Booms, Busts, and Mismatch in Capital Markets: Evidence from the Offshore Oil and Gas Industry

Job Market Paper

Please find the latest version here.

Nicholas Vreugdenhil*

November 22, 2018

Abstract

How efficiently do markets reallocate capital in booms and busts, and what are the effects of policies designed to smooth out fluctuations? This paper exploits a novel dataset of contracts and projects in the offshore oil and gas industry to examine the role of matching in shaping industry reallocation. Oil and gas companies undertake projects (wells) but do not own capital (drilling rigs), and so must search for capital in a decentralized market. The quality of the match matters: more efficient rigs are suited to drilling more complex projects. Moreover, search frictions arise due to the idiosyncratic nature of the projects. I find that booms - caused by increases in oil and gas prices - are associated with a sorting effect: booms increase the option value of searching for a better match which leads agents to avoid bad matches. This results in stronger sorting patterns in booms than busts, and less mismatch. I provide an identification strategy to disentangle changes in the composition of searching projects (demand) from the sorting effect. The strategy relies on inverting observed matches through a flexible search technology and acceptance sets to identify the composition of searching projects. I estimate a structural model of the industry that tractably incorporates rich dynamics in the distributions of searching agents. Comparing a model where agents are not selective in matching to the market benchmark, I find that the sorting effect increases welfare (measured in total profits) by 11.4%. Yet, frictionless matching would further increase welfare by 28.6%. Demand smoothing policies such as countercyclical tax credits - which are common in the industry - lead only to small increases in welfare.

*Department of Economics, Northwestern University. Email: nicholasvreugdenhil2013@u.northwestern.edu. Thank you to Robert Porter, Mar Reguant, and Gaston Illanes for their encouragement and advice. I would also like to thank Gaurab Aryal, Igal Hendel, Ken Hendricks, and Yuta Toyama for useful comments and suggestions. This paper has benefited from conversations with David Berger, Vivek Bhattacharya, Ying Fan, Phil Haile, Myrto Kalouptsidi, Jakub Kastl, Ryan Kellogg, Elena Krasnokutskaya, Bradley Larsen, Bill Rogerson, and Mark Satterthwaite. I thank participants in the Northwestern IO Seminar, Monash University, and the IIOC 2018. This research was supported by a grant from the Center for the Study of Industrial Organization at Northwestern. I acknowledge IHS and Rigzone for providing data.
1 Introduction

When markets surge in a boom or crash in a bust, firms adjust by reallocating capital. This reallocation process is central to understanding aggregate productivity, and has spurred large literatures in Industrial Organization, Labor, and Macroeconomics.\footnote{For a review of the literature see Eisfeldt and Shi (2018).} Although it is well established that fluctuations and productivity are broadly linked, the \textit{exact process} of capital reallocation within industries is not well understood.\footnote{Bartelsman et al. (2013) document empirically the role of reallocation.} This is largely due to the lack of producer-level data on covariates such as contracts, production, and relationships.\footnote{Collard-Wexler and De Loecker (2015) make a similar argument to motivate their study which uses micro-data to investigate reallocation in the US steel sector.} Filling this gap is important because the benefits of often-proposed policies - such as demand smoothing - hinge on the reallocation mechanism.

In this paper I focus on the role of \textit{matching} between physical capital and projects in shaping industry reallocation. Finding a good match is an important consideration in decentralized capital markets. However, these markets are often plagued by search frictions which can hinder firms from finding the best match for their capital (Gavazza (2016)). Despite this, little is known about how search frictions affect matching in booms and busts in real-world capital markets.

The goal of this paper is to answer the question: how efficiently do markets reallocate and match capital in booms and busts? I develop a framework to answer this question that combines elements of the search and matching literature and the firm dynamics literature. I apply the framework to study reallocation in the market for offshore oil and gas drilling rigs - an outstanding example of a cyclical decentralized capital market. Using a novel dataset of contracts and projects, I find that booms (which are caused by increases in oil and gas prices) are associated with a \textit{sorting effect}. The intuition is simple. Booms increase the option value of searching for a better match which raises the opportunity cost of being locked into a bad match. This leads agents to avoid bad matches in booms, resulting in stronger sorting patterns in booms than busts, and less mismatch.\footnote{Note that this result is not mechanical. Rather, it is an empirical question whether stronger sorting is optimal in booms. This is because the value of a match also increases in booms and therefore it may be optimal to be less selective.} I provide an identification strategy to disentangle changes in
Figure 1: Illustration of the sorting effect

Note: This figure contains a simple example of the sorting effect. Suppose that in both panels there are three rigs of each type, \{low, mid, high\}, and three wells of each type, \{0.25, 0.5, 0.75\}, where a higher number corresponds to a more complex well. Each panel plots an allocation of the nine wells to the nine rigs. In a bust all rigs drill similar wells resulting in a flat average match line. In a boom simple wells are allocated to low-efficiency rigs and complex wells are allocated to high-efficiency rigs, resulting in a more diagonal average match line. For a fixed number of rigs and wells, so long as the match value is supermodular in rig type, there will be higher total output in the boom allocation.

the composition of searching projects (demand) from the sorting effect. I use the framework to quantify the benefits of frictionless matching and the effects of a demand smoothing policy.

The market for offshore drilling rigs is an excellent setting for studying booms and busts because it is subject to large exogenous fluctuations in drilling activity caused by global oil and gas prices. Oil and gas companies, such as BP, Chevron, and ExxonMobil, undertake projects (wells) but do not own capital (drilling rigs). Instead, they must search for capital in a decentralized market. Capital can be ranked using an industry measure of efficiency and projects can be ranked using an engineering measure of complexity. The quality of the match matters: more efficient capital is suited to drilling more complex projects and this is reflected in sorting patterns in the industry.\(^5\)

\(^5\)The fact that agents care about the quality of the match - and not just whether they are matched or not - is an important difference between my setting and recent work in Industrial Organization on search markets such as taxis (Frechette et al. (2017), Buchholz (2018)) and bulk shipping (Brancaccio et al. (2017)) where agents are
Therefore, in the offshore drilling industry, stronger sorting corresponds to more efficient rigs matched to more complex wells, and less efficient rigs to simpler wells, as illustrated in Figure 1.

I focus on shallow water oil and gas drilling in the Gulf of Mexico in 2000-2009. I begin by documenting two main findings. First, there is positive assortive matching: more efficient drilling rigs tend to drill more complex wells. Second, booms are associated with matching patterns consistent with stronger sorting. In a bust (when oil and gas prices are low) all rigs drill relatively similar types of wells. In a boom high-efficiency rigs tend to match to more complex wells and low-efficiency rigs tend to match to simpler wells.

Although the reduced-form findings are consistent with stronger sorting in booms, to fully assess mismatch I need to estimate the composition of searching projects (demand). For example, if only simple projects enter in a bust then it would be optimal to assign high-efficiency capital only to simple projects. Therefore, I provide an identification strategy to disentangle changes in the composition of searching projects from the sorting effect. The strategy relies on inverting observed matches through a flexible search technology and acceptance sets to identify the composition of searching projects.

Next I estimate a model of the industry. In the model there are searching agents on both sides of the market. On one side of the market there are drilling rigs (capital) which are differentiated by efficiency. On the other side of the market there are projects (wells that need to be drilled that are owned by oil and gas companies). The model is dynamic with a period length of one month. In booms the option value of searching for a better match increases. This increases the opportunity cost of being locked into a bad match. Agents respond by avoiding bad matches in two ways. First, they can reject bad matches. Second, using the search technology, they can direct their search away from bad matches. Overall these two channels result in stronger sorting patterns and reduce mismatch.

I estimate the model in three steps. First, I estimate state transitions based on observed empirical relatively homogeneous.

6 The model allows for the possibility that stronger sorting is optimal in busts. This is because the value of a match also increases in booms (since oil companies receive a higher price a given quantity of oil and gas). Therefore, if the value of a match increases faster than the option value of searching, stronger sorting may be optimal in busts. Whether stronger sorting in booms is optimal is ultimately an empirical question.
frequencies. The second step is to compute parameters that underpin the value of a match using contract data and empirical policy functions. I compute the remaining parameters in the third step using Simulated Method of Moments (SMM).

I use the estimated model to conduct counterfactuals. Welfare is measured in total profits. First, I quantify how the sorting effect improves efficiency. I start from a myopic ‘no sorting’ world where rigs accept all matches. Moving to the market benchmark (and allowing for the sorting effect) increases welfare by 11.4%. The sorting effect is cyclical with most of the gains in the boom. Decomposing the total effect highlights the main tradeoff in the model: compared to the myopic model, agents in the market tend to drill less wells but the matches are higher quality. Overall, the gains from better matching outweigh the costs of fewer matches resulting in a net increase in welfare.

Next, I quantify the first-best which is an intermediary who can eliminate search frictions and perfectly match capital and projects. In addition to highlighting the effects of search frictions, this counterfactual gives an upper bound on the gains from recent advances in e-procurement in the industry. I find that the intermediary would increase welfare by around 28.6% compared to the market benchmark. The intermediary’s allocation features stronger sorting than the market benchmark in both busts and booms.

Finally, I consider a demand smoothing policy which would decrease natural gas price shocks by one standard deviation. This kind of intervention has precedent in the oil and gas industry: many producer incentives, such as tax credits, and royalty rates, are tied to oil and gas prices. I find that demand smoothing would cause large shifts in capital utilization from booms to busts. However, I find that the costs of mismatch increase almost linearly with oil and gas prices, and so the policy would increase overall welfare by only 1.9%.

Overall, this paper makes three main contributions. The first contribution is a novel dataset of a decentralized capital market that is subject to booms and busts. A major difficulty in studying firm-to-firm markets is that contracts are typically confidential. By contrast, in this paper I construct a dataset of the universe of contracts in the industry matched with rich micro data from the regulator on the characteristics of projects undertaken under these contracts. My

---

7The potential of the internet to reduce search frictions in the industry has been discussed by practitioners since as early as 2002: Rothgerber (2002).

8A literature estimates these unobserved transfers in certain settings (e.g. Villas-Boas (2007)).
analysis of the dataset presents a unique and detailed picture of how firms make decisions when they are faced with fluctuations.

Second, previous work typically uses a steady-state analysis to tractably incorporate two-sided heterogeneity in a search model. When there are fluctuations, however, the distributions of agents change through time. In this paper I use an estimation strategy that incorporates - for the first time in a random search model with fluctuations - two sided heterogeneity, distributions of searching agents that change over time, and Nash bargaining. The estimation strategy relies on the observation that the value of searching can be written in terms of data on contract prices and the probability of matching. My strategy is an extension of approaches in the Industrial Organization firm dynamics literature such as Kalouptsidi (2014) to cases where short-term contract data are available.

The third contribution is to solve a data limitation that often occurs in capital markets: searching agents on one side of the market are not observed. I show how the distribution of searching agents, as well as a more flexible search technology, can be identified from matches in the data when there is two-sided heterogeneity. The flexible search technology - partially directed search - nests typical assumptions of random search or directed search as special cases.

**Related literature**  This paper is related to four strands of literature. First it is related to the literature on capital reallocation. Eisfeldt and Shi (2018) provide a review of this literature. Recent work, such as Lanteri (2018), has tried to uncover the mechanisms by which markets reallocate capital. Several papers show that search frictions can help to fit economy-wide facts about capital utilization and productivity (see for example Ottonello (2017) and Dong et al. (2018) who both calibrate models with search frictions). This paper advances this literature by - for the first time - providing empirical evidence of how search frictions affect the inner workings of a real-world capital market in booms and busts.

Second, this paper is related to the literature in Industrial Organization that studies empirical firm dynamics in decentralized markets. My model and application contain both fluctuations

---

9 An exception is Lise and Robin (2017), who model non-stationary distributions of searching agents by assuming Bertrand wage competition.

10 Lentz and Moen (2017) consider a related setup. My approach differs because I need to deal with two-sided heterogeneity and fluctuations, which pose challenges for estimation.
and two-sided heterogeneity. Some recent papers incorporate fluctuations into search models with homogeneous agents (for example, Buchholz (2018), Frechette et al. (2017)). A related set of papers study how fluctuations affect long-run firm entry and exit decisions (Kalouptsidi (2014), Collard-Wexler (2013)). Other recent papers estimate search and matching models with two-sided heterogeneity in a stationary context (e.g. Gavazza (2016)). By contrast, my paper contains both fluctuations and heterogeneous agents and I study how the two interact in a decentralized firm-to-firm market.

Third, this paper is related to the literature on search and matching models. In recent work Hagedorn et al. (2017) show how prices can be used to identify the value of a match in a stationary search context. Lise and Robin (2017) estimate a model of sorting between workers and firms with random search and productivity fluctuations. For tractability they assume that a worker (which would correspond to a rig in my setting) is offered their outside option for new matches, and show that this implies that the value of unemployment is independent of the arrival rate and distribution of future matches. My model nests the possibility that workers have no bargaining power, but allows prices of new matches to depend on match quality and lets unemployed workers (rigs) take into account the arrival rate and distribution of future matches when making decisions. Another novel feature of my paper is that I show how to identify a more general search technology - partially directed search - using data on observed matches.

Finally, this paper is related to the economics literature about the oil and gas industry. When modeling the industry I build on some of the institutional features discussed in Kellogg (2014), Kellogg (2011), Corts and Singh (2004), and Corts (2008). For credible estimation my empirical strategy relies on having a measure of participants’ expected value of undertaking a project. In the context of the Gulf of Mexico an excellent proxy is available: participants’ beliefs about the value of drilling a well is related directly to lease bids (Porter (1995)).

---

11 See Rogerson et al. (2005) for a survey

12 My dataset can be compared to data used in previous studies of the offshore oil and gas industry. For example, Corts and Singh (2004) use a dataset with a limited number of covariates (water depth and if the well was exploratory/developmental). I have access to a much richer set of well characteristics. Further, their data is aggregated at the monthly level - my dataset is at the contract level.
2 Industry Description and Data

2.1 Overview of the offshore drilling industry

Offshore drilling is an important part of the global oil and gas industry and was valued at $43 Billion USD in 2010 (Kaiser and Snyder (2013)). I analyze a particular segment of this industry: shallow water drilling in the US Gulf of Mexico. Shallow water drilling is defined as drilling in less that 500ft of water.

The offshore drilling industry is a decentralized industry. Lease holders such as BP and Chevron do not own the equipment used to drill their wells. In order to drill a well a drilling rig must be procured from a drilling contractor. Both sides of the industry are unconcentrated with an HHI of 980 for rig owners and an HHI of 330 for well owners. Given that the concentration of this industry does not seem high enough for individual firms to exert substantial market power I model the decision problem as a single agent playing against industry aggregates.

