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

Nicholas Vreugdenhil*

October 28, 2022

Abstract

How efficiently do markets reallocate capital in booms and busts? Using a novel dataset of offshore drilling contracts I examine the role of matching in shaping industry reallocation. Oil companies search and 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 but that 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 the Editor, Chad Syverson, and four referees, for their thoughtful comments. Thank you to Gaurab Aryal, Esteban Aucejo, Vivek Bhattacharya, Hector Chade, Igal Hendel, Ken Hendricks, Andreas Kostøl, Nick Kuminoff, Tomas Larroucau, Alvin Murphy, 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 movements in aggregate productivity, and has spurred large literatures in Industrial Organization, Labor, and Macroeconomics.[1] Although it is well established that fluctuations and productivity are broadly linked, the exact process of capital reallocation within industries is not well understood.[2] [3] 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 markets where output is made by matching heterogeneous producers and heterogeneous factor inputs.[4] Here, booms can create thick factor markets that give the parties greater option value in the search process. The better matches and complementarities that result reduce factor misallocation and create procyclical productivity movements.

Overall, 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. I use the framework to quantify the benefits of an intermediary and the effects of a demand smoothing policy.

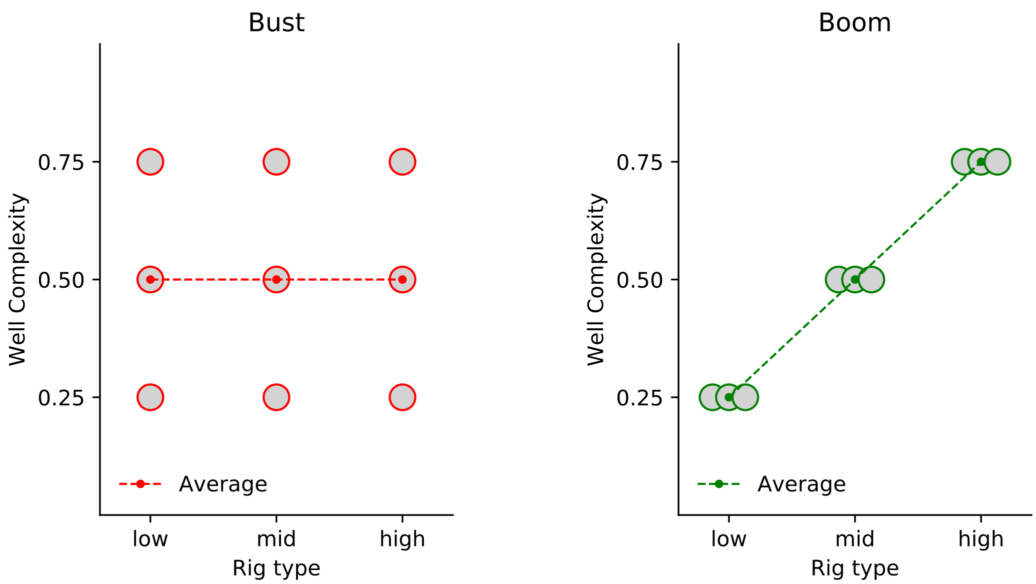
[1] For a review of the literature see Eisfeldt and Shi (2018).

[2] Bartelsman et al. (2013) document empirically the role of reallocation.

[3] This is largely due to the lack of producer-level data on covariates such as contracts, production, and relationships; 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.

[4] 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)).

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.

It is important to note that the fact that booms are associated with a sorting effect 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 (since oil companies receive a higher price for a given quantity of oil and gas) and therefore it may be optimal to be less selective. Overall, I find that the option value effect dominates the match value effect in the data, leading to pro-cyclical match quality.

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 undertake projects (wells) but do not own capital (drilling rigs). Instead, they must search for capital in a decentralized market. Capital can be ranked using an industry measure of efficiency and projects can be ranked using an engineering measure of complexity. The *quality* of the match matters: more efficient capital is suited to drilling more complex projects and this is reflected in sorting patterns in the industry.[5] Therefore, in the offshore drilling industry, stronger sorting corresponds to more efficient rigs matched to more complex wells, and less efficient rigs to simpler wells, as illustrated in Figure 1.

I focus on shallow water oil and gas drilling in the US Gulf of Mexico in 2000-2009. I begin by discussing several features that suggest search frictions are important in this industry.[6] I then document two main findings. First, there is positive assortative 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. Hence, the potential benefits to an intermediary or demand smoothing would be low since there is less mismatch. 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.[7] Agents respond by avoiding bad matches

[5]The fact that agents care about the quality of the match - and not just whether they are matched or not - is an important difference between my setting and recent work in Industrial Organization on search markets such as taxis (Frechette et al. (2019), Buchholz (2022)) and bulk shipping (Brancaccio et al. (2020)) where agents are relatively homogeneous.

[6]These include price dispersion for observationally equivalent matches, a large number of small firms on both sides of the market, the existence of brokers, the emergence of e-procurement, and an exercise documenting mismatch in the raw data.

[7]Recall that whether stronger sorting in booms is optimal is ultimately an empirical question in the model.

in two ways. First, they can reject bad matches. Second, using the search technology, they can direct their search away from bad matches. Overall these two channels result in stronger sorting patterns and reduce mismatch.

I estimate the model in two steps. First, I construct value functions based on empirical state transitions and data on prices, the probability of matching, and the probability of extending a contract. Second, I estimate the parameters using the simulated method of moments.

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 9.9%, or around $443 million dollars, over the 2000-2009 period. The sorting effect is cyclical with more efficiency gains in the boom. Decomposing the total effect highlights the main tradeoff in the model: compared to the 'no sorting effect' 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.[8] I find that the intermediary would achieve a welfare gain of 50.8% 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 drilling activity from booms to busts. However, the policy would increase overall welfare by only 16.6%. Given that eliminating price cycles entirely is an extreme example of a demand smoothing policy, the 16.6% benefit appears modest and suggests that demand smoothing policies (particularly more moderate interventions) would be somewhat ineffective.

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

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.[9] By contrast, in this paper I construct a dataset of the universe of contracts in the industry matched with rich micro data from the regulator on the characteristics of projects undertaken under these contracts. My 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.[10] Previous work on other markets in Industrial Organization has also faced the challenge of identifying demand from matches (e.g. taxis (Frechette et al. (2019), Buchholz (2022)) 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.[11] When there are fluctuations, however, the distributions of agents change through time. In this paper I use an estimation strategy that incorporates - for the first time in a random search model with fluctuations - two sided heterogeneity, distributions of searching agents that change over time, and Nash bargaining. The estimation strategy relies on the observation that the value of searching can be written in terms of data on contract prices and the probability of matching. My strategy is an extension of approaches in the Industrial Organization firm dynamics literature such as Kalouptsidi (2014) to cases where short-term contract data are available.

[9] A literature estimates these unobserved transfers in certain settings (e.g. Villas-Boas (2007)).

[10] Lentz and Moen (2017) consider a related setup. My approach differs because I need to deal with two-sided heterogeneity and fluctuations, which pose challenges for identification and estimation.

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

1.1 Related literature

This paper is related to five 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 - 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 (2022), Frechette et al. (2019)). A related set of papers study how fluctuations affect long-run firm entry and exit decisions (Kalouptsidi (2014), Collard-Wexler (2013)). Other recent papers estimate search and matching models with two-sided heterogeneity in a stationary context (e.g. Gavazza (2016)). By contrast, my paper contains *both* fluctuations and heterogeneous agents and I study how the two interact in a decentralized firm-to-firm market.

Third, this paper is related to the literature about the effects of the business cycle on labor search and matching. Many of these papers aim to rationalize empirical pro-cyclical productivity patterns in labor markets. For example, Barlevy (2002) uses a model with on-the-job search and match-specific quality to argue that busts are associated with a 'sullying effect'.[12] Similarly, Lise and Robin (2017) estimate a model of sorting between workers and firms with random search, productivity fluctuations, and two-sided vertical heterogeneity. There are some key differences between my framework and the models in these papers.[13] More broadly, however, my paper

[12] Also, see Moscarini and Postel-Vinay (2018) for a review of the cyclical job ladder literature.

[13] For example, for tractability Lise and Robin (2017) 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 -

illustrates the relevance of search and matching models in explaining procyclical productivity patterns outside the standard labor market context.

Fourth, this paper is related to the theoretical literature on dynamic matching. For example, Baccara et al. (2020) characterizes the optimal queuing mechanism when agents arrive sequentially. Although the theoretical framework is quite different from my paper, a focus of Baccara et al. (2020) is quantifying the gains to a centralized intermediary, which I similarly address in a counterfactual in this paper.

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).[14] For credible estimation my empirical strategy relies on having a measure of participants' expected value of undertaking a project. In the context of the Gulf of Mexico an excellent proxy is available: participants' beliefs about the value of drilling a well is related directly to lease bids (Porter (1995)).

2 Industry Description and Data

2.1 Overview of the offshore drilling industry

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

The offshore drilling industry is a decentralized industry. Lease holders such as BP and Chevron do not own the equipment used to drill their wells. In order to drill a well a drilling rig must be procured from a drilling contractor. Both sides of the industry are unconcentrated with an HHI

partially directed search - using data on observed matches.

[14] 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.

of 1239 for rig owners and an HHI of 335 for well owners.[15] 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 on changes in the gas price as the driver of exogeneous shocks in this industry.[16] 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

[15] I calculate the HHI with the definition of 'market share' as the proportion of total contracts.

[16] Furthermore, in the sample period the oil price is correlated with the natural gas price. I show in Appendix E.2 that simply tracking the natural gas price does not make any difference to the results.

