**	**0.69**	**0**	**0**	**0**	**0**
17	1967-01-24	0.52	0.46	0	0	0	0
18	**1969-09-10**	**0.95**	**0.85**	**0**	**0**	**0**	**0**
19	**1969-09-10**	**0.74**	**0.58**	**0**	**0**	**0**	**0**
20	1969-09-10	0.99	0.94	0	0	0	0
21	1982-09-28	0.86	0.82	0	0	0	0
22	1983-05-17	0.47	0.46	0	0	0	0
23	1984-05-22	0.67	0.60	0	0	0	0
24	1985-09-24	0.55	0.53	0	0	0	0
25	1986-02-25	0.48	0.46	0	0	0	0
			Total Bold	0	22.7	0	30.8
			Total All	227.9	700.0	7,059.2	1,228.9

Note: This is a list of clusters where the counterfactual probability of drilling, solo instead of joint with large companies only, is lower than the actual probability. The bolded rows are where the counterfactual probability is 10 percentage points lower. The counterfactual probability is calculated using parameters from Table 2.2. The revenues in this table are in year 2000 dollars.

Table A.3: **Clusters explored by L&L operators but may not have by L&F operators**

cluster	sale date	drilling probability		royalty			
				production		revenue	
		actual (joint)	counter (solo)	oil (MM bbl)	gas (MM mcf)	oil (MM $)	gas (MM $)
1	**1965-07-14**	**0.89**	**0.75**	**140.3**	**340.4**	**4,479.1**	**618.4**
2	1965-07-14	0.16	0.07	87.6	336.9	2,580.1	579.7
3	1961-12-19	0.24	0.16	15.2	10.7	277.5	11.5
4	1962-07-11	0.50	0.49	6.1	89.2	97.4	139.4
5	**1961-12-19**	**0.71**	**0.53**	**0**	**17.3**	**0**	**25.1**
6	**1959-12-10**	**0.80**	**0.56**	**0**	**5.4**	**0**	**5.7**
7	**1959-12-10**	**0.60**	**0.36**	**0**	**2.4**	**0**	**2.5**
8	**1959-12-10**	**0.55**	**0.31**	**0**	**0**	**0**	**0**
9	**1959-12-10**	**0.79**	**0.56**	**0**	**0**	**0**	**0**
10	1960-07-13	0.14	0.07	0	0	0	0
11	**1961-12-19**	**0.85**	**0.69**	**0**	**0**	**0**	**0**
12	**1962-07-11**	**0.46**	**0.34**	**0**	**0**	**0**	**0**
13	1962-07-11	0.68	0.67	0	0	0	0
14	**1962-07-11**	**0.44**	**0.34**	**0**	**0**	**0**	**0**
15	1963-05-08	0.27	0.21	0	0	0	0
16	1963-05-08	0.17	0.12	0	0	0	0
17	1963-05-08	0.12	0.08	0	0	0	0
18	**1963-12-11**	**0.58**	**0.45**	**0**	**0**	**0**	**0**
19	1964-12-09	0.99	0.93	0	0	0	0
20	**1965-07-14**	**0.72**	**0.48**	**0**	**0**	**0**	**0**
21	**1965-07-14**	**0.26**	**0.12**	**0**	**0**	**0**	**0**
22	**1965-07-14**	**0.68**	**0.47**	**0**	**0**	**0**	**0**
23	**1965-07-14**	**0.74**	**0.50**	**0**	**0**	**0**	**0**
24	**1967-01-24**	**0.53**	**0.39**	**0**	**0**	**0**	**0**
25	**1969-09-10**	**0.91**	**0.78**	**0**	**0**	**0**	**0**
26	**1969-09-10**	**0.69**	**0.49**	**0**	**0**	**0**	**0**
27	1969-09-10	0.96	0.87	0	0	0	0
28	1972-12-11	0.13	0.10	0	0	0	0
29	1982-05-26	0.40	0.31	0	0	0	0
30	1982-09-28	0.83	0.73	0	0	0	0
31	**1983-05-17**	**0.51**	**0.41**	**0**	**0**	**0**	**0**

