

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

Nicholas Vreugdenhil\*

August 27, 2020

## Abstract

How efficiently do markets reallocate capital in booms and busts? I exploit a novel dataset of offshore drilling contracts to examine the role of search and matching in shaping industry reallocation. Drilling companies match with capital (rigs) in a decentralized market. I find oil and gas booms increase the option value of searching which leads agents to avoid bad matches, reducing mismatch through a *sorting effect*. I provide an identification strategy to disentangle unobserved demand changes from the sorting effect. Estimating a model, I find substantial benefits to the sorting effect and an intermediary. However, demand smoothing policies are ineffective.

---

\*Department of Economics, Arizona State University. Email: nvreugde@asu.edu. Thank you to Robert Porter, Mar Reguant, and Gaston Illanes for their encouragement and advice. I would also like to thank Gaurab Aryal, Vivek Bhattacharya, Igal Hendel, Ken Hendricks, Bill Rogerson, Mark Satterthwaite, and Yuta Toyama for useful comments and suggestions. I thank seminar participants at ASU, Cornell, Duke Fuqua, IIOC, LSE (Management), Microsoft Research, Monash University, NYU, NYU (Stern), Penn State, UCL, and U Maryland AREC. 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.<sup>1</sup> Although it is well established that fluctuations and productivity are broadly linked, the exact process of capital reallocation within industries is not well understood.<sup>2</sup> This is largely due to the lack of producer-level data on covariates such as contracts, production, and relationships.<sup>3</sup> Filling this gap is important because the costs and benefits of often-proposed policies - such as demand smoothing - hinge on the reallocation mechanism.

In this paper I focus on the role of *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 *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.<sup>4</sup> I provide an identification strategy to disentangle changes in

---

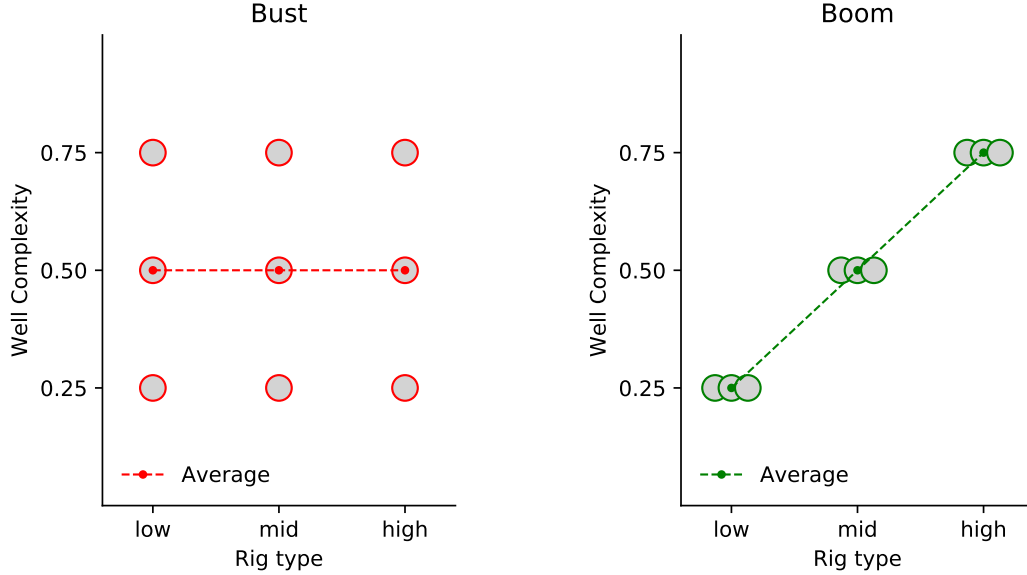
<sup>1</sup>For a review of the literature see [Eisfeldt and Shi \(2018\)](#).

<sup>2</sup>[Bartelsman et al. \(2013\)](#) document empirically the role of reallocation.

<sup>3</sup>[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.

<sup>4</sup>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.

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 an intermediary 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 numerous small ‘independents’, 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.<sup>5</sup> Therefore, in the offshore drilling industry, stronger sorting

<sup>5</sup>The fact that agents care about the quality of the match - and not just whether they are matched or not - is

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.<sup>6</sup> 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.

---

an important difference between my setting and recent work in Industrial Organization on search markets such as taxis (Frechette et al. (2017), Buchholz (2020)) and bulk shipping (Brancaccio et al. (2020)) where agents are relatively homogeneous.

<sup>6</sup>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 for 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.

I estimate the model in three steps. First, I estimate state transitions based on observed empirical 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 ‘no sorting’ world where rigs accept all matches and do not direct their search away from bad matches. Moving to the market benchmark (and allowing for the sorting effect) increases welfare by 11.2%, or around \$700 million dollars over the 2000-2009 period. 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 benefits of an intermediary who can reduce search frictions by offering an improvement in the search technology. In addition to highlighting the effects of search frictions, this counterfactual suggests potential gains from recent advances in e-procurement in the industry.<sup>7</sup> I find that the intermediary would increase welfare by around 28.2% compared to the market benchmark.

Finally, I consider a demand smoothing policy which would eliminate price cycles. 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. Furthermore, between 1954 - 1978 natural gas producer prices were fixed in the US for interstate trade. I find that demand smoothing would cause large shifts in drill activity from booms to busts. However, the policy would increase overall welfare by only 6.0%.

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.<sup>8</sup> By contrast, in this paper I construct a dataset of the universe of contracts in the industry matched with rich micro

---

<sup>7</sup>The potential of the internet to reduce search frictions in the industry has been discussed by practitioners since as early as 2002: [Rothgerber \(2002\)](#).

<sup>8</sup>A literature estimates these unobserved transfers in certain settings (e.g. [Villas-Boas \(2007\)](#)).

data from the regulator on the characteristics of projects undertaken under these contracts. My analysis of the dataset presents a unique and detailed picture of how firms make decisions when they are faced with fluctuations.

The second contribution is to solve a data limitation that often occurs in capital markets: demand (the distribution of searching wells) is not observed. I show how demand, as well as a more flexible search technology, can be identified from data on matches. The flexible search technology - partially directed search - nests typical assumptions of random search or directed search as special cases.<sup>9</sup> Previous work on other markets in Industrial Organization has also faced the challenge of identifying demand from matches (e.g. taxis ([Frechette et al. \(2017\)](#)), [Buchholz \(2020\)](#)) and bulk shipping ([Brancaccio et al. \(2020\)](#)). Unlike this previous work which has focused on settings where agents are relatively homogeneous, my method can be used where the *quality* of the match matters.

Third, previous work typically uses a steady-state analysis to tractably incorporate two-sided heterogeneity in a search model.<sup>10</sup> 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 [Kalouptsi \(2014\)](#) to cases where short-term contract data are available.

**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 \(2018\)](#) and [Dong et al. \(2020\)](#) who both calibrate models with search frictions). This paper advances this literature by

---

<sup>9</sup>[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 identification and estimation.

<sup>10</sup>An exception is [Lise and Robin \(2017\)](#), who model non-stationary distributions of searching agents by assuming Bertrand wage competition.

- 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 and two-sided heterogeneity. Some recent papers incorporate fluctuations into search models with homogeneous agents (for example, [Buchholz \(2020\)](#), [Frechette et al. \(2017\)](#)). A related set of papers study how fluctuations affect long-run firm entry and exit decisions ([Kalouptsi \(2014\)](#), [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.<sup>11</sup> 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 \(2008\)](#).<sup>12</sup> 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

---

<sup>11</sup>See [Rogerson et al. \(2005\)](#) for a survey

<sup>12</sup>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.

the value of drilling a well is related directly to lease bids ([Porter \(1995\)](#)).

## 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 than 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.<sup>13</sup> 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’. 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

---

<sup>13</sup>I calculate the HHI with the definition of ‘market share’ as the proportion of total contracts.



Table 1: Summary statistics for the dataset

Variable	Units	N	Mean	SD	10%	90%
Rig Price - New Contracts	1000s of USD/day	1896	54	33	22	101
Duration - New Contracts	Days	1896	81	85	40	160
Rig Price - Renegotiations	1000s of USD/day	818	45	25	24	75
Duration - Renegotiations	Days	818	83	85	35	153
Value	Millions of USD	2714	7.2	15	0.21	18
Complexity	Index	2714	0.83	0.41	0.39	1.35
Water Depth	Feet	2714	124	82	38	242
Monthly Utilization	% Rigs under contract	360	0.77	0.17	0.54	0.96

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 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 (for example, the rig has been scrapped). I use these data to compute how many 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.<sup>14</sup> In order to back out the quantity of hydrocarbons in the well, I then divide by average gas price (\$5.83).<sup>15</sup> My measure is a monotonic function of the expected oil and gas deposit size.

---

<sup>14</sup>This is motivated by the fact that offshore lease auctions are common value auctions and as the number of bidders  $n \rightarrow \infty$  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.

<sup>15</sup>The 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.

**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.<sup>16</sup> 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** Rigs are vertically differentiated. 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. 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.<sup>17</sup>

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 (around 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

---

<sup>16</sup>Details on the calculation of the Mechanical Risk Index can be found in Appendix A.2.

<sup>17</sup>The split is not quite exact because there are sometimes many rigs of exactly the same drilling depth.

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.<sup>18</sup>

## 2.3 Key features of the industry

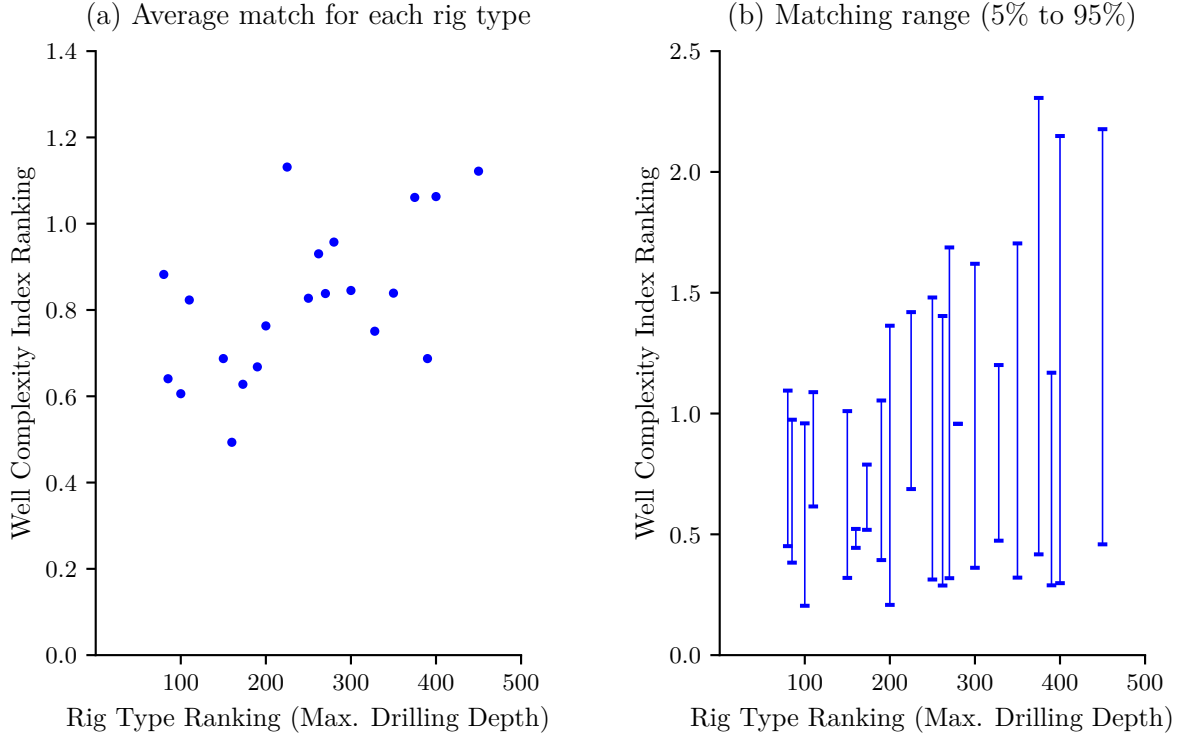
The offshore drilling industry is characterized by three key features: (1) sorting patterns; (2) booms and busts driven by oil and gas prices; (3) search frictions.

### Feature 1: Sorting patterns

Figure 2 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, 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'*<sup>19</sup>. Higher ranked rigs attract premium prices and are actively marketed as 'high-specification'.

Figure 2: Positive assortive matching: higher ranked rigs match with more complex wells



## Feature 2: Booms and busts

Figure 3 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.<sup>20</sup> Figure 3 shows that there is a strong correlation between gas prices and rig prices: rigs can command prices in excess of \$100 thousand per day when gas prices are high but this can fall to \$30 thousand 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.<sup>21</sup>

<sup>18</sup>I use 2 years as my cutoff for a ‘relationship’ because that is the definition used by Kellogg (2011).

<sup>19</sup><http://www.diamondoffshore.com/offshore-drilling-basics/offshore-rig-basics>

<sup>20</sup>According to the Energy Information Administration, total natural gas production in the Gulf of Mexico only accounts for around 5% of total production in the US: [https://www.eia.gov/special/gulf\\_of\\_mexico](https://www.eia.gov/special/gulf_of_mexico)

<sup>21</sup>From 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

Figure 3: Booms and busts

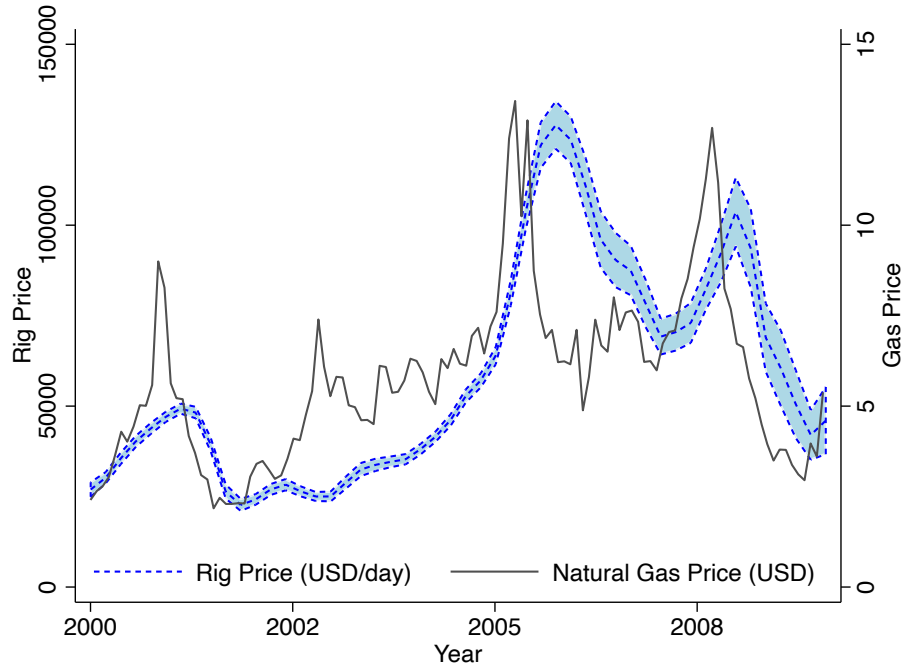
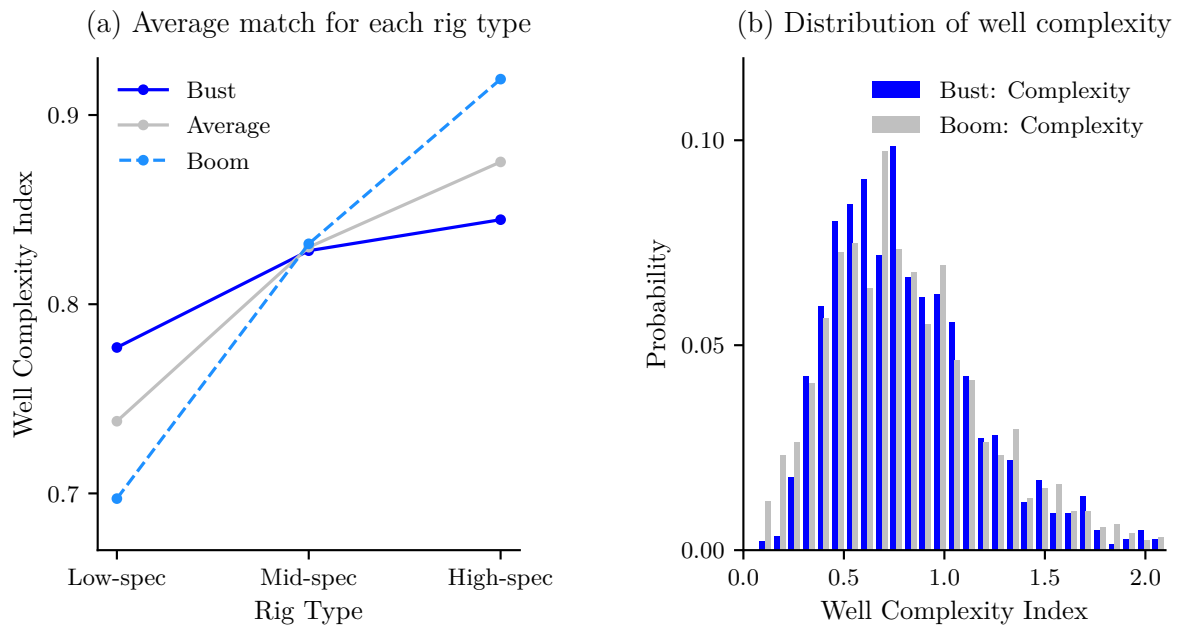


Figure 4: Matching patterns in booms and busts



**How booms and busts affect matching** Panel (a) of Figure 4 provides evidence consistent with stronger sorting in booms than busts. In 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 4 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 4. 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 (demand). Panel (b) of Figure 4 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. Despite this, the reduced form evidence is not conclusive. Therefore, as described further in the estimation section of the paper, I provide a new identification strategy to identify demand (the distribution of searching wells) from observed matches.