Table 1: Summary statistics for the dataset

Variable	**Units**	**N**	**Mean**	**SD**	**10%**	**90%**
Rig Price - New Contracts	1000s of USD/day	1733	62	35	27	111
Duration - New Contracts	Days	1733	65	76	27	120
Rig Price - Renegotiations	1000s of USD/day	922	52	26	28	81
Duration - Renegotiations	Days	922	68	59	29	126
Value	Millions of USD	2655	7.0	15	0.2	17
Complexity	Index	2655	0.87	0.51	0.34	1.46
Water Depth	Feet	2655	117	80	37	233
Monthly Utilization	% Rigs under contract	360	0.76	0.2	0.49	1.0

Note: Monthly utilization is for each of the 3 types of rig over 120 months (so = 360 observations in total).

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. In total there are 101 rigs. 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. Contracts are also often extended, and this is typically to drill a new well. A small number (11.0 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 87.6 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.[17] In order to back out the quantity of hydrocarbons in the well, I then divide by average gas price in the sample.[18] My measure is a monotonic function of the expected oil and gas deposit size.

Measuring well heterogeneity To rank wells I compute an engineering model of well complexity used in the industry called the 'Mechanical Risk Index'. The Mechanical Risk Index takes well covariates including depth, mud weight, horizontal displacement etc that describe the geological environment and transforms them into a one dimensional index of well complexity.[19] 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

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

[18]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.

[19]Details on the calculation of the Mechanical Risk Index can be found in Appendix A.2.

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. The split classifies 'low-efficiency' rigs as those with a maximum drilling depth of ≤ 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 ≥ 300 feet.[20]

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. Transportation costs in the offshore oil and gas industry mainly come in the form of lost drilling time. However, drilling rigs are extremely mobile and take around 1 day to move across the Gulf of Mexico. When compared to the average new contract length (around 65 days), a back-of-the-envelope calculation implies that choosing a far away rig over a nearby rig would increase costs by around 1.5% for the average contract. Since most rig moves are within-field I do not include distance to a well as a factor for rig choice in the model.

The second factor - past experience between a rig operator and a well owner - has been shown to be a consideration for rig choice in the *onshore* oil and gas industry (Kellogg (2011)). I capture repeated contracting in my model by allowing for contract extensions. However, for new contracts, I assume that agents' decisions about who to match with are independent of past experience. This modeling assumption seems to be supported by the data: I find 66 percent of new contracts are between a rig-owner pair who have not worked together in the previous 2 years.[21]

[20] The split is not quite exact because there are sometimes many rigs of exactly the same drilling depth.

[21] I use 2 years as my cutoff for a 'relationship' because that is the definition used by Kellogg (2011).

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.

2.3.1 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 10% and 90% 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'*[22]. Higher ranked rigs attract premium prices and are actively marketed as 'high-specification'. Note that in Table 10 in the Appendix I perform a hedonic regression of prices on match characteristics.

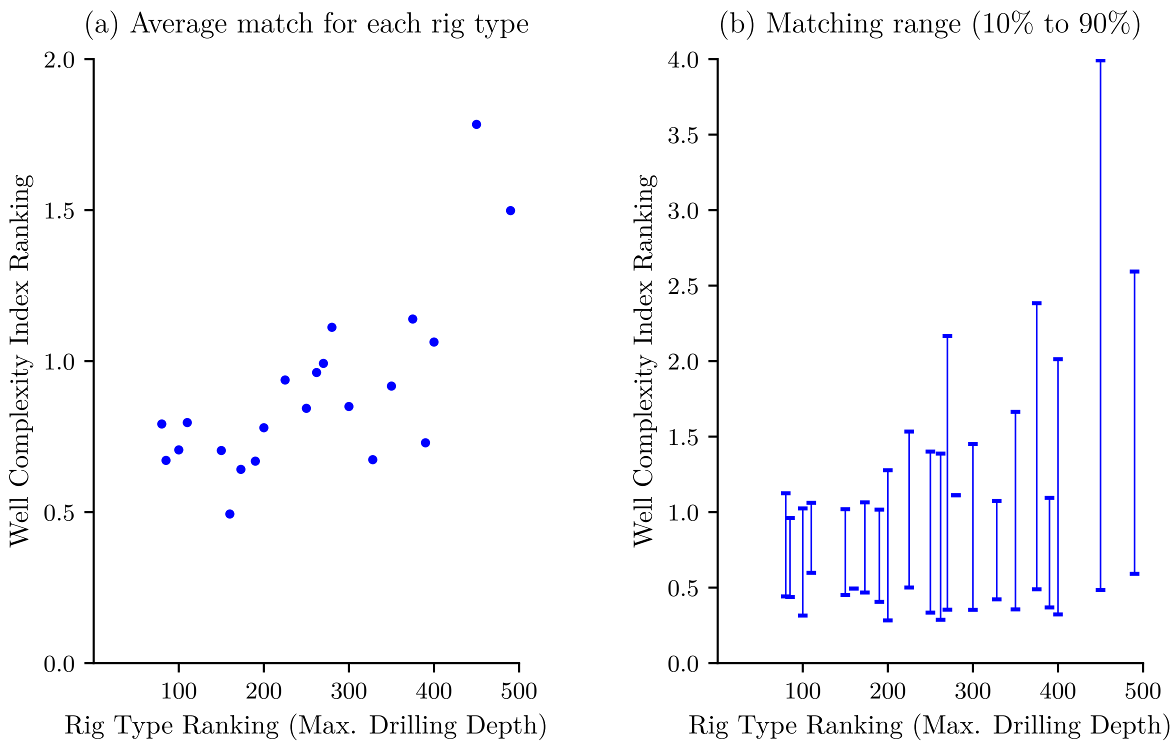
2.3.2 Feature 2: Booms and busts

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.[23] 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

[22] http://www.diamondoffshore.com/offshore-drilling-basics/offshore-rig-basics

[23] 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

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



Note: This figure shows the sorting patterns in the data. Rigs are constructed in discrete increments of maximum drilling depth and so each point on the x-axis might correspond to many unique rigs.

busts are a key factor in how they make decisions about prices and utilization.[24] (Note that I document rig utilization by type and whether the market is in a boom or bust in Table 8 in Appendix D.4.)

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 the three types of rigs. Figure 4 shows a rotation in the average match line between rig rankings and well complexity rankings.

[24]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 time and subsequently resume full or near full utilization when business cycles change'.

Figure 3: Booms and busts

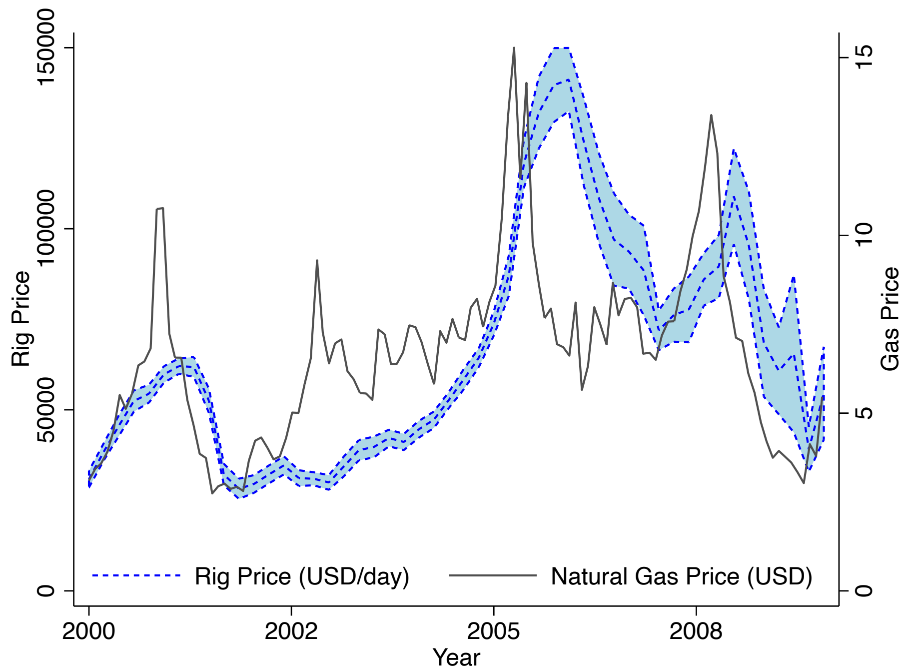
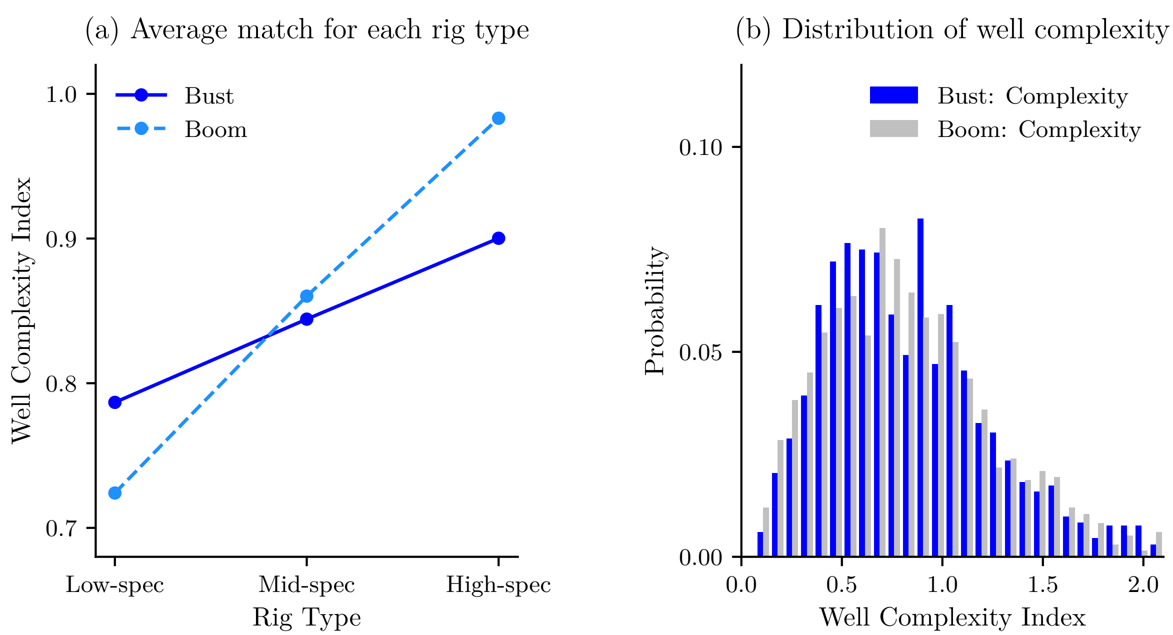


Figure 4: Matching patterns in booms and busts



Note: The left panel shows that when a bust (in the solid line) turns to a boom (the dotted line), the market features stronger sorting patterns. I verify in Table 9 in the Appendix that these effects are statistically significant, and that these patterns are robust to controls for observables.