What is a drilling rig (capital)? Shallow wells are drilled using ‘jackup rigs’ and an example is pictured in Figure 2. Jackup rigs are barges fitted with long support legs that can be raised or lowered. In order to drill a well a jackup rig first moves to a well site. Upon arrival the rig then extends (‘jacks down’) its legs into the seabed for stability and commences drilling. The rig drills 24 hours a day until the well is completed. Once the well drilling is completed the well is connected to an undersea pipe where the oil and gas flows back to a refinery on land. The rig then ‘jacks up’ its legs, leaves the well site, and moves on to the next drilling job.

What is a well (a project)? Oil and gas producers own leases which are tracts of the seabed where they can drill a well to extract oil and gas. In this paper I use the terms drilling a ‘well’ and drilling a ‘lease’ interchangeably. Wells produce both oil and natural gas in different quantities. In the shallow water of the US Gulf of Mexico wells tend to contain more natural gas so I focus on changes in the gas price as the driver of exogeneous shocks in this industry. In the sample period the oil price is almost perfectly correlated with the natural gas price and so just using the natural gas price does not make any difference to the results. Once a well has been drilled

\[\text{I calculate the HHI with the definition of ‘market share’ as the proportion of total contracts.}\]
an operator extracts oil and gas at maximum capacity for the lifetime of the well (Anderson et al. (2018)) unless external factors such hurricanes or internal production problems intervene.

2.2 Data

Overview I construct a new and novel dataset by exploiting a number of rich, proprietary datasets of firm-to-firm contracts matched with the characteristics of wells drilled under each contract. Descriptive statistics for the industry are in Table 1. I focus on the subset of data for the years 2000-2009. The year 2000 is the earliest year for one of the contract datasets and so it is the earliest year I have a full picture of the industry. In 2010 the now infamous Deepwater Horizon oil spill triggered a new and tighter regulatory environment. Therefore I focus on the years before 2010.

Contract data The contract data come from two sources: IHS and Rigzone. The Rigzone dataset contains all offshore drilling contracts worldwide. The Rigzone dataset has detailed information on the status of rigs currently drilling and if they are not drilling whether they are available or off the market (‘cold stacked’ or scrapped). I use these data to compute how many
Table 1: Summary statistics for the dataset

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>10%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contracts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rig Price - New Contracts</td>
<td>1000s of USD/day</td>
<td>1766</td>
<td>54</td>
<td>34</td>
<td>23</td>
<td>100</td>
</tr>
<tr>
<td>Duration - New Contracts</td>
<td>Days</td>
<td>1766</td>
<td>77</td>
<td>79</td>
<td>40</td>
<td>145</td>
</tr>
<tr>
<td>Rig Price - Renegotiations</td>
<td>1000s of USD/day</td>
<td>984</td>
<td>47</td>
<td>26</td>
<td>24</td>
<td>75</td>
</tr>
<tr>
<td>Duration - Renegotiations</td>
<td>Days</td>
<td>984</td>
<td>78</td>
<td>73</td>
<td>31</td>
<td>147</td>
</tr>
<tr>
<td><strong>Wells</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value</td>
<td>Millions of USD</td>
<td>4475</td>
<td>8.4</td>
<td>21.5</td>
<td>0.2</td>
<td>22.6</td>
</tr>
<tr>
<td>Complexity</td>
<td>Index</td>
<td>4475</td>
<td>0.80</td>
<td>0.44</td>
<td>0.33</td>
<td>1.36</td>
</tr>
<tr>
<td>Water Depth</td>
<td>Feet</td>
<td>4475</td>
<td>125</td>
<td>85</td>
<td>37</td>
<td>245</td>
</tr>
<tr>
<td><strong>Rigs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly Utilization</td>
<td>% Rigs under contract</td>
<td>360</td>
<td>0.81</td>
<td>0.08</td>
<td>0.70</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Rigs are available at a point in time in the US Gulf of Mexico. I also have access to the Rigzone order book which contains information about the technological capabilities, ownership history, and age of each rig. The IHS contract dataset has slightly more detailed information on whether the contract is new or a renegotiation and so I merge this dataset with the well data.

Contracts follow a simple form: rig owners are paid a fixed price ‘dayrate’ for the length of the contract. Using this price data will be central to my estimation strategy. Contracts can differ in their length and I treat differences in the duration of contracts as one of the characteristics of a project. For example, a deep well will take longer to drill than a shallow well. A small number (10.6 percent) of contracts are ‘turnkey’ contracts which means that the rig operator, rather than the well owner, is responsible for additional costs if there are cost overruns such as a well blowout. Of these turnkey contracts 86.5 percent were drilled by a single operator (ADTI). The proportion of turnkey contracts in my sample is smaller than in Corts and Singh (2004), who study the industry in an earlier period (July 1998-October 2000). Therefore, due to the small number of turnkey contracts, and the fact that in my dataset their use is driven by a single operator, I do not model the choice of contract form explicitly as in Corts and Singh (2004).
Well data  The well data come from the Bureau of Safety and Environmental Enforcement (BSEE). The well permit data contain detailed information about the characteristics of each well including depth, location, mud weight, oil and gas produced, etc.

In addition I have lease bid data from which I can estimate participants’ beliefs about the value of drilling a well because it is related directly to lease bids (Porter (1995)). To do this I take the highest bid for the corresponding lease. In order to back out the quantity of hydrocarbons in the well, I then divide by average gas price ($5.71). My measure is a monotonic function of the expected oil and gas deposit size.

Measuring well heterogeneity  To rank wells I compute an engineering model of well complexity used in the industry called the ‘Mechanical Risk Index’. The Mechanical Risk Index takes well covariates including depth, mud weight, horizontal displacement etc that describe the geological environment and transforms them into a one dimensional index of well complexity. More complex wells (for example, a deep well that needs to bend around a difficult geological formation) are more costly to drill because there is a higher probability of encountering a problematic formation. Costs are typically in the form of extra materials when the rig encounters a problem. A higher ranking on the index corresponds to a more complex well that is more difficult to drill.

Measuring rig heterogeneity  A natural ranking for capital (drilling rigs) is their maximum drilling depth in water which ranges from 85 ft to 450 ft. This is a good proxy for many other characteristics of rig efficiency including age and technology. This ranking is also used in the industry and rig owners market rigs that can drill in deeper water as ‘high-specification’ rigs. Due to a limited sample size for the estimation I aggregate rigs into three classes: low, mid, and high efficiency rigs. These classes correspond to splitting the rig ranking into 3 quantiles.

---

14This is motivated by the fact that offshore lease auctions are common value auctions and as the number of bidders $n \to \infty$ the maximum bid converges to the expected value of oil and gas in the prospect. Although in practice the number of bidders is finite, see Haile et al. (2010) for evidence that ex-post returns in shallow water OCS auctions are not excessive.

15The motivation is that (as I show later in the paper) oil and gas prices are mean reverting and so the expected gas price when the lease is eventually drilled will be approximately the average gas price in the sample.

16Details on the calculation of the Mechanical Risk Index can be found in the Appendix.
The split classifies ‘low-efficiency’ rigs as those with a maximum drilling depth of \( \leq 200 \) feet, ‘mid-efficiency’ rigs as those with a maximum drilling depth of \( > 200 \) feet and \( < 300 \) feet, and ‘high-efficiency’ rigs as those with a maximum drilling depth of \( \geq 300 \) feet.\textsuperscript{17}

One might ask whether rigs are also differentiated by other factors. Two possible factors are: (i) the distance between a rig and a particular well, and (ii) past experience between a rig operator and a well owner. The first factor - distance - is unlikely to be an issue for within-field rig moves. Drilling rigs are extremely mobile and take around 1 day to move across the Gulf of Mexico. When compared to the average contract length (80 days), a back-of-the-envelope calculation implies that choosing a far away rig over a nearby rig would increase costs by around 1.25% for the average contract. Since most rig moves are within-field I do not include distance to a well as a factor for rig choice in the model.

The second factor - past experience between a rig operator and a well owner - has been shown to be a consideration for rig choice in the onshore oil and gas industry (Kellogg (2011)). I capture repeated contracting in my model by allowing for contract extensions. However, for new contracts, I assume that agents’ decisions about who to match with are independent of past experience. This modeling assumption seems to be supported by the data: I find 70 percent of new contracts are between a rig-owner pair who have not worked together in the previous 2 years.\textsuperscript{18}

2.3 Key features of the industry

The offshore drilling industry is characterized by three key features: (1) sorting patterns; (2) booms and busts driven by oil and gas prices; (3) search frictions.

Feature 1: Sorting patterns

Figure 3 illustrates the pattern of positive assortive matching in the data. It shows that better rigs tend to drill more complex wells on average. In addition I plot the 5% and 95% quantile of well complexity observed in the sample. The figure shows that although there is positive sorting,

\textsuperscript{17}The split is not quite exact because there are sometimes many rigs of exactly the same drilling depth.

\textsuperscript{18}I use 2 years as my cutoff for a ‘relationship’ because that is the definition used by Kellogg (2011).
Figure 3: Positive assortive matching: higher ranked rigs match with more complex wells

(a) Average match for each rig type

(b) Matching range (5% - 95%)

Note: Many rigs may share the same maximum drilling depth and so each point represents all rigs that share the same drilling depth rather than an individual rig.

there is not perfect segmentation in this industry: even the highest-ranked rigs still drill simple wells.

The observed sorting patterns imply that the match between rig technology and the well complexity matters. Qualitative evidence from the industry provides more detail about how agents make decisions about who to match with. For example, the website of Diamond Offshore, a rig owner, states: ‘Oil companies (“operators”) select rigs that are specifically suited for a particular job, because each rig and each well has its own specifications and the rig must be matched to the well’\textsuperscript{19}. Higher ranked rigs attract premium prices and are actively marketed as ‘high-specification’.
Figure 4: Natural gas price fluctuations drive rig prices

Note: The rig price/day is in US Dollars and is plotted using a local polynomial regression with a 95% confidence interval.

Figure 5: Matching patterns in booms and busts

(a) Average match for each rig type

(b) Distribution of well complexity

Note: In Panel (a), rigs are broken up into 3 quantiles based on the maximum drilling depth (recall that rig types are fixed and not dependent on whether the market is in a boom or a bust). A ‘boom’ is defined as an above average natural gas price (> $5.71) and ‘bust’ is a below average natural gas price. Panel (b) plots the density of well complexity in a boom and a bust.
Feature 2: Booms and busts

Figure 4 displays how fluctuations in the natural gas price affect rig prices in the industry. I assume that agents in this industry take the natural gas price as given which seems a reasonable assumption given that the output of each well owner is a small fraction of global production. Figure 4 shows that there is a strong correlation between gas prices and rig prices: rigs can command prices in excess of $100000 per day when gas prices are high but this can fall to $30000 per day when gas prices are low. Industry participants say that booms and busts are a key factor in how they make decisions about prices and utilization.

How booms and busts affect matching

Panel (a) of Figure 5 provides evidence consistent with stronger sorting in booms than busts. To produce the Figure I split the data up into two bins: a gas price above average which I label a ‘boom’ and a gas price below average which I label a ‘bust’. I then plot the average match in the raw data across 3 quantiles of rigs. Figure 5 shows a rotation in the average match line between rig rankings and well complexity rankings. Here, less efficient rigs are matched to simpler wells in booms than busts, and more efficient rigs are more likely to be matched to complex wells in booms than busts.

There are two possible explanations for the matching patterns in Panel (a) of Figure 5. One explanation is stronger sorting: capital is better matched in booms. However, since the distribution of searching wells is not observed, these patterns may also arise from changes in the composition of searching wells. Panel (b) of Figure 5 shows that the distribution of matches in booms and busts is very similar, which suggests that there is not a dramatic shift in the composition of searching wells in booms and busts. In addition, I allow my structural model to flexibly account for both changes in composition and changes in sorting, and see which combination of the two mechanisms best fits the data.

19http://www.diamondoffshore.com/offshore-drilling-basics/offshore-rig-basics
20According to the Energy Information Administration, total natural gas production is the Gulf of Mexico only accounts for around 5% of total production in the US: https://www.eia.gov/special/gulf_of_mexico
21From page 21 of the 2015 annual report of a rig owner (ENSCO): ‘The offshore drilling industry historically has been highly cyclical and it is not unusual for rigs to be unutilized or underutilized for significant periods of time and subsequently resume full or near full utilization when business cycles change’.
Feature 3: Search frictions

In order to drill a well, well owners contact drilling contractors with the particular specifications of each project. After successfully finding an available rig that is able to drill the project specifications a price is then individually negotiated. Search frictions arise because of the idiosyncratic nature of the projects and the fact that this industry is in constant flux: for example, individual rig availability is constantly shifting. Furthermore there are a large number of agents on both sides of the market.

Price dispersion  Next I show suggestive evidence for search frictions in the data by showing that different prices are paid for observationally equivalent matches. I regress prices on rig characteristics, well characteristics, and contract characteristics. I demean prices by the monthly average price. I run the following regression on new contracts:

\[ \hat{p}_{it} = X' \beta + \tilde{p}_{it} \]

Where \( \hat{p}_{it} \) are the demeaned prices for match \( i \) at month \( t \) and \( \tilde{p}_{it} \) are residual prices (that is, the residual after regressing prices on the covariates). I use the following covariates \( X \), as well as interactions between rig types and well characteristics:

\[ X = \{ \text{rig type, well complexity, well water depth, well value, gas price, rig utilization, month FE}s, \text{ year FE}s, \text{ contractor FE}s, \text{ rig owner FE}s, \text{ contract duration} \} \]

In Table 2 I report the unexplained variation \( 1 - R^2 \), the standard deviation of residual prices \( \tilde{p}_{it} \), and the standard deviation of all prices \( \hat{p}_{it} \). In panel (a) ‘rig-type’ is the aggregated classes (i.e. using \{high-spec, mid-spec, low-spec\}); in panel (b) ‘rig-type’ is the disaggregated rig classes (i.e. by maximum drilling depth).

Despite controlling for detailed match and contract characteristics Table 2 illustrates there is a high amount of unexplained price variation: 45% of price variation is unexplained when using the aggregated rig types and 40% of price variation is unexplained when using the finer disaggregated rig types. Similarly, the standard deviation of residual prices is 11 thousand USD/day when using aggregated rig types and 10 thousand USD/day when using disaggregated rig types. The high unexplained price variation in the data is consistent with a model of search frictions where the ‘law of one price’ does not hold.\(^{22}\)

---

\(^{22}\)One recent paper that documents a similar magnitude of price dispersion in a firm-to-firm search market is
Table 2: Evidence of price dispersion

<table>
<thead>
<tr>
<th></th>
<th>(a)</th>
<th>(b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using aggregated rig types</td>
<td>0.45</td>
<td>0.40</td>
</tr>
<tr>
<td>$SD(\hat{p}_{it})$</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>$SD(\hat{p}_{it})$</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

Note: Standard deviations are measured in thousands of US dollars per day.