### 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.

**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:

$$\hat{p}_{it} = \mathbf{X}'\beta + \tilde{p}_{it}$$

---

time and subsequently resume full or near full utilization when business cycles change’.

Table 2: Evidence of price dispersion

	(a)	(b)
	Using aggregated rig types	Using disaggregated rig types
$1 - R^2$	0.43	0.38
$SD(\tilde{p}_{it})$	11	10
$SD(\hat{p}_{it})$	16	16

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

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  $\mathbf{X}$ , as well as interactions between rig types and well characteristics:

$$\mathbf{X} = \{\text{rig type, well complexity, well water depth, well value, gas price, rig availability, month FEs, year FEs, contractor FEs, rig owner FEs, contract duration}\}$$

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: 43% of price variation is unexplained when using the aggregated rig types and 38% 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.<sup>22</sup>

<sup>22</sup>One recent paper that documents a similar magnitude of price dispersion in a firm-to-firm search market is Salz (2017). Using similar descriptive regressions that control for buyer and seller characteristics Salz documents an unexplained price variation  $1 - R^2$  of 0.54 for non-brokered contracts in the trade waste industry.



### 3 The Model: Sequential Search with Booms and Busts

#### 3.1 Environment

**Agents** Agents are capital owners (owners of rigs) and projects (potential wells). The characteristics of project type  $x$  are:

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

where  $x_{\text{complexity}}$  is the complexity of a project,  $x_{\text{value}}$  is the value of oil and gas, and  $\tau$  is the duration of the project in months. I do not directly include the water depth in these characteristics because it is part of the well complexity index  $x_{\text{complexity}}$ .

There are  $K_t$  *potential projects* in each period, which are undrilled leases in the US Gulf of Mexico. The dependence on  $t$  is used to 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  $h(x)$  - the probability density of potential projects. Capital differs in its efficiency and the types are  $y \in \{\text{low, mid, high}\}$ . 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 objects at the time  $t$  state  $s_t$ . Within each period the timing is:

1. *Entry.* The set of potential projects is comprised of  $K_t$  draws from  $h(x)$ . Each potential project chooses whether to enter and search for capital. Denote the probability that type  $x$  enters (and searches for a match) at time  $t$  as  $e_t(x)$ , and the distribution of projects that enter and search ('searching projects') as  $f_t(x) = e_t(x)h(x) / \int_z e_t(z)h(z)dz$ .
2. *Search.* Searching projects contact available capital using a *search technology*. Denote the probability that a type  $x$  project targets type  $y$  capital at time  $t$  with the search technology as  $\omega_t(y|x)$ .
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 in the next period.<sup>23</sup>

4. *Contract extensions.* Existing matches are (possibly) extended with probability  $\eta$ .

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

$$v_t(x, y) = m_{0,y} + m_{1,y} \cdot x_{\text{complexity}} + m_{2,y} \cdot \mathbb{E}_t[\beta^\tau g_{t+\tau} x_{\text{value}}] \quad (1)$$

and so the total value of a contract is  $\sum_{k=0}^{\tau-1} \beta^k v_t(x, y)$ . The above equation can be broken down into two main components. The first component is the match value added:  $m_{0,y} + m_{1,y} \cdot x_{\text{complexity}}$ . This is defined on a per-period basis and so is summed over the entire  $\tau$  length contract. Here  $m_{0,y}$  and  $m_{1,y}$  are coefficients that vary with rig type  $y \in \{low, mid, high\}$ . Importantly, this equation captures complementarities between rig type and well type through  $m_{1,y}$ . For example, if high-efficiency capital is well-suited to undertaking complex projects then  $m_{1,y}$  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 ( $m_{1,high} > m_{1,mid} > m_{1,low}$ ) (Becker (1973)).

The second component is  $m_{2,y} \cdot \mathbb{E}_t[\beta^\tau g_{t+\tau} x_{\text{value}}]$ . This component captures the expected total value of oil and gas that is produced after the contract (which is  $\tau$  periods long) has elapsed. The variable  $g_{t+\tau}$  is the gas price after the contract has been completed. Therefore the dollar value of oil and gas in the well is  $g_{t+\tau} x_{\text{value}}$ . Since the covariate  $x_{\text{value}}$  is the ex-ante value of hydrocarbons in the well (proxied by the maximum bid in the lease auction), I include a parameter  $m_{2,y}$  which is defined as the weight that agents put on this ex-ante proxy when making decisions. The expected value of oil and gas scales with the length of the contract since longer contracts may involve constructing multiple similar wells over the same oil and gas deposit.

---

<sup>23</sup>The assumption that well owners exit immediately if unmatched is based on the fact that well owners tend to wait until the end of their lease to drill a well and so cannot continue to search. Previous literature suggests that well owners do this because 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)). If the lease elapses without drilling taking place then the well owner forfeits the rights to the lease, which leads to drilling at the end of the lease (this fact is also documented for the onshore oil and gas industry in Herrnstadt et al. (2020)).

**Summary** Agents make three main choices in the model (with the rest of the model determined endogenously in equilibrium). The first choice is the project entry decision. The second choice is the project targeting decision. The third choice is whether to match if agents successfully contact each other (and then 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 which 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 to a project owner from matching with a particular rig  $y$  is:

$$\Pi_t^{project}(x, y) = \underbrace{\sum_{k=0}^{\tau-1} \beta^k \left[ v_t(x, y) - p_t(x, y) \right]}_{\text{Value of the initial contract}} + \underbrace{\beta^\tau \mathbb{E}_t \left[ \eta \Pi_{t+\tau}^{project}(x, y) \right]}_{\text{Extension value}} \quad (2)$$

The project owner's value of matching  $\Pi_t^{project}(x, y)$  can be decomposed in the following way. For each period of the  $\tau$  length contract the project owner receives the match value  $v_t(x, y)$  minus the price  $p_t(x, y)$  to hire the capital. The contract will be extended with probability  $\eta$  which I assume is not dependent on the state.

#### 3.2.1 Entry and targeting

The timing is: 1. potential projects choose whether to enter 2. if a project enters then it targets its search.

### Step 1: Entry

A potential project with characteristics  $x$  will enter if:

$$\sum_{y \in \{low, mid, high\}} \omega_t(y|x) \pi_t(y|x) - c + \epsilon_t^{\text{entry}} \geq \epsilon_t^{\text{no entry}} \quad (3)$$

where  $c$  is the entry cost and  $\epsilon_t^{\text{entry}}, \epsilon_t^{\text{no entry}}$  are drawn from an IID type-I extreme value distribution (note that these draws are taken for each potential project but that I suppress this dependence). The component  $\omega_t(y|x)$  is the probability that a type  $x$  project targets type  $y$  capital and the component  $\pi_t(y|x)$  is the expected benefit to targeting capital type  $y$  (I discuss how these are constructed in more detail in the next section). Therefore, the first term on the left-hand-side is the expected benefit of entering. 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. Integrating over  $\epsilon_t^{\text{entry}}$  and  $\epsilon_t^{\text{no entry}}$  yields the probability that a project of type  $x$  enters:

$$e_t(x) = \frac{\exp\{\sum_{y \in \{low, mid, high\}} \omega_t(y|x) \pi_t(y|x) - c\}}{1 + \exp\{\sum_{y \in \{low, mid, high\}} \omega_t(y|x) \pi_t(y|x) - c\}} \quad (4)$$

### Step 2: (Partially directed) search

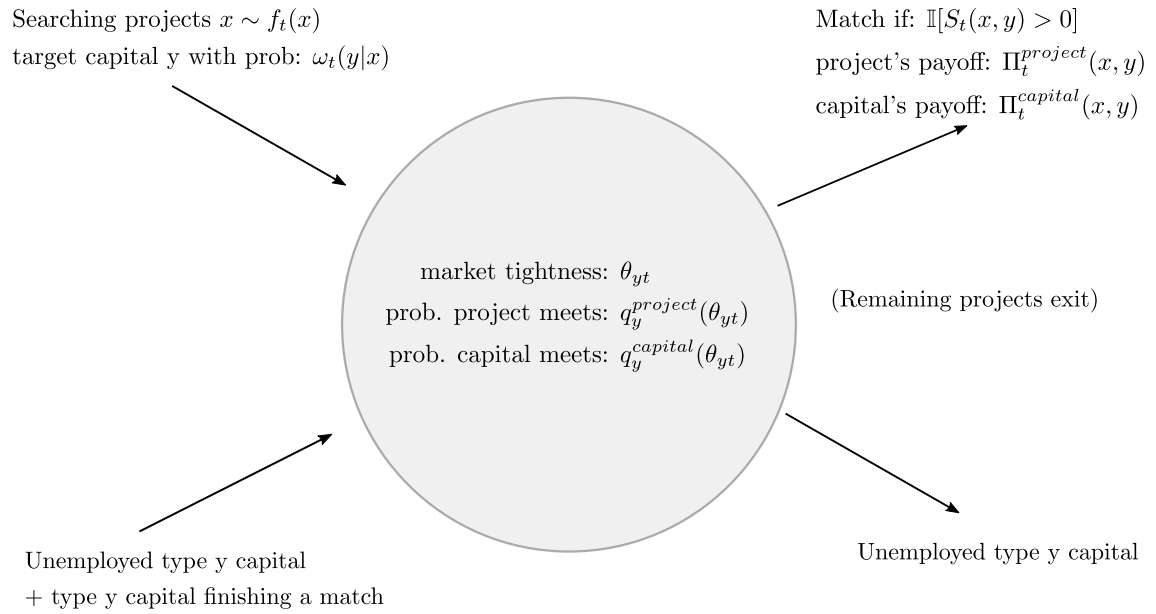
Each type of capital  $y \in \{low, mid, high\}$  is located in a submarket. Figure 5 provides a diagram of the meeting process. Meetings are determined randomly in each submarket. Therefore the expected value of a project  $x$  to targeting capital type  $y$  is:

$$\pi_t(y|x) = q_y^{\text{project}}(\theta_{yt}) \Pi_t^{\text{project}}(x, y) \quad (5)$$

The probability of successfully contacting capital in submarket  $y$  is  $q_y^{\text{project}}(\theta_{yt})$ , which is increasing in  $\theta_{yt}$ . 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

Figure 5: Matching within each submarket



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 the probability of matching is dependent on the market tightness  $\theta_{yt}$ . Finally, agents choose whether to match.

search technology - partially directed search - where the probability of targeting a submarket is given by:

$$\omega_t(y|x) = \frac{n_{yt} \exp\left(\gamma \pi_t(y|x)\right)}{\sum_{k \in \{low, mid, high\}} n_{kt} \exp\left(\gamma \pi_t(k|x)\right)} \quad (6)$$

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  $\gamma \rightarrow \infty$ .*

At the extreme case of random search  $\gamma = 0$ , the targeting weight  $\omega_t(y|x) = share_y$ , where  $share_y$  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  $y = \operatorname{argmax}_y \{\pi_t(y|x)\}$  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 time  $t$ :

**Proposition 1.** *Demand for capital type  $y$  at state  $s_t$  is given by:*

$$d_t(x|y) = q^{capital}(\theta_{yt}) \cdot \frac{\omega_t(y|x)e_t(x)h(x)}{\int_z \omega_t(y|z)e_t(z)h(z)dz} \quad (7)$$

$$d_t(\emptyset|y) = 1 - q^{capital}(\theta_{yt}) \quad (8)$$

where  $d_t(x|y)$  is the probability that capital type  $y$  will be contacted by project type  $x$ , and  $d_t(\emptyset|y)$  is the probability that capital type  $y$  is not contacted by any project. The probability of entry  $e_t(x)$ , targeting weights  $\omega_t(y|x)$ , and submarket tightness  $\theta_{yt}$ , are determined in equilibrium by Equations (2) - (6)

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 may increase when the market moves from a bust to a boom if the number of potential project draws  $K_t$  increases in a boom. Second, the *distribution* of trading opportunities  $d_t(x|y)$  will change due to different projects entering and different targeting behavior.

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 determined 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.

If capital is contacted by a project it faces the following tradeoff. *Accept the match* - and be unable to match for the duration of the contract - or *search again* for a better match:

$$\left\{ \underbrace{\Pi_t^{capital}(x, y)}_{\text{Accept match}}, \underbrace{b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y)}_{\text{Search again}} \right\} \quad (9)$$

Here  $\Pi_t^{capital}(x, y)$  is the profit from matching. The value of searching again for a better match is  $b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y)$ : the capital is not used for 1 period which results in a payoff  $= b_y$  but then the capital can search again with expected value  $\beta \mathbb{E}_t V_{t+1}^{capital}(y)$ . The term  $b_y$  is defined as an 'off-contract saving' which incorporates cost savings such as depreciation, accommodation for workers, and fuel, which are not incurred when a rig is unemployed.<sup>24</sup> It is dependent on the rig type  $y$  so that different types of rigs may have different costs. After one period of unemployment

---

<sup>24</sup>I model this as a 'cost saving' but it could also be modeled as an extra on-contract cost. Since the capital owner makes their decision about whether to match based on the relative value of the two options, this modeling choice is without loss of generality.

the rig can search again in period  $t + 1$  because it is not locked into a contract. The profit from matching is:

$$\Pi_t^{capital}(x, y) = \underbrace{\sum_{k=0}^{\tau-1} \beta^k p_t(x, y)}_{\text{Contract value}} + \beta^\tau \mathbb{E}_t \left[ \underbrace{\eta \Pi_{t+\tau}^{capital}(x, y)}_{\text{Extension}} + \underbrace{(1 - \eta) V_{t+\tau}^{capital}(y)}_{\text{No extension}} \right] \quad (10)$$

The profit from matching  $\Pi_t^{capital}(x, y)$  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  $\Pi_{t+\tau}^{capital}(x, y)$  if the contract is extended. If the contract is not extended then the rig will be available to search again and will receive  $V_{t+\tau}^{capital}(y)$ .