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. I verify in Table 9 in the Appendix that these effects are statistically significant across the boom-bust cycle, and that these patterns are robust to controls for observables.

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. Consistent with this evidence, in Appendix D.1, I also show that there are stronger synergies between high-efficiency rigs and complex wells in booms, as reflected in prices.[25]

2.3.3 Feature 3: Search frictions

I begin with a description of how matches are formed in practice and discuss a number of institutional features that highlight the importance of search frictions in the market.

In the model, I model matching as an arms-length search and matching process, which is a simplified version of how rigs are selected in practice.[26] In practice, oil company engineers will first determine the well design, write up the details, and initially solicit rigs, sometimes with the aid of specialized rig brokers. The rig selection process rarely ends there: offshore rig companies stress that the process of obtaining an offshore rig can be relatively unstructured, with further discussions between the parties and that ultimately "our contracts to provide offshore drilling services are individually negotiated" Transocean (2015). Since my dataset only contains data on the eventual outcomes of this process and no information about interim discussions, I use a reduced form of how this process occurs in practice, ending with one rig ultimately selected and

[25]One may also ask if drilling speed (that is, how long it takes a rig to finish drilling a well) also increases in booms. I document that this is not the case in Table 11 in Appendix D.4.

[26]I emphasize that - relative to previous work that models matching in markets with two-sided vertical heterogeneity - my empirical framework is also relatively more flexible when estimating the search technology. Concretely, I allow for parameters that govern the degree to which agents can target their search towards their best match.

negotiated with.[27]

Next, I discuss several institutional features which suggest search frictions in the industry. The first feature is the emergence of e-procurement in the industry.[28] These recent technological improvements to the search and matching process suggest that in the earlier period of this study, there were potential gains to better matching that were unrealized.

Second, both rig owners and well owners often enlist the help of a fragmented group of brokers to help find a match.[29] As discussed by Brancaccio et al. (2022), the very existence of brokers has been used in a variety of settings and papers as evidence of search frictions.[30]

Finally, the industry is unconcentrated on both sides with a large number of agents simultaneously trying to match with each other in an uncoordinated fashion. In addition, since this industry is constantly in flux, participants may not have good information about the status of other rigs.[31] As is argued in the macroeconomics search literature, modeling each of these sources of frictions and heterogeneity explicitly would "introduce intractable complexities" into the model Petrongolo and Pissarides (2001). Instead, I include a reduced-form matching function that allows for realistic frictions in the search process.

Price dispersion Next I show suggestive evidence for search frictions in the data by showing that different prices are paid for observationally equivalent matches. Since price will also vary with market conditions I de-mean prices by the average price in each month. I regress these de-meaned prices on rig characteristics, well characteristics, and contract characteristics. I run

[27] I provide a similar discussion for the deepwater market in Vreugdenhil (2021).

[28] For an early discussion about the potential benefits to e-procurement in the industry see Rothgerber (2002). Raghothamarao (2016) discusses how advances in e-procurement are being used in the oil and gas industry.

[29] Examples of companies that offer brokering services for offshore rig chartering include Clarksons, Bassoe Offshore, and Pareto Offshore.

[30] One may ask whether brokers might in fact eliminate search frictions entirely. Arguably, this is not the case in the offshore oil and gas industry: for example, in a discussion with a brokerage firm, I discovered that they were developing their own e-procurement platform where oil and gas companies can directly match with rigs.

[31] For example, which rigs have not yet signed a contract but are in a late stage of negotiations, or if a rig suddenly needs maintenance, or other idiosyncratic heterogeneity.

Table 2: Evidence of price dispersion

	(a) Using aggregated rig types	(b) Using disaggregated rig types
$1 - R^2$	0.37	0.31
$SD(\tilde{p}_{it})$	11	10
$SD(\hat{p}_{it})$	18	18

Note: Standard deviations are measured in thousands of US dollars per day. The dependent variable $\hat{p}_{it}$ is prices de-meaned by the average price in each month. The residual from the price regression is denoted by $\tilde{p}_{it}$.

the following regression on new contracts:

$$\hat{p}_{it} = \mathbf{X}'\beta + \tilde{p}_{it} \tag{1}$$

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 a third order polynomial of the state variables (gas price and rig availability of each aggregated rig type), and interactions between rig types and a third-order polynomial of well characteristics and a third-order polynomial of contract duration:

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

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, mid, low}); 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: 0.37 of total price variation is unexplained when using the aggregated rig types and 0.31 of total 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.[32]

Documenting instances of mismatch I document instances of mismatch in the data in the following way. I take the observed matches in each period and reallocate the rigs and wells in these matches optimally using a linear sum assignment algorithm. For example, if there is a new match between a high-efficiency rig and a simple well, and simultaneously a new match between a low-efficiency rig and a complex well, the algorithm will reallocate these matches.[33]

I report the results from the above procedure in Table 3. In this table I split up the benefits to better matching into the bust versus the boom. I also test for differences in the average match value in the bust compared to the boom. There are two main findings. First, there are benefits to better matching across the cycle. In a bust the average increase in match value is \$0.298 million; for comparison, the average payment to a rig for a new contract is around \$3 million. In a boom the average increase in match value is \$0.147 million (but this difference is not significant).[34] Second, this return is counter-cyclical: the returns to better matching are higher in the bust than the boom by \$0.151 million, which is consistent with the sorting effect.

[32]One recent paper that documents a similar magnitude of price dispersion in a firm-to-firm search market is Salz (2022). 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. Similarly, Brancaccio et al. (2022) find that unexplained price dispersion is around 30% in the bulk shipping market, which they argue is consistent with search frictions.

[33]Whether this example of reallocation - which involves more assortatively matching the two sides of the market - is truly optimal depends on the underlying match value. Therefore, in order to compute the optimal matches, as well as quantify the degree of mismatch, I use the match values I later estimate in the model in this exercise. These match values convert a given rig-well match into a dollar figure. Beyond these match values I place no additional assumptions on the data, using just the empirical matches and available rigs.

[34]Note that the fact that the difference is not significant in booms is consistent with the sorting effect because it implies that agents are relatively well-matched.

Table 3: Documenting mismatch in the data

	Change in Match Value (Millions USD)		
	Bust	Boom	Difference: Bust vs Boom
Optimal Match vs Empirical Match	0.298	0.147	0.151
T-test	0.001***	0.109	0.023**

Note: This table documents instances of mismatch in the data by reallocating empirical matches to optimal matches. In order to compute optimal matches between available rigs and matched wells I solve a linear sum assignment problem in each period. I measure the degree of mismatch moving to optimal matching using the estimated match values from the model. I report the average change in millions of USD at the contract level (for comparison, the average payment to a rig for a new contract is around $3 million USD). I split the results into the improvements to matching in the bust, the improvements to matching in the boom, and the difference in the boom/bust change. I also report the p-values from a t-test for the difference in mean match values moving from the empirical matches to optimal matching, where: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

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 a project are:

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

where $x_{\text{complexity}}$ is the complexity of a project, x_{quantity} is the quantity of hydrocarbons (oil and gas), and τ 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 draws of *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 $f(x)$ - the probability density of potential projects.

Capital differs in its efficiency $y \in Y = \{\text{low, mid, high}\}$. 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.[35] 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. *Contract extensions.* Existing matches are extended with probability $\eta_t(x, y)$ which is dependent on the state as well as the value of a particular match.

2. *Entry.* The set of potential projects is comprised of K_t draws from a distribution $f(x)$. Each potential project chooses whether to enter and search for capital.

3. *Search.* Each type of capital is located in a submarket. Meetings are determined probabilistically within each submarket as a function of the market tightness θ_{yt} (the ratio of

[35] I verify in Table 16 in the Appendix that halving the period length does not affect the results substantially.

available capital to the mass of searching projects). Figure 5 provides a diagram of the search and matching process. Projects are able to direct their search towards a particular capital submarket using a *search technology*. Denote the probability that a type-x project targets type-y capital at time t with the search technology as $\omega_{yt}(x)$.

4. *Matching.* If a project owner contacts a capital owner then agents choose whether to match. Prices are determined by Nash bargaining and since this implies perfectly transferable utility a match will be accepted if the total match surplus is positive ($S_t(x,y) \geq 0$). If capital is matched then it cannot match for the duration of the contract (τ periods). If agents choose to not match then projects exit the market immediately and the capital is available to match in the next period.[36] Note that in the Appendix (Section E.3) I run robustness checks around the assumption of myopic projects and find that relaxing this assumption does not substantially change the results.