(To continue on next page)

Clusters explored by L&L operators but may not have by L&F operators (Cont.)

cluster	sale date	drilling probability		royalty			
				production		revenue	
		actual (joint)	counter (solo)	oil (MM bbl)	gas (MM mcf)	oil (MM $)	gas (MM $)
32	1983-05-17	0.32	0.23	0	0	0	0
33	**1984-05-22**	**0.68**	**0.53**	**0**	**0**	**0**	**0**
34	**1985-09-24**	**0.52**	**0.41**	**0**	**0**	**0**	**0**
35	**1986-02-25**	**0.45**	**0.32**	**0**	**0**	**0**	**0**
36	**1988-01-26**	**0.45**	**0.35**	**0**	**0**	**0**	**0**
37	1989-01-24	0.11	0.09	0	0	0	0
38	1991-01-29	0.61	0.53	0	0	0	0
39	1991-01-29	0.43	0.42	0	0	0	0
40	1992-01-22	0.37	0.36	0	0	0	0
41	1996-08-20	0.16	0.14	0	0	0	0
42	1996-12-18	0.09	0.07	0	0	0	0
43	1998-06-24	0.51	0.45	0	0	0	0
			Total Bold	140.3	365.5	4,479.1	651.7
			Total All	249.2	802.3	7,434.1	1,382.3

Note: This is a list of clusters where the counterfactual probability, joint operation between large and small instead of joint with large only company drilling, is lower than the actual probability. The bolded rows are where the counterfactual probability is 10 percentage points lower. The counterfactual probability is calculated using parameters from Table 2.2. The revenues in this table are in year 2000 dollars.

Appendix B

CHAPTER 3 APPENDIX

B.1 Analytical Results for the Model

Proof of Proposition 1. We conjecture and verify later that the optimal investment is linear in state variables A_t, $\overline{A}_t$, and Q_t

$$i_t^* = x + y_1 \cdot A_t + y_2 \cdot \overline{A}_t + z \cdot Q_t. \tag{B.1}$$

Equation (3.11) then implies

$$P(A_t, \overline{A}_t, Q_t) = (kx + ky_1 \cdot A_t + ky_2 \cdot \overline{A}_t + kz \cdot Q_t + P_r)/p. \tag{B.2}$$

By applying the law of iterated expectations on (3.11), firms derive

$$P_r + k \cdot i_t^* = p \cdot \frac{\mathbb{E}_f[\Pi_{t+1} + P(A_{t+1}, \overline{A}_{t+1}, Q_{t+1})|A_t, \overline{A}_t, Q_t]}{1+r}. \tag{B.3}$$

Equations (3.4), (3.5), (3.7), and (3.8) allow us to write (B.3) out as

$$\begin{aligned}
& P_r + kx + ky_1 \cdot A_t + ky \cdot \overline{A}_t + kz \cdot Q_t \\
= \quad & p \cdot \frac{M\{\alpha\overline{A}_t + (1-\alpha)\overline{A} + \rho_f[A_t - \alpha\overline{A}_t - (1-\alpha)\overline{A}]}{1+r} \\
+ \quad & \frac{-B(Q_t + px + py_1 \cdot A_t + py_2 \cdot \overline{A}_t + pz \cdot Q_t)\} - C - \delta P_r}{1+r} \\
+ \quad & \frac{kx + (ky_1 + \rho_A ky_2) \cdot \{\alpha\overline{A}_t + (1-\alpha)\overline{A} + \rho_f[A_t - \alpha\overline{A}_t - (1-\alpha)\overline{A}]\}}{1+r} \\
+ \quad & \frac{ky_2 \cdot (1-\rho_A)\overline{A}_t}{1+r} \\
+ \quad & \frac{kz \cdot (Q_t + px + py_1 \cdot A_t + py_2 \cdot \overline{A}_t + pz \cdot Q_t) + P_r}{1+r}.
\end{aligned} \tag{B.4}$$