3 The Model: Sequential Search with Booms and Busts

3.1 Environment

**Agents** Agents are capital owners (owners of rigs) and projects (potential wells). Project types are:

$$x = (x_{\text{complexity}}, x_{\text{value}}, x_{\text{water}}, \tau)$$

where $x_{\text{complexity}}$ is the complexity of the well, $x_{\text{value}}$ is the value of oil and gas, $x_{\text{water}}$ is the water depth of the well, and $\tau$ is the duration of the project in months.

There are $D_t = d_0 + d_1 g_t + d_t t$ potential projects in each period, where $g_t$ is the natural gas price, $t$ is a time trend, and $d_0, d_1, d_t$ are parameters. These potential projects are undrilled leases in the US Gulf of Mexico. Leases are tracts of the seabed owned by oil and gas companies where they can potentially drill a well. The dependence on $g_t$ and the time trend are used to flexibly capture the fact that the number of potential projects may be changing over time. For example, an increase in the gas price may induce drillers to revisit old prospects, or to be more likely to explore new tracts. Each of these potential projects has characteristics drawn from $x \sim \text{Lognormal}(\mu, \Sigma)$ where $\mu$ is a 1x4 vector and $\Sigma$ is a 4x4 covariance matrix. I choose the lognormal distribution in order to capture the heavy right tails of the project covariates.

Salz (2017). Using similar descriptive regressions that control for buyer and seller characteristics he documents an unexplained price variation $1 - R^2$ of 0.54 for non-brokered contracts in the trade waste industry.

---

23 As previously discussed, I only use the natural gas price to track if the market is in a boom or bust (rather than the oil price). The two prices are almost perfectly correlated in my sample and so just using the natural gas price does not make any difference to the results.
Capital differs in its efficiency and the types are \( y \in \{\text{low}, \text{mid}, \text{high}\} \). Capital is either available to match or under contract. Only available capital can match with a project.

**Timing** The model is dynamic and one period in the model is one month. To keep notation concise, I let the subscript \( t \) represent the state \( s_t \). Within each period the timing is:

1. **Entry.** Each of the potential projects \( D_t \) draws from \( \text{Lognormal}(\mu, \Sigma) \) chooses whether to enter and search for capital. This results in a distribution of searching projects and the probability density is denoted by \( f_t(x) \).

2. **Search.** Searching projects contact available capital.

3. **Matching.** If a project owner contacts a capital owner then agents choose whether to match. Prices are determined by Nash bargaining. If capital is matched then it cannot match for the duration of the contract (\( \tau \) periods). If agents choose to not match then projects exit the market immediately and the capital is available to match next period.\(^{24}\)

4. **Contract extensions.** Existing matches are (possibly) extended.

**Value of a match** I use the following specification for the value of a match:

\[
 \left( \sum_{k=0}^{\tau-1} \beta^k m(x, y) + \beta^\tau \mathbb{E}_t [g_{t+\tau} x_{\text{value}}] \right) \mathbb{I}[x_{\text{water}} < y_{\text{water}}]
\]

Overall, the match value added can be interpreted as a cost function. The match value added can be broken down into three components. First, there is the term \( \mathbb{I}[x_{\text{water}} < y_{\text{water}}] \). This ensures that rigs only drill wells that are below their maximum drilling depth \( y_{\text{water}} \).

The second component is \( m(x, y) = m^y_0 + m^y_1 x_{\text{complexity}} \), which is the match value added. The term \( m(x, y) \) is defined on a per-period basis and so is summed over the entire \( \tau \) length contract.

\(^{24}\)Previous literature has modeled the decision to drill a well as a real options problem: oil and gas companies are waiting for the gas price to cross a threshold that makes drilling profitable (Kellogg (2014)), or they are waiting to see if the drilling of neighboring leases reveals good information about a project (Hendricks and Kovenock (1989), Hendricks and Porter (1996)). However, explicitly modeling the dynamic decision of each driller to drill over time, as well as the arrival and exit of potential wells, would be a complicated problem to compute and estimate. Instead I model the decision to drill a well as a static entry decision problem that occurs in each period.
Here $m_0^y$ and $m_1^y$ are coefficients that vary with rig type $y \in \{l, m, h\}$. Importantly, this equation captures complementaries between rig type and well type through $m_1^y$. For example, if high-efficiency capital is well-suited to undertaking complex projects then $m_1^h$ will be high. These parameters will determine how beneficial positive sorting is for welfare. For example, in a static setting with no search frictions, positive sorting is the optimal allocation if the match value is supermodular ($m_1^h > m_1^m > m_1^l$) (Becker (1973)).

The third component is $\beta \tau \mathbb{E}_t [g_{t+\tau} x_{\text{value}}]$. After the contract (which is $\tau$ periods long) has elapsed the well will be completed and will produce oil and gas. The expected total value of this oil and gas is $\beta \tau \mathbb{E}_t [g_{t+\tau} x_{\text{value}}]$. Here the covariate $x_{\text{value}}$ is the quantity of hydrocarbons in the well and $g_{t+\tau}$ is the gas price. Therefore the dollar value of oil and gas in the well is $g_{t+\tau} x_{\text{value}}$.

**Summary** Agents make two main choices in the model (with the rest of the parts determined endogenously in equilibrium). The first choice is the project entry decision, which determines the distribution of entered projects in each period $f_t(x)$. The second choice is whether to match if agents successfully contact each other (and whether to extend a contract when an existing match ends). Overall the model focuses on the dynamic tradeoff for capital owners.

### 3.2 Demand for capital: Project entry and search

The focus of this section is to characterize demand for capital. Demand for capital at each state $s_t$ is a distribution of *trading opportunities* $q_t(x|y)$, where $q_t(x|y)$ is the probability that a type $x$ project will contact type $y$ capital.\footnote{I denote the probability that capital will not be contacted by $q_t(\emptyset|y)$.} Demand for capital results from project entry and search behavior.

**Payoffs**

First I consider the profits to a project of type $x$ matching with capital of type $y$. Intuitively, the profit will depend on the per-period match value and the per-period capital price. In addition, because contracts can be extended, agents will take these future contract extensions into account as well when matching. Overall, the value from matching with a particular rig $y$
The project owner’s value of matching $\Pi_t^*(x, y)$ can be decomposed in the following way. For each period of the $\tau$ length contract the project owner receives the match value added $m(x, y)$ minus the price $p_t(x, y)$ to hire the capital. When the initial contract has ended the well will produce oil and gas. The contract will be extended with probability $e_t(x, y)$ where:

$$e_t(x, y) = \eta 1[S_t(x, y) > 0]$$

Here $\eta$ is the mean of a Bernoulli random variable that is a match-specific shock. The $\eta$ term captures the decision to renew the contract in a reduced form way. For example, if drilling reveals good information then agents may choose to renew the contract. If the total surplus of a match $S_t(x, y)$ is negative - for example if the gas price falls and the project is no longer profitable to drill - then the contract will not be extended.\(^26\)

**Step 1: Entry**

A potential project with characteristics $x$ will enter and search if the expected benefits are higher than the entry cost:

$$\sum_y \omega_t(y|x) \max\{\pi_t(y|x), 0\} \geq c$$

where $c$ is the entry cost. The entry cost $c$ takes into account the cost of submitting a permit (which includes a detailed project design) to the regulator, amongst other things. The left-hand-side is the expected benefit of entering. The term $\omega_t(y|x)$ is the probability that a type $x$ project searches for type $y$ capital. The term $\pi_t(y|x)$ is the expected value of searching for type $y$ capital. I formally define these two terms in Equations (4) and (5), respectively. The expression $\max\{\pi(y|x), 0\}$ takes into account that the project may not match if the expected surplus is negative, and instead will receive a payoff of 0.\(^27\)

\(^{26}\)I will formally define surplus in Equation 10.

\(^{27}\)Recall that projects exit immediately if unmatched and therefore will receive a payoff of 0.
Figure 6: Matching within each submarket

Searching projects $x \sim f_t(x)$ target capital $y$ with prob: $\omega_t(y|x)$

Match if: $\mathbb{I}[S_t(x,y) > 0]$
project's payoff : $\Pi^y_t(x,y)$
capital's payoff : $\Pi^e_t(x,y)$

Remaining projects exit

Unemployed type $y$ capital
Unemployed type $y$ capital

# capital : $n_{yt}$
# projects : $n^e_{yt}$
market tightness : $\theta_{yt}$
prob. project meets : $\alpha_0 \theta_{yt}^{\alpha_1}$

Notes: This figure illustrates how capital and projects match. At the beginning of each period there is a distribution of searching projects $f_t(x)$ and available capital $n_{yt}$. The searching projects first choose which type of capital to target. Meetings are determined randomly within each submarket and are dependent on the market tightness $\theta_{yt}$. Finally, agents choose whether to match.

Step 2: (Partially directed) search

Each type of capital $y \in \{l,m,h\}$ is located in a submarket. Figure 6 provides a diagram of the meeting process. Meetings are determined randomly in each submarket. Therefore the expected value of searching in submarket $y$ is:

$$\pi_t(y|x) = \alpha_0 \theta_{yt}^{\alpha_1} \Pi^e_t(x,y)$$

(4)

The probability of successfully contacting capital in submarket $y$ is: $\alpha_0 \theta_{yt}^{\alpha_1}$. Here $\alpha_0, \alpha_1$ are meeting efficiency parameters. The term $\theta_{yt}$ is the equilibrium submarket tightness for rig type $y$ at time $t$ (the number of available capital of type $y$ divided by the number of projects that target type $y$ capital). For a fixed number of searching projects, more type $y$ capital implies that each project in the submarket will be more likely to meet capital. Note that submarket tightness is not directly observed in the data (since I do not observe searching projects) and therefore will need to be computed as an equilibrium object.
Potential projects need to choose which submarket to search in. The choice of submarket depends on the characteristics of the project and the search technology. I allow for a flexible search technology - partially directed search - where the probability of targeting a submarket is given by:

\[
\omega_t(y|x) = \frac{n_{yt} \exp \left( \gamma \pi_t(y|x) \right)}{\sum_{k \in \{l,m,h\}} n_{kt} \exp \left( \gamma \pi_t(k|x) \right)}
\]  

(5)

Here \(n_{yt}\) is the number of available rigs of type \(y\) at time \(t\). The term \(\gamma\) is a ‘targeting parameter’ that indexes how precisely a project can target capital. Higher values of \(\gamma\) imply that projects are more likely to enter the submarket that contains the optimal match. This search technology is more flexible than typical assumptions of random search or directed search which are used in search models. At the extremes this specification nests random search (where projects contact capital completely at random) and directed search (where projects can perfectly identify the best match):

**Lemma 1.** The targeting parameter \(\gamma\) nests random search at \(\gamma = 0\) and directed search at \(\gamma \to \infty\).

At the extreme case of random search \(\gamma = 0\), the targeting weight \(\omega_t(y|x) = s_y\), where \(s_y\) is the market share of type \(y\) capital. In other words, under random search the probability of contacting a type \(y\) rig is just its market share. At the other extreme case of directed search, \(\omega_t(y|x) = 1\) if \(y = \arg\max_y \{\pi_t(y|x)\}\) and \(\omega_t(y|x) = 0\) otherwise. That is, a project targets its search towards its best match. The setup is similar to Lentz and Moen (2017), who show partially directed search can be identified in a labor market application with homogeneous workers, a steady state, and data on observed searching worker and firm transitions. By contrast, my application has heterogeneity on both sides of the market, fluctuations, and searching projects are not observed, which poses challenges for estimation.

**Summary: Characterizing demand for capital**

Putting it all together, the decision to enter and the search technology characterize demand for capital at each state \(s_t\).

---

\(^{28}\)Recall the simplified notation that a subscript \(t\) implies the object is dependent on state \(s_t\).
Proposition 1. Demand for capital type $y$ at state $s_t$ is given by:

$$q_t(x|y) = \alpha_0^{\theta_{yt}^{-1}} \frac{\omega_t(y|x)f_t(x)}{\int_z \omega_t(y|z)f_t(z)dz}$$  \hspace{1cm} (6)$$

$$q_t(\emptyset|y) = 1 - \alpha_0^{\theta_{yt}^{-1}}$$  \hspace{1cm} (7)$$

where $q_t(x|y)$ is the probability that capital type $y$ will be contacted by project type $x$, $q_t(\emptyset|y)$ is the probability that capital type $y$ is not contacted by any project. The distribution of entered projects $f_t(x)$, targeting weights $\omega_t(y|x)$, and submarket tightness $\theta_{yt}$, are determined in equilibrium by Equations (1) - (5).

Proposition 1 allows for considerable flexibility in how demand changes in booms and busts along two dimensions. First, the probability of capital finding a project will increase when the market moves from a bust to a boom. This occurs because the number of potential projects increases (since an increase in the gas price increases the number of draws $D_t = d_0 + d_1 g_t + d_t t$ and I find when estimating the model that $d_1 > 0$) and also because more potential projects enter due to an increase in the value of oil and gas they contain. Second, the distribution of trading opportunities $q_t(x|y)$ will change due to different projects entering and different targeting behavior. For example, projects with lower quantities of oil and gas $x_{value}$ may enter in booms.

Given demand for capital, I now turn to the capital owners’ problem.

### 3.3 Capital owners’ problem: Choosing whether to match

In this section I characterize how matches are formed which is pinned down by the capital owners’ problem. The capital owners’ problem is to choose whether to match and occurs in Step 3 after projects have entered and contacted capital owners. Capital owners are forward looking so that given the current state the capital owner can also forecast future trading opportunities $\{q_{t+k}(x|y)\}_{k \geq 1}$.

If capital is contacted by a project it faces the following tradeoff. 

\begin{equation*}
\begin{cases}
\text{Accept match} & \Pi_t^y(x, y) + 0 + \beta E_t V_{t+1}^y(y) \\
\text{Search again} & \end{cases}
\end{equation*}
Here $\Pi_t^y(x,y)$ is the profit from matching. The value of searching again for a better match is $0 + \beta \mathbb{E}_{t} V_{t+1}^y(y)$: the capital is not used for 1 period (which results in a payoff = 0) but can search again in period $t+1$ because it is not locked into a contract. The profit from matching is:

$$\Pi_t^y(x,y) = \sum_{k=0}^{\tau-1} \beta^k p_t(x,y) + \beta^\tau \mathbb{E}_t \left[ e_{t+\tau}(x,y) \Pi_{t+\tau}^y(x,y) + (1 - e_{t+\tau}(x,y)) V_{t+\tau}^y(y) \right]$$

The profit from matching $\Pi_t^y(x,y)$ can be decomposed as follows. The rig will first receive the value of the contract, which is the per period price $p_t(x,y)$ for $\tau$ periods. When the contract is complete the rig owner receives $\Pi_{t+\tau}^y(x,y)$ if the contract is extended. If the contract is not extended then the rig will be available to search again and will receive $V_{t+\tau}^y(y)$.