The value of searching (before projects contact capital) is:

$$V_t^{capital}(y) = \underbrace{\int_z d_t(z|y) \max \left\{ \Pi_t^{capital}(z, y), b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y) \right\} dz}_{\text{Exp. Value Of A Meeting}} + \underbrace{d_t(\emptyset|y) (b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y))}_{\text{No Meeting}} \quad (11)$$

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

## Bargaining

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

$$p_t = \operatorname{argmax}_{p_t} \left[ \Pi_t^{capital}(x, y) - b_y - \beta \mathbb{E}_t V_{t+1}^{capital}(y) \right]^\delta \left[ \Pi_t^{project}(x, y) \right]^{1-\delta} \quad (12)$$

Note that prices  $p_t$  are embedded in the value of matching for capital  $\Pi_t^{capital}(x, y)$  and projects  $\Pi_t^{project}(x, y)$ . The outside option for the capital is to search again the following period for another match, with value  $b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y)$ . 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:*

$$A_t(y) = \left\{ x : S_t(x, y) \geq 0 \right\} \quad (13)$$

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

$$S_t(x, y) = \Pi_t^{project}(x, y) + \Pi_t^{capital}(x, y) - b_y - \beta \mathbb{E}_t V_{t+1}^{capital}(y) \quad (14)$$

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

$$s_t = [g_t, n_{low,t}, n_{mid,t}, n_{high,t}] \quad (15)$$

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 ( $\tau_k = 0$ ) or it is in the final period of a contract and can match in the following period ( $\tau_k = 1$ ).

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

$$s_t = R_0 + R_1 s_{t-1} + \epsilon_t \quad (16)$$

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:

$$R_0 = \begin{bmatrix} r_g^0 \\ r_{low}^0 \\ r_{mid}^0 \\ r_{high}^0 \end{bmatrix}, \quad R_1 = \begin{bmatrix} r_g^1 & 0 & 0 & 0 \\ r_{low}^1 & r_{low}^2 & r_{low}^3 & r_{low}^4 \\ r_{mid}^1 & r_{mid}^2 & r_{mid}^3 & r_{mid}^4 \\ r_{high}^1 & r_{high}^2 & r_{high}^3 & r_{high}^4 \end{bmatrix} \quad (17)$$

In the matrix  $R_1$  I set the coefficients  $r_g^2 = r_g^3 = r_g^4 = 0$ . 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.<sup>25</sup>

**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).

---

<sup>25</sup>As previously discussed, this assumption seems reasonable given that total natural gas production in the Gulf of Mexico is a small fraction of global natural gas production.

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 again 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  $\{n_{yt}\}_{y \in \{low, mid, high\}}$ , demand for capital  $d_t(x|y)$ , targeting weights  $\omega_t(y|x)$ , entry probability  $e_t(x)$ , a distribution of searching projects  $f_t(x)$ , submarket tightness  $\{\theta_{yt}\}_{y \in \{low, mid, high\}}$ , and agents' state transition beliefs, that satisfy at each state  $s_t$ :*

1. *The distribution of searching projects  $f_t(x)$ , the targeting weights  $\omega_t(y|x)$ , the entry probability  $e_t(x)$ , and submarket tightness  $\{\theta_{yt}\}_{y \in \{low, mid, high\}}$ , determined by Equations (2) - (6)*
2. *Demand for capital  $d_t(x|y)$  determined by Equations (7) - (8)*
3. *Equilibrium prices  $p_t(x, y)$  determined by Nash bargaining: Equation (12)*
4. *Agents optimally choose whether to accept/wait if matched using Equation (13)*
5. *Updating rule for the distribution of capital  $\{n_{yt}\}_{y \in \{low, mid, high\}}$*
6. *Beliefs about the future evolution of states given by Equation (16)*

## 4 Estimation Strategy

### 4.1 Overview of the estimation

I calibrate the monthly discount factor as  $\beta = 0.99$ . There are five sets of parameters:

$$\lambda = \{\mathbf{r}, \lambda_0, \lambda_1, \lambda_2, \lambda_3\}$$

These parameters are: the state transitions  $\mathbf{r} = \{R_0, R_1, \sigma_\epsilon\}$ , the parameters  $\lambda_0 = \{\eta, b\}$ , the other parameters that characterize the profits from match  $\lambda_1 = \{m, \delta\}$ , the parameters that

characterize entered projects  $\lambda_2 = \{c, h(\cdot), \gamma\}$ , and the parameters that characterize the meeting technology and number of draws  $\lambda_3 = \{q_y^{capital}(\cdot), q_y^{project}(\cdot), K_t\}$ . I estimate these parameters in four steps.

1. Estimate the state transitions  $\mathbf{r}$  using maximum likelihood.
2. Compute the  $\lambda_0$  parameters, and then estimate the value functions.
3. Estimate the parameters  $\lambda_1$  that govern the value of a match between capital and projects.
4. 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 steps 2-4 in detail.

#### 4.1.1 Summary of the challenges and contributions

There are two main challenges and contributions. The first challenge is that demand (the distribution of searching wells) is not observed. This challenge has important implications for quantifying welfare and mismatch. For example, in a bust I observe high-efficiency rigs are allocated to simpler wells. Is this mismatch (there is a complex well searching that the high-efficiency rig should be allocated to), or is this due to demand (no complex wells are searching and so it is optimal to allocate the high-efficiency rig to the simple well)? Motivated by this challenge, I provide an identification strategy to recover unobserved demand (the distribution of searching projects) from observed matches. Broadly, the idea is to reverse-engineer the search and matching process by which searching wells form into observed matches. There are two main components to the process. First, there are acceptance sets, which are defined as matches where the value of accepting a match is higher than the outside option of rejecting the match and searching again. Second, there is the search technology (partially directed search). I show in this section how these components can be identified from the data.

The second challenge is that value functions - particularly the value of searching  $V_t^{capital}(y)$  - 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 Kalouptside (2014) to cases where short-term contract data are available.<sup>26</sup>

## Step 2: Estimating the value functions

Step 1 is straightforward so I begin by discussing step 2 in detail. The value of searching can be written non-parametrically in terms of data on matches, data on prices, and data on the probability of extending a contract. The intuition is as follows. For a capital owner, next period they may be matched with some type of project  $x$ . 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 capital is not matched then it receives  $b_y$ , the state is updated, and then it searches again. Therefore, capital owners' payoffs can be constructed just using the data (plus external information on the cost parameters  $\{b_{low}, b_{mid}, b_{high}\}$ ). 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 nor does it require data on the composition of searching projects (which I do not observe).

The following Proposition formalizes the above intuition (where I leave the proof to the Appendix).

**Proposition 3.** *Assume that the  $\{b_y\}_{y \in \{low, mid, high\}}$  values are known. Then, sufficient policy functions to estimate the value of searching  $V_t^{capital}(y)$  are:*

- $\bar{p}_t(\tau, y)$ : the average price at a given contract length  $\tau$ , capital type  $y$ , and state  $s_t$
- $\mathbb{P}_t(\tau = n|y)$ : the probability of rig type  $y$  matching with a  $\tau$  length contract (where  $n = 0$  denotes no match)
- $\eta$ : the extension probability

---

<sup>26</sup>Kalouptside (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.

*Proof.* See Appendix B.1. ■

Proposition 3 uses data in the form of policy functions and all of these policy functions are observed directly in the data. For example, the extension probability  $\eta$  is the average probability that a contract is extended in the data. I calibrate the capital owner’s marginal cost  $b_y$  using external data as  $b_{low} = b_{mid} = \$25$  thousand US dollars per day and  $b_{high} = \$50$  thousand US dollars per day.<sup>27</sup>

Putting it all together, I use the following steps to compute the value functions:

1. Estimate the policy functions.
2. Compute the value of search  $V_t^{capital}(y)$  using forward simulation.

In Appendix C.1 I provide more details about how I construct the policy functions. In Appendix C.2 I detail the forward simulation algorithm.

### Step 3: Estimating the value of a match

In this section I describe how to identify and estimate the parameters which characterize the value of a match. Combined with  $V_t^{capital}(y)$ , these elements characterize the surplus of a match and therefore the acceptance sets of agents  $A_t(y)$ . There are 8 unknown parameters: the bargaining parameter  $\delta$ , and the match value parameters  $m_2$  and  $\{m_{0,y}, m_{1,y}\}_{y \in \{low, mid, high\}}$ .

Taking the first order condition of the Nash Bargaining equation and rearranging results in the following Lemma (where I leave the proof to the Appendix).

---

<sup>27</sup>Specifically, I rely on Kaiser et al. (2013) that puts non-high specification rig operating costs at around \$45 thousand USD/Day when on-contract. Off-contract (when rigs are ‘warm-stacked’), a large rig owner Ensco (2015) puts operating expenses at around \$20 thousand USD/Day. Since the  $b_y$  terms are defined as the difference in costs between drilling and not drilling, I calibrate  $b_{low}$  and  $b_{mid}$  to the difference between these values at \$25 thousand USD/Day. Further, Kaiser et al. (2013) puts marginal costs for high-specification Jackup rigs at around twice that of other Jackup rigs (e.g. a rig owner Transocean reports expenses of \$87 thousand/day for high-specification jackup rigs vs \$46 thousand/day for other jackup rigs). Therefore, I set  $b_{high}$  to be twice the value of  $b_{low}$  and  $b_{mid}$  which is \$50 thousand USD/Day.

**Lemma 2.** (*Price equation*) Prices can be written as:

$$\begin{aligned}
\sum_{k=0}^{\tau-1} \beta^k p_t(x, y) &= \underbrace{\delta S_t(x, y)}_{\text{Surplus split}} \\
&\quad - \underbrace{\beta^\tau \mathbb{E}_t \left[ \eta \left( b_y + \beta V_{t+\tau+1}^{\text{capital}}(y) + \delta S_{t+\tau}(x, y) \right) + (1 - \eta) V_{t+\tau}^{\text{capital}}(y) \right]}_{\text{Extension adjustment}} \\
&\quad + \underbrace{b_y + \beta \mathbb{E}_t V_{t+1}^{\text{capital}}(y)}_{\text{Outside option}}
\end{aligned} \tag{18}$$

*Proof.* See Appendix B.2. ■

Lemma 2 is useful because it isolates prices on the left-hand side in terms of objects which are known (such as extensions and capital's value of searching), the (unknown) bargaining weight  $\delta$ , and the match surplus  $S_t(x, y)$  which can be constructed from  $m_{0,y}, m_{1,y}, m_2$  using the following Lemma.

**Lemma 3.** (*Surplus equation*) The match surplus can be written as:

$$\begin{aligned}
S_t(x, y) &= \underbrace{\sum_{k=0}^{\tau-1} \beta^k v_t(x, y)}_{\text{Initial contract}} \\
&\quad + \underbrace{\beta^\tau \mathbb{E}_t \left[ \eta \left( \sum_{k=0}^{\tau-1} \beta^k v_{t+\tau}(x, y) + \dots \right) + (1 - \eta) V_{t+\tau}^{\text{capital}}(y) \right]}_{\text{Extension}} \\
&\quad - \underbrace{b_y - \beta \mathbb{E}_t \left[ V_{t+1}^{\text{capital}}(y) \right]}_{\text{Outside option}}
\end{aligned} \tag{19}$$

where the match value is:

$$v_t(x, y) = m_{0,y} + m_{1,y} \cdot x_{\text{complexity}} + m_{2,y} \cdot \beta^\tau \mathbb{E}_t [g_{t+\tau} x_{\text{value}}]$$

*Proof.* Substitute the expressions for the value functions for  $\Pi_t^{\text{capital}}$  and  $\Pi_t^{\text{project}}$  into Equation (14) and then iterate forward. ■

I prove the following Lemma which shows how price variation can be used to identify the match value parameters up to the bargaining weight:

**Lemma 4.** *Using Equation (18) and Equation (19), the components  $\delta m_2$  and  $\{\delta m_{0,y}, \delta m_{1,y}\}_{y \in \{low, mid, high\}}$  can be identified from data on contracts and prices.*

*Proof.* See Appendix B.3. ■

I now discuss the intuition behind Lemma 4. The component  $\delta m_{1,y}$  - which governs the degree of complementarity between rig specification and well complexity - is identified through variation in matches with different  $x_{complexity}$  for a given rig type. The component  $\delta m_2$  - which governs how much weight agents put on their beliefs about the value of oil and gas in a well  $x_{value}$  - is identified by the degree to which matches with higher  $x_{value}$  pass through to prices. Since the other components are identified, the final component  $\delta m_{0,y}$  is identified by a constructive argument: match predicted prices from the model with the level of prices in the data.

To recover the match value parameters from Lemma 4 I need an estimate of the bargaining parameter. I use a strategy similar in spirit to Brancaccio et al. (2020). I focus on the year 2005 when the market was approximately in a steady-state with future expectations (computed in Stage 1) for the state evolution close to the current state and the gas price near its long-run average. I compute a non-stochastic steady state of the model. Then, using external data from the annual reports of the largest oil and gas companies in the Gulf of Mexico, I compute total revenue and operating margins. Intuitively, the bargaining weight is set so that the split of surplus between capital owners and project owners is consistent with these operating margins.<sup>28</sup>

I implement an estimation procedure based on the above identification strategy. For a guess of the match value parameters, I compute the predicted match surplus for each contract in the data using Equation (19) and forward simulation. Then, I compute predicted prices using Equation (18). I use non-linear least squares to search for the parameter guess that most closely fits the

---

<sup>28</sup>More concretely, in Appendix B.4 I derive a steady state of the model and rearrange it in terms of the bargaining weight  $\delta$ :

$$\delta = \frac{\sum_{k=0}^{\tau-1} \beta^k p(x, y) + \beta^\tau (1 - \eta) V^{capital}(y) - (1 - \beta^\tau \eta)(b_y + \beta V^{capital}(y))}{r(x, y) + \beta^\tau (1 - \eta) V^{capital}(y) - (1 - \beta^\tau \eta)(b_y + \beta V^{capital}(y))} \quad (20)$$

where I drop  $t$  subscripts and denote the total revenue from a contract as  $r(x, y) = \sum_{k=0}^{\tau-1} \beta^k v_t(x, y)$ . I obtain margins from the 2005 annual reports of the three largest non-major oil and gas companies operating in the Gulf of Mexico (I do not use the majors since it is difficult to disentangle their deepwater vs shallow water operations). Then, I convert margins to per-contract revenue using  $r(x, y) = p(x, y)/(1 - \text{margin})$ . I compute the above equation for each contract in 2005 and then take the average to recover the bargaining parameter.



observed price data.

#### Step 4: Estimating the remaining parameters

The remaining components to compute in Step 4 are:

- The targeting parameter  $\gamma$
- The distribution of potential projects  $h(\cdot)$
- The entry cost  $c$
- How potential well draws  $K_t$  change over time
- The meeting technology  $q_y^{capital}(\cdot)$  and  $q_y^{project}(\cdot)$

I identify these components using two sets of moments. The first set are moments related to the distribution of observed matches and I use these to identify the targeting parameter  $\gamma$  and  $h(\cdot)$ . The second set are moments related to the evolution of capital availability and I use these to identify the meeting technology and the number of well draws  $K_t$ . I calibrate the entry cost  $c$  using industry studies that decompose drilling expenditure into entry costs (‘pre-spud costs’) vs other costs. Using the average total payment to a rig owner as my measure for other drilling costs, I calibrate  $c = 1.41$  million USD.<sup>29</sup>

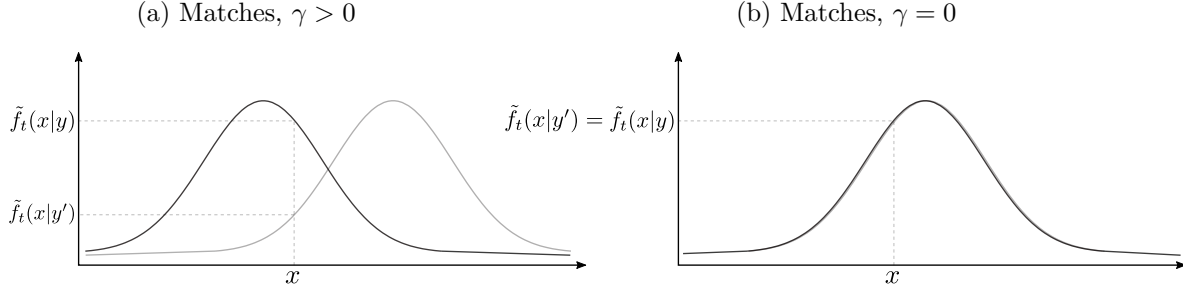
#### Identification

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.