The fact that both sides of (B.4) are linear functions of A_t, $\overline{A}_t$, and Q_t verifies the conjecture in (B.1). Matching terms in a sequential order then solves for x, y_1, y_2, and z. First, matching terms for Q_t gives the solution of z in (3.13). Then matching terms for A_t and $\overline{A}_t$, we obtain

$$\begin{aligned} ky_1 &= pM\frac{\rho_f - Bpy_1}{1+r} + \frac{(ky_1 + \rho_A ky_2)\rho_f + kzpy_1}{1+r}, \\ ky_2 &= pM\frac{(1-\rho_f)\alpha - Bpy_2}{1+r} + \frac{(ky_1 + \rho_A ky_2)(1-\rho_f)\alpha}{1+r} \\ &+ \frac{ky_2 \cdot (1-\rho_A) + kzpy_2}{1+r}. \end{aligned} \tag{B.5}$$

Notice that y_1 and y_2 are interrelated because the evolution of $\overline{A}_t$ is driven by past realizations of A_t. Solving these two simultaneous equations then leads to (3.14). Finally, matching the constant term gives (3.15). ■

Proof of Corollary 1. Conditional on knowing A_{t-1}, $\overline{A}_{t-1}$, and Q_{t-1}, I_{t-1} and therefore Q_t are both determined. In this case, the movements of H_t and I_t are only caused by the realization of the random shock ε_t. That is

$$\begin{aligned} I_t &= x + (y_1 + \rho_A y_2) \cdot H_t + y_2(1-\rho_A) \cdot \overline{A}_{t-1} + (z + y_1 B + \rho_A y_2 B) \cdot Q_t \\ &= x + (y_1 + \rho_A y_2) \cdot H_t + y_2(1-\rho_A) \cdot \overline{A}_{t-1} \\ &+ (z + y_1 B + \rho_A y_2 B) \cdot f(A_{t-1}, \overline{A}_{t-1}, Q_{t-1}). \end{aligned} \tag{B.6}$$

So the coefficient for regressing I_t on H_t, both conditional and unconditional on A_{t-1}, $\overline{A}_{t-1}$, and Q_{t-1}, is $\beta_0 = y_1 + \rho_A y_2$.

We now consider the coefficient of regressing I_{t+1} on H_t. Conditional on A_{t-1}, $\overline{A}_{t-1}$, and Q_{t-1}, the realization of ε_t determines H_t and I_t, which further determine Q_{t+1}.

Then the realization of ε_{t+1} determines A_{t+1}, $\overline{A}_{t+1}$, and I_{t+1}

$$
\begin{aligned}
I_{t+1} &= x + y_1 \cdot A_{t+1} + y_2 \cdot \overline{A}_{t+1} + z \cdot Q_{t+1} \\
&= x + (y_1 + \rho_A y_2) \cdot [\overline{A} + \rho_0(A_t - \overline{A}) + \varepsilon_{t+1}] + y_2(1 - \rho_A)\overline{A}_t + z \cdot [Q_t + pI_t] \\
&= x + (y_1 + \rho_A y_2)\varepsilon_{t+1} + (y_1 + \rho_A y_2)(1 - \rho_0)\overline{A} \\
&+ [(y_1 + \rho_A y_2)\rho_0 + y_2(1 - \rho_A)\rho_A]H_t + y_2(1 - \rho_A)^2\overline{A}_{t-1} \\
&+ \{z + B[(y_1 + \rho_A y_2)\rho_0 + y_2(1 - \rho_A)\rho_A]\}f(A_{t-1}, \overline{A}_{t-1}, Q_{t-1}) \\
&+ zp[x + (y_1 + \rho_A y_2)H_t + y_2(1 - \rho_A)\overline{A}_{t-1} \\
&+ (z + y_1 B + \rho_A y_2 B)f(A_{t-1}, \overline{A}_{t-1}, Q_{t-1})] \\
&= x + (y_1 + \rho_A y_2)\varepsilon_{t+1} + (y_1 + \rho_A y_2)(1 - \rho_0)\overline{A} \\
&+ [(y_1 + \rho_A y_2)\rho_0 + y_2(1 - \rho_A)\rho_A + zp(y_1 + \rho_A y_2)]H_t + y_2(1 - \rho_A)^2\overline{A}_{t-1} \\
&+ \{z + B[(y_1 + \rho_A y_2)\rho_0 + y_2(1 - \rho_A)\rho_A]\}f(A_{t-1}, \overline{A}_{t-1}, Q_{t-1}) \\
&+ zp[x + y_2(1 - \rho_A)\overline{A}_{t-1} + (z + y_1 B + \rho_A y_2 B)f(A_{t-1}, \overline{A}_{t-1}, Q_{t-1})].
\end{aligned}
\tag{B.7}
$$