The expression for the profits from matching for capital implicitly sets costs borne by the capital owner to zero. Although the rig does incur operational expenses each period such as labor costs, these are not dependent on whether the rig is drilling. Since the cost is incurred regardless of whether the rig is drilling it is not a factor in agents’ decisions and so without loss of generality I set it to zero.

Given the state $s_t$, the value of searching before at the start of a period before projects contact capital is:

$$V_t^y(y) = \int x q_t(x|y) \max \left\{ \Pi_t^y(x,y), \beta \mathbb{E}_t V_{t+1}^y(y) \right\} dx + q_t(\emptyset|y) \beta \mathbb{E}_t V_{t+1}^y(y)$$

The first term is the expected value of a meeting: capital meets a particular project type with probability $q_t(x|y)$ and it will choose whether or not to match with it. If capital is not contacted by a project (which happens with probability $q_t(\emptyset|y)$) then it will be unemployed for one period but will be available the following period.

[29] For example, the 2005 annual report for Diamond Offshore, a rig owner, states: "Operating expenses generally are not affected by changes in dayrates and may not be significantly affected by short-term fluctuations in utilization. For instance, if a rig is to be idle for a short period of time, few decreases in operating expenses may actually occur since the rig is typically maintained in a prepared or ready-stacked state with a full crew."
Bargaining

If capital and a project match then prices are determined by generalized Nash bargaining where \( \delta \in [0, 1] \) is the bargaining weight:

\[
p_t = \operatorname{argmax}_{p_t} \left[ \Pi^y_t(x, y) - \beta \mathbb{E}_t V^y_{t+1}(y) \right]^\delta \left[ \Pi^x_t(x, y) \right]^{1-\delta}
\]

Note that prices \( p_t \) are embedded in the value of matching for capital \( \Pi^y_t(x, y) \) and projects \( \Pi^x_t(x, y) \). The outside option for the capital is to search again the following period for another match \( \beta \mathbb{E}_t V^y_{t+1}(y) \). Since the project will exit immediately if it is not matched, the project’s outside option is 0.

Summary: Characterizing matching

I assume Nash bargaining which implies transferable utility. Under this assumption there is a simple characterization of the decision to accept or reject a match:

**Proposition 2.** Under Nash bargaining, agents’ decision to accept or reject a match is whether the total surplus is positive:

\[
A_t(y) = \left\{ x : S_t(x, y) \geq 0 \right\}
\]

Here \( A_t(y) \) is the acceptance set - all \( x \) where a match with \( y \) will be accepted. The total surplus is:

\[
S_t(x, y) = \Pi^y_t(x, y) + \Pi^x_t(x, y) - \beta \mathbb{E}_t V^y_{t+1}(y)
\]

The above proposition says that the decision to accept or reject a match is mutual and dependent on whether the total surplus of a match is positive.

### 3.4 Transitions and states

**Transitions**

Denote the number of periods left on a contract as \( k \). If a contract is not extended then the number of periods remaining on the contract counts down by 1. Unemployed rigs who do not
find a new match remain unemployed with $k = 0$. Matches that will expire next period ($k = 1$) are possibly extended. Rigs that are available next period ($k = 0, k = 1$) will possibly find a new match.

**States**

The detailed industry state in each period is the price in dollars for natural gas $g_t$, the distribution of current matches, and the distribution of unemployed rigs. Modeling firms as keeping track of the entire industry state would be computationally difficult due to the curse of dimensionality. I assume instead that firms keep track of their own state and some moments of the industry state. This is similar to a moment-based Markov Equilibrium (Ifrach and Weintraub (2017)). I assume these moments that characterize an agent’s beliefs about state $s_t$ are:

$$s_t = [g_t, n_{lt}, n_{mt}, n_{ht}]$$

Here $n_{y,t}$ is the number of available rigs of type $y$ at time $t$, and $g_t$ is the natural gas price at time $t$. A rig is ‘available’ to match if it is either unemployed ($k = 0$) or it is in the final period of a contract and can match in the following period ($k = 1$).

I model agents’ beliefs about equilibrium industry state transitions as an AR(1) process:

$$s_t = R_0 + R_1 s_{t-1} + \epsilon_t$$

I assume that rig transitions are deterministic so the only stochastic component to the model is the gas price error term, which implies that $\Sigma = \text{Diag}(\sigma_\epsilon, 0, 0, 0)$. I write the elements of $R_0, R_1$ as:

$$R_0 = \begin{bmatrix} r^0_g \\ r^0_l \\ r^0_m \\ r^0_h \end{bmatrix}, \quad R_1 = \begin{bmatrix} r^1_g & 0 & 0 & 0 \\ r^1_l & r^2_l & r^3_l & r^4_l \\ r^1_m & r^2_m & r^3_m & r^4_m \\ r^1_h & r^2_h & r^3_h & r^4_h \end{bmatrix}$$

In the matrix $R_1$ I set the coefficients $r^2_g = r^3_g = r^4_g = 0$. That is, while changes in the natural gas price cause changes in rig availability in the Gulf of Mexico, rig availability in the Gulf of Mexico does not affect the global natural gas price. This assumption seems reasonable given that total natural gas production in the Gulf of Mexico is a small fraction of global natural gas production.
Discussion of state choice  I need to choose which moments of the industry state agents keep track of. I choose the natural gas price and rig availability because these statistics are commonly reported in the annual reports of rig owners and are used by firms who track the industry to describe the state of the market. Drillers respond to an increase in the natural gas price by drilling more projects. Rig availability falls after a sustained increase in gas prices which means that agents differentiate between a short term increase in natural gas prices (high gas prices and high rig availability) versus a long-term increase in gas prices (high gas prices and low rig availability).

I experiment with including natural gas futures prices but over the 2000-2009 sample these prices are not statistically significant or economically significant, once first order lags of the natural gas price are taken into account. I also experiment with including second order lags of the state variables but these are also not significant once first order lags are included.

3.5 Equilibrium

Definition 1. Equilibrium is a set of prices $p_t(x,y)$, capital availability $(n_{lt}, n_{mt}, n_{ht})$, demand for capital $q_t(x|y)$, targeting weights $\omega_t(y|x)$, a distribution of searching projects $f_t(x)$, submarket tightness $(\theta_{lt}, \theta_{mt}, \theta_{ht})$, and agents’ state transition beliefs, that satisfy at each state $s_t$:

1. The distribution of searching projects $f_t(x)$, the targeting weights $\omega_t(y|x)$, and submarket tightness $(\theta_{lt}, \theta_{mt}, \theta_{ht})$, determined by Equations (1) - (5)

2. Demand for capital $q_t(x|y)$ determined by Equations (6) - (7)

3. Optimal extensions of contracts satisfying Equation (2)

4. Equilibrium prices $p_t(x,y)$ determined by Nash bargaining: Equation (8)

5. Agents optimally choose whether to accept/wait if matched using Equation (9)

6. Updating rule for the distribution of capital $(n_{lt}, n_{mt}, n_{ht})$

7. Beliefs about the future evolution of states given by Equation (12)
4 Estimation Strategy

Overview of the estimation

I calibrate the monthly discount factor as $\beta = 0.99$. There are four sets of parameters that I need to estimate:

$$\lambda = \{r, \lambda_1, \lambda_2, \lambda_3\}$$

These parameters are: the state transitions $r = \{R_0, R_1, \sigma\}$, the parameters that characterize the profits from match $\lambda_1 = \{m, \delta, \eta\}$, the parameters that characterize entered projects $\lambda_2 = \{\mu, \Sigma, c, \gamma\}$, and the parameters that characterize the meeting technology and number of draws $\lambda_3 = \{\alpha, d\}$. I estimate these parameters in three steps.

1. Estimate the state transitions $r$ using maximum likelihood.
2. Estimate the parameters $\lambda_1$ that govern the value of a match between capital and projects.
3. Estimate the parameters that characterize entered projects $\lambda_2$ and the parameters $\lambda_3$ using the method of simulated moments.

Step 1 is straightforward so I only discuss step 2 and step 3 in detail.

Summary of the challenges and contributions There are two main challenges and contributions in the estimation strategy. The first challenge is that value functions - particularly the value of searching $V_t^y(y)$ - are complicated to solve. The complexity comes from the fact that there are distributions of agents on both sides of the market and these distributions are changing through time. Previous work avoids this complexity typically by using a steady state analysis (an exception is Lise and Robin (2017)). In this paper I incorporate rich and complex dynamics by using prices to non-parametrically estimate the value function for searching. My strategy is an extension of approaches in the Industrial Organization firm dynamics literature such as Kalouptsidi (2014) to cases where short-term contract data are available.\(^\text{30}\)

\(^{30}\)Kalouptsidi (2014) uses data on second-hand sales to estimate value functions with the observation that the resale price of a ship equals the value of a ship.
The second contribution is to show that both a more flexible meeting technology (partially directed search) and the unobserved distribution of searching projects can be identified from observed matches.

**Step 2: Estimating the value of a match**

Step 1 is straightforward so I begin by discussing step 2 in detail. In this section I describe how to estimate the parameters which characterize the value of a match. Overall, I use the Nash bargaining solution combined with forward simulation of the value functions. There are 8 parameters to estimate: the bargaining parameter \( \{\delta, m^y_0, m^y_1\} \) for \( y \in \{l, m, h\} \), and the extension parameter \( \eta \). I first estimate \( \{\delta, m^y_0, m^y_1\} \) and then back out \( \eta \) from extension probabilities.

**Estimating the bargaining weights and the match value added**

Writing out the Nash bargaining problem:

\[
p_t = \arg\max_{p_t} \left[ \Pi_t^y(x, y, p_t | \lambda) - \beta \mathbb{E}_t V_{t+1}^y(y | \lambda) \right]^{\delta} \left[ \Pi_t^x(x, y, p_t | \lambda) \right]^{1-\delta}
\]

(14)

Recovering \( \lambda_1 \) from the above equation appears challenging. In particular, prices are dependent on the outside option of searching for a better match, \( \mathbb{E}_t V_{t+1}^y(y | \lambda) \), which is in turn dependent on the distribution of searching projects \( f_t(x) \) which is not known at this stage.

I simplify this by observing that the value of searching can be written in terms of data on matches, data on prices, and data on the probability of extending a contract. The intuition is as follows. From the point of view of a capital owner, next period they may be matched with some type of project \( y \). The distribution of these matches is observed directly in the data. If the capital is matched then it receives a price (which is observed in the data), and then the contract may be extended (with a probability observed in the data). If the rig is not matched then the rig searches again and the state is updated. Therefore, agents payoffs can be constructed just using the data. One notable feature of this approach is that it only requires data on observed matches that were accepted. It does not require data on matches that were rejected (agents get 0 payoff regardless of whether they chose to reject a project or fail to meet a project). Nor does it require data on the composition of searching projects.
The following Proposition formalizes the above intuition (where I leave the proof to the Appendix).

**Proposition 3.** The value of search $V_t^y(y|\lambda)$ can be computed directly from the data. Furthermore, the term $\Pi_t^y(x,y,p_t|\lambda)$ can be computed from the data and the term $\Pi_t^y(x, y, p_t|\lambda)$ can be computed from the data and the parameters $\{m_0^y, m_1^y\}$.

Proposition 3 uses data in the form of policy functions. These policy functions are prices $\hat{p}_t(x, y)$, the pdf of matches for each $y$: $\hat{f}_t(x|y)$, and the probability of extending a match $\hat{e}_t(x, y)$. All of these policy functions are observed directly in the data. Using Proposition 3 the Nash bargaining problem is only dependent on the parameters $\{\delta, m_0^y, m_1^y\}$. Taking the first order condition and rearranging results in the following Lemma.

**Lemma 2.** The Nash bargaining solution can be written as:

$$\tilde{p}_t(x, y) = \delta \left[ m_0^y + m_1^y x_{\text{complexity}} + g_t(x, y) \right]$$

(15)

Where $\tilde{p}_t(x, y)$ are ‘adjusted prices’ and $g_t(x, y)$ is a known function of $x$ and $y$ (the exact functions are in the Appendix).

I leave the proof of Lemma 2, as well as the exact form of $g_t(x, y)$ and $\tilde{p}_t(x, y)$, to the Appendix. Intuitively, adjusted prices $\tilde{p}_t(x, y)$ take into account the value of the current contract and also a discount for the benefits of future contract extensions. The term $g_t(x, y)$ captures how different values of oil and gas pass through to the price of a contract. Putting it all together, I use the following steps to compute $\{\delta, m_0^y, m_1^y\}$:

1. Estimate the policy functions.
2. Compute the value of search $V_t^y(y|\lambda)$ using forward simulation.
3. For each contract observed in the data compute $\tilde{p}_t(x, y)$ and $g_t(x, y)$ using Lemma 2. Then use non-linear least squares and Equation 15 to estimate $\{\delta, m_0^y, m_1^y\}$.

In the Appendix I provide more details about how I construct the policy functions in and the forward simulation algorithm.
Identification Identification follows from Lemma 2. First, the bargaining parameter can be identified by variation in matches with different $x_{value}$ within a period. Intuitively, the bargaining parameter indexes how changes in $x_{value}$ pass through to prices. For example, consider two projects $x, x'$ with different $x_{value}$ where $g_t(x', y) > g_t(x, y)$. If $\delta$ is high then capital has higher bargaining power and prices will increase by more when moving from $x$ to $x'$ than if $\delta$ is low. Once $\delta$ is known, $m^y_1$ and $m^y_2$ are identified by variation in prices for matches with different $x_{complexity}$.

Estimating the extension parameter

The extension parameter $\eta$ can now be estimated. First, note the following Lemma, which holds because all components of the match surplus $S_t(x, y)$ are observable or can be estimated as a policy function:

**Lemma 3.** The match surplus $S_t(x, y)$ can be constructed using $\{m^y_0, m^y_1\}$ and Proposition 3.

The extension parameter $\eta$ is identified by the average probability of extending a match averaged over the empirical distribution of matches:

$$P(extend) = \eta \times \left[ \frac{\sum_{x,y,t} 1[S_t(x,y) > 0]}{N} \right]$$

Here $P(extend)$ is the average extension probability over the sample and $N$ is the total number of observations. Note that this is not conditional on the capital or project type. The right-hand side is the match specific shock, which we are interested in estimating, multiplied by an ‘average positive surplus’ term that is computed by simulating the surplus for each match observed in the data. Intuitively, $\eta$ is identified by matching the average extension probability in the data to the extension probability predicted by the model.