---

<sup>29</sup>Specifically, I rely on [Hossain \(2015\)](#) which puts pre-spud drilling costs at around 18% of total expenses. Using this number, and setting other expenses to the mean total payment per contract to a rig (including extensions) (\$6.42 million), I calibrate the entry cost as  $c = (0.18/0.82) \times 6.42 = 1.41$  million dollars.

Figure 6: Identification of the targeting parameter  $\gamma$



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  $x$ .

**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?<sup>30</sup> 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 6. Both panels show the observed distributions of matches within a period conditional on a capital type  $y$ :  $\tilde{f}_t(x|y)$ . 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 type. 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

<sup>30</sup>As I confirm in the estimation, complex projects are best suited to matching with high-efficiency capital.

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  $\gamma$ , and the distribution of searching projects  $f_t(x)$ , are separately identified from the distribution of observed matches  $\tilde{f}_t(x|y)$  for  $y \in \{l, m, h\}$ .*

*Proof.* See Appendix B.5. ■

Note that since the acceptance sets can be constructed independently in Step 3 of the estimation, I can verify that the ‘overlapping acceptance set’ assumption holds in Proposition 4. I now discuss the intuition behind Proposition 4. 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  $t$ :

$$\ln \left( \tilde{f}_t(x|y) \right) - \ln \left( \tilde{f}_t(x|y') \right) = \gamma \left( \pi_t(x|y) - \pi_t(x|y') \right) + \ln \left( q_y^{project}(\theta_{yt}) / q_{y'}^{project}(\theta_{y't}) \right) \quad (21)$$

The above equation shows that the targeting parameter  $\gamma$  indexes how sensitive the probability of matching is to the expected value of matching  $\pi_t(x|y)$ .<sup>31</sup> 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** Once the targeting parameter has been pinned down, Proposition 4 shows that the distribution of searching projects  $f_t(x)$  can be identified by inverting observed matches through the search technology and acceptance sets.

---

<sup>31</sup>The  $\theta$  terms are also dependent on  $\gamma$  but can be removed by differencing again over  $y$ .

Then, given that the entry cost is calibrated, the distribution of potential projects  $h(\cdot)$  can be identified since I can map potential projects directly into observed matches through the entry decision and the search technology.

The above identification arguments do not hinge on a particular distribution of potential projects  $h(\cdot)$ . However, for estimation, I place the following parametric assumptions on  $h(\cdot)$ :

1. The value of hydrocarbons is a function of the well complexity:

$$x_{\text{value}} = \rho_0 + \rho_1 x_{\text{complexity}} + \rho_2 x_{\text{complexity}}^2 \quad (22)$$

where  $\rho_0$ ,  $\rho_1$ , and  $\rho_2$ , are parameters.

2. Contract durations are for either 2, 3, or 4 months, and are distributed independently of the other covariates with probability weights  $(\tau_2, \tau_3, \tau_4)$ , where  $\tau_4 = 1 - \tau_2 - \tau_3$ . Although the assumption that contract durations are distributed independently of the other covariates is restrictive, the entry condition generates complex changes in the distribution of searching wells over the cycle including correlations between contract duration and well complexity and value.

3. Well complexity is distributed as a two-component gaussian mixture model:

$$x_{\text{mri}} \sim \lambda \cdot \mathcal{N}(\mu_0, \sigma_0^2) + (1 - \lambda) \cdot \mathcal{N}(\mu_1, \sigma_1^2) \quad (23)$$

where the parameters  $\{\lambda, \mu_0, \mu_1, \sigma_0, \sigma_1\}$  need to be estimated.

**Identifying the remaining parameters** The remaining parameters are those that characterize the meeting technology and number of potential project draws. I assume that the meeting technology is in the following form:

$$q_y^{\text{project}}(\theta_{yt}) = a_{y,0} \left( 1 - \exp(-a_{y,0} \theta_{yt}) \right)$$

$$q_y^{\text{capital}}(\theta_{yt}) = \frac{1}{a_{y,0} \theta_{yt}} \left( 1 - \exp(-a_{y,0} \theta_{yt}) \right)$$

There are 6 parameters to be estimated in the above equations:  $\{a_{y,0}, a_{y,1}\}$  for  $y \in \{low, mid, high\}$ . The parameters  $a_{y,0}$  are identified by matching the mean utilization for each rig type. The meeting efficiency parameters  $a_{y,1}$  are identified by matching the variance of utilization for each rig type.

Finally, I use the specification that there are  $K_t = k_0 + k_1 g_t$  potential projects in each period, where  $g_t$  is the natural gas price and  $k_0$  and  $k_1$  are parameters.<sup>32</sup> The potential project draw parameters  $k_0, k_1$  are identified by the covariance between the gas price and capital utilization for each rig type.

### 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 first run an OLS regression using Equation (22) to recover the parameters  $\rho_0, \rho_1, \rho_2$ . Then, 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 use the average match of high and low specification capital in booms and busts (4 moments) which I denote  $m_1(\boldsymbol{\lambda}) = \{m_{y,T}(\boldsymbol{\lambda})\}_{y \in \{low, high\}, T \in \{boom, bust\}}$ . The difference between these moments is:

$$m_{high,T}(\boldsymbol{\lambda}) - m_{low,T}(\boldsymbol{\lambda}) = \frac{1}{\#T} \sum_{t \in T} \left[ \int_{A_t(h)} x_1 f_t(x_1 | y = h) dx_1 - \int_{A_t(l)} x_1 f_t(x_1 | y = l) dx_1 \right] \quad (24)$$

where  $\#T$  denotes the number of periods in  $T \in \{boom, bust\}$ . 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 (21), and this will be reflected in changes in  $m_{high,T}(\boldsymbol{\lambda}) - m_{low,T}(\boldsymbol{\lambda})$ . Under random search ( $\gamma = 0$ ) the only channel in the model for differences in matching patterns  $m_{high,T}(\boldsymbol{\lambda}) - m_{low,T}(\boldsymbol{\lambda}) > 0$  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.

The above moments are also useful in pinning down the parameters in the distribution of potential wells  $\{\mu_0, \mu_1, \sigma_0, \sigma_1, \lambda\}$ . In addition, I also include the mean well complexity matched by

---

<sup>32</sup>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.

mid-specification capital in booms and busts (2 moments), the standard deviation of well complexity matches (1 moment), and the probability of observing a 2 month and 3 month contract (2 moments). I denote the vector of these moments as  $m_2(\boldsymbol{\lambda})$ .

To identify the parameters  $\{a_{y,0}, a_{y,1}\}_{y \in \{low, mid, high\}}$  and  $k_0, k_1$  I include the moments related to patterns of capital utilization over the boom-bust cycle which I denote  $m_3(\boldsymbol{\lambda})$ . 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 18 simulated moments:

$$m_s(\boldsymbol{\lambda}) = (m_1(\boldsymbol{\lambda}), m_2(\boldsymbol{\lambda}), m_3(\boldsymbol{\lambda}))'$$

I fit the simulated moments to the empirical moments  $m_d$  by minimizing the following objective function:

$$(m_d - m_s(\boldsymbol{\lambda}))' \Omega (m_d - m_s(\boldsymbol{\lambda})) \tag{25}$$

Here  $\Omega$  is the weighting matrix which I set as a diagonal matrix. I set the weights for the  $m_1(\boldsymbol{\lambda})$  moments at 10, the weights for the standard deviation of well complexity at 1/10, and the other weights at 1.<sup>33</sup> 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 Appendix C.3.

---

<sup>33</sup>These weights are chosen to ensure that the model closely replicates moments such as the sorting patterns, which are of primary interest.

## 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):

$$R_0 = \begin{bmatrix} r_g^0 \\ r_{low}^0 \\ r_{mid}^0 \\ r_{high}^0 \end{bmatrix} = \begin{bmatrix} 0.64 \text{ (0.3)} \\ 4.33 \text{ (1.18)} \\ 5.7 \text{ (1.36)} \\ 9.2 \text{ (1.91)} \end{bmatrix} \quad \sigma_\epsilon = 1.02 \text{ (0.05)}$$

$$R_1 = \begin{bmatrix} r_g^1 & 0 & 0 & 0 \\ r_{low}^1 & r_{low}^2 & r_{low}^3 & r_{low}^4 \\ r_{mid}^1 & r_{mid}^2 & r_{mid}^3 & r_{mid}^4 \\ r_{high}^1 & r_{high}^2 & r_{high}^3 & r_{high}^4 \end{bmatrix} = \begin{bmatrix} 0.89 \text{ (0.03)} & 0 & 0 & 0 \\ -0.25 \text{ (0.1)} & 0.86 \text{ (0.06)} & -0.06 \text{ (0.07)} & 0.08 \text{ (0.07)} \\ -0.48 \text{ (0.12)} & 0.11 \text{ (0.06)} & 0.58 \text{ (0.08)} & 0.15 \text{ (0.07)} \\ -0.52 \text{ (0.17)} & -0.06 \text{ (0.09)} & 0.09 \text{ (0.1)} & 0.37 \text{ (0.08)} \end{bmatrix}$$

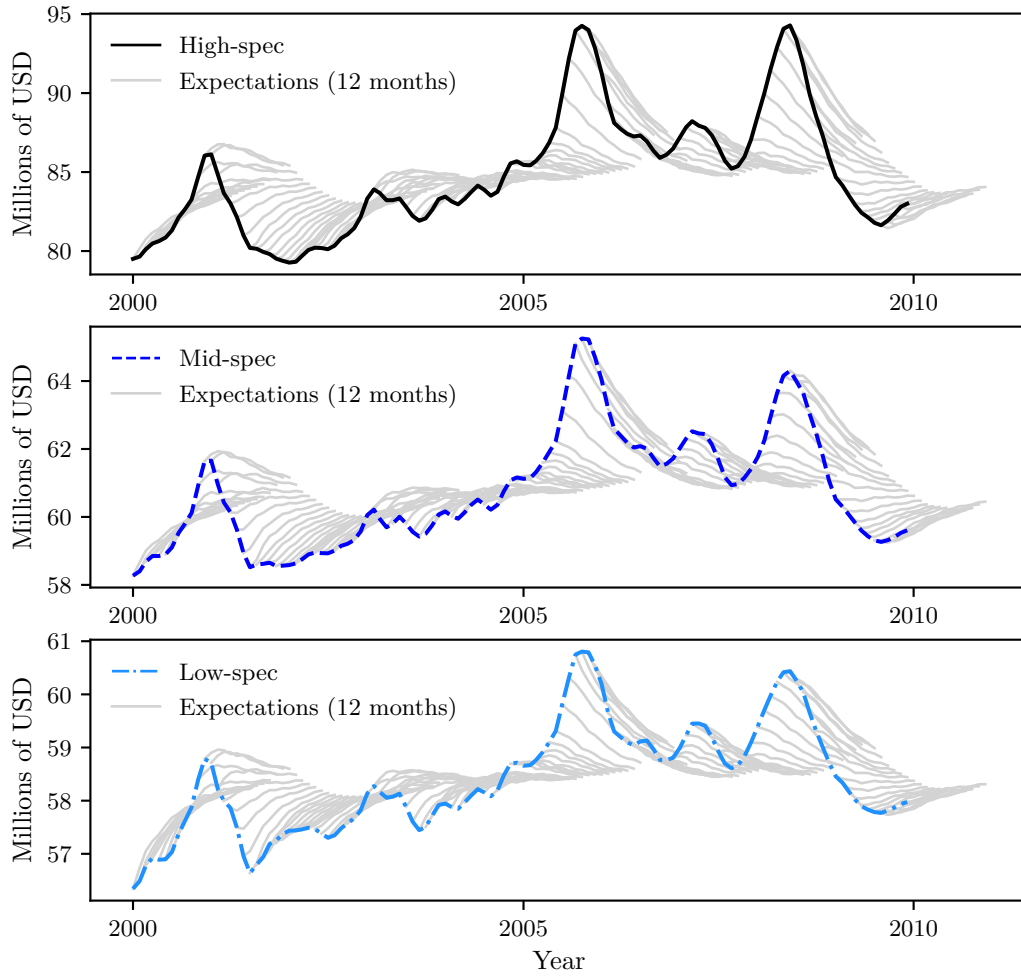
Since all the eigenvalues of  $R_1$  lie within the unit circle, the transition matrix is stationary.

### 5.2 Estimating the rig's value of searching

I compute the rig's value of searching non-parametrically using prices and agents' policy functions. First, I estimate the policy functions for prices, extensions, and matches (see Appendix C.1 for the details). For the extension probability I estimate a value of  $\eta = 0.3$  by taking the average probability that a contract is extended in the data. I then use forward simulation to recover the rig's value of searching using the algorithm in Appendix C.2.

Figure 7 illustrates the outcome of the procedure which is to construct the discounted sum of the average values to get the value of searching  $V_t^{capital}(y)$ . The light blue, dark blue, and black lines correspond to the value of searching in the current period  $V_t^{capital}(y)$ . 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 12 months - for example,  $\mathbb{E}_t V_{t+2}^{capital}(y)$ ,  $\mathbb{E}_t V_{t+3}^{capital}(y)$  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

Figure 7: The rig's value of searching  $V_t^{capital}(y)$



Notes: Points on the graph are plotted monthly. The gray lines correspond to the 12-month future beliefs of the value of searching.



are mean reverting.

### 5.3 Estimating the value of a match

Table 8 contains the results for the value of a match. I provide the estimates both in absolute dollar amounts (per day) and also as a proportion of the average dayrate for new contracts (\$54 thousand USD/Day). Recall that the bargaining weight is calibrated with external revenue data using the procedure described in the estimation section.

I now discuss the parameter estimates. First, the estimates for the  $m_{0,y}$  terms imply that, for very simple wells, low-specification rigs generate higher match values than mid- and high-specification rigs. This occurs because - from the perspective of a well-owner drilling a simple well - high-specification rigs are over-built, featuring complicated on-board technology that is difficult and costly to monitor.

Next, consider the estimates for  $m_{1,y}$ . The sign of these parameters is theoretically ambiguous. This is because there are both costs and benefits to matching different rigs with different well types and the parameters  $m_{1,y}$  are reduced-form expressions for how the *total* match value changes with well complexity. For example, a low-specification rig drilling more complex wells may reduce the match value through higher costs in the form of blowouts or extra materials after a drilling incident. By contrast, a high-specification rig may increase the match value since a rig operator can more effectively target oil and gas deposits (and this benefit may be more pronounced in complex wells which are more likely to contain deposits that are difficult to target). Consistent with this logic, I find that  $m_{1,low} < 0$  and  $m_{1,mid}, m_{1,high} > 0$ . Furthermore, the empirical ordering that  $m_{1,high} > m_{1,mid} > m_{1,low}$  implies that the match value is supermodular. Therefore, positive sorting would be efficient in a static model with no search frictions.

Finally, I find that the value weight  $m_2 = 8.2$ . The interpretation is that for a \$1 million increase in expectations about the value of oil and gas in a well (proxied by the lease bid), the total match value increases by \$8.2 thousand USD/Day. Scaling up this per-day figure over an average new contract duration of 81 days implies that agents weight a \$1 million increase in this lease bid proxy to a \$0.66 million increase in the total match value.

From these value of a match estimates, and estimates of capital's value of searching, I can

Figure 8: Match value estimates

Variable	Symbol	Value (\$1000s/Day)	Error (\$1000s/Day)	Value (vs Av. Price)	Error (vs Av. Price)
Match value					
High	$m_{0,high}$	37.8	6.3	0.7	0.12
	$m_{1,high}$	32.0	4.5	0.59	0.08
Mid	$m_{0,mid}$	48.2	6.1	0.89	0.11
	$m_{1,mid}$	23.5	6.2	0.44	0.11
Low	$m_{0,low}$	78.5	6.9	1.45	0.13
	$m_{0,low}$	-26.6	8.3	-0.49	0.15
Gas value weight	$m_2$	8.2	2.3	0.15	0.04
Bargaining parameter					
Bargaining parameter	$\delta$	0.4	n.a.		