So the coefficient for regressing I_{t+1} on H_1, both conditional and unconditional on A_{t-1}, $\overline{A}_{t-1}$, and Q_{t-1}, is $\beta_1 = (y_1 + \rho_A y_2)\rho_0 + y_2(1 - \rho_A)\rho_A + zp(y_1 + \rho_A y_2)$. ■

Appendix C

CHAPTER 4 APPENDIX

C.1 Estimation Procedure for Equilibrium Bids

Let n denote the number of potential bidders and n' denote the number of actual bidders, who submitted bids above the reserve price and are observed in the data. We have the following scenarios.

n = 1

If there is only one potential bidder, then his equilibrium bidding strategy is:

$$b^*(x) = \begin{cases} r \text{ if } x \geq x^*(r) \\ 0 \text{ if } x < x^*(r) \end{cases}$$

That is, the sum of differences between predicted and actual bids from this set of auctions are fixed. Since no parameters need to be estimates for these auctions, I will focus on estimating parameters using auctions with two or more potential bidders.

n ≥ 2

Step 1: Estimate $v_n(x, x)$ without a reserve price

- Write $v_n(x, x)$ explicitly. Let $i = 1$ and the 2nd highest bidder be $j = 2$. Then

$$\begin{aligned} v_n(x,x) &= \mathbb{E}[U_i|X_1 = x, X_2 = x, X_j < x \; \forall j \geq 3, N = n] \\ &= \underbrace{\int_{\underline{x}}^{x} \dots \int_{\underline{x}}^{x}}_{n-1} \mathbb{E}[U_i|X_1, X_{-1}] dF(X_3, ..., X_n|X_1 = x, X_2 = x, X_k < x; \theta) \\ &\quad \text{for } k \geq 3 \\ &= \frac{1}{Prob(X_k < x, k \geq 3|X_1 = x, X_2 = x; \theta)} * \\ &\quad \int_{\underline{x}}^{x} \dots \int_{\underline{x}}^{x} \mathbb{E}[U_i|X_1, X_{-1}] dF(X_3, ..., X_n|X_1 = x, X_2 = x; \theta) \end{aligned}$$

  Hence we want to get $E[U_i|X_1, ..., X_n]$ and $F(\cdot)$.

- Get an analytic form for $E[U_i|X_1, ..., X_n]$. Given our assumptions, $(\tilde{U}_1, ..., \tilde{U}_n, \tilde{X}_1, ...\tilde{X}_n)$ is distributed jointly normal with mean $\vec{\mu}$ and variance covariance matrix $\Sigma = \begin{bmatrix} \Sigma_{\mathbf{u}} & \Sigma_{\mathbf{ux}} \\ \Sigma_{\mathbf{ux}}^T & \Sigma_{\mathbf{x}}. \end{bmatrix}$. Then given the log-normality assumption, we have

$$\mathbb{E}[U_i|X_1, ..., X_n] = exp(\mathbb{E}[\tilde{U}_i|\tilde{X}_1, ..., \tilde{X}_n] + \frac{1}{2}Var[\tilde{U}_i|\tilde{X}_1, ..., \tilde{X}_n]) \;\; i = 1, ..., n$$