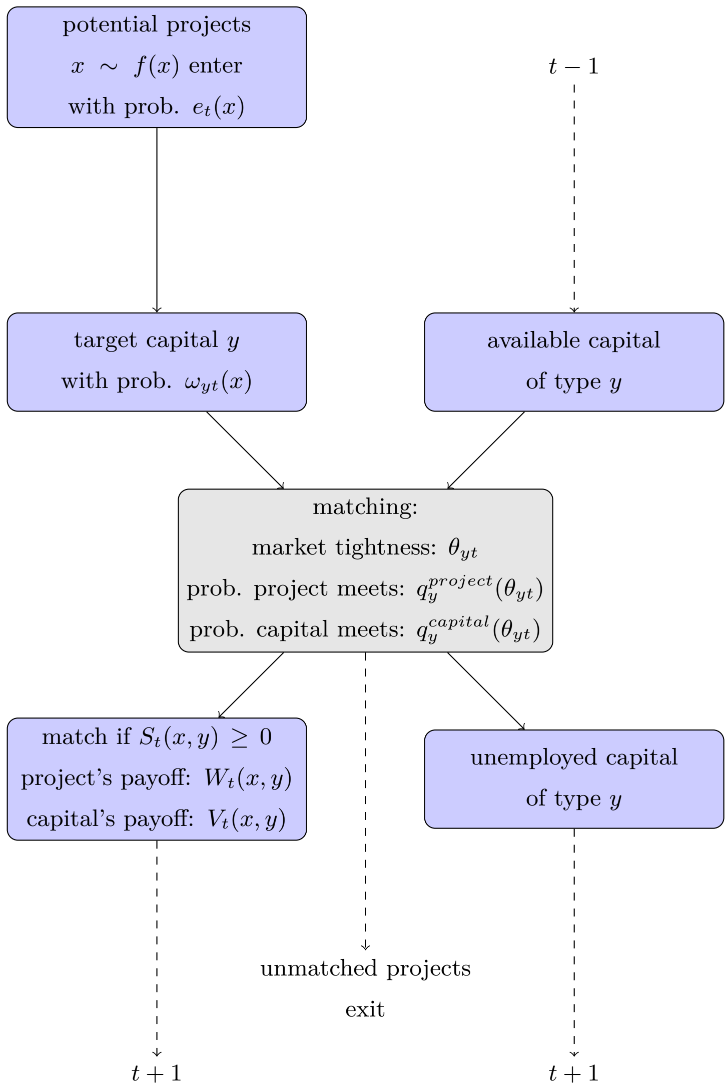
Value of a match I use the following specification for the per-period value of a match k periods after the time t that the match is created:

$$v_{t,k}(x,y) = m_{0,y} + m_{1,y} \cdot x_{\text{complexity}} + m_2 \cdot \mathbb{E}_t\big[g_{t+k}\big] x_{\text{quantity}} \tag{2}$$

and so the total value of a contract is $\sum_{k=0}^{\tau-1} \beta^k v_{t,k}(x,y)$. The above equation can be broken down into two main components. The first component is the match value added: $m_{0,y} + m_{1,y} \cdot x_{\text{complexity}}$. Here $m_{0,y}$ and $m_{1,y}$ are coefficients that vary with rig type. This equation captures complementarities between rig type and well type through $m_{1,y}$. 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. On the other hand, 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)).

[36]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)).

Figure 5: Search and 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 and available capital n_{yt}. The searching projects first choose which type of capital to target. Meetings are determined randomly within each submarket and are dependent on the market tightness θ_{yt}. Finally, agents choose whether to match based on whether the total surplus of a match is positive ($S_t(x, y) \geq 0$).

The second component is $m_2 \cdot \mathbb{E}_t\left[g_{t+k}\right]x_{\text{quantity}}$. This component captures the expected total value of oil and gas that is produced. The variable g_{t+k} is the gas price at the period $t+k$. Since the covariate x_{quantity} is the ex-ante quantity of hydrocarbons in the well (proxied by the maximum bid in the lease auction and then converted to a quantity as discussed in the previous section), I include a parameter m_2 which is defined as the weight that agents put on this ex-ante proxy when making decisions.

Note that the quantity of oil and gas extracted from the well does not depend on the rig type. Rather, the quantity of hydrocarbons that a well produces is dependent on geological features. Consistent with this assumption, the focus of industry practitioners and the engineering literature is how rig choices and well design affect extraction *costs* e.g. Hossain (2015).

The expected value of oil and gas scales with the length of the contract since longer contracts usually involve constructing multiple similar wells over the same oil and gas deposit. One may ask whether longer contracts may instead correspond to a single well being drilled over many periods, in which case the value $\mathbb{E}_t\left[g_{t+k}\right]x_{\text{quantity}}$ would accrue at the end of the contract when this single well is completed. However, the data suggest that longer contracts are typically multiple wells in the same oil and gas formation. For example, the average well in the sample takes around 22 days from when the drill enters the sea floor to when it reaches the target depth (from the 'spud date' to 'depth date') and an additional few days to 'complete' (cap and run in the production tube), which is approximately one month in total and one period in the model.

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. Overall the model focuses on the dynamic tradeoff for capital owners.

Comparison to previous industrial organization work on search and matching models The key difference I need to contend with in my setting is two-sided vertical heterogeneity. This contrasts with previously studied markets like taxis (e.g. Buchholz (2022)) and bulk shipping (e.g. Brancaccio et al. (2020)) where agents are relatively homogeneous. Due to this feature the model departs from the past industrial organization literature in search and matching models in two main ways. First, I allow for search to be (partially) directed, where heterogeneous projects can target the type of capital that they are best suited to match with. Second, I account for the fact that matches can be rejected and so agents have acceptance sets.[37]

My framework also shares some elements with previous work. Most notably, once projects have decided which type of capital to target, meetings take place within a sub-market in a similar way to how they would within an individual location in a taxi market, or a port in the bulk shipping market.[38]

3.2 Demand for capital

Payoffs First I consider the profits to a type-x project matching with type-y capital. 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 is:

$$W_t(x,y) = \underbrace{\sum_{k=0}^{\tau-1} \beta^k \Big[v_{t,k}(x,y) - p_t(x,y) \Big]}_{\text{Value of the initial contract}} + \underbrace{\beta^\tau \mathbb{E}_t \Big[\eta_{t+\tau}(x,y) W_{t+\tau}(x,y) \Big]}_{\text{Extension value}} \tag{3}$$

The project owner's value of matching $W_t(x,y)$ can be decomposed in the following way. For each period of the τ length contract the project owner receives the match value $v_{t,k}(x,y)$ minus the price $p_t(x,y)$ to hire the capital. The contract will be extended with probability $\eta_{t+\tau}(x,y)$ which is dependent on the state at time $t+\tau$ and the value of the match.[39]

Partially directed search I first discuss how search operates once entry has occurred and then I turn to the entry decision. In the search process, potential projects choose which capital submarket to search in. The choice of submarket depends on the characteristics of the project

[37] Another key difference with existing work is that I allow for contract extensions. While match extensions are a regular occurrence in decentralized capital markets, they are not present in, for example, taxi markets.

[38] In this way, Figure 5 which sets out how meetings take place within a sub-market, can be compared to a similar figure in Buchholz (2022).

[39] In Appendix D.2 I show that the assumption that the extended contract has the same duration as the initial contract is reasonable.

and the search technology. For a type-x project denote the (expected) value of searching in the type-y capital submarket as $\pi_{yt}(x) = q_y^{project}(\theta_{yt})W_t(x,y)$. I allow for a flexible search technology - partially directed search - where the probability that a type-x project targets type-y capital is:[40]

$$\omega_{yt}(x) = \frac{n_{yt} \exp\left(\gamma_0 \left[\pi_{yt}(x) - \gamma_1 1\left[S_t(x,y) < 0\right]\right]\right)}{\sum\limits_{k \in Y} n_{kt} \exp\left(\gamma_0 \left[\pi_{kt}(x) - \gamma_1 1\left[S_t(x,k) < 0\right]\right]\right)} \tag{4}$$

Here, γ_0 and γ_1 are 'targeting parameters' that index how precisely a project can target capital. I allow for targeting to be responsive to both whether the match will be rejected (through the parameter γ_1) as well as the overall quality of the match (through the parameter γ_0).[41] In Appendix B.1 I show how the reduced-form search technology in Equation (4) can be micro-founded from individual project targeting choices.[42]

This search technology is more flexible than the typical assumptions of random search or directed search which are used in search models. At the extremes this specification nests random search (at $\gamma_0 = 0$, where projects contact capital completely at random) and directed search (as $\gamma_0 \to \infty$, where projects can perfectly identify the best match).[43]

Entry Before making the targeting decision each potential project chooses whether to enter. This decision is: $\max \left\{ \sum_{k \in Y} \omega_{kt}(x)\pi_{kt}(x) - c + \epsilon_t^{\text{entry}}, \epsilon_t^{\text{no entry}} \right\}$. Here, c is the entry cost and $\epsilon_t^{\text{entry}}$, $\epsilon_t^{\text{no entry}}$ are drawn from an i.i.d. type-1 extreme value distribution. The first term in the maximization 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

[40]Here, I denote $1[S_t(x,y) < 0]$ as an indicator function for whether the total surplus of a match is negative, which would result in the match being rejected.

[41]Whether a match could be rejected might be more salient to capital owners than other features of the match - I allow the data to determine whether this is the case.

[42]I also derive a related formula for the search technology if rigs target wells in Section E.4 and show that the results are relatively unchanged if this alternative search technology is used instead.

[43]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.

things. The resulting conditional choice probability that a project enters is:

$$e_t(x) = \frac{\exp\Big(\sum_{k\in Y}\omega_{kt}(x)\pi_{kt}(x) - c\Big)}{1+\exp\Big(\sum_{k\in Y}\omega_{kt}(x)\pi_{kt}(x) - c\Big)} \tag{5}$$

Demand for capital Aggregating up the individual project entry and targeting decisions results in the demand for capital. Denote $h_{yt}(x)$ as the probability that type-y capital will be contacted by a type-x project. This is given by:

$$h_{yt}(x) = q_y^{capital}(\theta_{yt}) \cdot \frac{\omega_{yt}(x)e_t(x)f(x)}{\int_z \omega_{yt}(z)e_t(z)f(z)dz} \tag{6}$$

and the probability that capital is not contacted by any project is $h_{yt}(\emptyset) = 1 - q_y^{capital}(\theta_{yt})$.