Note: The n.a. denotes a calibrated value. The parameter value and standard error are provided both in USD per day, and as a proportion of the average price of a new contract (\$54 thousand USD/Day). Parameters are computed using the procedure described in the estimation section: the bargaining parameter is calibrated using external revenue data and the other parameters are estimated using contract price data and non-linear least squares. Water depth controls are included in the regression. The interpretation of the match value is the total value (incorporating both costs and benefits) of the match and therefore the sign of  $m_{1,y}$  is theoretically ambiguous. The value weight is the weight that agents place on the corresponding lease bid (which is a proxy for expectations about the value of oil and gas in a deposit) when determining rig prices.

construct the surplus of a match and therefore acceptance sets. I plot the acceptance sets over time in Figure 14 of Appendix D. Note that in Figure 14 it appears that there is a region of ‘overlapping acceptance sets’, which is the assumption that is required to identify the targeting parameter  $\gamma$  in Proposition 4.

## 5.4 The remaining parameters

Table 3 contains the estimated parameters from the Simulated Method of Moments. Overall the parameters seem reasonable. The estimated targeting parameter  $\gamma$  is 2.32. To get a sense of where this lies between random search and directed search I consider the probability that a complex well (I set  $x_{\text{complex}} = 2.0$ ) targets its optimal match (which is a high-specification rig) at approximately the average state.<sup>34</sup>

$$\omega_t(y = \text{high} | x = \text{complex}) = \begin{cases} 0.39 & \text{random search} \\ 0.44 & \text{estimated model} \\ 1 & \text{directed search} \end{cases}$$

The above example indicates the search technology that best rationalizes the data is closer to random search than directed search. Since imposing assumptions on the search technology has welfare implications, it is nevertheless still necessary to estimate this targeting parameter flexibly.

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.

## 5.5 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 sorting patterns in Figure 9. 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

---

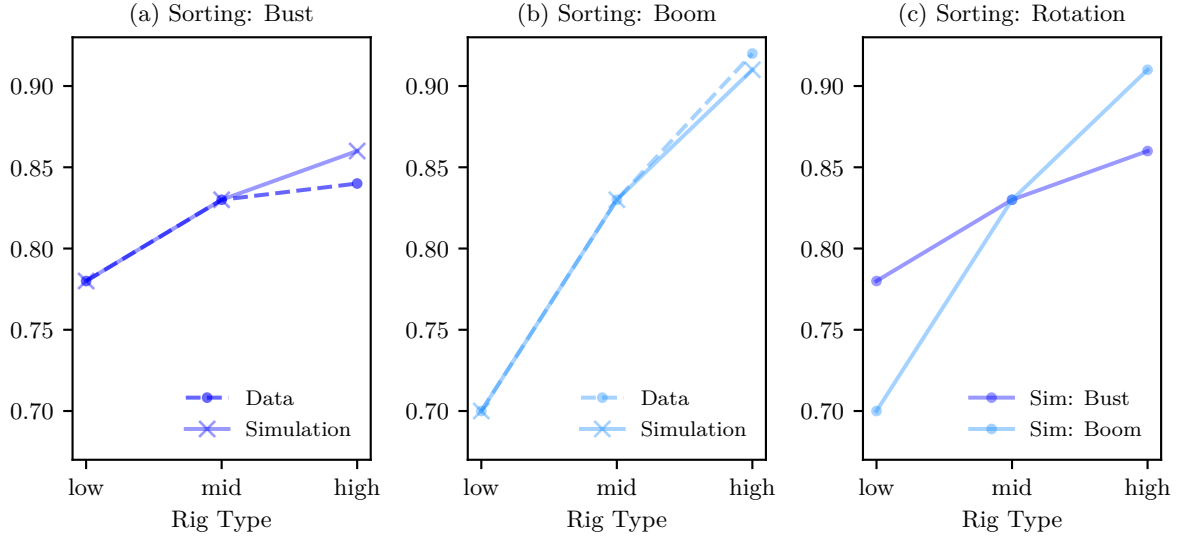
<sup>34</sup>Specifically, I choose the state halfway through the sample at January 2005 which is also between a boom and bust.

Table 3: Estimated remaining parameters

Variable	Symbol	Value	Symbol	Value
Demand	$\mu_0$	0.53	$\tau_2$	0.58
		(0.27, 0.65)		(0.57, 0.6)
	$\mu_1$	2.12	$\tau_3$	0.24
		(2.06, 2.15)		(0.23, 0.26)
	$\sigma_0$	0.48	$\rho_0$	0.029
		(0.35, 0.55)		(0.015, 0.044)
	$\sigma_1$	0.002	$\rho_1$	0.033
		(0.001, 0.03)		(0.004, 0.061)
Entry Cost	$\lambda$	0.9	$\rho_2$	-0.014
		(0.8, 0.93)		(-0.027, -0.002)
	$c$	\$1.4 Million		
		n.a.		
	$k_0$	70.7	$k_1$	14.1
		(70.6, 73.0)		(14.0, 19.3)
	$a_{low,0}$	0.55	$a_{mid,1}$	2.11
		(0.49, 0.72)		(1.31, 3.09)
Meeting Technology	$a_{low,1}$	1.97	$a_{high,0}$	1.54
		(1.04, 3.01)		(1.43, 3.49)
	$a_{mid,0}$	0.7	$a_{high,1}$	0.88
		(0.62, 0.83)		(0.41, 1.0)
Targeting Parameter	$\gamma$	2.32		
		(1.03, 3.88)		

Note: All parameters are estimated using the simulated method of moments, except for estimates for  $\rho_0, \rho_1, \rho_2$  which are computed using the OLS regression of  $x_{\text{value}} = \rho_0 + \rho_1 x_{\text{mri}} + \rho_2 x_{\text{mri}}^2$ , and the entry cost  $c$  which has a *n.a.* term in the confidence interval to denote a calibrated value. Confidence intervals (95%) in brackets. The confidence intervals are computed using 100 bootstrap replications, except for estimates for  $\rho_0, \rho_1, \rho_2$  which are computed using the standard errors from the OLS regression.

Figure 9: Data vs simulation



high-specification rigs are matched to more complex wells. Appendix D provides a complete comparison of how the simulated moments fit the data.

## 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  $T = \{2000 : 1, \dots, 2009 : 12\}$ , the total welfare is:

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

where in the ‘total value’ I subtract  $b_y$  for each period a rig of type  $y$  is under contract. For each of the counterfactuals I decompose the total effect into three components:

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

- **Quantity effect:** 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 Appendix C.

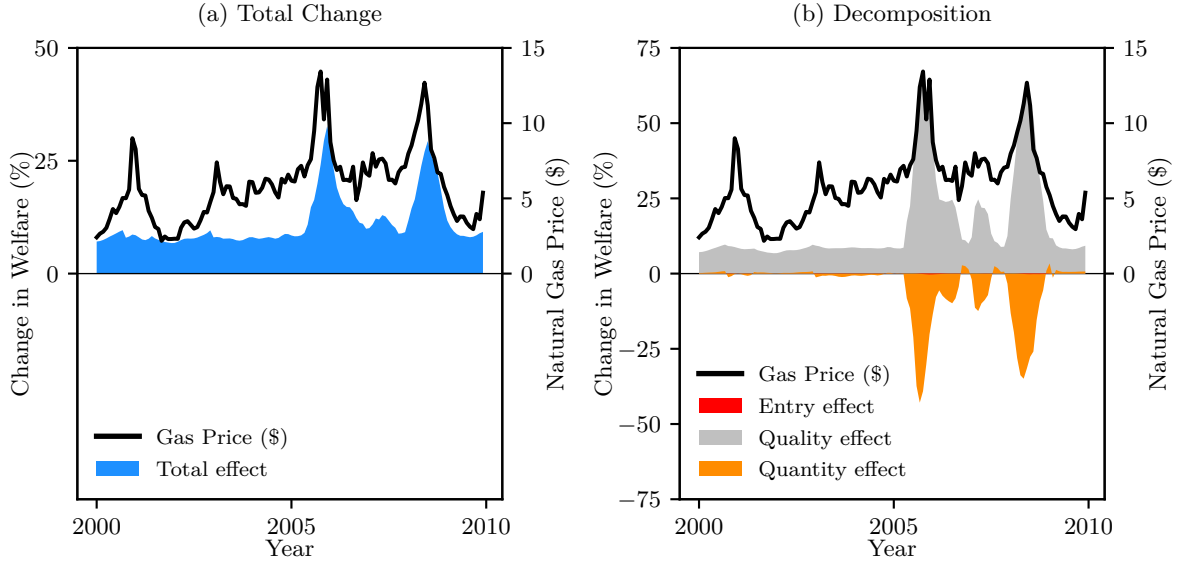
## 6.1 Quantifying the sorting effect

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 simulate an equilibrium that shuts down the two channels by which agents can be selective. First, I extend the acceptance sets to include all matches with positive match value, which prevents agents from rejecting matches based on changes in the outside option. Second, I set the targeting parameter  $\gamma = 0$  which shuts down the channel of agents using the search technology to selectively avoid rigs with high outside options.

I simulate the model using the empirical natural gas price. Starting from the 'no sorting effect' counterfactual, I compute the change in welfare when moving to the market benchmark. I keep the composition of searching wells the same in the 'no sorting effect' counterfactual as in the market benchmark.

I plot the results in Figure 10. Panel (a) plots the total change in welfare (joint profits). Welfare with the sorting effect is greater in every period and the total increase is 11.2%. In USD this is around 700 million over the 2000-2009 sample period. The effect is cyclical: the welfare increase in a boom is 14.6% compared to around 7.9% in a bust.

Figure 10: No sorting counterfactual results



(c) Summary of changes

	Boom	Bust	Average
Quality Effect	24.3%	7.9%	16.0%
Quantity Effect	-9.7%	0.1%	-4.8%
Entry Effect	0.0%	0.0%	0.0%
Total	14.6%	7.9%	11.2%

Note: This figure shows the change in welfare when moving from the ‘no sorting’ counterfactual to the market baseline. I keep the composition of searching wells fixed to the market benchmark resulting in an entry effect = 0. The welfare in dollars at the market baseline is 6.25 billion.

Panels (b) and (c) decompose how the sorting effect increases welfare: there are less matches (which by itself decreases welfare by -4.8%), but the remaining matches are of higher quality because agents are more selective (which increases welfare by 16.0%). Overall, the match quality effect dominates, which results in a net increase in welfare.

## 6.2 An intermediary that reduces search frictions

Next, I study the potential gains from an intermediary. This counterfactual highlights the effects of search frictions in the industry while also assessing the potential gains from recent advances in e-procurement in the industry.<sup>35</sup> I need to take a stand on the exact nature of the intermediary in the marketplace. I choose to set the targeting parameter  $\gamma = 20$  which is approximately 10 times the estimated value of the current targeting parameter  $\gamma$ . Conceptually, this intermediary can be thought of as an ‘Uber for rigs’ which introduces a vast improvement in the search technology of the industry. Similar to online marketplaces (like Uber) I allow free entry into the marketplace with the new search technology. I leave details on the implementation algorithm to Appendix C.5, which includes recomputing agents’ value functions and beliefs over the state transition.

Figure 11 illustrates the change in welfare due to an intermediary. Panel (a) shows that the intermediary increases welfare by around 28.2% over 2000-2009 compared to the market baseline. In USD this is around \$1.8 billion over the 2000-2009 sample period. This increase in welfare is slightly counter-cyclical: according to Panel (c) the welfare increase in booms is around 27.7% versus 28.7% in busts.

Panel (b) decomposes the effect of the intermediary. Overall, match quality is higher in every period and the total improvement in match quality is 30.0%. The magnitude of the quality effect is counter-cyclical: in booms agents are already quite selective and reject low quality matches (through the sorting effect). Therefore match quality gains are smaller in the boom (27.1%) than in the bust (33.1%).

The model produces interesting non-linear behavior for the quantity and entry effects. Note that how an intermediary will affect the quantity of matches and entry behavior is theoretically ambiguous. On one hand, better targeting may cause fewer matches to be rejected and therefore a higher quantity effect (which, by a similar argument, may also induce more entry). However, better targeting may reduce the quantity of matches through an equilibrium *congestion externality*: wells precisely target their best match, but congest the submarkets with high match values, resulting in a lower equilibrium probability of matching.<sup>36</sup> Overall, at the height of the boom

---

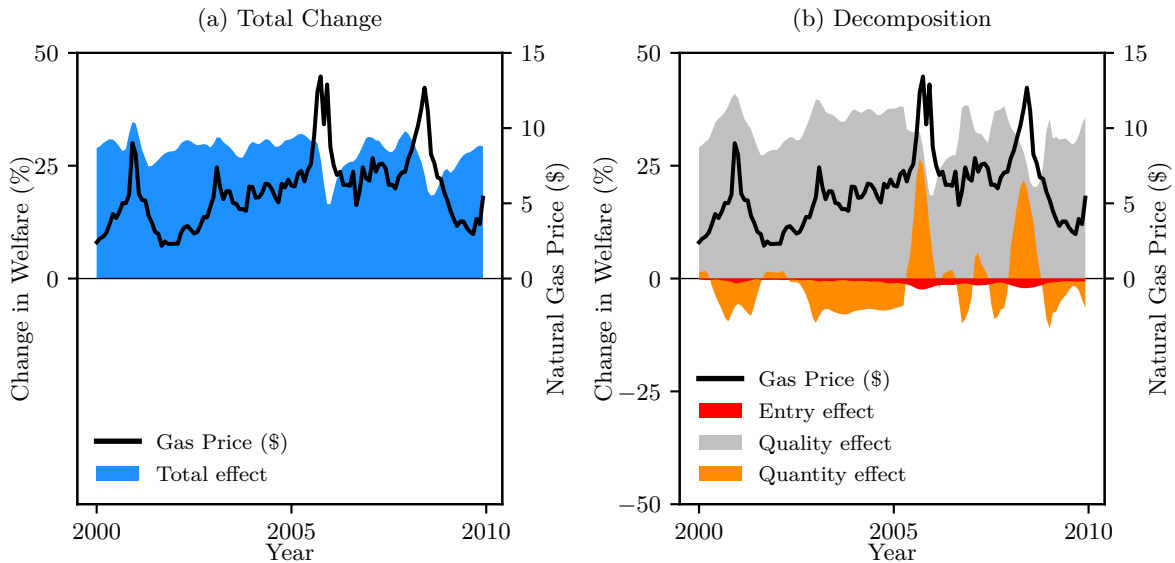
<sup>35</sup>For an early discussion about the potential benefits to e-procurement in the industry see Rothgerber (2002).

<sup>36</sup>Since agents also take into account the submarket tightness when making their targeting decision, how changes



(i.e. when the gas price is around \$10 or higher) the quantity effect is positive through agents avoiding matches that would be rejected. Conversely, at more moderate gas prices, the quantity effect is sometimes negative, indicating a congestion externality. Wells take these effects into account when making their entry decision: for example, there is a slight increase in entry during the height of the boom (which leads to a negative entry effect on welfare since more entry is costly).

Figure 11: Intermediary counterfactual results



(c) Summary of changes

	Boom	Bust	Average
Quality Effect	27.1%	33.1%	30.0%
Quantity Effect	2.0%	-3.9%	-0.9%
Entry Effect	-1.4%	-0.5%	-0.9%
Total	27.7%	28.7%	28.2%

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 6.25 billion. The entry effect corresponds to the total change in entry costs and so will be negative when there is more entry.

in the targeting parameter interact with the congestion externality is itself theoretically ambiguous. This is apparent in the decomposition graph at the lowest point of the bust (where the gas price is around \$2-3 dollars). Here, there is a very slight increase in the quantity of matches.