Step 3: Estimating the remaining parameters

The remaining parameters to compute in step 3 are the parameters that characterize the distribution of searching projects $\lambda_2 = \{\mu, \Sigma, c, \gamma\}$, and the parameters that characterize the meeting technology and number of potential project draws $\lambda_3 = \{\alpha, d\}$. I identify these parameters using two sets of moments. The first set are moments related to the distribution of observed matches.
and I use these to identify $\lambda_2$. The second set are moments related to the evolution of capital availability and I use these to identify $\lambda_3$.

I first reduce the number of parameters I need to estimate by assuming that the off-diagonal elements in $\Sigma$ are 0.\footnote{The model fits the data well given this assumption, as shown in Figure 9 in Section 5.} I can directly calibrate $\{\mu_{\text{value}}, \mu_{\text{water}}, \sigma_{\text{value}}, \sigma_{\text{water}}\}$ from the lease data: potential projects are undrilled leases and the lease data contain information about the value of a project (through the winning bid) and the water depth. However, the lease data do not contain information about the complexity of the well $x_{\text{complexity}}$ or the duration of the contract $\tau$. Even if I knew these parameters I do not observe which potential projects chose to enter and search (which is governed by the entry cost $c$).

Identification

I first discuss identification of each of these parameters from data on the observed distribution of matches and capital utilization. I then use these identification arguments to motivate the inclusion of particular moments of the data in the estimation.

Identifying the targeting parameter Identification of the targeting parameter is challenging because searching projects $f_t(x)$ are not observed. To fix ideas, consider the following example. Suppose that an econometrician observes many matches between high efficiency capital and complex projects. Should the econometrician conclude that there are many complex projects searching? Or is it evidence that agents are good at targeting the match they are best suited to?\footnote{As I confirm in the estimation, complex projects are best suited to matching with high-efficiency capital.} The answer will affect the model primitives and therefore welfare: if an econometrician attributes these matches to a good targeting technology then the market may seem quite efficient. If the matches are attributed to the distribution of searching wells (rather than the targeting technology), the market may appear inefficient.

I show that the targeting parameter can be identified from observed matches. For intuition, consider Figure 7. Both panels show the observed distributions of matches within a period conditional on a capital type $y$: $\hat{f}_t(x|y)$. Under random search where $\gamma = 0$ (the right panel) the probability that a project meets a particular type of capital is not dependent on project...
Figure 7: Identification of the targeting parameter $\gamma$

(a) Matches, $\gamma > 0$

(b) Matches, $\gamma = 0$

Note: This figure gives intuition about how the targeting parameter $\gamma$ can be identified. Under partially directed search (Panel (a), where $\gamma > 0$) the probability of observing some type of project $x$ will depend on the surplus of the match and so the distribution of observed matches will be different ($\tilde{f}_t(x|y) \neq \tilde{f}_t(x|y')$) and proportional to the targeting weights. Under random search (Panel (b), where $\gamma = 0$), the distribution of observed matches will be the same and not dependent on $x$.

Therefore, for a given project type $x$, different types of capital $y, y'$ should have the same probability of matching: $\tilde{f}_t(x|y) = \tilde{f}_t(x|y')$. Under partially directed search (the left panel), the probability that a particular project matches with a type of capital is now dependent on the project type $x$. Therefore if I pick some $x$ the probability of matching with different types of capital may not be the same: $\tilde{f}_t(x|y) \neq \tilde{f}_t(x|y')$.

The above intuition relies on observing similar projects matching with different types of capital. When agents reject bad matches, as is the case in my application, I will only observe a project $x$ matching capital $y$ if the project is within the acceptance set of capital $y$. Therefore, to compare matching probabilities across capital, I need to assume that the acceptance sets overlap. I formalize this idea in the following Proposition:

**Proposition 4.** Assume there is a region where the acceptance sets overlap so that there are three project types $x, x', x'' \in \bigcap_y A_t(y)$. Then the targeting parameter $\gamma$ is identified from the distribution of observed matches $\tilde{f}_t(x|y)$ for $y \in \{l, m, h\}$.

Since the acceptance sets can be constructed independently in Part 2 of the estimation, I can test that the ‘overlapping acceptance set’ assumption holds. I leave the proof of Proposition 4 to the Appendix but discuss the intuition here. Denote the empirical probability of observing a match involving a project $x$ conditional on capital type $y$ at time $t$ by $\tilde{f}_t(x|y)$. Comparing the
probability of the same type of project $x$ matching different capital $y, y'$ at $t$:

$$\ln \left( \tilde{f}_t(x|y) \right) - \ln \left( \tilde{f}_t(x|y') \right) = \gamma \left( \pi_t(x|y) - \pi_t(x|y') \right) + \ln \left( \theta_t^{\alpha_1-1} / \theta_t'^{\alpha_1-1} \right)$$

(16)

The above equation says that the targeting parameter $\gamma$ indexes how sensitive the probability of matching is to the expected value of matching $\pi_t(x|y)$.

In the extreme case of random search, $\gamma = 0$, differences in the expected value of matching between rigs have no effect on the probability of matching. In addition, under random search, the RHS = 0 and all types of capital match will match a given project with the same probability. Alternatively, for $\gamma > 0$, the probability of matching is now sensitive to the expected value of matching $\pi_t(x|y)$.

**Identifying the distribution of potential projects and the entry cost** Once the targeting parameter has been pinned down, the distribution of searching projects $f_t(x)$ can be identified by inverting observed matches through the search technology. Since $f_t(x)$ is determined by an entry condition, the entry cost $c$ can be identified by differences in the value of projects that enter at different points in time. For example, projects with low quantities of hydrocarbons might not enter in a bust, but may be willing to pay the entry cost in a boom when the gas price is high. Once the entry cost is identified, the distribution of potential projects $(\mu, \Sigma)$ can be identified since I can map potential projects directly into observed matches through the entry decision and the search technology.

**Identifying the remaining parameters $\lambda_3$** The remaining parameters are those that characterize the meeting technology and number of potential project draws $\lambda_3 = \{\alpha, d\}$. Using Equation 16, the matching efficiency parameter $\alpha_1$ can be identified by differences in matching probabilities for different capital types for the same project that are driven by changes in $\theta_{yt}$.

The parameter $\alpha_0$ is identified mainly by matching the variance of utilization: low values of $\alpha_0$ imply that the meeting technology does not vary much with $\theta_{yt}$. The parameters $d_0, d_1$ are identified by the covariance between the gas price and capital utilization.

\[33\)The $\theta$ terms are also dependent on $\gamma$ but can be removed by differencing again over $y$.\]
Estimation using Simulated Method of Moments

The data requirements to implement the identification strategy above are demanding. For example, nonparametric estimation based on the identification argument in Proposition 4 would require observing many similar projects drilled by different types of capital in the same period. I instead estimate the model using the Simulated Method of Moments. To do this I match moments of the model to the data and I use the identification strategy to motivate which moments to use in the estimation.

Moments used in the estimation  First, I need to include moments to identify the $\lambda_2$ parameters that characterize the distribution of searching projects. I use difference in the average match between high and low specification capital, aggregated over booms and busts, when the period $T \in \{\text{boom, bust}\}$:

$$m_{1,T}(\lambda) = \frac{1}{\#T} \sum_{t \in T} \left[ \int_{A_t(h)} x_1 f_t(x_1 | y = h) dx_1 - \int_{A_t(l)} x_1 f_t(x_1 | y = l) dx_1 \right]$$

(17)

Here $\#T$ denotes the number of periods in $T \in \{\text{boom, bust}\}$. These moments are sensitive to the targeting parameter $\gamma$. Specifically, as the targeting parameter increases, the difference between the probabilities $f_t(x_1 | y = h) - f_t(x_1 | y = l)$ increases, as in Equation 16, and this will be reflected in changes in the moments $m_{1,T}(\lambda)$. Under random search ($\gamma = 0$) the only channel in the model for differences in matching patterns ($m_{1,T}(\lambda) > 0$) is differences in the acceptance sets. Under partially directed search ($\gamma > 0$), differences in matching patterns can also occur due to targeting behavior within acceptance sets.

To pin down $(\mu, \Sigma)$ I include the following moments which I denote $m_2(\lambda)$. I use the mean well complexity matched by mid-capital in booms and busts (2 moments). I also use the mean well complexity matched by low and high capital in busts (2 moments). Finally, I use the mean and variance of contract duration in observed matches (2 moments).

To identify $\lambda_3 = \{\alpha, d\}$ I include the moments related to patterns of capital utilization over the boom-bust cycle which I denote $m_3(\lambda)$. Specifically, I use the mean utilization for each capital type (3 moments). I also use the covariance of utilization and the gas price for each capital type (3 moments), and the variance of utilization for each capital type (3 moments).
Estimation details  I stack the 17 simulated moments:

\[ m_s(\lambda) = (m_{1,bust}(\lambda), m_{1,boom}(\lambda), m_2(\lambda), m_3(\lambda))' \]

I fit the simulated moments to the empirical moments \( m_d \) by minimizing the following objective function:

\[ (m_d - m_s(\lambda))' \Omega (m_d - m_s(\lambda)) \]

Here \( \Omega \) is the weighting matrix which I set as \( \Omega = \text{diag}(m_d'm_d)^{-1} \). The objective function is highly non-linear so I perform the optimization by grid-search. I construct the grid using a quasi-random Sobol sequence in order to efficiently span the parameter space. In order to compute these parameters I simulate the full model using the simulated method of moments from January 2000-December 2009. I provide details on the simulation algorithm in the Appendix.

5 Results

5.1 Computing the state transitions

I compute the transition probabilities using a separate Maximum Likelihood estimation for each state variable. The results are (standard errors in brackets):

\[
R_0 = \begin{bmatrix}
r_0^g \\
r_0^l \\
r_0^m \\
r_0^h \\
\end{bmatrix} = \begin{bmatrix}
0.64 (0.03) \\
4.06 (1.43) \\
5.52 (1.53) \\
8.90 (1.80) \\
\end{bmatrix}
\]

\[ \sigma_e = 1.0 (0.05) \]

\[
R_1 = \begin{bmatrix}
r_1^g \\
r_1^l \\
r_1^m \\
r_1^h \\
\end{bmatrix} = \begin{bmatrix}
0.89 (0.03) & 0 & 0 & 0 \\
-0.24 (0.11) & 0.83 (0.06) & -0.03 (0.07) & 0.09 (0.07) \\
-0.47 (0.12) & 0.17 (0.06) & 0.52 (0.08) & 0.12 (0.08) \\
-0.47 (0.14) & -0.02 (0.08) & 0.02 (0.09) & 0.36 (0.09) \\
\end{bmatrix}
\]

Since all the eigenvalues of \( R_1 \) lie within the unit circle, the transition matrix is stationary.
5.2 Estimating the value of a match

The rig’s value of searching I compute the rig’s value of searching non-parametrically using prices and agents’ policy functions. After estimating the policy functions for prices, extensions, and matches, I use forward simulation to recover the rig’s value of searching.

This procedure is illustrated in Figure 8. Overall, I compute the value of searching at a given state by forward simulating the future state evolution and matching paths. In Panel (a) I plot a single example of the simulation in the gray line. Starting from the state \((g = 15, n_l = 5, n_m = 5, n_h = 5)\) (a boom) I forward simulate one possible path for a rig’s contracts. The x-axis is the months from the initial state. The gray line moves around with matches and extensions and sometimes drops to zero if the rig is unemployed. I average over many of these simulations and Panel (b) illustrates this procedure. In gray are 10 example simulations and the black line is the average of 2000 simulations. The graph illustrates some of the rich non-linear dynamics that the model can account for: it appears that the average of all these simulations initially rises and then slowly falls towards a steady state. A high initial gas prices causes available capital to fall which raises prices in the short term. In the long-term the gas price reverts to the mean which causes rig prices to fall.

Panel (c) illustrates the final step of the procedure which is to construct the discounted sum of the average values to get the value of searching \(V_t^y(y)\). The light blue, dark blue, and black lines correspond to the value of searching in the current period \(V_t^y(y)\). The value functions increase in booms and fall in busts which is consistent with there being more matching opportunities when the gas price is high. The gray lines correspond to agents’ forecast of the value of searching over the next 18 months - for example, \(E_t V_{t+2}^y(y), E_t V_{t+3}^y(y)\) etc. The gray lines tend to converge towards the mean which indicates that agents have mean-reverting expectations about the value of searching. This is not surprising because the state transitions are mean reverting.

The match value function \(m\) and bargaining parameter \(\delta\) I plot the fitted match value function in Figure 9. When estimating the parameters, I impose the constraint that the match value must be (weakly) increasing for each rig type. In practice this constraint only binds for the low-specification rig. When computing the match values in Figure 9, I ensure that the fitted values are ‘within-sample’ by restricting the well complexity domain to \([0.3, 1]\) for low-specification rigs, \([0.3, 1.3]\) for mid-specification rigs, and \([0.3, 1.6]\) for high-specification rigs,
Figure 8: Computing the rig’s value of searching $V_t^g(y)$

(a) One simulation

(b) Averaging over many simulations

(c) The value of searching $V_t^g(y)$

Notes: Panel (a): The x-axis is the number of months after a given state $s_t$. The gray line is a single example simulation. Panel (b): Gray lines correspond to 10 example simulations. The black line corresponds to the average of these simulations: the average price per day. Panel (c): Points on the graph are plotted monthly. The gray lines correspond to the 18-month future beliefs of the value of searching.
Note: I restrict the domain to [0.3, 1] for low-specification rigs, [0.3, 1.3] for mid-specification rigs, and [0.3, 1.6] for high-specification rigs, which corresponds to approximately the 10% and 90% quantiles of observed matches for each rig type. Therefore the match value function is ‘in-sample’ for all rig types.

Comparing the estimated match value added across rig types, the data suggest that drilling rigs are relatively undifferentiated at low values of well complexity. By contrast, rigs appear to be highly differentiated when drilling complex wells. The idea that rig types are relatively undifferentiated for simple projects aligns with the common wisdom in the industry. Furthermore, the fact that high-specification rigs have a comparative advantage in drilling complex wells is important to producing an equilibrium with positive sorting. Specifically, the match value added is supermodular which implies that positive sorting would be efficient in a static model with no search frictions.
5.3 The remaining parameters

Table 3 contains the estimated parameters from the Simulated Method of Moments. I initially draw potential wells from a multivariate log-normal (and the corresponding $\mu$ and $\Sigma$ are reported), and then I scale complexity by 2, water depth by 500, well value by 1 million, and $\tau$ by 4. I also scale the targeting parameter by $1/(10 \text{ million})$ to ensure that it is of a similar scale to the other parameters in the optimization (the targeting weights are multiplied by the match surplus which is measured in millions of dollars).