The above setup 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 $h_{yt}(x)$ 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

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 after one period: $\max\Big\{V_t(x,y), \beta\mathbb{E}_t U_{t+1}(y)\Big\}$. Here, $V_t(x,y)$ is the profit from matching and is given by:

$$V_t(x,y) = \sum_{k=0}^{\tau-1}\beta^k p_t(x,y) + \beta^\tau \mathbb{E}_t\Big[\eta_{t+\tau}(x,y)V_{t+\tau}(x,y) + (1-\eta_{t+\tau}(x,y))U_{t+\tau}(y)\Big] \tag{7}$$

The profit from matching $V_t(x,y)$ 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 τ periods. When the contract is complete the rig owner receives $V_{t+\tau}(x,y)$ if the contract is extended. If the contract is not extended then the rig will be available to search again and will receive $U_{t+\tau}(y)$.

The value of searching is:

$$U_t(y) = \underbrace{\int_z \max\Big\{V_t(z,y), \beta\mathbb{E}_t U_{t+1}(y)\Big\} h_{yt}(z)dz}_{\text{Exp. Value Of A Meeting}} + h_{yt}(\emptyset)\underbrace{\beta\mathbb{E}_t U_{t+1}(y)}_{\text{No Meeting}} \tag{8}$$

The first term is the expected value of a meeting: capital meets a particular project type with probability $h_{yt}(z)$ and it will choose whether or not to match with it. If capital is not contacted by a project (which happens with probability $h_{yt}(\emptyset)$) 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:

$$p_t(x,y) = \text{argmax}_{p_t(x,y)} \left[V_t(x,y) - \beta\mathbb{E}_t U_{t+1}(y)\right]^{\delta} \left[W_t(x,y)\right]^{1-\delta} \tag{9}$$

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

Total surplus and contract extensions I assume Nash bargaining which implies transferable utility and therefore that the decision to accept or reject a match is mutual and dependent on whether the total surplus of a match is positive: $A_{yt} = \big\{x : S_t(x,y) \geq 0\big\}$. Here, the total surplus of a match is given by: $S_t(x,y) = W_t(x,y) + V_t(x,y) - \beta\mathbb{E}_t U_{t+1}(y)$.

I assume that the contract will be extended with the following probability: $\eta_t(x,y) = \eta 1[S_t(x,y) \geq 0]$. The parameter η is related to the probability that drilling on the original prospect reveals good information that induces the contract to be extended. The term $1[S_t(x,y) \geq 0]$ is an indicator function that allows for extensions to be rejected if, at the time t state, the surplus of the original match is now negative.

3.4 Transitions and states

Transitions At the start of each period, rigs are either unemployed or are currently matched. If a rig is currently matched denote τ_k as the number of periods remaining on its contract.

Matches with $\tau_k = 0$ are possibly extended. Rigs that are unemployed or whose contracts are not extended are available to match. Rigs which do not find a new match become unemployed. At the end of each period, τ_k counts down by 1.

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:

$$s_t = [g_t, n_{low,t}, n_{mid,t}, n_{high,t}] \tag{10}$$

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 either enters the period unemployed or if there are zero periods remaining on its contract and the match is not extended.

I model agents' beliefs about equilibrium industry state transitions as an $AR(1)$ process: $s_t = R_0 + R_1 s_{t-1} + \epsilon_t$. I assume that rig transitions are deterministic so the only stochastic component in the model is the gas price error term, which implies that $\Sigma = Diag(\sigma_\epsilon, 0, 0, 0)$. I write the elements of R_0, R_1 as:

$$R_0 = \begin{bmatrix} r_g^0 \\ r_{low}^0 \\ r_{mid}^0 \\ r_{high}^0 \end{bmatrix}, \quad R_1 = \begin{bmatrix} r_g^g & 0 & 0 & 0 \\ r_{low}^g & r_{low}^{low} & r_{low}^{mid} & r_{low}^{high} \\ r_{mid}^g & r_{mid}^{low} & r_{mid}^{mid} & r_{mid}^{high} \\ r_{high}^g & r_{high}^{low} & r_{high}^{mid} & r_{high}^{high} \end{bmatrix} \tag{11}$$

In the matrix R_1 I set the coefficients $r_g^{low} = r_g^{mid} = r_g^{high} = 0$. 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.[44]

[44] 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.

Discussion of state choice I need to choose which moments of the industry state agents keep track of. I choose the natural gas price and rig availability because these statistics are commonly reported in the annual reports of rig owners and are used by firms who track the industry to describe the state of the market. Drillers respond to an increase in the natural gas price by drilling more projects. Rig availability falls after a sustained increase in gas prices which means that agents differentiate between a short term increase in natural gas prices (high gas prices and high rig availability) versus a long-term increase in gas prices (high gas prices and low rig availability).

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

3.5 Equilibrium

Equilibrium is defined as a set of prices $p_t(x, y)$, capital availability $\{n_{yt}\}_{y \in \{low, mid, high\}}$, demand for capital $h_{yt}(x)$, targeting weights $\omega_{yt}(x)$, entry probability $e_t(x)$, submarket tightness $\{\theta_{yt}\}_{y \in Y}$, and agents' state transition beliefs, that satisfy at each state s_t:

1. The targeting weights $\omega_{yt}(x)$, the entry probability $e_t(x)$, and submarket tightness $\{\theta_{yt}\}_{y \in Y}$, determined by Equations (3) - (5)
2. Demand for capital $h_{yt}(x)$ determined by Equation (6)
3. Agents optimally choose whether to accept/wait if matched using Equations (7) and (8).
4. Equilibrium prices $p_t(x, y)$ determined by Nash bargaining: Equation (9)
5. Updating rule for the distribution of capital $\{n_{yt}\}_{y \in Y}$
6. Beliefs about the future evolution of states given by an AR(1) process with components defined in Equation (11)

4 Estimation and Identification

4.1 Overview

I begin by laying out all the parameters in Table 4 as well the empirical specification of model objects. I then discuss how I estimate the model using data on contracts and whether a rig is matched in two steps: in the first step I construct value functions from objects in the data; in the second step I use simulated method of moments. Finally, I discuss moment choice and identification.

Empirical specification: meeting technology For the meeting technology within each capital submarket I use the following parametric forms:[45]

$$q_y^{capital}(\theta_{yt}) = \min\{1 - \exp(-a_y/\theta_{yt}), 1/\theta_{yt}\} \tag{12}$$

$$q_y^{project}(\theta_{yt}) = \min\{\theta_{yt}(1 - \exp(-a_y/\theta_{yt})), 1\} \tag{13}$$

Empirical specification: demand I place the following parametric assumptions on the distribution of potential wells $f(x)$:

1. The quantity of hydrocarbons is a third-dimensional polynomial of the well complexity: $x_{\text{quantity}} = \rho_0 + \rho_1 x_{\text{complexity}} + \rho_2(x_{\text{complexity}})^2 + \rho_3(x_{\text{complexity}})^3$ where ρ_0, ρ_1, ρ_2, and ρ_3, are parameters. I run an OLS regression to recover the parameters $\rho_0, \rho_1, \rho_2, \rho_3$.
2. Contract durations are for either 2, 3, or 4 months, and are distributed independently of the other covariates with probability weights (τ_2, τ_3, τ_4), where $\tau_4 = 1 - \tau_2 - \tau_3$. Although

[45] This meeting technology can be derived as an approximation to an urn-ball matching function with a large number of agents. This results in the total number of matches as $M = \min\{Y(1 - \exp(-aX/Y)), X, Y\}$, where a is an index of matching efficiency, X is the total number of projects, and Y is the number of available capital (see Petrongolo and Pissarides (2001) for a derivation). Then, the probability of a project matching is $M/X = \min\{\theta(1 - \exp(-a/\theta), 1\}$ and the probability of capital matching is $M/Y = \min\{1 - \exp(-a/\theta), 1/\theta\}$ where $1/\theta = X/Y$. Note that when $a > 1$ - and if market tightness $1/\theta = X/Y$ is sufficiently high so that there are many projects searching compared to available capital - it is possible that all capital might be matched ($M = Y$). Therefore, I need to bound the meeting technology to prevent the model predicting more matches than there are available capital.

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, $x_{\text{complexity}}$, and x_{quantity}.

3. Well complexity is distributed as a truncated normal: $x_{\text{complexity}} \sim TN(\mu, \sigma)$ where the parameters μ, σ need to be estimated. I choose the minimum of the truncated normal as $x_{\text{complexity}} = 0$ and the maximum as $x_{\text{complexity}} = 2.15$.

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.[46]

Calibrated parameters In the steps that follow I calibrate the discount parameter to $\beta = 0.99$ (recall that one period is one month). 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 $c = 1.32$ million USD.[47]

4.2 Estimation

Step 1: Constructing the value functions I first construct the value of searching $U_t(y)$, which serves as a key input into the estimation of parameters in later steps. Note that I proceed in the opposite way to the usual approach in most applications, where value functions are constructed conditional on the parameters. Instead, my strategy is an extension of recent approaches in the Industrial Organization firm dynamics literature such as Kalouptsidi (2014) to cases where short-term contract data are available. (Kalouptsidi (2014) uses data on second-hand sales to estimate value functions with the observation that the resale price of a ship equals the value of a ship.)

[46] As previously discussed, I only use the natural gas price to track if the market is in a boom or bust.

[47] 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 to a rig (including extensions), which is around \$6 million, I calibrate the entry cost as $c = (0.18/0.82) \times 6 = 1.32$ million dollars.