After deriving the details of Σ, we have:

$$\Sigma_{\mathbf{u}} = \begin{bmatrix} \sigma_v^2 + \sigma_a^2 & \sigma_v^2 & ... & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 + \sigma_a^2 & ... & \sigma_v^2 \\ ... & ... & ... & ... \\ \sigma_v^2 & \sigma_v^2 & ... & \sigma_v^2 + \sigma_a^2 \end{bmatrix}$$

$$\Sigma_{\mathbf{ux}} = \begin{bmatrix} \sigma_v^2 + \sigma_a^2 & \sigma_v^2 & ... & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 + \sigma_a^2 & ... & \sigma_v^2 \\ ... & ... & ... & ... \\ \sigma_v^2 & \sigma_v^2 & ... & \sigma_v^2 + \sigma_a^2 \end{bmatrix}$$

$$\Sigma_{\mathbf{x}} = \begin{bmatrix} \sigma_v^2 + \sigma_a^2 + \sigma_e^2 & \sigma_v^2 & ... & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 + \sigma_a^2 + \sigma_e^2 & ... & \sigma_v^2 \\ ... & ... & ... & ... \\ \sigma_v^2 & \sigma_v^2 & ... & \sigma_v^2 + \sigma_a^2 + \sigma_e^2 \end{bmatrix}$$

Note that $(\tilde{U}_i, \tilde{X}_1, ..., \tilde{X}_n)$ is distributed jointly normal as well i.e. $(\tilde{U}_i, \tilde{X}_1, ..., \tilde{X}_n) \sim N(\vec{\mu}_{n+1}, \Sigma_i)$, where $\Sigma_i = \begin{bmatrix} \sigma_v^2 + \sigma_a^2 & (\Sigma_{\mathbf{ux}}^i)^T \\ \Sigma_{\mathbf{ux}}^i & \Sigma_{\mathbf{x}} \end{bmatrix}$ such that $\Sigma_{\mathbf{ux}}^i$ is the i^{th} column of $\Sigma_{\mathbf{ux}}$. Then the joint normality assumption gives us

$$E[\tilde{U}_i|\tilde{X}_1, ..., \tilde{X}_n] = \mu + (\Sigma_{\mathbf{ux}}^i)^T \Sigma_{\mathbf{x}}^{-1}(\tilde{x} - \vec{\mu}_n)$$
$$Var[\tilde{U}_i|\tilde{X}_1, ..., \tilde{X}_n] = \sigma_v^2 + \sigma_a^2 - (\Sigma_{\mathbf{ux}}^i)^T \Sigma_{\mathbf{x}}^{-1} \Sigma_{\mathbf{ux}}^i$$

- **If n = 2**

$$v_n(x, x) = \mathbb{E}[U_i|X_1 = x, X_2 = x] = exp[\mu + (\Sigma_{\mathbf{ux}}^i)^T \Sigma_{\mathbf{x}}^{-1}(\tilde{x} - \vec{\mu}_n) + \frac{1}{2}\{\sigma_v^2 + \sigma_a^2 - (\Sigma_{\mathbf{ux}}^i)^T \Sigma_{\mathbf{x}}^{-1} \Sigma_{\mathbf{ux}}^i\}]$$

where $\tilde{x} = (\log x, \log x)$.

- **If n ≥ 3**: Simulate $\tilde{X}$ satisfying the truncation restriction. Let $\tilde{X}_{+2} = (\tilde{X}_1, \tilde{X}_2)$ and $\tilde{X}_{3+} = (\tilde{X}_3, ..., \tilde{X}_n)$. Since $(\tilde{X}_1, ..., \tilde{X}_n)$ is distributed jointly normal, then

$$\begin{bmatrix} \tilde{X}_{3+} \\ \tilde{X}_{+2} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{n-2} \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{3+} & \Sigma_{3+,+2} \\ \Sigma_{3+,+2}^T & \Sigma_{+2} \end{bmatrix} \right),$$

where $\Sigma_{3+,+2} = \underbrace{\begin{bmatrix} \sigma_v^2 & \sigma_v^2 \\ ... & ... \\ \sigma_v^2 & \sigma_v^2 \end{bmatrix}}_{n-2 \text{ rows}}$. And hence