Overall the parameters seem reasonable. The entry cost is $0.11 \text{ Million}$ which seems a reasonable value given that drawing up oil well plans, and applying for a permit, can be costly in terms of time and resources. The potential well draw parameters $(d_0, d_1, d_t)$ imply that the elasticity of the number of potential project draws with respect to the gas price is approximately 1. Furthermore, the trend of potential project draws is slightly decreasing over time $(d_t = -0.29)$. This reflects the fact that the shallow water of the US Gulf of Mexico is an old oil and gas field and much of the oil and gas has already been mapped and extracted.

The estimated targeting parameter $\gamma$ is 1.23. To get a sense of where this lies between random search and directed search I consider the probability that a complex well (I set $x_{\text{complex}} = 1.3$) targets its optimal match (which is a high-specification rig) at approximately the average state:\footnote{I calculate this approximate elasticity by taking the number of potential projects at the average gas price ($5.71$) in the sample and at the average period in the sample ($t = 60$) which is: $14 + 5.71 \times 17 - 0.29 \times 60 = 94$. Here, a $1$ (18\%) increase in the gas price would cause an 18\% increase in draws of potential projects, resulting in an elasticity of 1.}

$$\omega_t(y = \text{high}|x = \text{complex}) = \begin{cases} 0.27 & \text{random search} \\ 0.48 & \text{estimated model} \\ 1 & \text{directed search} \end{cases}$$

Note that this is the probability that a well targets a particular rig; the probability that the well meets a rig will also depend on the submarket tightness (and is equal to $\alpha_0 \theta_{yt}^\alpha \omega_t(y|x)$). The above example indicates the search technology that best rationalizes the data is somewhere between random search and directed search, and validates the need to estimate the search technology flexibly.\footnote{Specifically, I choose the state halfway through the sample at January 2005 which is also between a boom and bust.}
For the remaining parameters (the matching efficiency and the mean and standard deviation of potential projects) it is difficult to interpret them in isolation. Therefore, I see how closely the model fits the data overall.

5.4 Replication exercise

To evaluate the within-sample fit I run the simulated model over the gas price sequence over 2000-2009 and compare the simulated model to the data. I plot the fit to rig utilization and sorting patterns in Figure 10. The model replicates the data well. The sorting patterns predicted by the model are close to those in the data: in booms low-specification rigs are matched to simpler wells and high-specification rigs are matched to more complex wells.

Predicted rig utilization patterns match the empirical utilization patterns. The model matches the ordering between rig vintages observed in the data with high-specification rigs utilized more than mid and low-specification rigs. Furthermore, the model captures the pattern of lower utilization in busts and higher utilization in booms. In particular, the model successfully captures utilization patterns during the bust in 2001-2002, the boom in 2005-2008, and the downturn in 2009.
Figure 10: Data vs simulation

(a) Sorting: Bust  
(b) Sorting: Boom  
(c) Sorting: Rotation

(d) Utilization: High type  
(e) Utilization: Mid type  
(f) Utilization: Low type
6 Counterfactuals

I now use the model to perform several counterfactuals. In the counterfactuals my measure of welfare is the total value of wells drilled minus entry costs. Denoting $Y$ as the set of capital in the market, and letting $T = \{2000 : 1, \ldots, 2009 : 12\}$, the total welfare is:

$$
\Pi(Y) = \sum_{t \in T} \left( \left\{ \text{Total value of the projects undertaken by } Y \text{ at time } t \right\} - \left\{ \text{#projects entered at } t \right\} \ast c \right)
$$

For each of the counterfactuals I decompose the total effect into three components:

- **Match quality**: The change in the value of matches for the set of rigs which are matched in both the market baseline and counterfactual.

- **Number of matches**: The value of the new rigs that are matched in the counterfactual (or the loss in value if more rigs are unmatched in the counterfactual).

- **Entry cost saving**: The change in the total entry cost in the counterfactual compared to the market baseline.

I recompute the value functions in the counterfactuals. In addition, since state transitions will change, I also need to recompute agents’ beliefs about state transitions. I leave the computational details to the Appendix.

6.1 Quantifying the sorting effect

I first quantify how stronger sorting in booms increases welfare. Recall that the sorting effect arises because the option value of searching increases in a boom compared to a bust, and therefore agents are more selective in matching in booms than busts. Consequently, to quantify the sorting effect, I shut down the incentive to strategically reject bad matches by setting the option value of searching to zero:

$$
E_t V_{t+1}^S(y) = 0
$$
Figure 11: Quantifying the sorting effect: results

(a) Total change

(b) Decomposition

(c) Summary of changes

<table>
<thead>
<tr>
<th></th>
<th>Bust</th>
<th>Boom</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match quality</td>
<td>13.4%</td>
<td>20.2%</td>
<td>17.3%</td>
</tr>
<tr>
<td># Matches</td>
<td>-6.9%</td>
<td>-5.1%</td>
<td>-5.9%</td>
</tr>
<tr>
<td>Entry cost saving</td>
<td>-0.4%</td>
<td>1.1%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Total</td>
<td>6.5%</td>
<td>15.1%</td>
<td>11.4%</td>
</tr>
</tbody>
</table>

Note: This figure shows the change in welfare moving from the ‘no sorting’ counterfactual to the market baseline (where the sorting effect is present). Welfare at the baseline is $8.4 billion. In this counterfactual I fix the composition of wells to isolate the sorting effect. For completeness I include the entry cost saving if the composition of wells were allowed to adjust but this is not included in the totals. There is more output in booms and so changes in booms will be weighted more heavily than busts in the average column.

In this counterfactual world, agents will myopically accept all matches because there is no (perceived) benefit to searching for a better match. Therefore, for example, if a high-efficiency rig is contacted by a simple well then these agents will always match, regardless of other possible matches.

Computationally, after calculating the new value functions (the algorithm is in the Appendix), I simulate the model using the empirical natural gas price. Starting from the myopic ‘no sorting’ counterfactual, I compute the change in welfare when moving to the market benchmark. In addition (for this counterfactual only) in each period I use the same distribution of searching wells in the counterfactual as in the baseline model. This is to isolate the sorting effect from
changes in the composition of searching wells.\textsuperscript{36}

I plot the results in Figure 11. Panel (a) plots the total change in welfare (joint profits). Welfare when agents strategically reject bad matches is greater in every period and the total increase is 11.4\%. In USD this is around $1 billion over the 2000-2009 sample period. The effect is cyclical: the welfare increase in a boom is 15.1\% compared to around 6.5\% in a bust.

Panels (b) and (c) decompose how the sorting effect increases welfare: there are less matches (which by itself decreases welfare by 5.9\%), but the remaining matches are of higher quality because agents are more selective (which increases welfare by 17.3\%). Overall, the match quality effect dominates, which results in a net increase in welfare. In the Appendix I plot counterfactual sorting patterns which confirm that sorting patterns are flatter in both booms and busts in the ‘no sorting’ counterfactual.

\textbf{6.2 An intermediary that eliminates search frictions}

Next, I study how closely the market allocation approximates the first best allocation. I consider a counterfactual world where an intermediary assigns the optimal match within each period. Conceptually this is an ‘Uber for rigs’. The intermediary counterfactual highlights the effects of search frictions, while also providing an upper bound on the gains from recent advances in e-procurement in the industry.\textsuperscript{37}

Computing the optimal set of matches is a ‘generalized assignment problem’ and I implement a version of the Hungarian algorithm (\textit{Kuhn (1955)}) to compute the optimal matches at each state. I leave details on the implementation algorithm to the Appendix.

Figure 12 illustrates the change in welfare due to an intermediary. Panel (a) shows that the intermediary increases welfare by around 28.6\% over 2000-2009 compared to the market baseline. In USD this is around $2.4 billion over the 2000-2009 sample period. This increase in welfare is cyclical and concentrated in the boom: according to Panel (c) the welfare increase in booms is around 36.2\% versus only 18.3\% in busts.

\textsuperscript{36}For completeness I also include the entry cost saving if wells were allowed to enter and exit. Ultimately this has little effect on the result because similar numbers of wells enter in both the baseline and counterfactual.

\textsuperscript{37}For an early discussion about the potential benefits to e-procurement in the industry see \textit{Rothgerber (2002)}. 

45
Figure 12: Intermediary counterfactual results

(a) Total change

(b) Decomposition

(c) Summary of changes

<table>
<thead>
<tr>
<th></th>
<th>Bust</th>
<th>Boom</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match quality</td>
<td>11.3%</td>
<td>22.8%</td>
<td>17.9%</td>
</tr>
<tr>
<td># Matches</td>
<td>4.0%</td>
<td>5.0%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Entry cost saving</td>
<td>3.0%</td>
<td>8.4%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Total</td>
<td>18.3%</td>
<td>36.2%</td>
<td>28.6%</td>
</tr>
</tbody>
</table>

Note: This figure shows the change in welfare when moving from the market baseline to the intermediary counterfactual. The welfare in dollars at the market baseline is $8.4 billion. Note that there is more output in the booms and so changes in the booms will be weighted more heavily than busts in the average column.

Panel (b) decomposes the effect of the intermediary. Overall there is a gain from more matches of 4.6%. Wells know for certain if they will be matched if they enter (there is no search - the intermediary simply allocates a well to a rig or leaves it unmatched). Overall this has the effect of less wells entering in the intermediary’s allocation which causes an entry cost saving of 6.1%. The largest effect of the intermediary comes from an increase in match quality of 17.9%. In the Appendix I also plot counterfactual sorting patterns which show that the intermediary’s allocation features stronger sorting than the market baseline.

Given the gains from an intermediary, why is there not one in the market? The model suggests two reasons. First, a hypothetical intermediary’s profits would be highly cyclical, which would
require smoothing profits over many years. Second, the market is extremely fragmented: the largest oil and gas company accounts for only a 6% market share. Therefore any intermediary would need to coordinate the needs of many small firms on both sides of the market, which could be costly. Despite these difficulties, recent advances in technology and e-procurement (using the internet to share information and find matching partners) are slowly being incorporated into the industry. This suggests that some of the gains to improving the search process may soon be realized.

### 6.3 Effects of a demand smoothing policy

I now consider the effects of a demand smoothing policy. There is a long history in the oil and gas industry of policies designed to smooth out the disruptive effects of the boom-bust cycle. Between 1954 and 1978 natural gas producer prices were fixed in the United States for interstate trade. Today, the oil and gas industry is a global market which makes fixed price regulation impossible. However, many producer incentives, tax credits and royalty rates are tied to oil and gas prices. For example, the Federal Marginal Well Tax Credit is only available when the oil prices is below $18 per barrel. The Federal Enhanced Oil Recovery Credit is only available if oil prices are below $28 per barrel. The Bureau of Ocean Energy Management (BOEM) sets oil and gas price thresholds each year above which oil and gas producers do not receive royalty relief. The consequence of these counter-cyclical policies is to ‘smooth’ out the prices that producers face, increasing oil and gas prices in the bust and decreasing them in the booms.

To understand the effects of these policies on drilling behavior I consider a counterfactual demand smoothing policy that results in a one standard deviation decrease in the variance of the natural gas price. I am agnostic in the counterfactual about the exact implementation of taxes and subsidies that result in the smoother gas price, and I instead focus on the net benefits to the industry. The value functions need to be recomputed since agents’ beliefs about the future state evolution will change, and I describe the computational algorithm in the Appendix.

The results are depicted in Figure 13. Panel (a) shows that the smoothing policy results in large shifts in drilling activity from the boom to the bust. Welfare would be around 8.9%

---

38. Raghothamarao (2016) discusses how advances in e-procurement are being used in the oil and gas industry.

39. Potter et al. (2017) summarizes the tax credits oil and gas producers receive in low-price environments.
Figure 13: Demand smoothing counterfactual results

(a) Total change

(b) Decomposition

(c) Summary of changes

<table>
<thead>
<tr>
<th></th>
<th>Bust</th>
<th>Boom</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match quality</td>
<td>14.9%</td>
<td>-13.4%</td>
<td>-1.4%</td>
</tr>
<tr>
<td># Matches</td>
<td>8.6%</td>
<td>-2.4%</td>
<td>2.3%</td>
</tr>
<tr>
<td>Entry cost saving</td>
<td>-7.1%</td>
<td>6.9%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Total</td>
<td>16.4%</td>
<td>-8.9%</td>
<td>1.9%</td>
</tr>
</tbody>
</table>

Note: This figure shows the change in welfare when moving from the market baseline to the demand smoothing counterfactual. The welfare in dollars at the market baseline is $8.4 billion. Note that there is more output in the booms and so changes in the booms will be weighted more heavily than busts in the average column.

lower in booms but this is offset by welfare being 16.4% higher in the bust. Despite these large changes, the overall effect of the smoothing policy is to only increase output by 1.9%. Given the counterfactual demand smoothing policy is quite large (1 standard deviation), and the total effects of the policy are modest, this suggests that demand smoothing policies are somewhat ineffective in improving welfare.

Panel (b) decomposes the economic forces underlying this result. Match quality increases in a bust but decreases by more in a boom, leading to an net change in welfare of -1.4%. There

\[^{40}\text{Note that output in booms is higher than busts so changes in the booms will be weighted more heavily than busts in the total effect.}\]
are more matches in the bust but this is partially offset by a decrease in matches in the boom, leading to an average increase in welfare of 2.3%. Similarly, less wells enter in the boom which saves on the entry cost but this is offset by more wells entering in the bust. Overall, this suggests that the costs of mismatch are relatively linear with respect to the gas price, and so increases in welfare in a bust are offset by decreases in a boom.

Collard-Wexler (2013) finds qualitatively similar results for demand smoothing in the ready-to-mix cement industry: smoothing results in large changes in industry structure, but a small improvement in welfare. Although the market structure of the ready-to-mix cement industry differs from offshore drilling, these results suggest that understanding the industry structure is important for predicting the effects of demand smoothing policies.

7 Conclusion

A large literature has established that firms adjust to booms and busts by reallocating capital and that this process drives aggregate productivity. But much less is known about how firms reallocate capital in practice. Research in this area is needed because the effects of commonly proposed policies such as demand smoothing hinge on the reallocation mechanism.

In this paper I shed light on one such mechanism: matching. I develop a new framework that combines elements of the sequential search literature and firm dynamics literature. The framework incorporates distributions of searching agents that change over time, two-sided heterogeneity leading to sorting, and a more flexible search technology. I show how the framework can be tractably estimated by extending approaches from Industrial Organization. I apply the framework to a novel contract dataset in the market for offshore drilling rigs. I argue that booms are associated with a sorting effect. I use the framework to quantify the sorting effect, as well as the value of an intermediary and the effects of a demand-smoothing policy.

Overall this paper presents a unique picture of the inner workings of a decentralized capital market that is affected by booms and busts. My estimation strategy may be of some interest to economists working on search markets in other industries. Overall, my results show that matching is an important reallocation channel in booms and busts for capital markets, and that this has significant implications for policy design.
References


Appendices

A Data

A.1 Dataset Construction

The data construction process combines several datasets. The contract datasets are:

- IHS contract dataset
- Rigzone contract dataset
- Rigzone order book

The well datasets are all from the regulator (BSEE). These are:

- The borehole dataset
- The permit dataset
- The lease dataset

I begin by merging the borehole, permit and lease datasets on the unique API well number to obtain a single dataset with all well characteristics at the well level which I call the ‘well database’. 