Table 4: Overview of how the model components are computed

Object	Parameters	Method
Discount rate, monthly	β	Calibrated: Preliminary
Entry cost	c	Calibrated: Preliminary
State transition beliefs	$R_0, R_1, \sigma_\epsilon$	Estimated: Step 1
Bargaining weight	δ	Calibrated: Step 1
Demand distribution	$\tau_2, \tau_3, \tau_4, \mu, \sigma, \rho_0, \rho_1, \rho_2, \rho_3$	Estimated: Step 2
Demand draws	k_0, k_1	Estimated: Step 2
Match value	$\{m_{0,y}, m_{1,y}\}_{y\in Y}, m_2$	Estimated: Step 2
Extension parameter	η	Estimated: Step 2
Targeting parameters	γ_0, γ_1	Estimated: Step 2
Meeting Technology	$\{a_y\}_{y\in Y}$	Estimated: Step 2

Note: This table provides an overview of the parameters to be estimated or calibrated.

The intuition is as follows. First, I estimate the beliefs over the state transitions using maximum likelihood and the data on empirical state transitions. With these state transitions in hand, 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.[48] I provide a more formal proof that the value function can be constructed directly from objects in the data in Appendix B.2. In Appendix C.1 I provide more details about how I construct the objects that the value functions can be built from as well as detailing the forward simulation algorithm.

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). An additional benefit of computing the value functions nonparametrically from the data in the above way is that the value of searching $U_t(y)$ is complicated to solve in this application. The complexity comes from

[48] Concretely, 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 the state is updated, and then it searches again.

the fact that there are distributions of agents on both sides of the market and these distributions are changing through time. By contrast, previous work avoids this complexity typically by using a steady state analysis (an exception is Lise and Robin (2017)).

Finally, using the constructed value functions, I calibrate 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 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.[49]

Step 2: Simulated Method of Moments I simulate the model from January 2000 to December 2009. The simulation algorithm computes the equilibrium shares of potential wells that enter and target each type of rig, given the value functions which were computed in Step 1. Specifically, in each period, I iterate over the entry decision of wells and the targeting weights to compute a within-period equilibrium. I then update the model to the next period using these entry and targeting decisions which define the probability of each type of rig being matched to a particular well type. I provide complete details on the simulation algorithm, as well as the implementation of the simulated method of moments, in Appendix C.4.

4.3 Identification and Choice of Moments

I now discuss the intuition behind how the parameters are identified and the choice of moments. I leave a more rigorous discussion of identification to Appendix B.5.

[49] More concretely, I use the following 'steady state' equation for the bargaining weight δ:

$$\delta = 1 - \frac{margin \cdot \bar{p}}{\bar{p} - (1 - margin) \cdot a}$$

where I drop t subscripts, the 'bar' denotes the steady state average value in the sample, and $a = \frac{1}{\sum_{s=0}^{\tau-1} \beta^s} \mathbb{E}_t \left[\beta U - \beta^{\tau}(1-\eta)U - \beta^{\tau+1}\eta U \right]$. I show that this equation can be derived from the price equation in Appendix B.4. I then 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). I compute the above equation for each contract in 2005 and then take the average to recover the bargaining parameter.

Recall that a major challenge is separately identifying changes in the composition of potential wells from the sorting effect. In practice this corresponds to identifying acceptance sets - as well as the meeting technology - separately from the underlying distribution of searching wells. As an outline, I first show that the match value parameters can be identified from prices for different matches. Combined with the value functions from Step 1, these match values then pin down the acceptance sets. Then, I show how the meeting technology can be identified. Finally, I argue that the distribution of potential wells can be reverse engineered from observed matches using the meeting technology, acceptance sets, and the (calibrated) entry cost.

Identifying the match value parameters Rearranging the Nash bargaining solution, prices can be written as a function of the bargaining parameter δ, the parameters that underlie the value of a match, and an additional object $a_t(x, y)$ that can be constructed from the data, in the following way:[50]

$$p_t(x,y) = (1-\delta)a_t(x,y) + \delta m_{0,y} + \delta m_{1,y} x_{\text{complexity}} + \delta \left[\frac{\sum_{k=0}^{\tau-1} \beta^k \mathbb{E}_t[g_{t+k}]}{\sum_{k=0}^{\tau-1} \beta^k} \right] x_{\text{quantity}} \tag{14}$$

Therefore, using the bargaining parameter computed previously and the above result, I run the following auxiliary regression:

$$p_t(x,y) - (1-\delta)a_t(x,y) = \hat{\beta}_{0,y} + \hat{\beta}_{1,y} x_{\text{complexity}} + \hat{\beta}_2 g_t x_{\text{quantity}} \tag{15}$$

This regression provides a 'reduced-form' relationship between prices and the underlying rig-well match in the contract.[51] I include the coefficients from this regression as moments.[52] Intuitively, the $m_{0,y}$ parameters are identified by the $\beta_{0,y}$ moments, the $m_{1,y}$ parameters are identified by the $\beta_{1,y}$ moments, and the β_2 moments identify m_2.

In addition to the moments from the auxiliary regression, I also include two additional price moments that correspond to the price difference between a high-specification rig and a mid-specification rig, and the difference between a mid-specification rig and a low-specification rig.

[50] I derive this result in Appendix B.3. The precise form for $a_t(x,y)$ is: $a_t(x,y) = \frac{1}{\sum_{s=0}^{\tau-1} \beta^s} \mathbb{E}_t \left[\beta U_{t+1}(y) - \beta^\tau (1 - \eta_{t+\tau}(x,y)) U_{t+\tau}(y) - \beta^{\tau+1} \eta_{t+\tau}(x,y) U_{t+\tau+1}(y) \right]$.

[51] For simplicity, I just use the gas price in the first period of the contract in this regression, rather than the average gas price over the entire contract ($\sum_{k=0}^{\tau-1} \beta^k \mathbb{E}_t[g_{t+k}] / \sum_{k=0}^{\tau-1} \beta^k$).

[52] I implement this regression by setting $\beta_{0,low}$ as the intercept.

These moments ensure that the model also generates the ordering in average prices as rig-specification increases.

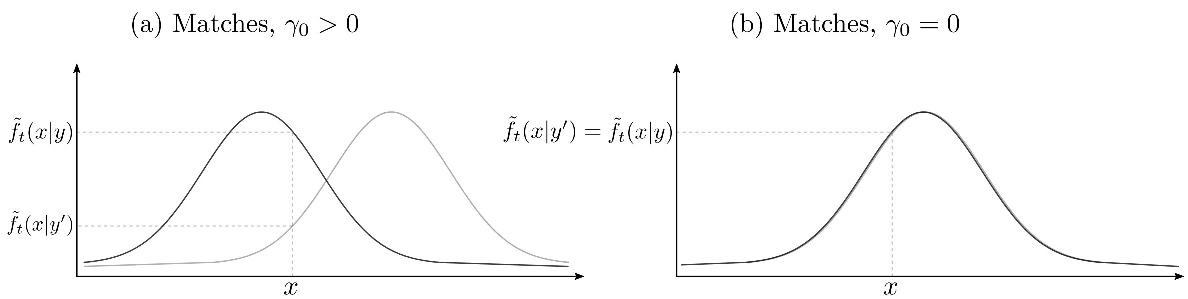
Identifying the extension parameter With the match value parameters identified and the rig's value of searching computed from Step 1, I can also construct the match surplus. I detail the match surplus computational algorithm in the Appendix C.2. Recall that the probability of extending a contract is $\eta 1[S_t(x,y) \geq 0]$: an extension occurs if the match surplus is positive and a draw from a Bernoulli distribution with parameter η. Since the match surplus is known, η can then be identified by matching the model-predicted probability of extension to the empirical probability of extension. Therefore, I include as a moment the average probability of extending a contract over the sample period.

Identifying the targeting parameters For intuition about how the targeting parameter γ_0 can be identified from observed matches, 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 $\gamma_0 = 0$ (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 capital may not be the same: $\tilde{f}_t(x|y) \neq \tilde{f}_t(x|y')$.[53]

Following the above intuition, I use the average match of low, mid, and high specification capital in booms and busts (6 moments). These moments represent the sorting patterns in the data and are sensitive to the targeting parameter γ_0. For example, consider the sorting patterns across the different capital types in a bust. Under random search ($\gamma_0 = 0$), the main channel in the model for differences in matching patterns between the capital types is through differences in the acceptance sets.

[53]This 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 also require that there is some region where the acceptance sets overlap. Recall that the acceptance sets are defined by the region where total surplus of a match is positive, and I can construct total surplus from the match value parameters as well as the rig's value of searching.

Figure 6: Identification of the targeting parameters



Note: This figure gives intuition about how the targeting parameter γ_0 can be identified. Under partially directed search (Panel (a), where $\gamma_0 > 0$) 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 $\gamma_0 = 0$), the distribution of observed matches will be the same and not dependent on x.

I include a moment corresponding to the average utilization in 2006 for high-specification rigs. This helps to pin down the targeting parameter γ_1: if the targeting parameter is $\gamma_1 = 0$ then wells do not avoid matches that will be rejected. In 2006 this would then lead to a utilization rate that is too low, given that in this year high-specification rigs have very constrained acceptance sets. Higher values of γ_1 allow for wells to avoid these matches, resulting in fewer rejections and a higher capital utilization.

Identifying the meeting technology parameters In Appendix B.5 I show that the market tightness terms θ_{yt} can be identified given the above parameters. Then, the meeting technology parameters a_y - which relate how θ_{yt} affects the probability of capital matching - can then be identified from the probability of matching of each type in a single period t.

The potential project draw parameters k_0 and k_1 (where the total number of draws $K_t = k_0 + k_1 g_t$) can be identified from variation in the probability of capital matching *across* periods. For example, higher values of k_1 generate a higher covariance between natural gas prices and the probability of capital matching. Similarly, the value of k_0 is pinned down by the probability of matching when natural gas prices are relatively low.