$$\tilde{X}_{3+}|\tilde{X}_{+2} \sim TN(\mu_{n-2} + \Sigma_{3+,+2}(\Sigma_{+2})^{-1}(\tilde{X}_{+2} - \mu_2), \Sigma_{3+} - \Sigma_{3+,+2}(\Sigma_{+2})^{-1}\Sigma_{3+,+2}^T; \\ \underline{\vec{\tilde{x}}} \leq \tilde{X}_{3+} \leq \vec{\tilde{x}}).$$

Based on this, let

$$\begin{aligned} \vec{\mu}^* &= \mu_{n-2} + \Sigma_{3+,+2}(\Sigma_{+2})^{-1}(\tilde{X}_{+2} - \mu_2), \\ \Sigma^* &= \Sigma_{3+} - \Sigma_{3+,+2}(\Sigma_{+2})^{-1}\Sigma_{3+,+2}^T, \end{aligned}$$

and let $m = n - 2$. Also let $(\Sigma^*)^{1/2}$ denote the lower-triangular Cholesky factorization of Σ^*, with elements:

$$\begin{bmatrix} s_{11} & 0 & ... & 0 & 0 \\ s_{21} & s_{22} & ... & 0 & 0 \\ ... & ... & s_{ii} & 0 & 0 \\ s_{m1} & s_{m2} & ... & s_{mm-1} & s_{mm} \end{bmatrix}.$$

We can rewrite

$$\begin{gathered} \tilde{X}_{3+}|\tilde{X}_{+2} = \vec{\mu}^* + (\Sigma^*)^{1/2}\vec{Z} \sim N(\vec{\mu}^*, \Sigma^*) \; s.t. \\ \frac{\underline{\tilde{x}} - \mu_1^*}{s_{11}} \leq z_1 \leq \frac{\tilde{x} - \mu_1^*}{s_{11}} \\ \frac{\underline{\tilde{x}} - \mu_2^* - s_{21}z_1}{s_{22}} \leq z_2 \leq \frac{\tilde{x} - \mu_2^* - s_{21}z_1}{s_{22}} \\ ... \\ \frac{\underline{\tilde{x}} - \mu_m^* - \sum_{i=1}^{m-1} s_{mi}z_i}{s_{mm}} \leq z_2 \leq \frac{\tilde{x} - \mu_m^* - \sum_{i=1}^{m-1} s_{mi}z_i}{s_{mm}} \end{gathered}$$

We can use the **truncnorm** package from R to draw these z_m's.

Step 2: Identify $x^*(r)$

- Since $x^*(r) \equiv inf\{x : \mathbb{E}[U_i|X_i = x, Y_i < x, N = n] \geq r\}$, define $v_n(x, y) = \mathbb{E}[U_i|X_i = x, Y_i < x, N = n]$, where $y < x$. We can use similar methodology to estimate $v_n(x, y)$ as $v_n(x, x)$, where instead of $\tilde{X}_{+2}$ and $\tilde{X}_{3+}$, we have $\tilde{X}_1$ and

$\tilde{X}_{2+}$. We use the same methodology to simulate the signals of $n-1$ (instead of $n-2$) bidders. In this case, we have:

$$\begin{bmatrix}\tilde{X}_{2+}\\ \tilde{X}_1\end{bmatrix} \sim N\Bigg(\begin{bmatrix}\mu_{n-1}\\ \mu\end{bmatrix}, \begin{bmatrix}\Sigma_{2+} & \Sigma_{2+,1}\\ \Sigma^T_{2+,1} & \sigma_v^2+\sigma_a^2+\sigma_e^2\end{bmatrix}\Bigg),$$

where $\Sigma_{2+,1} = \underbrace{\begin{bmatrix}\sigma_v^2\\ \dots\\ \sigma_v^2\end{bmatrix}}_{n-1 \text{ rows}}$ And hence

$$\begin{aligned}\tilde{X}_{2+}|\tilde{X}_1 \sim TN(&\mu_{n-1}+\Sigma_{2+,1}(\sigma_v^2+\sigma_a^2+\sigma_e^2)^{-1}(\tilde{X}_1-\mu),\\ &\Sigma_{2+}-\Sigma_{2+,1}(\sigma_v^2+\sigma_a^2+\sigma_e^2)^{-1}\Sigma^T_{2+,1};\\ &\underline{\tilde{x}} \leq \tilde{X}_{2+} \leq \tilde{x})\end{aligned}$$