Next I create the ‘contract database’. To keep the analysis focused on one market I analyze jackup rigs that drilled wells between 01 January 2000 and 31 December 2009. That is, I remove contracts drilled by deepwater semisubmersibles, drillships, and fixed platforms. There are 3593 contracts for jackup rigs in total. I further remove 17 ‘workover’ rigs (these correspond to 365 contracts so the dataset is reduced to 3228 contracts), whose main purpose is to reenter wells, typically for maintenance. These rigs rarely drill new wells and so are not in direct competition with drilling rigs. I identify workover rigs as any rig offered by the drilling company Nabors as well as rigs whose status is ‘workover’ in the Rigzone contract > 80% of the time.
Sometimes the rig name differs between the contract data and the well data due to ownership changes and so I first map rig names between the two databases using the Rigzone order book (which has previous rig names), and the websites maritime-connector.com and marine-traffic.com. I also use these websites plus the Rigzone order book to find the maximum drilling depth of each rig. I merge 100% of rig names in the IHS contract dataset to the well dataset in this way.\footnote{Since there is a unique rig ID in the permit database and the contract database is for all contracts offered in the industry, this procedure also matches the names of 100% of wells.}

In total there are 3228 contracts and 5202 wells in the full datasets for the years 2000-2009. To do the analysis I require a dataset of contracts merged with wells. I merge wells to rigs, matching on the rig name and if the initial drilling date (the ‘spud date’), or the final drilling date (the ‘depth date’), are within the start and end dates of the contract. The procedure successfully results in matching 2394 contracts and 4698 of wells. I further impute the characteristics of 224 contracts if the contract was an extension/renegotiation. I also impute the contracts of 83 wells if the well was drilled subsequently to a merged well by the same rig-well owner pair. In total 2618 contracts (81 percent) and 4781 wells (92 percent) are matched. Why are some contracts and wells unmatched? The most likely explanation is that the contract data contains the expected contract start and end date rather than the actual contract start and end date. Sometimes there are unforeseen delays with drilling a well which can affect the actual end date or start date of a contract.

Sometimes a drilling contract will contain two or more wells. Therefore, I collapse multi-well contracts by taking the mean well complexity, the mean well water depth, and the mean well value. The resulting and final dataset that I use for estimation is at the contract level.

**A.2 Computing the Mechanical Risk Index**

This section draws directly from Kaiser (2007). The Mechanical Risk Index was developed by Conoco engineers in the 1980s (Kaiser (2007)). The idea behind the index is to collapse the many dimensions that a well can differ on into a one-dimensional ranking of well complexity. Well complexity is directly related to the cost of drilling a well: these wells run an increased risk of technical issues which may require new materials or result in blowouts. Figure 14 contains
Figure 14: Diagram of a simple vs complex well

(a) A simple well

(b) A complex well

Note: This figure gives an example of a simple well design and a complex well design. Simple wells will rank low on the Mechanical Risk Index whereas complex wells will rank high. Panel (a) illustrates a simple well - in this case it is just a short vertical hole. Panel (b) illustrates a complex well. In this case there are many connected sections and curves. A more complex well design increases the risk the rig will encounter a problem and high-specification rigs are more suited to drilling these types of wells. Source: https://directionaldrilling.wordpress.com/

The Mechanical Risk Index is computed by first computing ‘component factors’:

\[
\phi_1 = \left( \frac{TD + WD}{1000} \right)^2 \\
\phi_2 = \left( \frac{VD}{1000} \right)^2 \left( \frac{TD + HD}{VD} \right) \\
\phi_3 = (MW)^2 \left( \frac{WD + VD}{VD} \right) \\
\phi_4 = \phi_1 \sqrt{NS + \frac{MW}{(NS)^2}}
\]

Here TD is total depth in feet, WD is water depth in feet, VD is vertical depth in feet, MW is mud weight in ppg, NS is the number of strings.

Next ‘key drilling factors’ are computed. These are: \( \psi_1 = 3 \) if there is a horizontal sections; \( \psi_2 = 3 \) if there is a J-curve; \( \psi_3 = 2 \) if there is an S-curve; \( \psi_4 \) if there is a subsea well; \( \psi_5 = 1 \) if there is an \( H_2S/CO_2 \) environment; \( \psi_6 = 1 \) if there is a hydrate environment; \( \psi_7 = 1 \) if there is a depleted sand section; \( \psi_8 = 1 \) if there is a salt section; \( \psi_9 = 1 \) if there is a slimhole, \( \psi_{10} = 1 \)
if there is a mudline suspension system installed; \( \psi_{11} = 1 \) if there is coring; \( \psi_{12} = 1 \) if there is shallow water flow potential; \( \psi_{13} = 1 \) if there is riserless mud to drill shallow water flows; \( \psi_{14} = 1 \) if there is a loop current.

The Mechanical Risk Index is then computed as:

\[
MRI = \left( 1 + \frac{\sum_j \psi_j}{10} \right) \sum_i \psi_i
\]

In my data I have excellent information for all wells on \( TD, WD, VD, HD \) using the BSEE permit data and the BSEE borehole data. I have data for \( MW, NS \) for a subset of wells and I impute the remainder based on geological proximity and well depth - based on the fact that geological conditions are usually similar for nearby wells.

Computing the ‘key drilling factors’ \( \psi_j \) presents a greater challenge because the data are either not recorded (e.g. if there is shallow water flow potential) or would need to be imputed from well velocity surveys (e.g. if there is an S-curve). Rather than guess I set all \( \psi_j = 0 \). The implication for the index is that there will be a less accurate measure of complexity which will result in measurement error.

B Proofs

B.1 Proof of Proposition 3

The aim is to show that:

1. The value of search \( V_t^y(y|\lambda) \) and the value of a match for capital \( \Pi_t^y(x,y,p_t|\lambda) \) can be computed from the data

2. The term \( \Pi_t^y(x,y|\lambda) \) can be computed from the data and the parameters \( \{m^y_0, m^y_1\} \).

**Part 1:** Writing out the value of searching for rig \( y \) at time \( t \):

\[
V_t^y(y|\lambda) = \int_x q_t(x|y,\lambda) \max \left\{ \Pi_t^y(x,y|\lambda), \beta E_t V_{t+1}^y(y|\lambda) \right\} dx + q_t(\emptyset|y,\lambda) \beta E_t V_{t+1}^y(y|\lambda)
\]

\[
= \int_x \tilde{f}_t(x|y) \Pi_t^y(x,y|\lambda) dx + \tilde{f}_t(\emptyset|y) \beta E_t V_{t+1}^y(y|\lambda)
\]

(19) 56
Here \( f_t(x|y) \) is the distribution of observed matches for type \( y \) capital at state \( s_t \). Also writing out capital’s value of searching:

\[
\Pi^y_t(x, y|\lambda) = \sum_{s=0}^{\tau-1} \beta^s \bar{p}_t(x, y) + \beta^T E_t \left[ \bar{e}_{t+s}(x, y) \Pi^{y}_{t+s}(x, y|\lambda) + (1 - \bar{e}_{t+s}(x, y) V^{y}_{t+s}(y|\lambda)) \right]
\] (20)

Here \( \bar{p}_t(x, y) \) is the price of an \((x, y)\) match at state \( s_t \) and \( \bar{e}_t(x, y) \) is the probability of renegotiating an \((x, y)\) match at state \( s_t \). Observe that Equation 19 and Equation 20 do not contain any parameters in their per-period payoffs after writing them in terms of the policy functions. Therefore both the value of search and the value of a match for capital are not dependent on the parameters. Therefore:

\[
V^y_t(y) = \int f_t(x|y) \Pi^y_t(x, y) dx + f_t(0|y) \beta E_t V^{y}_{t+1}(y)
\]

\[
\Pi^y_t(x, y) = \sum_{s=0}^{\tau-1} \beta^s \bar{p}_t(x, y) + \beta^T E_t \left[ \bar{e}_t(x, y) \Pi^{y}_{t+s}(x, y) + (1 - \bar{e}_t(x, y) V^{y}_{t+s}(y)) \right]
\]

**Part 2:** Next I need to show that the project’s value of searching \( V^x_t(x|\lambda) \) is only dependent on \( \{m_0^y, m_1^y\} \). Writing out the project’s value of searching:

\[
\Pi^x_t(x, y|\lambda) = \sum_{s=0}^{\tau-1} \beta^s [p_t(x, y) - m_t(x, y)] + \beta^T E_t \bar{e}_t(x, y) \Pi^{x}_{t+s}(x, y|\lambda)
\]

Since the per-period payoff \( m_t(x, y) \) is only dependent on \( \{m_0^y, m_1^y\} \), it follows that \( \Pi^x_t(x, y|\lambda) \) is also just a function of \( \{m_0^y, m_1^y\} \).

**B.2 Proof of Lemma 2**

Applying Proposition 3 to the Nash bargaining problem:

\[
p_t = \arg \max_{p_t} \left[ \Pi^y_t(x, y, p_t) - \beta E_t V^{y}_{t+1}(y) \right]^\delta \left[ \Pi^x_t(x, y, p_t|\lambda_1) \right]^{1-\delta}
\]

Writing out the above objective function and taking the first-order condition with respect to \( p_t \), the total value of the contract is:

\[
\sum_{s=0}^{\tau-1} \beta^s p_t(x, y) = (1 - \delta) \left[ \beta^T E_t \left[ e_{t+s}(x, y) \Pi^{y}_{t+s}(x, y, p_{t+s}) + (1 - e_{t+s}(x, y)) V^{y}_{t+s}(y) \right] - \beta E_t V^{y}_{t+s}(y) \right] + \delta \left[ \sum_{s=0}^{\tau-1} \beta^s m(x, y) + \beta^T E_t [g_{t+s}x_{value}] + \beta^T E_t \Pi^{x}_{t+s}(x, y, p_{t+s}|\lambda_1) \right]
\]
Rearranging the above equation it can be written as in Lemma 2:

\[ \tilde{p}_t(x, y) = \delta \left[ m_0^y + m_1^y x_{\text{complexity}} + g_t(x, y) \right] \]

Where:

\[ \tilde{p}_t(x, y) = \sum_{s=0}^{\tau-1} \beta^s p_t(x, y) - A \]
\[ g_t(x, y) = \frac{C - A}{B} \]

And:

\[ A = \beta^t \mathbb{E}_t \left[ e_{t+\tau}(x, y) \Pi_{t+\tau}^y(x, y, p_{t+\tau}) + (1 - e_{t+\tau}(x, y)) V_{t+\tau}^y(y) \right] - \beta \mathbb{E}_t V_{t+1}^y(y) \]
\[ B = \mathbb{E}_t \left[ \sum_{j=0}^{\tau-1} \beta^j \left[ \sum_{k=0}^{\infty} \beta^{k\tau} \prod_{s=0}^{k} e_{t+s\tau}(x, y) \right] \right] \]
\[ C = \mathbb{E}_t \left[ \sum_{k=0}^{\infty} \beta^{k\tau} \prod_{s=0}^{k} e_{t+s\tau}(x, y) g_{t+s\tau}, x_{\text{value}} \right] \]

B.3 Proof of Proposition 4

The aim is to show that the targeting parameter \( \gamma \) and the distribution of searching wells \( f_t(x) \) are separately identified. Note that the parameters \( \alpha_0, \alpha_1 \) will be identified from utilization patterns and so I take them as known for this proof. The observed distribution of projects that type \( y \) capital matches with is:

\[ \tilde{f}_t(x|y) = \begin{cases} 
\alpha_0 \theta_{yt}^{\alpha_1-1} w_t(y|x) f_t(x) & \text{if } x \in A_t(y) \\
1 - \alpha_0 \theta_{yt}^{\alpha_1-1} \int_{z \in A_t(y)} w_t(y|z) f_t(z) dz & \text{if } x = \emptyset
\end{cases} \]

At this stage the market tightness \( \theta_{yt} \), the weights \( w_t(y|x) \) (which are a function of the targeting parameter \( \gamma \)), and the distribution of searching projects \( f_t(x) \) are not known. I do the proof in four steps:

1. Identify the targeting weights \( \omega_t(y|x) \)
2. Identify the market tightness \( \theta_{yt} \)
3. Identify the targeting parameter \( \gamma \)
4. Identify the distribution of searching projects \( f_t(x) \)
**Part 1** I show how the targeting weights $\omega_t(y|x)$ can be identified. Rewriting the equation for the targeting weights:

$$\omega_t(y|x) = \frac{n_{yt} \exp \left( \gamma \pi_t(y|x) \right)}{\sum_{k \in \{l,m,h\}} n_{kt} \exp \left( \gamma \pi_t(k|x) \right)} \quad \tag{21}$$

I first show how $\gamma \pi_t(y|x)$ in the above equation can be identified. Comparing the probability of the same type of project $x$ matching different capital $y, y'$ at $t$:

$$\ln \left( \tilde{f}_t(x|y) \right) - \ln \left( \tilde{f}_t(x|y') \right) = \ln \left( \omega_t(y|x)/\omega_t(y'|x) \right) + \ln \left( \theta_{yt}^{\alpha_1-1}/\theta_{y't}^{\alpha_1-1} \right) \quad \tag{22}$$

$$= \gamma \left( \pi_t(x|y) - \pi_t(x|y') \right) + \ln \left( \theta_{yt}^{\alpha_1-1}/\theta_{y't}^{\alpha_1-1} \right) \quad \tag{23}$$

$$= \gamma \alpha_0 \theta_{yt}^{\alpha_1} \Pi_t^x(y,x) + \gamma \alpha_0 \theta_{y't}^{\alpha_1} \Pi_t^x(y,y') + \ln \left( \theta_{yt}^{\alpha_1-1}/\theta_{y't}^{\alpha_1-1} \right) \quad \tag{24}$$

Here the third line follow from substituting in the expression $\pi_t(y|x) = \alpha_0 \theta_{yt}^{\alpha_1} \Pi_t^x(y,x)$. Since $\Pi_t^x(y,x)$ is known I can identify the term $\gamma \alpha_0 \theta_{yt}^{\alpha_1}$ for each $y \in \{\text{low, mid, high}\}$. Therefore $\gamma \pi_t(y|x)$ can be constructed for any match $(x,y)$. Then I can recover the weights $\omega_t(y|x)$ because they are just a function of $\gamma \pi_t(y|x)$.