Following the above arguments, I include moments related to patterns of capital utilization over the boom-bust cycle.[54] Specifically, I 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). These moments pin down the potential well draws: higher values of k_1 increase the covariance between gas prices and utilization; higher values of k_0 tend to reduce the variance of utilization across time. I also include the mean utilization of each capital type (3 moments): given the potential well draw parameters these pin down the meeting technology parameters a_y for $y \in \{low, mid, high\}$.

Identifying the distribution of potential wells The distribution of potential wells $f(x)$ can be identified by inverting the distribution of observed matches through the (now identified) targeting weights and the entry condition. In practice, to pin down the parameters in the distribution of potential wells I include the variance of well complexity matches (1 moments), as well as the previously discussed moments relating to the mean well-complexity matches for each rig type in booms and busts. I also include the probability of observing a 2 month and 3 month contract (2 moments).

[54]The empirical probabilities of matching are quite noisy across time so I instead use capital utilization. Capital utilization increases as the probability of matching increases.

5 Results

Constructing the value functions The results for the beliefs over the state transitions are (standard errors in brackets):

$$R_0 = \begin{bmatrix} r_g^0 \\ r_{low}^0 \\ r_{mid}^0 \\ r_{high}^0 \end{bmatrix} = \begin{bmatrix} 0.81\ (0.36) \\ 6.95\ (1.86) \\ 4.62\ (2.14) \\ 7.02\ (1.72) \end{bmatrix} \qquad \sigma_\epsilon = 1.15\ (0.06)$$

$$R_1 = \begin{bmatrix} r_g^g & 0 & 0 & 0 \\ r_{low}^g & r_{low}^{low} & r_{low}^{mid} & r_{low}^{high} \\ r_{mid}^g & r_{mid}^{low} & r_{mid}^{mid} & r_{mid}^{high} \\ r_{high}^g & r_{high}^{low} & r_{high}^{mid} & r_{high}^{high} \end{bmatrix} = \begin{bmatrix} 0.88\ (0.04) & 0 & 0 & 0 \\ -0.41\ (0.15) & 0.71\ (0.08) & 0.07\ (0.09) & -0.03\ (0.08) \\ -0.34\ (0.15) & 0.2\ (0.09) & 0.48\ (0.09) & 0.21\ (0.08) \\ -0.4\ (0.15) & -0.2\ (0.07) & 0.23\ (0.08) & 0.48\ (0.08) \end{bmatrix}$$

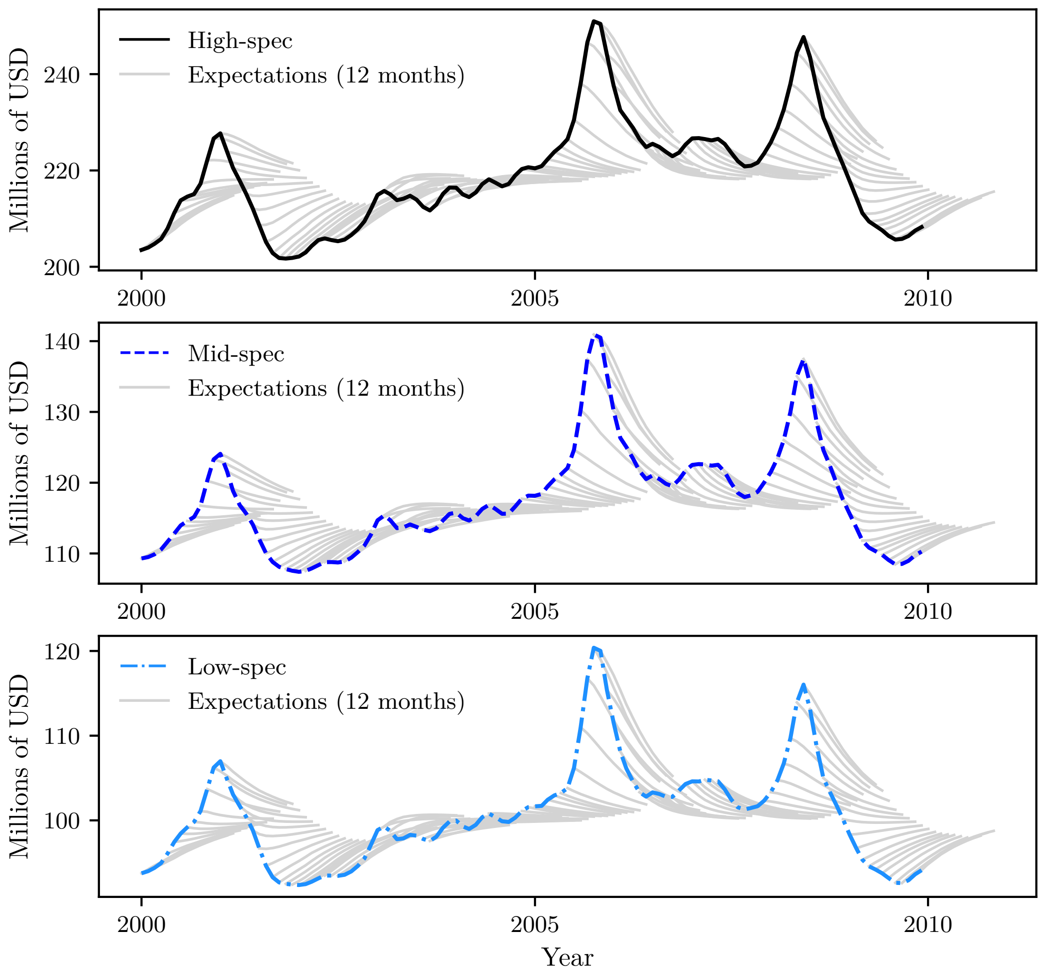
Since all the eigenvalues of R_1 lie within the unit circle, the transition matrix is stationary.

Figure 7 illustrates the outcome of the forward simulation procedure to obtain $U_t(y)$. The light blue, dark blue, and black lines correspond to the value of searching in the current period $U_t(y)$. 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, $\mathbb{E}_t U_{t+2}(y), \mathbb{E}_t U_{t+3}(y)$ etc. The gray lines tend to converge towards the mean which indicates that agents have mean-reverting expectations about the value of searching. This is not surprising because the state transitions are mean reverting.

Estimating the value of a match Table 5 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. 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 that mid- and high-

Figure 7: The rig's value of searching $U_t(y)$



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

Table 5: Match value estimates

Variable	Symbol	Value ($1000s/Day)	95% CI ($1000s/Day)	Value (vs Av. Price)	95% CI (vs Av. Price)
Match value					
High	$m_{0,high}$	56.3	(51.8, 64.8)	0.95	(0.88, 1.1)
	$m_{1,high}$	37.6	(28.4, 43.5)	0.64	(0.48, 0.74)
Mid	$m_{0,mid}$	73.6	(67.2, 80.3)	1.25	(1.14, 1.36)
	$m_{1,mid}$	0.6	(0.5, 0.9)	0.01	(0.008, 0.015)
Low	$m_{0,low}$	88.5	(77.7, 92.8)	1.5	(1.31, 1.57)
	$m_{1,low}$	-38.3	(-40.9, -33.7)	-0.65	(-0.69, -0.57)
Gas value weight	m_2	11.6	(7.1, 16.2)	0.2	(0.12, 0.27)

Variable	Symbol	Value	Error
Bargaining parameter	δ	0.37	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 dayrate. 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.

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 $m_{1,y}$. Recall - as mentioned in Section 3.1 - that 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. 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 $m_2 = 11.6$. 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 \$11.6 thousand USD/Day. Scaling up this per-day figure over a month implies that agents weight a \$1 million increase in this lease bid proxy to a \$0.348 million increase in the total match value.

From these value of a match estimates, and estimates of capital's value of searching, I can construct acceptance sets. I plot the acceptance sets over time in Figure 12 of Appendix D. The acceptance sets for low-specification rigs and high-specification rigs shrink in booms (e.g. most of the time period 2005 - 2007) compared to busts. The way that these acceptance sets shrink depends primarily on the sign of $m_{1,y}$. For example, $m_{1,low} < 0$, so low-specification rigs are not well suited to matching with complex wells; this translates into these rigs rejecting more complex wells in booms. Likewise, $m_{1,high} > 0$ and so these rigs are suited to drilling more complex wells; these rigs therefore reject simple wells in booms. Finally, since $m_{1,mid}$ has a small magnitude (but $m_{0,mid}$ is high), these mid-specification rigs are not poorly suited to drilling any well types but at the same time do not have a notable advantage at drilling any well type. Overall, therefore, although the value of waiting for a better match still increases for these mid-specification rigs, they choose to accept all wells.[55] The implication of mid-specification rigs accepting all wells is that their matches do not change substantially in booms and busts; this is consistent with the empirical sorting patterns in Figure 4.

[55]Note that the empirical value for $m_{1,y}$ is primarily identified by the moment $\hat{\beta}_{1,y}$ which is how responsive prices are to well complexity. The magnitude of $\hat{\beta}_{1,mid}$ is much lower than $\hat{\beta}_{1,low}$ and $\hat{\beta}_{1,high}$. Therefore, since prices are less responsive to well complexity for mid-specification rigs, so is the match value.