- We can find the x such that $v_n(x,y)=r$. Then this x is $x^*(r)$.

Step 3: Calculate $L(x^*(r)|x,n)$.

- Find the closed form for $f_{Y_i|X_i,N}(s|s,n)/F_{Y_i|X_i,N}(s|s,n)$.
  - Since $\tilde{X}_i|\tilde{v}$ is independent of $\tilde{X}_j|\tilde{v}$, then $X_i|v$ is also independent of $X_j|v$. Now we can show $f(X_2,...,X_n|X_1=x)$ in closed form.

$$\begin{aligned}f(X_2,...,X_n|X_1) &= \int_{-\infty}^{\infty} f(X_2,...,X_n,v|X_1)dv\\ &= \int_{-\infty}^{\infty} f(X_2,...,X_n|v,X_1)f(v|X_1)dv\\ &= \int_{-\infty}^{\infty} \frac{f(X_1,...,X_n|v)}{f(X_1|v)}\frac{f(X_1|v)f(v)}{f(X_1)}\\ &= \int_{-\infty}^{\infty} \Pi_{i=1}^n f(X_i|v)f(v)\frac{1}{\int_{-\infty}^{\infty} f(X_1|v)f(v)dv}dv\end{aligned}$$

Then $F(Y_1 \leq \tilde{s}|X_1=s) = Prob(X_2 \leq \tilde{s}, X_3 \leq \tilde{s}, ..., X_n \leq \tilde{s}|X_1=s)$. So

$$\begin{aligned}F(\tilde{s}|s) &= \int_{-\infty}^{\infty}\Big(\underbrace{\int_{\underline{x}}^{\tilde{s}}\dots\int_{\underline{x}}^{\tilde{s}}}_{n-1}\Pi_{i=2}^n f(X_i|v)dX_2dX_3...dX_n\Big)\cdot\\ &\qquad \frac{f(X_1=s|v)f(v)}{\int_{-\infty}^{\infty} f(X_1|v)f(v)dv}dv\\ &= \int_{-\infty}^{\infty}\Pi_{i=2}^n F_{X_i|v}(\tilde{s})\frac{f(X_1=s|v)f(v)}{\int_{-\infty}^{\infty} f(X_1|v)f(v)dv}dv\end{aligned}$$

Note that

$$F_{X_i|v}(\tilde{s}) = Prob(X_i \leq \tilde{s}|v) = Prob(e_i^{\tilde{X}} \leq s|v) = Prob(\tilde{X}_i \leq \log s|v)$$
$$\equiv H(\log s) \sim N(\log v + \bar{a}, \sigma_a^2 + \sigma_e^2)$$

And since $\tilde{V} \sim N(v, \sigma_v^2)$, then

$$f(v) = \frac{1}{v\sqrt{2\pi}\sigma_v} e^{-\frac{(\log v - v)^2}{2\sigma_v^2}}$$

Assuming $\bar{a} = 0$ in the prior, then we can replace $\log v + \bar{a}$ by $\log v$ and v by μ.