**Part 2:** Next I show that the equilibrium market tightness $\theta_{yt}$ is identified. Rearranging Equation 22 and taking logs:

$$\ln \left( \theta_{yt}^{\alpha_1-1} \right) - \ln \left( \theta_{y't}^{\alpha_1-1} \right) = \ln \left( \frac{\tilde{f}_t(x|y)}{\tilde{f}_t(x|y')} \right) \frac{\omega_t(y'|x)}{\omega_t(y|x)} \quad \tag{25}$$

Denote the term on the right-hand side as $h_t^{y,y'}(x)$ and note that it can be constructed from the data using Part 1. Then evaluating the above equation at three points $x, x', x''$:

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \ln(\theta_{yt}^{\alpha_1-1}) \\ \ln(\theta_{mt}^{\alpha_1-1}) \\ \ln(\theta_{lt}^{\alpha_1-1}) \end{bmatrix} = \begin{bmatrix} h_t^{hm}(x) \\ h_t^{hl}(x') \\ h_t^{ml}(x'') \end{bmatrix} \quad \tag{26}$$

Therefore I can identify the meeting probability terms $\theta_{yt}^{\alpha_1-1}$ from the data. Since the parameter $\alpha_1$ is known, I can also recover $\theta_{yt}$.

**Part 3:** Next I show how the targeting parameter $\gamma$ can be identified. Rewriting Equation 23:

$$\ln \left( \tilde{f}_t(x|y) \right) - \ln \left( \tilde{f}_t(x|y') \right) = \gamma \left( \pi_t(x|y) - \pi_t(x|y') \right) + \ln \left( \theta_{yt}^{\alpha_1-1}/\theta_{y't}^{\alpha_1-1} \right)$$
Note that the expected value of project type $x$ targeting rig type $y$, $\pi_t(x|y)$, can be constructed from Part 1 and Part 2. Therefore I can recover $\gamma$ from variation in $\pi_t(x|y)$.

**Part 4:** Finally I show how to identify the distribution of searching projects $f_t(x)$. Within an acceptance set the probability of observing $x$ given $y$ is:

$$\tilde{f}_t(x|y) = \alpha_0 \theta_{yt}^{\alpha_1 - 1} w_t(y|x) f_t(x)$$

(27)

Since the left-hand side is data and $\alpha_0 \theta_{yt}^{\alpha_1 - 1} w_t(y|x)$ is known from Part 1 and Part 2, I can recover $f_t(x)$ for all $x \in A_t(y)$. If $c > 0$ then a project will only enter if it is at least one acceptance set. Therefore $f_t(x)$ is identified over its full support.

## C Computation

### C.1 Policy functions

To perform the forward simulation I need to know prices $\tilde{p}_t(x, y)$, the extension probability $\tilde{e}_t(x, y)$, and the distribution of matches $\tilde{f}_t(x|y)$ at each state $s_t$.

**Prices** To estimate $\tilde{p}_t(x, y)$ I regress observed prices on a second order polynomial of $(x, y, s_t)$.

**Extension probability** To estimate the probability of extending a contract $\tilde{e}_t(x, y)$ I use a linear probability model with the independent variables a second order polynomial of $(x, y, s_t)$.

**Matches** I approximate $\tilde{f}_t(x|y)$ with the following discretized distribution. I discretize the four project covariates plus the possibility of not being matched into the following bins:

$$\{\text{No Match}\} \cup \left\{ \begin{array}{l} x_{\text{duration}} \leq 3, x_{\text{duration}} \geq 4 \end{array} \right\} \times \left\{ \begin{array}{l} x_{\text{value}} < 250000, x_{\text{value}} \geq 250000 \end{array} \right\}$$

$$\times \left\{ \begin{array}{l} x_{\text{complexity}} < 0.75, x_{\text{complexity}} \geq 0.75 \end{array} \right\} \times \left\{ \begin{array}{l} x_{\text{water depth}} < 100, x_{\text{water depth}} \geq 100 \end{array} \right\}$$
The bins for all the covariates correspond to being above/below the median. The above set of alternatives has $1 + 2^4 = 17$ elements. I run a multinomial logit using the above set of alternatives using a second order polynomial of $(y, s_t)$ interactions as covariates.

C.2 Simulated Method of Moments

Algorithm

1. Initialize the state.

2. Draw $D_t$ projects from $x \sim \text{Lognormal} (\mu, \Sigma)$. Construct an initial guess of $f_t^0(x)$ by guessing whether each project $x$ enters.

3. For each well $x$, compute surplus $S_t(x, y)$ for $y \in \{\text{low, mid, high}\}$ by forward simulation.

4. While $\text{dist}(f_t^k(x), f_{t+1}^k(x)) > \epsilon_1$:
   
   a. For $x \in f_t^k(x)$, guess a targeting weight for each rig type: $\omega_{t,0}^k(y|x)$
   
   b. While $|\theta_{y^t}^{k,j+1} - \theta_{y^t}^{k,j}| > \epsilon_2$:
      
      i. Compute $\theta_{y^t}^{k,j}$ from each the targeting weights $\omega_{t,j}^k(y|x)$
      
      ii. Update the targeting weights $\omega_{t,j}^k(y|x)$ using $\theta_{y^t}^{k,j}$
   
   c. Compute contact rates $q_t^k(y|x)$ using equilibrium $\omega_t^k(y|x), \theta_{y^t}^k$
   
   d. Compute the expected value of entering for each project $x$ using $q_t^k(y|x)$
   
   e. $f_{t+1}^k(x) \leftarrow \text{All } x \text{ where expected value of entering } > c$

5. Compute matches between available capital and searching projects $f_t(w)$

6. Update the state at time $t + 1$.

Explanation of the algorithm

**Step 1:** I begin by initializing the detailed state which is a list of every rig in the market, whether it is currently matched with a particular well or unmatched, and the natural gas price. I initialize the detailed state by burning in the algorithm for 24 months using the empirical natural gas price at January 2000.
Step 2: Next I draw $D_t$ projects from the distribution of potential projects $Lognormal(\mu, \Sigma)$. take a guess of which of these projects will enter which results in an initial guess of searching wells $f_t^0(x)$.

Step 3: For each project draw I compute the total surplus from matching with each capital type. Note that the total surplus is known from Part 2 of the estimation strategy and therefore no value function iteration is necessary.

Step 4: I iterate over the distribution of searching projects and targeting weights using a nested fixed point method. For each guess of searching projects $f_t^k(x)$ I iterate to find an equilibrium submarket tightness $\theta_{yt}$ for $y \in \{low, mid, high\}$.

The equilibrium submarket tightness $\theta_{yt}$ pins down the expected value of targeting each type of capital for each well. Therefore I can construct the targeting weights. Using these targeting weights I can then find the expected value of entering for each project. I then update the distribution of searching projects $f_t^{k+1}(x)$ using the well entry condition.

After convergence I am left with an equilibrium distribution of searching projects $f_t(x)$ and equilibrium submarket tightness $\theta_{yt}$.

Step 5: I compute matches. To do this I first back out the probability that each type of capital is contacted by a particular project $q_t(y|x)$, using $f_t(x)$, $\theta_{yt}$, and equilibrium targeting weights. I then simulate meetings for each unit of available capital based on these contact rates. Meetings will turn into matches if $S_t(x, y) > 0$.

Step 6: Finally I update the state by computing contract extensions and updating matches. I update the natural gas price using the empirical natural gas price.
C.3 Counterfactual value function computation algorithm

In the counterfactuals, it is necessary to re-compute the value functions for match surplus and the value of searching. I compute separate value functions for each capital type.\footnote{There are six value functions to compute in total: the value of match surplus and the value of searching for each of the 3 capital types.} The state space dimension for the value of searching is 4-dimensional: one dimension for the gas price and three dimensions for the number of available rigs. The dimension of the surplus value is 8-dimensional: four dimensions for each well covariate and four dimensions for the state. I approximate the value function using multidimensional interpolation of Chebyshev functions. Interpolation over the full grid of nodes would run into the curse of dimensionality so instead I use a sparse grid of order 3 which reduces the number of nodes to compute the value function in each iteration from hundreds of thousands of nodes to a few hundred nodes. Note that agents’ state transition beliefs over available rigs also need to be recomputed because they will change with the counterfactual.\footnote{The natural gas price is exogenous and therefore I treat agents’ beliefs about the natural gas price as fixed.}

The solution algorithm requires iterating each value function to a fixed point. The computationally intensive part is computing the set of matches and so to speed up the computation time I perform this as an outer loop, nesting many iterations of the surplus value function as an inner loop.

The algorithm is as follows:

1. Guess agents’ state transition beliefs.
2. Guess an initial value for each value function node.
3. Fix capital’s value of searching $V^y(x, y, s)$ and begin the inner loop, repeating the following until convergence:
   
   (a) For each node $(x, s)$, interpolate the value of searching $V^y$ and the surplus value $S$.
   
   (b) For each node $(x, s)$ do one update of the surplus value $S$ using Equation (33)
   
   (c) Using these updated nodes, fit the interpolation coefficients of the surplus value $S$.
4. Update capital’s value of searching $V^y(x, y, s)$ by:
(a) For each node \( s \), draw \( D \) potential projects.

(b) Compute the surplus of each project by interpolating the surplus function \( S \).

(c) Compute the assignment of capital to projects. Note that this depends on the counterfactual. For the demand-smoothing counterfactual use the estimated search technology. For the intermediary use the generalized assignment algorithm.

(d) Repeat (a)-(c) several times.

(e) Update \( V^R(r, w, s) \) once using Equation (34).

5. Check for convergence of \( V^u \) and \( S \).

6. Return to step (3) if convergence has not taken place.

7. After convergence of the value functions, simulate the model and update agents’ state transition beliefs.

C.4 Counterfactual: Intermediary

**Setup** In every period the intermediary observes available capital of each type \( Y_j \) where \( j \in \{l, m, h\} \), the set of potential projects \( X \), and the state \( s_t \). The intermediary’s problem is to allocate projects to capital to maximize surplus:

\[
\max_{a_{ij}} \sum_{x, y_j} S_t(x, y_j) a_{ij}
\]

subject to:

\[
a_{ij} \in \{0, 1\} \quad (28)
\]

\[
\sum_j a_{ij} = 1 \quad (29)
\]

\[
\sum_i a_{ij} \leq Y_j \text{ for } j \in \{l, m, h, \emptyset\} \quad (30)
\]

\[
\delta S_t(x, y_j) \geq c \quad \text{Project entry constraint} \quad (31)
\]

\[
S_t(x, y_j) \geq 0 \quad \text{Positive surplus constraint} \quad (32)
\]

The match surplus embeds the continuation value of capital not matching and continuing to search. Note that the project entry constraint in Equation (31) ensures that the decision to enter is individually rational for each project.

64
**Value functions**  Denote \( f_t^*(x|y) \) as the distribution of projects an intermediary allocates to capital type \( y \) at state \( s_t \). The value functions for total surplus and capital’s value of searching are:

\[
S_t(x, y) = \Pi_t^y(x, y) + \Pi_t^y(x, y) - \beta E_t V_{t+1}^y(y)
\]

\[
= \sum_{k=0}^{\tau-1} \beta^k m(x, y) + \beta^\tau E_t \left[ g_{t+\tau} x_{\text{value}} \right] + \beta^\tau E_t \left[ c_{t+\tau}(x, y) S_{t+\tau}(x, y) + V_{t+\tau}^y(y) \right] - \beta E_t V_{t+1}^y(y)
\]

(33)

\[
V_t^y(y) = \int x \Pi_t^y(x, y) f_t^*(x|y) dx + f_t^*(\emptyset|y) \beta E_t V_{t+1}^y(y)
\]

(34)

Where \( \Pi_t^y(x, y) = (1-\delta) S_t(x, y) \).
Table 4: Fit of the simulation to the moments used in the estimation

<table>
<thead>
<tr>
<th>Moment</th>
<th>Empirical Value</th>
<th>Simulated Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean: Complexity, Boom</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>Mid</td>
<td>0.83</td>
<td>0.81</td>
</tr>
<tr>
<td>High/Low difference</td>
<td>0.065</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>Mean: Complexity, Bust</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.71</td>
<td>0.73</td>
</tr>
<tr>
<td>Mid</td>
<td>0.83</td>
<td>0.82</td>
</tr>
<tr>
<td>High/Low difference</td>
<td>0.20</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Mean: Duration</strong></td>
<td>2.64</td>
<td>3.40</td>
</tr>
<tr>
<td><strong>Variance: Duration</strong></td>
<td>0.62</td>
<td>0.71</td>
</tr>
<tr>
<td><strong>Mean: Utilization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.63</td>
<td>0.60</td>
</tr>
<tr>
<td>Mid</td>
<td>0.77</td>
<td>0.79</td>
</tr>
<tr>
<td>High</td>
<td>0.93</td>
<td>0.85</td>
</tr>
<tr>
<td><strong>Variance: Utilization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>Mid</td>
<td>0.016</td>
<td>0.007</td>
</tr>
<tr>
<td>High</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td><strong>Covariance: Utilization and Gas Price</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.13</td>
<td>0.20</td>
</tr>
<tr>
<td>Mid</td>
<td>0.17</td>
<td>0.11</td>
</tr>
<tr>
<td>High</td>
<td>0.04</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Note: This table contains the moments used in the simulated method of moments step. I report both the value observed in the data and the simulated moments at the optimal parameters.
Figure 15: Industry states

(a) Gas prices

(b) Number of available rigs

Note: This figure contains the empirical states used in the estimation. The raw empirical sequence of available rigs is noisy because of the fine aggregation at a monthly level. Therefore after estimating the transitions I smooth the number of available rigs using a local polynomial smoother with a bandwidth equal to 1 month.
Figure 16: Acceptance sets for 3 month contracts

(a) Low-spec  
(b) Mid-spec  
(c) High-spec 

Note: Acceptance sets are defined as the region where total surplus is positive: $1[S_t(x, y) > 0]$. The blue region is the acceptance set in a boom (defined as a $(g, n_l, n_m, n_h) = (15, 10, 6, 5)$). The gray region (plus the blue region) is the acceptance set in a bust (defined as a $(g, n_l, n_m, n_h) = (3, 15, 12, 8)$). I assume that $x_{\text{water depth}} = 100$ so all rig types can drill the well. Acceptance sets are graphed for each rig type. The vertical axis is $x_{\text{value}}$, the quantity of hydrocarbons in a well (in millions).
Figure 17: Counterfactual utilization and sorting patterns

(a) No sorting: $\Delta$ utilization

(b) No sorting: Sorting patterns

(c) Intermediary: $\Delta$ utilization

(d) Intermediary: Sorting patterns

(e) Smooth demand: $\Delta$ utilization

(f) Smooth demand: Sorting patterns

Note: This figure contains counterfactual utilization and sorting patterns. The left column plots the change in capital utilization (the proportion of capital under contract) between the market baseline and the counterfactual. The right column plots counterfactual sorting patterns in booms and busts (where a boom is defined as an above average (> $5.71) gas price).