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 somewhat 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.<sup>37</sup> 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, 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.<sup>38</sup> 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.

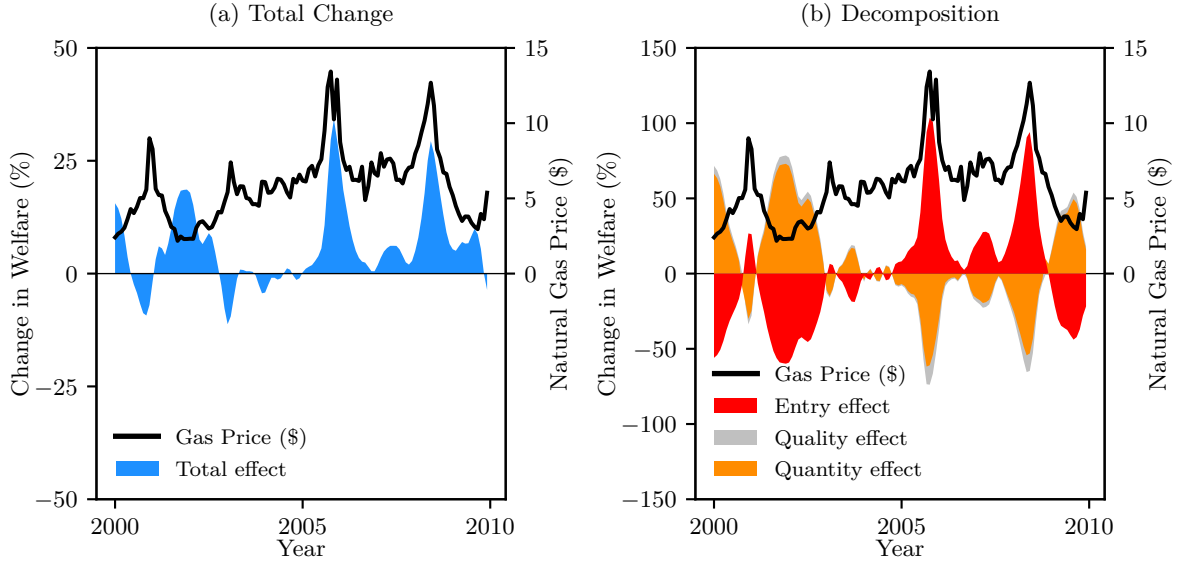
To understand the effects of these policies on drilling behavior I consider a counterfactual demand smoothing policy that results in the natural gas price being held at its long-run average. I am agnostic in the counterfactual about the exact implementation of taxes and subsidies that result in the smoother 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 Appendix C.5.

---

<sup>37</sup>Raghothamarao (2016) discusses how advances in e-procurement are being used in the oil and gas industry.

<sup>38</sup>Potter et al. (2017) summarizes the tax credits oil and gas producers receive in low-price environments.

Figure 12: Demand smoothing counterfactual results



(c) Summary of changes

	Boom	Bust	Average
Quality Effect	-3.3%	2.9%	-0.3%
Quantity Effect	-17.9%	31.2%	6.4%
Entry Effect	28.5%	-29.5%	-0.2%
Total	7.3%	4.6%	6.0%

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 6.25 billion. The entry effect corresponds to the total change in entry costs and so will be negative when there is more entry.

The results are depicted in Figure 12. Panel (a) shows that the smoothing policy results in large shifts in drilling activity. The direction of the total change is non-linear in the gas price. At the lowest point of the bust (when the gas price is around \$2-3) and at the height of the boom (when the gas price is around \$10) demand smoothing results in an increase in welfare. At more moderate gas prices (around \$5) the effect is close to zero or even negative.

Panel (b) illustrates the determinants of this non-linear total change. First, consider the effects of demand smoothing on the quality effect and the quantity effect. The changes here are straightforward: demand smoothing increases both the quality and quantity of matches in a bust

(it increase the quantity because there are more potential well draws and it increases the quality due to the sorting effect). Similarly, demand smoothing lowers the gas price in a boom and therefore reduces the quality and quantity of matches. The effects on the sign of the entry effect are also intuitive: more potential wells results in more entry in a bust (resulting in a negative entry effect through the decrease in profits due to higher total entry costs of 4.6%) and less entry in a boom. Much less straightforward - and what accounts for the non-linear total change - is which of these effects dominate (that is, whether the change in the quality and quantity effects combined outweighs the change in the entry effect). At the lowest point of the bust (when the gas price is around \$2-3) the quality and quantity improvements do outweigh the entry effect. In addition, at the height of the boom in the market baseline there appears to be inefficient entry which is reduced through demand smoothing. However, at more moderate gas prices around the long-run average, demand smoothing may cause small increase in inefficient entry that outweigh changes in the quality and quantity effects.

Given the counterfactual demand smoothing policy is extreme (it completely removes gas price cycles) - and despite dramatic changes in entry and the quantity of matches - the overall effect of the smoothing policy is relatively modest at 6.0%. This suggests that demand smoothing policies are somewhat ineffective in improving welfare.

[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 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

- Anderson, S. T., Kellogg, R. and Salant, S. W. (2018), ‘Hotelling under Pressure’, *Journal of Political Economy* **126**(3), 984–1026.
- Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2013), ‘Cross-Country Differences in Productivity: The Role of Allocation and Selection’, *American Economic Review* **103**(1), 305–34.
- Becker, G. S. (1973), ‘A Theory of Marriage: Part I’, *Journal of Political Economy* **81**(4), 813–846.
- Brancaccio, G., Kalouptsi, M. and Papageorgiou, T. (2020), ‘Geography, Transportation, and Endogenous Trade Costs’, *Econometrica* **88**, 657–691.
- Buchholz, N. (2020), ‘Spatial Equilibrium, Search Frictions and Efficient Regulation in the Taxi Industry’, *Working Paper*.
- Collard-Wexler, A. (2013), ‘Demand Fluctuations in the Ready-Mix Concrete Industry’, *Econometrica* **81**(3), 1003–1037.
- Collard-Wexler, A. and De Loecker, J. (2015), ‘Reallocation and Technology: Evidence from the US Steel Industry’, *American Economic Review* **105**(1), 131–71.
- Corts, K. S. (2008), ‘Stacking the Deck: Idling and Reactivation of Capacity in Offshore Drilling’, *Journal of Economics And Management Strategy* **17**(2), 271–294.
- Corts, K. S. and Singh, J. (2004), ‘The Effect of Repeated Interaction on Contract Choice: Evidence from Offshore Drilling’, *The Journal of Law, Economics, and Organization* **20**(1), 230–260.
- Dong, F., Wang, P. and Wen, Y. (2020), ‘A Search-Based Neoclassical Model of Capital Reallocation’, *European Economic Review* **128**.
- Eisfeldt, A. and Shi, Y. (2018), ‘Capital Reallocation’, *Annual Review of Financial Economics* **10**.
- EnSCO (2015), ‘Investor Presentation’. Retrieved From: <http://s1.q4cdn.com/651804090/files/presentations/2015/10292015-3Q15-Investor-Presentation.pdf>.
- Frechette, G., Lizzeri, A. and Salz, T. (2017), ‘Frictions in a Competitive Regulated Market: Evidence from Taxis’, *Working Paper*.

- Gavazza, A. (2016), ‘An Empirical Equilibrium Model of a Decentralized Asset Market’, *Econometrica* **84**(5), 1755–1798.
- Hagedorn, M., Law, T. H. and Manovskii, I. (2017), ‘Identifying Equilibrium Models of Labor Market Sorting’, *Econometrica* **85**(1), 29–65.
- Haile, P., Hendricks, K. and Porter, R. (2010), ‘Recent U.S. Offshore Oil and Gas Lease Bidding: A Progress Report’, *International Journal of Industrial Organization* **28**(4), 390 – 396.
- Hendricks, K. and Kovenock, D. (1989), ‘Asymmetric Information, Information Externalities, and Efficiency: The Case of Oil Exploration’, *The RAND Journal of Economics* **20**(2), 164–182.
- Hendricks, K. and Porter, R. H. (1996), ‘The Timing and Incidence of Exploratory Drilling on Offshore Wildcat Tracts’, *The American Economic Review* **86**(3), 388–407.
- Herrnstadt, E. M., Kellogg, R. and Lewis, E. (2020), ‘The Economics of Time-Limited Development Options: The Case of Oil and Gas Leases’, *NBER Working Paper* (No. 27165).
- Hossain, M. E. (2015), ‘Drilling Cost Estimation for Hydrocarbon Wells’, *Journal of Sustainable Energy Engineering* **3**(1).
- Ifrach, B. and Weintraub, G. Y. (2017), ‘A Framework for Dynamic Oligopoly in Concentrated Industries’, *The Review of Economic Studies* **84**(3), 1106–1150.
- Kaiser, M. (2007), ‘A Survey of Drilling Cost and Complexity Estimation Models’, *International Journal of Petroleum Science and Technology* **1**(1), 1–22.
- Kaiser, M. J., Snyder, B. and Pulsipher, A. G. (2013), ‘Offshore Drilling Industry and Rig Construction Market in the Gulf of Mexico’, *Coastal Marine Institute Report* p. 259.
- Kaiser, M. and Snyder, B. (2013), *The Offshore Drilling Industry and Rig Construction in the Gulf of Mexico*, Springer-Verlag, London.
- Kalouptsi, M. (2014), ‘Time to Build and Fluctuations in Bulk Shipping’, *American Economic Review* **104**(2), 564–608.
- Kellogg, R. (2011), ‘Learning by Drilling: Interfirm Learning and Relationship Persistence in the Texas Oilpatch’, *The Quarterly Journal of Economics* **126**(4), 1961–2004.

- Kellogg, R. (2014), ‘The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling’, *American Economic Review* **104**(6), 1698–1734.
- Lanteri, A. (2018), ‘The Market for Used Capital: Endogenous Irreversibility and Reallocation over the Business Cycle’, *American Economic Review* **108**(9), 2383–2419.
- Lentz, R. and Moen, E. R. (2017), ‘Competitive or Random Search?’, *Working Paper* .
- Lise, J. and Robin, J.-M. (2017), ‘The Macrodynamics of Sorting between Workers and Firms’, *American Economic Review* **107**(4), 1104–35.
- Ottonello, P. (2018), ‘Capital Unemployment’, *Working Paper* .
- Porter, R. H. (1995), ‘The Role of Information in U.S. Offshore Oil and Gas Lease Auctions’, *Econometrica* **63**(1), 1–27.
- Potter, K., Shirley, D., Manos, I. and Muraoka, K. (2017), ‘Tax Credits and Incentives for Oil and Gas Producers in a Low-Price Environment’, *Journal of Multistate Taxation and Incentives* **27**(2).
- Raghothamarao, V. (2016), ‘Strategic Supply Chain and Procurement Best Practices in Oil and Gas’, *Supply and Demand Chain Executive Magazine* .
- Rogerson, R., Shimer, R. and Wright, R. (2005), ‘Search-Theoretic Models of the Labor Market: A Survey’, *Journal of Economic Literature* **43**(4), 959–988.
- Rothgerber, U. (2002), ‘E-Procurement Still Evolving For E and P Industry’, *Oil and Gas Journal* **100**(16).
- Salz, T. (2017), ‘Intermediation and Competition in Search Markets: An Empirical Case Study’, *Working Paper* .
- Villas-Boas, S. B. (2007), ‘Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data’, *The Review of Economic Studies* **74**(2), 625–652.



# 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 *marinetraffic.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.<sup>39</sup>

In total there are 3131 contracts and 5553 wells in the full datasets for the years 2000-2009. The procedure successfully results in matching 2374 contracts and 5049 wells. I further impute the characteristics of 340 contracts if the contract was an extension/renegotiation. I also impute the contracts of 82 wells if the well was drilled subsequently to a merged well by the same rig-well owner pair. In total 2714 contracts ( 87 percent) and 5131 wells ( 92 percent) are matched.

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 13 contains an example of a complex versus a simple well.

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

$$\begin{aligned}\phi_1 &= \left( \frac{TD + WD}{1000} \right)^2 \\ \phi_2 &= \left( \frac{VD}{1000} \right)^2 \left( \frac{TD + HD}{VD} \right) \\ \phi_3 &= (MW)^2 \left( \frac{WD + VD}{VD} \right) \\ \phi_4 &= \phi_1 \sqrt{NS + \frac{MW}{(NS)^2}}\end{aligned}$$

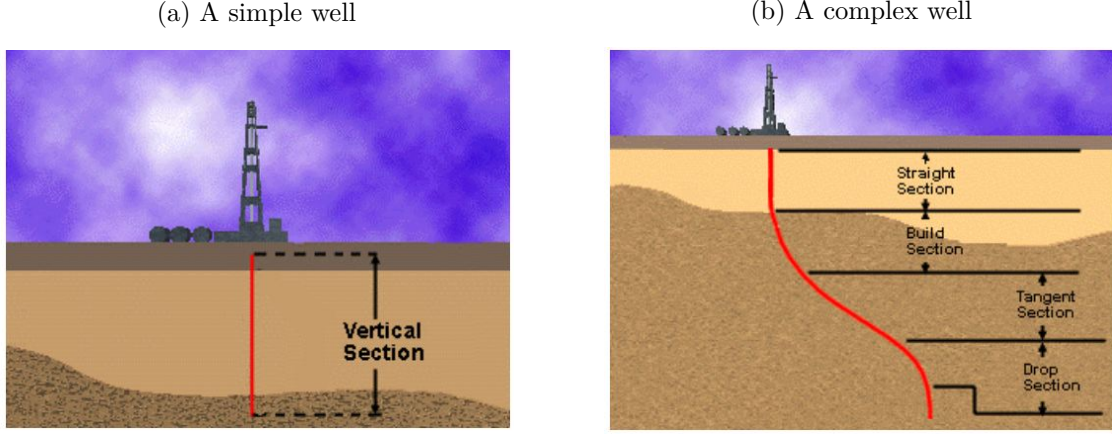
Here TD is total depth in feet, WD is water depth in feet, VD is vertical depth in feet, MW is mud weight in ppg, NS is the number of strings.

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

---

<sup>39</sup>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.

Figure 13: Diagram of a simple vs 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/>

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 the value of search  $V_t^{capital}(y)$  can be computed from the data and the  $b_y$  parameter for  $y \in \{low, mid, high\}$ .