Hence,

$$\begin{aligned} F(\tilde{s}|s) = & \int_{-\infty}^{\infty} (H(\log \tilde{s}))^{n-1} \frac{1}{\int_{-\infty}^{\infty} f(X_1 = s|v) f(v) dv} f(v) f(X_1 = s|v) dv \\ = & \int_{-\infty}^{\infty} (H(\log \tilde{s}))^{n-1} \frac{1}{\frac{1}{s} \int_{-\infty}^{\infty} f(\tilde{X}_1 = \log s|v) f(v) dv} \cdot \\ & f(v) \frac{f(\tilde{X}_1 = \log s|v)}{s} dv \\ = & \int_{-\infty}^{\infty} (H(\log \tilde{s}))^{n-1} \frac{1}{\int_{-\infty}^{\infty} f(\tilde{X}_1 = \log s|v) f(v) dv} f(v) h(\log s) dv \end{aligned}$$

and correspondingly,

$$\begin{aligned} f(\tilde{s}|s) = & (n-1) \int_{-\infty}^{\infty} (H(\log \tilde{s}))^{n-2} h(\log \tilde{s}) \frac{1}{\tilde{s}} \frac{1}{\int_{-\infty}^{\infty} f(\tilde{X}_1 = \log s|v) f(v) dv} \cdot \\ & f(v) h(\log s) dv \end{aligned}$$

Hence,

$$\frac{f(s|s)}{F(s|s)} = \frac{(n-1) \int_{-\infty}^{\infty} (H(\log s))^{n-2} (h(\log s))^2 / s f(v) dv}{\int_{-\infty}^{\infty} (H(\log s))^{n-1} h(\log s) f(v) dv}$$

- Then $L(x^*(r)|x, n) = exp(-\int_{x^*(r)}^{x} \frac{f_{Y_i|X_i,N}(s|s,n)}{F_{Y_i|X_i,N}(s|s,n)} ds)$

Step 4: Calculate $\int_{x^*(r)}^{x} L(\alpha|x, n) \frac{d}{d\alpha} v_n(\alpha, \alpha) d\alpha \equiv g(x)$.

- Based on Step 3, $L(\alpha|x, n) = exp(-\int_{\alpha}^{x} \frac{f_{Y_i|X_i,N}(s|s,n)}{F_{Y_i|X_i,N}(s|s,n)} ds)$

- $\frac{d}{d\alpha}v_n(\alpha,\alpha)$ can be approximated numerically

$$\frac{d}{d\alpha}v_n(\alpha,\alpha) = \frac{v(\alpha+\delta*\alpha,\alpha+\delta*\alpha)-v(\alpha,\alpha)}{\delta*\alpha}$$

  where $\delta = 10^{-6}$ after trying different values of $\delta = 10^{-4}, 10^{-6}, 10^{-8}$, and 10^{-10} yielded no significant differences.

Step 5: Combine all elements for equilibrium bidding strategy for observed bidders

Simulate a series of $\{x_1, ..., x^S\}$ based on the log normal distribution. Keep only the x's such that $x^s \geq x^*(r)$. Sort the remaining x's and identify the x_τ's that correspond to each quantile in the truncated distribution.

Then for each x_τ, the equilibrium bidding strategy for bidder i in auction t is hence:

$$b^*_{it}(x_\tau) = L(x^*(r_t)|x_\tau,n)[r_t - v_n(x^*_t(r_t),x^*_t(r_t))] + v_n(x_\tau,x_\tau) - \int_{x^*_t(r_t)}^{x_\tau} L(\alpha|x_\tau,n)\frac{d}{d\alpha}v_n(\alpha,\alpha)d\alpha$$

Since bidders are symmetric by assumption, then $b^*_t(x_\tau) = b^*_{it}(x_\tau)$ for all observed bidders $i = 1, ..., n'$.

Step 6: Find the parameters that minimize the quantile objective function

For each auction $t = 1, ..., T$, and for each observed bidder $i = 1, ..., n'$, we try to match the shape of the predicted bid with the actual bid. Hence the estimator $\hat{\theta}$ minimizes the following objective function:

$$Q(\theta) = \sum_{t=1}^{T}\sum_{i=1}^{n'}\sum_{\tau=1}^{\mathcal{T}}\rho_k[b_{it} - b^*_i(x_\tau;\theta,r_t)]$$

where $\rho_k(\cdot)$ is defined as

$$\rho_\tau(x) = [\tau - \mathbf{1}(x \leq 0)]x.$$