Writing out the value of searching for rig  $y$  at time  $t$ :

$$V_t^{capital}(y) = \int_x d_t(x|y) \max \left\{ \Pi_t^{capital}(x, y), b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y) \right\} dx + d_t(\emptyset|y) \cdot \left( b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y) \right) \quad (26)$$

$$= \int_{x \in A_t(y)} d_t(x|y) \Pi_t^{capital}(x, y) dx + \left( \int_{x \notin A_t(y)} d_t(x|y) dx + d_t(\emptyset|y) \right) \cdot \left( b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y) \right) \quad (27)$$

$$= \sum_{n \in T} \mathbb{P}_t(\tau = n|y) \cdot \mathbb{E}_t \mathbb{E}_{x_{-\tau}} \left[ \Pi_t^{capital}(x, y) \middle| \tau = n \right] + \mathbb{P}_t(\tau = 0|y) \cdot \left( b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y) \right) \quad (28)$$

where, in the third equality,  $T$  is the set of possible contract lengths (I later set  $T = \{2, 3, 4\}$  months) and  $\mathbb{E}_{x_{-\tau}}[\cdot | \tau = n]$  is an expectation taken over all the covariates in a potential project (except  $\tau$ ) conditional on  $\tau = n$  (i.e. the expectation is over the covariates  $x_{\text{value}}$  and  $x_{\text{complexity}}$ ). Recall that value of a match to the capital owner is:

$$\Pi_t^{capital}(x, y) = \sum_{k=0}^{\tau-1} \beta^k p_t(x, y) + \beta^\tau \mathbb{E}_t \left[ \eta \Pi_{t+\tau}^{capital}(x, y) + (1 - \eta) V_{t+\tau}^{capital}(y) \right] \quad (29)$$

Taking conditional expectations of both sides of Equation (29) with respect to the covariates (except  $\tau$ ):

$$\begin{aligned} \mathbb{E}_{x_{-\tau}} \left[ \Pi_t^{capital}(x, y) \middle| \tau = n \right] &= \sum_{k=0}^{n-1} \beta^k \mathbb{E}_{x_{-\tau}} \left[ p_t(x, y) \middle| \tau = n \right] \\ &\quad + \beta^n \mathbb{E}_t \left( \eta \mathbb{E}_{x_{-\tau}} \left[ \Pi_{t+n}^{capital}(x, y) \middle| \tau = n \right] + (1 - \eta) \mathbb{E}_{x_{-\tau}} \left[ V_{t+n}^{capital}(y) \middle| \tau = n \right] \right) \end{aligned} \quad (30)$$

Note that the per-period payoff  $\mathbb{E}_{x_{-\tau}} \left[ p_t(x, y) \middle| \tau = n \right] = \bar{p}_t(y, \tau = n)$ , which is the average price (conditional on the rig type, state, and contract length), and is directly observed in the data. Furthermore,  $\mathbb{E}_{x_{-\tau}} \left[ V_{t+n}^{capital}(y) \middle| \tau = n \right] = V_{t+n}^{capital}(y)$ . Therefore, substituting in for  $\bar{p}_t(y, \tau = n)$  and writing out future values of  $\Pi^{capital}$  the above equation can be written as:

$$\begin{aligned} \mathbb{E}_{x_{-\tau}} \left[ \Pi_t^{capital}(x, y) \middle| \tau = n \right] &= \sum_{k=0}^{n-1} \beta^k \bar{p}_t(y, \tau = n) \\ &\quad + \beta^n \mathbb{E}_t \left[ \eta \left( \sum_{k=0}^{n-1} \beta^k \bar{p}_{t+n}(y, \tau = n) + \dots \right) + (1 - \eta) V_{t+n}^{capital}(y) \right] \end{aligned} \quad (31)$$

Equation (31) shows that the term  $\mathbb{E}_{x_{-\tau}} \left[ \Pi_t^{capital}(x, y) \middle| \tau = n \right]$  can be constructed from the extension probability  $\eta$ , future predictions of the average price  $\bar{p}$ , and future values of  $V^{capital}$ . Substituting Equation (31) into Equation (28):

$$V_t^{capital}(y) = \sum_{n \in T} \mathbb{P}_t(\tau = n|y) \cdot \left\{ \begin{aligned} & \sum_{k=0}^{n-1} \beta^k \bar{p}_t(\tau = n|y) + \beta^n \mathbb{E}_t \left[ \eta \left( \sum_{k=0}^{n-1} \beta^k \bar{p}_{t+n}(\tau = n|y) + \dots \right) + (1 - \eta) V_{t+n}^{capital}(y) \right] \\ & \left. \right\} + \mathbb{P}_t(\tau = 0|y) \cdot \left( b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y) \right) \end{aligned} \quad (32)$$

Notice that the above expression implies that  $V^{capital}(y|s_t)$  can be written in terms of the average price, the probability of matching different length contracts, the extension probability, the value  $b_y$ , and future values of  $V^{capital}$ , which proves the result.

## B.2 Proof of Lemma 2

Under Nash bargaining the surplus is split in the following way:

$$\Pi_t^{capital}(x, y, p_t(x, y)) - b_y - \beta \mathbb{E}_t V_{t+1}^{capital}(y) = \delta S_t(x, y) \quad (33)$$

Substituting in Equation (10) into  $\Pi_t^{capital}(x, y, p_t(x, y))$ :

$$\sum_{k=0}^{\tau-1} \beta^k p_t(x, y) + \beta^\tau \mathbb{E}_t \left[ \eta \Pi_{t+\tau}^{capital}(x, y, p_{t+\tau}(x, y)) + (1 - \eta) V_{t+\tau}^{capital}(y) \right] - \beta \mathbb{E}_t V_{t+1}^{capital}(y) = \delta S_t(x, y) \quad (34)$$

Iterating Equation (33) forward  $\tau$  periods, substituting for  $\Pi_{t+\tau}^{capital}$ , and rearranging to get the LHS in terms of prices:

$$\begin{aligned} \sum_{k=0}^{\tau-1} \beta^k p_t(x, y) &= \delta S_t(x, y) - \beta^\tau \mathbb{E}_t \left[ \eta \left( b_y + \beta V_{t+\tau+1}^{capital}(y) + \delta S_{t+\tau}(x, y) \right) + (1 - \eta) V_{t+\tau}^{capital}(y) \right] \\ &\quad + b_y + \beta \mathbb{E}_t V_{t+1}^{capital}(y) \end{aligned} \quad (35)$$

## B.3 Proof of Lemma 4

### Identification of $m_{1,y}$

Denote  $\Delta$  as an operator that takes the difference between two matches which differ only in the project complexity (that is, holding the capital type and other covariates fixed). Taking first differences of

Equation (35) with respect to  $x_{\text{complexity}}$ :

$$\sum_{k=0}^{\tau-1} \beta^k \Delta p_t(x, y) = \Delta \left[ \delta S_t(x, y) - \eta \beta^\tau \mathbb{E}_t \delta S_{t+\tau}(x, y) \right] \quad (36)$$

$$= \delta \sum_{k=0}^{\tau-1} \beta^k \left[ (m_{1,y} + \eta \beta^\tau m_{1,y} + \eta^2 \beta^{2\tau} m_{1,y} \dots) - (\eta \beta^\tau m_{1,y} + \eta^2 \beta^{2\tau} m_{1,y} \dots) \right] \Delta x_{\text{complexity}} \quad (37)$$

$$= \delta m_{1,y} \cdot \sum_{k=0}^{\tau-1} \beta^k \Delta x_{\text{complexity}} \quad (38)$$

Equation (38) can be further simplified to:  $\Delta p_t(x, y) = \delta m_{1,y} \cdot \Delta x_{\text{complexity}}$ . Since we observe the difference in prices  $\Delta p_t(x, y)$  and the difference in complexity  $\Delta x_{\text{complexity}}$ ,  $m_{1,y}$  is identified up to the bargaining parameter  $\delta$ .

### Identification of $m_2$

By a similar argument to identification of  $m_{1,y}$ , denote  $\Delta'$  as an operator that takes the difference between two matches which differ only in the project value (that is, holding the capital type and other covariates fixed). Taking first differences of Equation (35) with respect to  $x_{\text{value}}$ :

$$\sum_{k=0}^{\tau-1} \beta^k \Delta' p_t(x, y) = \Delta' \left[ S_t(x, y) - \eta \beta^\tau \mathbb{E}_t S_{t+\tau}(x, y) \right] \quad (39)$$

$$= \delta m_2 \cdot \sum_{k=0}^{\tau-1} \beta^k \left[ (\beta^\tau \mathbb{E}_t [g_{t+\tau}] + \eta \beta^{2\tau} \mathbb{E}_t [g_{t+2\tau}] + \eta^2 \beta^{3\tau} \mathbb{E}_t [g_{t+3\tau}] \dots) - (\eta \beta^{2\tau} \mathbb{E}_t [g_{t+2\tau}] + \eta^2 \beta^{3\tau} \mathbb{E}_t [g_{t+3\tau}] \dots) \right] \Delta' x_{\text{value}} \quad (40)$$

$$= \delta m_2 \cdot \sum_{k=0}^{\tau-1} \beta^{\tau+k} \mathbb{E}_t [g_{t+\tau}] \Delta' x_{\text{value}} \quad (41)$$

The above Equation can be further simplified to:  $\Delta' p_t(x, y) = \delta m_2 \cdot \beta^\tau \mathbb{E}_t [g_{t+\tau}] \Delta' x_{\text{value}}$ . Since we observe the difference in prices  $\Delta' p_t(x, y)$  and the difference in complexity  $\Delta' x_{\text{value}}$ ,  $m_2$  is identified up to the bargaining parameter  $\delta$ .

### Identification of $m_{0,y}$

All the parameters that govern Equation (35) (including the parameters that govern the match surplus) are now identified, with the exception of  $\delta$  and  $m_{0,y}$ . The parameter  $m_{0,y}$  is then identified by a constructive argument: match predicted prices from the model with the level of prices in the data.

## B.4 Steady state prices and the bargaining parameter $\delta$

For this section, since I solve for a steady state, I drop  $t$  subscripts. Further, I denote the total revenue from a contract as  $r(x, y) = \sum_{k=0}^{\tau-1} \beta^k v_t(x, y)$ . Using the surplus split under Nash Bargaining:

$$\Pi^{capital}(y) - b_y - \beta V^{capital}(y) = \delta S(x, y) = \delta [\Pi^{capital}(y) + \Pi^{project}(y) - b_y - \beta V^{capital}(y)] \quad (42)$$

The left-hand-side is:

$$\begin{aligned} \Pi^{capital}(y) - b_y - \beta V^{capital}(y) &= \sum_{k=0}^{\tau-1} \beta^k p(x, y) \\ &\quad + \beta^\tau \left[ \eta \sum_{k=0}^{\tau-1} \beta^k p(x, y) + (1 - \eta) V^{capital}(y) \right] \\ &\quad + \eta \beta^{2\tau} \left[ \eta \sum_{k=0}^{\tau-1} \beta^k p(x, y) + (1 - \eta) V^{capital}(y) \right] + \dots \\ &\quad - b_y - \beta V^{capital}(y) \end{aligned} \quad (43)$$

$$= \sum_{k=0}^{\tau-1} \beta^k p(x, y) + \beta^\tau \cdot \frac{\eta \sum_{k=0}^{\tau-1} \beta^k p(x, y) + (1 - \eta) V^{capital}(y)}{1 - \beta^\tau \eta} - b_y - \beta V^{capital}(y) \quad (44)$$

The surplus term  $S(x, y)$  on the right-hand-side is:

$$\begin{aligned} S(x, y) &= r(x, y) \\ &\quad + \beta^\tau [\eta r(x, y) + (1 - \eta) V^{capital}(y)] \\ &\quad + \eta \beta^{2\tau} [\eta r(x, y) + (1 - \eta) V^{capital}(y)] + \dots \\ &\quad - b_y - \beta V^{capital}(y) \end{aligned} \quad (45)$$

$$= r(x, y) + \beta^\tau \cdot \frac{\eta r(x, y) + (1 - \eta) V^{capital}(y)}{1 - \beta^\tau \eta} - b_y - \beta V^{capital}(y) \quad (46)$$

Substituting Equation (44) and Equation (46) into Equation (42) and rearranging in terms of  $\delta$  generates the result that:

$$\delta = \frac{\sum_{k=0}^{\tau-1} \beta^k p(x, y) + \beta^\tau (1 - \eta) V^{capital}(y) - (1 - \beta^\tau \eta)(b_y + \beta V^{capital}(y))}{r(x, y) + \beta^\tau (1 - \eta) V^{capital}(y) - (1 - \beta^\tau \eta)(b_y + \beta V^{capital}(y))} \quad (47)$$

## B.5 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. The observed distribution of projects that type  $y$  capital matches with is:

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

At this stage the meeting technology  $q_y^{project}(\theta_{yt})$ , 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 complete the proof in four steps:

1. Identify the targeting weights  $\omega_t(y|x)$
2. Identify the meeting technology  $q_y^{project}(\theta_{yt})$
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:

$$\omega_t(y|x) = \frac{n_{yt} \exp\left(\gamma \pi_t(y|x)\right)}{\sum_{k \in \{l, m, h\}} n_{kt} \exp\left(\gamma \pi_t(k|x)\right)} \quad (48)$$

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  $t$ :

$$\ln\left(\tilde{f}_t(x|y)\right) - \ln\left(\tilde{f}_t(x|y')\right) = \ln\left(\omega_t(y|x)/\omega_t(y'|x)\right) + \ln\left(q_y^{project}(\theta_{yt})/q_{y'}^{project}(\theta_{y't})\right) \quad (49)$$

$$= \gamma\left(\pi_t(x|y) - \pi_t(x|y')\right) + \ln\left(n_{yt}/n_{y't}\right) + \ln\left(q_y^{project}(\theta_{yt})/q_{y'}^{project}(\theta_{y't})\right) \quad (50)$$

$$= \gamma q_y^{project}(\theta_{yt}) \Pi_t^{project}(x, y) + \gamma q_{y'}^{project}(\theta_{y't}) \Pi_t^{project}(x, y') + \ln\left(n_{yt}/n_{y't}\right) + \ln\left(q_y^{project}(\theta_{yt})/q_{y'}^{project}(\theta_{y't})\right) \quad (51)$$

Here the third equality follows from substituting the expression:

$$\pi_t(y|x) = q_y^{project}(\theta_{yt}) \Pi_t^{project}(x, y) \quad (52)$$

Differencing Equation (51) over two points  $x$  and  $x'$ :

$$\begin{aligned} & \left( \ln\left(\tilde{f}_t(x|y)\right) - \ln\left(\tilde{f}_t(x|y')\right) \right) - \left( \ln\left(\tilde{f}_t(x'|y)\right) - \ln\left(\tilde{f}_t(x'|y')\right) \right) = \\ & \gamma q_y^{project}(\theta_{yt}) \cdot \left( \Pi_t^{project}(x, y) - \Pi_t^{project}(x', y) \right) + \gamma q_{y'}^{project}(\theta_{y't}) \cdot \left( \Pi_t^{project}(x, y') - \Pi_t^{project}(x', y') \right) \end{aligned} \quad (53)$$

In the above equation the LHS is data and the  $\Pi_t^{project}$  terms are known. Therefore, I can identify  $\gamma q_y^{project}(\theta_{yt})$  for each  $y \in \{low, mid, high\}$ . Hence,  $\gamma \pi_t(x|y)$  can be constructed for any match  $(x, y)$  using Equation (52) and  $\gamma q_y^{project}(\theta_{yt})$ . Finally, 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 meeting probability  $q_y^{project}(\theta_{yt})$  is identified. Rearranging Equation (49):

$$\ln \left( q_y^{project}(\theta_{yt}) \right) - \ln \left( q_{y'}^{project}(\theta_{y't}) \right) = \ln \left( \frac{\tilde{f}_t(x|y)}{\tilde{f}_t(x|y')} \frac{\omega_t(y'|x)}{\omega_t(y|x)} \right) \quad (54)$$

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

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \ln \left( q_h^{project}(\theta_{ht}) \right) \\ \ln \left( q_m^{project}(\theta_{mt}) \right) \\ \ln \left( q_l^{project}(\theta_{lt}) \right) \end{bmatrix} = \begin{bmatrix} h_t^{hm}(x) \\ h_t^{hl}(x') \\ h_t^{ml}(x'') \end{bmatrix} \quad (55)$$

Since the matrix on the LHS is full rank and the RHS is known, the meeting probability terms  $q_y^{project}(\theta_{yt})$  are identified from the data.

**Part 3:** Next I show how the targeting parameter  $\gamma$  can be identified. Rewriting Equation (50):

$$\ln \left( \tilde{f}_t(x|y) \right) - \ln \left( \tilde{f}_t(x|y') \right) = \gamma \left( \pi_t(x|y) - \pi_t(x|y') \right) + \ln \left( n_{yt}/n_{y't} \right) + \ln \left( q_y^{project}(\theta_{yt})/q_{y'}^{project}(\theta_{y't}) \right)$$

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. Furthermore, since the number of available rigs  $n_{yt}$  is observed, the term  $\ln \left( n_{yt}/n_{y't} \right)$  is known. 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:

$$\tilde{f}_t(x|y) = q_y^{project}(\theta_{yt}) w_t(y|x) f_t(x) \quad (56)$$

Since the left-hand side is data and  $q_y^{project}(\theta_{yt}) w_t(y|x)$  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 in at least one acceptance set. Therefore  $f_t(x)$  is identified over its full support.

## C Computation

### C.1 Policy functions

I estimate the policy functions in Proposition 3 as follows.

**Prices** To estimate  $\bar{p}_t(y, \tau)$  for each  $y \in \{low, mid, high\}$  I regress observed prices on first and second order polynomial combinations of the state vector plus an extra term for the contract length.

**Extension probability** Since  $\eta$  is assumed to be exogenous, I estimate  $\eta$  as the average probability of extending a contract in the data.

**The probability of matching** To estimate  $\mathbb{P}_t(\tau = n|y)$  I estimate a separate multinomial logit equation for each  $y \in \{low, mid, high\}$ . The dependent variable alternatives are matching with contract lengths  $\tau \in \{0, 2, 3, 4\}$  where  $\tau = 0$  denotes the probability of not matching. The independent variables are first and second order polynomial combinations of the state vector.

## C.2 Algorithm: computing capital's value of searching

The algorithm that I use is based on the proof of Proposition 3 in Appendix B.1.

Overall, I forward simulate the  $\mathbb{E}_{x_{-\tau}}[V_{t+n}^{capital}(y)|\tau = n]$  terms. I use these plus value function iteration to compute  $V^{capital}(y|s_t)$ . I linearly interpolate the value of searching over the state space using a grid with 5 nodes for each dimension (so there are  $5^4 = 625$  grid nodes in total). I perform the algorithm separately for each  $y \in \{low, mid, high\}$ .

1. Guess an initial value at each node of the state space.
2. Generate 200 simulations of the state evolution for 100 future periods using agents' computed beliefs from Stage 1 of the estimation.
3. For each node  $s$  in the state space grid:
  - (a) For each of the future state evolution draws, and for each contract duration  $n \in \{2, 3, 4\}$ , compute the evolution of average prices  $\bar{p}$  and the value of searching  $V^{capital}$ .
  - (b) For each of the future state evolution draws, forward simulate capital's value of matching  $\mathbb{E}_{x_{-\tau}}[V_{t+n}^{capital}(y)|\tau = n]$  according to Equation (31) using the computed  $\bar{p}$  and  $V^{capital}$ .<sup>40</sup>
  - (c) Average over these future state evolutions to compute  $\mathbb{E}_{x_{-\tau}}[V_{t+n}^{capital}(y)|\tau = n]$ .
4. Using Equation (28), update value functions at each node of the state space. This incorporates:

---

<sup>40</sup>Note that this forward simulation only needs to be performed over the state evolution and the expectation with respect to  $x_{-\tau}$  is constructed directly from policy functions using Equation (31).

- The simulated  $\mathbb{E}_{x-\tau} \left[ V_{t+n}^{capital}(y) \middle| \tau = n \right]$  weighted by the probability of matching each contract length  $\mathbb{P}_t(\tau = n|y)$ .
  - The value of being unemployed for one period  $b_y + \beta \mathbb{E} V^{capital}(y|s_{t+1})$  weighted by the probability of not matching  $\mathbb{P}_t(\tau = 0|y)$ .<sup>41</sup>
5. Repeat from Step 3 until convergence (where convergence is defined as ensuring that the difference in the value functions at each node between successive iterations is sufficiently small).

### C.3 Algorithm: simulated method of moments

In this section I describe the algorithm used to simulate the market which I use to estimate the parameters in the simulated method of moments. I first describe how to compute the equilibrium within each period (a month) for a given set of the parameters. I then describe how to nest the computation of the per-period equilibrium to simulate the market over the period 2000-2009.

#### C.3.1 Computing the per-period equilibrium

The state that agents take into account when computing their value functions is  $s_t = (g_t, n_{low,t}, n_{mid,t}, n_{high,t})$  where  $n_{yt}$  denotes the number of rigs of type  $y$  that are available to match the following period (i.e. the number of rigs where  $\tau_k = 0$  or  $\tau_k = 1$  periods left on their contract). Given  $s_t$  the per-period equilibrium is computed using the following algorithm:

1. Guess the share of potential well draws  $K_t = k_0 + k_1 g_t$  that choose to enter and target a rig of type  $y$ . Denote this share as  $share_{yt}^i$  where the share of wells that do not enter is  $1 - share_{lt}^i - share_{mt}^i - share_{ht}^i$ . The variable  $i$  denotes the iteration (so the guess initializes at  $i = 0$ ).
2. Get the submarket tightness  $\theta_{yt}^i$  using:

$$\theta_{yt}^i = \frac{n_{yt}}{share_{yt}^i \cdot K_t} \quad (57)$$

and note that this pins down the expected surplus of project type  $x$  to targeting capital type  $y$ :

$$\pi_t^i(y|x) = q_y^{project}(\theta_{yt}^i) \Pi_t^{project}(x, y) \quad (58)$$

3. Update the shares using:

$$share_y^{k+1} = \int_z w_t^i(z|y) e_t^i(z) h(z) dz \quad (59)$$

---

<sup>41</sup>Here, I again compute this expectation by forward simulating the value of searching at  $s_{t+1}$  over the state evolution from  $s_{t+1}$ .

where the weights are defined using Equation (6):

$$\omega_t^i(y|x) = \frac{n_{yt} \exp(\gamma \pi_t^i(y|x))}{\sum_{j \in \{l, m, h\}} n_{kt} \exp(\gamma \pi_t^i(j|x))} \quad (60)$$

and the entry condition is defined using Equation (4):

$$e_t^i(x) = \frac{\sum_{y \in \{low, mid, high\}} \omega_t^i(y|x) \pi_t^i(y|x) - c}{1 + \sum_{y \in \{low, mid, high\}} \omega_t^i(y|x) \pi_t^i(y|x) - c} \quad (61)$$

4. Repeat steps 3-4 until the targeting shares converge. Denote the equilibrium shares as  $share_{yt}^*$ , the equilibrium weights as  $w_t^*(z|y)$ , and the equilibrium probability of entry as  $e_t^*(z)$ .
5. The distribution and total number of matches (amongst other things) can now be computed from this targeting equilibrium and the acceptance sets. Recall that the acceptance set is the set of projects  $x$  where the match surplus with capital type  $y$  is positive at the time  $t$  state. To simplify the computation I use the empirical state at time  $t$  (rather than the model predicted state) to compute the surplus and the acceptance sets. For example, the probability a rig of type  $y$  is matched can be computed as:

$$\mathbb{P}_t(\text{capital type } y \text{ is matched}) = \frac{\int_{z \in A_t(y)} q_y^{project}(\theta_{yt}^*) w_t^*(z|y) e_t^*(z) h(z) dz}{n_{yt}} \quad (62)$$

### C.3.2 Computing the market over the 2000-2009 period

I nest the preceding algorithm to compute the market over the 2000-2009 period. I use the empirical evolution of the gas price  $g_t$  (since this is assumed to be exogenous) but I update the number of available rigs in accordance with the model equilibrium. To do this I introduce a ‘detailed state’ which is the current natural gas price  $g_t$  and a distribution, for each capital type  $y$ , over how many periods remain on the contract. Therefore the detailed state is a 5x3 dimensional matrix (plus the gas price  $g_t$ ), where each column is a rig type  $y \in \{low, mid, high\}$  and each row is the proportion of capital of type  $y$  with  $\tau_k \in \{0, 1, 2, 3, 4\}$  periods left of the contract (where  $\tau_k = 0$  periods left denotes unemployed capital).

1. Starting from a guess of the detailed state for January 2000 (which includes the empirical natural gas price), repeatedly compute the per-period equilibrium to burn in an initial state for the market.
2. Compute the equilibrium at the current state using the per-period equilibrium algorithm described in the preceding section.
3. For this equilibrium, compute the equilibrium probabilities of each type of rig matching with each duration contract. Using these equilibrium values (and the probability that a contract is extended) to update the ‘detailed state’.

4. Compute the state that agents use to compute value functions from the ‘detailed state’.
5. Repeat from step 2-4 over the monthly natural gas price evolution from January 2000 - December 2009.

## C.4 Algorithm: no sorting counterfactual

Note that no extra computation needs to be performed for the no sorting counterfactual. This is because the three components that depend on the value function in the model are constrained to not depend on the value function in the counterfactual. These three components are as follows. First, the entry decision is assumed to be the same in the no sorting equilibrium as in the market baseline (this is to avoid compositional effects). Second, the targeting parameter is  $\gamma = 0$  which implies that projects do not take into account value functions when making their targeting decisions. Finally, acceptance sets are computed as if the future value of searching  $E_t V_{t+k}^{capital}(y) = 0$  for  $k \in \{1, 2, \dots\}$  (that is, agents do not reject matches due to high future outside options).

## C.5 Algorithm: intermediary

In this section I describe the algorithm I use to compute the market equilibrium after an increase in the targeting parameter from an intermediary. I need to recompute value functions. In addition, I also need to recompute agents’ beliefs about the evolution of the state space (since these beliefs were computed when estimating the model from the empirical state evolution which will change in the counterfactual). An additional assumption that I make is to restrict agents’ beliefs about the state evolution to the following transitions in  $R_1$ :<sup>42</sup>

$$R_1 = \begin{bmatrix} r_g^1 & 0 & 0 & 0 \\ r_{low}^1 & r_{low}^2 & 0 & 0 \\ r_{mid}^1 & 0 & r_{mid}^3 & 0 \\ r_{high}^1 & 0 & 0 & r_{high}^4 \end{bmatrix} \quad (63)$$

That is, when forming their beliefs about the evolution of  $n_{yt}$ , I assume that agents only consider the gas price and the previous value of that capital type  $n_{yt-1}$ . The reason is mainly computational: in the following algorithm I run an AR(1) regression treating the model-predicted state evolution as ‘data’, rather than the empirical state evolution. Occasionally the model predicts sequences of available capital that are highly collinear when comparing across capital types which leads to unstable estimates of some parameters.<sup>43</sup>

---

<sup>42</sup>I fully re-estimate the  $R_0$  component of the state transition.

<sup>43</sup>Of course, these sequences are also strongly correlated in the data.

Overall, the algorithm uses an inner loop and an outer loop. In the inner loop I hold the state evolution beliefs fixed and compute the equilibrium. In the outer loop I iterate over the state evolution beliefs.

1. Denote the state evolution beliefs in the  $j$ -th outer loop iteration as  $R_0^j, R_1^j$ . Initialize  $R_0^0, R_1^0$  using the empirical beliefs.
  - (a) Denote the value of searching in the  $k$ -th inner loop iteration as  $V^{capital,k}(y)$ . Initialize  $V^{capital,0}(y)$  using the value functions computed from the data (i.e. the value functions previously used to estimate the model).
  - (b) Compute predicted policy functions (prices and the probability of matching) at the counterfactual parameters and under the state evolution beliefs  $R_0^j, R_1^j$ .
  - (c) Update the value functions to  $V^{capital,k+1}(y)$  using the algorithm that computes the value of searching from policy functions in Section C.2
  - (d) Check for convergence at each node of  $V^{capital,k+1}(y)$ .<sup>44</sup> If there is not convergence, repeat steps (a)-(d).
2. Simulate the model over the (exogenous) natural gas price evolution in 2000-2009.
3. Using an  $AR(1)$  regression over the simulated data, update agents' beliefs  $R_0^{j+1}, R_1^{j+1}$ .
4. Check for convergence in  $R_0^{j+1}, R_1^{j+1}$ . If these have not converged, repeat from step 1.

## C.6 Algorithm: demand smoothing

I compute the demand smoothing equilibrium in a similar way to the intermediary counterfactual in Appendix C.5 but with different assumptions on the state evolution (and hence on agents' beliefs about the state evolution). Specifically, I assume that agents believe the state evolution is in a steady state at the mean natural gas price over 2000-2009 ( $\bar{g}$ ) which implies that the transitions that need to be computed in the outer loop are in the following form:

$$R_0 = \begin{bmatrix} \bar{g} \\ r_{low}^0 \\ r_{mid}^0 \\ r_{high}^0 \end{bmatrix}, \quad R_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \sigma_\epsilon = 0 \quad (64)$$

---

<sup>44</sup>Recall that I linearly interpolate the value functions over a grid of the state space.

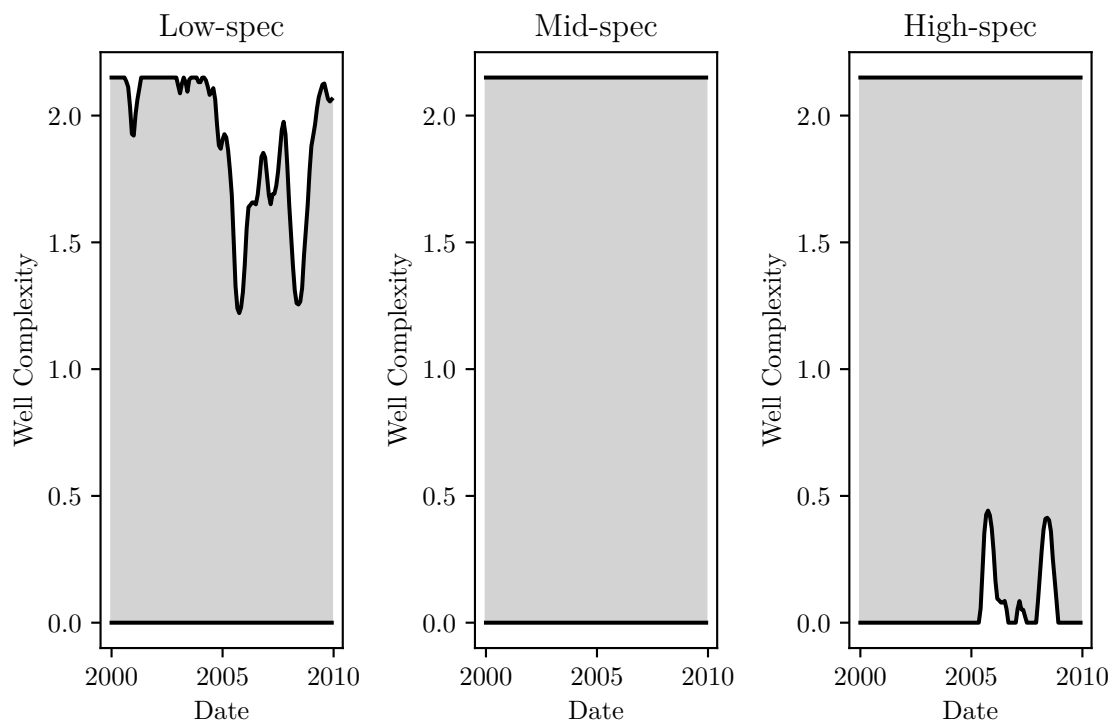
## D Additional Graphs and Tables

Table 4: Fit of the simulation to the moments used in the estimation

<b>Moment</b>	<b>Empirical Value</b>	<b>Simulated Value</b>
Mean: Complexity, Boom		
Low	0.7	0.7
Mid	0.83	0.83
High	0.92	0.91
Mean: Complexity, Bust		
Low	0.78	0.78
Mid	0.83	0.83
High	0.84	0.86
Variance: Complexity	0.17	0.35
Probability: Duration		
2 Month Contract	0.58	0.58
3 Month Contract	0.25	0.25
Mean: Utilization		
Low	0.63	0.63
Mid	0.76	0.77
High	0.92	0.91
Variance: Utilization		
Low	0.028	0.005
Mid	0.016	0.007
High	0.004	0.003
Covariance: Utilization and Gas Price		
Low	0.13	0.14
Mid	0.17	0.17
High	0.04	0.07

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 14: Acceptance sets for 3 month contracts over time



Note: This table contains the acceptance sets computed for an (approximately) average length contract of  $\tau = 3$  months. The vertical distance between the two black lines (shaded gray) is the acceptance set at a given date. I plot the acceptance sets for the 3 rig types computed at the empirical state value and with the minimum bound 0 and the maximum bound 2.15 (2.15 is the maximum well complexity that a low-specification rig undertakes in the data and so these acceptance sets are ‘in sample’ for all rig types). Judging only by the vertical distance, the low-spec rig acceptance set appears to shrink more than for high-spec rigs in booms. However, for the high-spec rig, this rejection occurs in a region where the density of searching wells is higher. Therefore, the high-spec rig may still reject more matches in total.