Table 6: Estimated demand and meeting technology parameters

Variable	Symbol	Value	Symbol	Value
Demand	μ	0.68	τ_2	0.69
		$(0.56, 0.77)$		$(0.66, 0.71)$
	σ	0.72	τ_3	0.2
		$(0.61, 0.89)$		$(0.18, 0.22)$
	ρ_0	0.021	ρ_1	0.039
		$(0.01, 0.033)$		$(0.013, 0.065)$
	ρ_2	-0.024	ρ_3	0.003
		$(-0.04, -0.009)$		$(0.001, 0.005)$
Entry Cost	c	$1.32 Million		
		n.a.		
Potential Project Draws	k_0	22.4	k_1	8.2
		$(16.3, 29.5)$		$(6.7, 10.3)$
Meeting Technology	a_{low}	0.44	a_{mid}	0.56
		$(0.32, 0.53)$		$(0.43, 0.68)$
	a_{high}	14.02		
		$(9.74, 17.43)$		
Targeting Parameters	γ_0	0.69	γ_1	1.29
		$(0.34, 0.86)$		$(0.77, 1.85)$
Extension Parameter	η	0.37		
		$(0.36, 0.4)$		

Note: All parameters are estimated using the simulated method of moments, except for estimates for $\rho_0, \rho_1, \rho_2, \rho_3$ which are computed using the OLS regression of $x_{\text{quantity}} = \rho_0 + \rho_1 x_{\text{complexity}} + \rho_2 x^2_{\text{complexity}} + \rho_3 x^3_{\text{complexity}}$, 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 200 bootstrap replications, except for estimates for $\rho_0, \rho_1, \rho_2, \rho_3$ which are computed using the standard errors from the OLS regression.

Demand and meeting technology parameters Table 6 contains the estimated parameters from the simulated method of moments. Overall the parameters seem reasonable. The estimated targeting parameter γ_0 is 0.69. To get a sense of where this lies between random search and directed search I consider the probability that a complex well (I set $x_{\text{complexity}} = 2.0$) targets its optimal match (which is a high-specification rig) at approximately the average state:[56]

$$\omega_{y=\text{high},t}(x = \text{complex}) = \begin{cases} 0.23 & \text{random search} \\ 0.33 & \text{estimated model} \\ 1 & \text{directed search} \end{cases}$$

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.

In terms of the targeting parameter γ_1 I find that it is positive but relatively small. Note that even if γ_1 was very large (that is, even if projects were able to completely direct their search away from rejected matches), the sorting effect as well as the quantity-quality tradeoff would still be present. For example, suppose that in booms complex projects are able to completely redirect their search away from low-efficiency capital (where the match would be rejected). This improves the match quality of the potential matches in the low efficiency capital submarket. However it comes as the expense of the meeting probability: there are now fewer projects searching in the low-efficiency submarket, which leads to a reduced probability of low-efficiency capital matching.

For the matching efficiency parameters, the values of a_{low}, a_{mid} indicate that matching in the corresponding submarkets is not perfectly efficient. In contrast, the value of a_{high} is quite high and quite close to frictionless matching in the high-efficiency capital submarket.

For the remaining parameters (e.g. 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.

[56]Specifically, I choose the state halfway through the sample at January 2005 which is also between a boom and bust.

Model fit Table 12 in Appendix D provides a complete comparison of how the simulated moments fit the data. The model replicates the data well. I also provide out-of-sample fit exercises in Section D.6.

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

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

where OPEX is the total operating expenses of the rigs, which I set to be \$32 thousand per day per rig.[57] For each of the counterfactuals I decompose the total effect into three components:

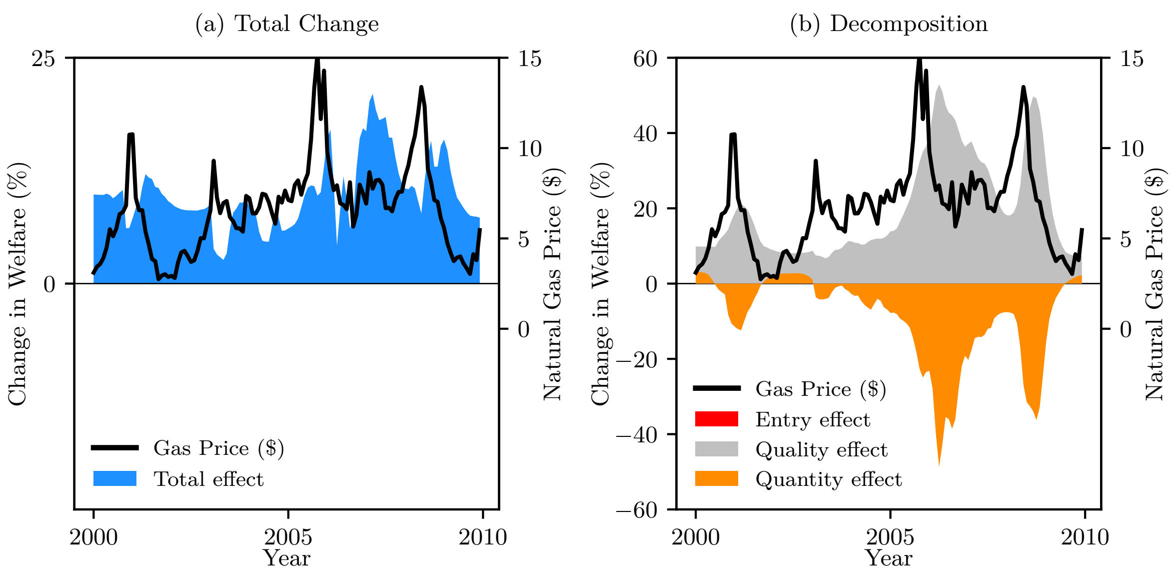
- **Quality effect**: The change in the value of matches keeping the number of matched rigs fixed in both the baseline and counterfactual.
- **Quantity effect**: The value of the new rigs that are matched in the counterfactual (or the loss in value if more rigs are unmatched in the counterfactual).
- **Entry cost saving**: The change in the total entry cost in the counterfactual.

I recompute the value functions in the counterfactuals where necessary. 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.

[57]This figure comes from Kaiser and Snyder (2013), as the expenses for an operating Jackup rig in the US Gulf of Mexico 2010-2011 as reported in the Hercules Offshore annual report. Hercules Offshore is a firm that owns and leases out drilling rigs. Operating expenses include, for example, routine rig maintenance.

6.1 Quantifying the sorting effect

Figure 8: No sorting counterfactual results



(c) Summary of changes

	Boom	Bust	Average
Quality Effect	14.9%	4.6%	19.5%
Quantity Effect	-9.1%	-0.5%	-9.6%
Entry Effect	0.0%	0.0%	0.0%
Total	5.8%	4.1%	9.9%

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 4.9 billion.

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 parameters $\gamma_0 = \gamma_1 = 0$ 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 8. 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 9.9%. The effect is cyclical: the welfare increase in a boom is 5.8% compared to around 4.1% in a bust.

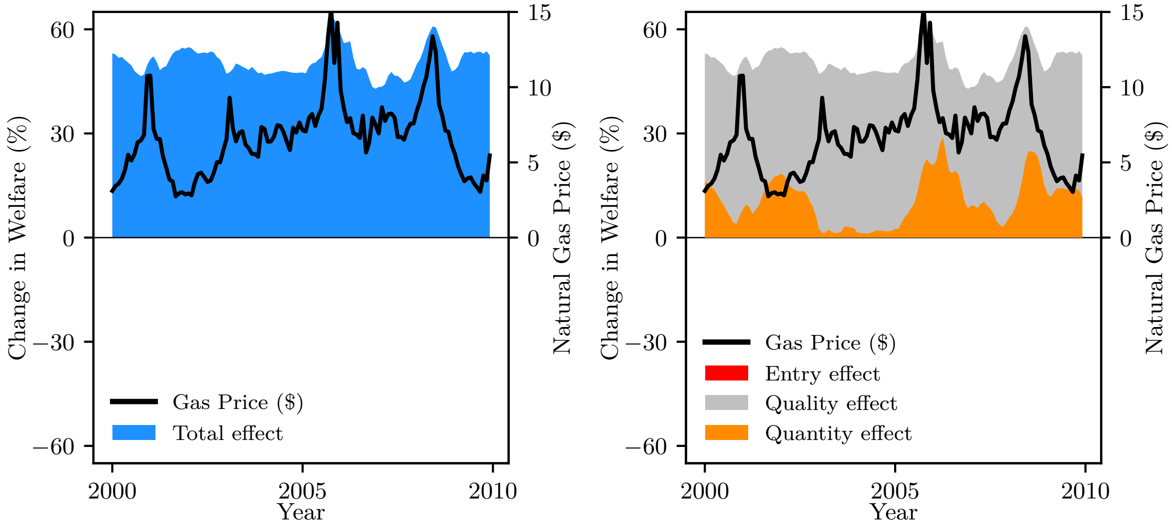
Panels (b) and (c) decompose how the sorting effect increases welfare: there are less matches (which by itself decreases welfare by -9.6%), but the remaining matches are of higher quality because agents are more selective (which increases welfare by 19.5%). Overall, the match quality effect dominates, 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. I need to take a stand on the exact nature of the intermediary in the marketplace. In summary, I use a 'greedy matching algorithm' which matches the set of entered wells to a submarket of available rig types in each period.[58] The algorithm is 'greedy' because it only considers the static match value when choosing which submarket to match each well to. Conceptually, this intermediary can be thought of as an 'Uber for rigs' which introduces a vast improvement in the search technology of the industry. To facilitate a comparison with the sorting effect counterfactual, I also keep the composition of searching wells fixed. I leave details on the implementation algorithm to Appendix C.6.

[58] In this way, this intermediary can be thought of as optimizing the targeting decision of wells.

Figure 9: Intermediary counterfactual results



(c) Summary of changes

	Boom	Bust	Average
Quality Effect	22.7%	17.4%	40.1%
Quantity Effect	5.9%	4.8%	10.7%
Entry Effect	0.0%	0.0%	0.0%
Total	28.6%	22.2%	50.8%

Note: This figure shows the change in welfare when moving from the market baseline to the intermediary counterfactual. 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 4.9 billion.

Figure 9 illustrates the change in welfare due to an intermediary. Panel (a) shows that the intermediary increases welfare by around 50.8% over 2000-2009.[59] This increase in welfare is slightly cyclical: according to Panel (c) the welfare increase in booms is around 28.6% versus 22.2% in busts.

Panel (b) decomposes the effect of the intermediary. Overall, match quality is higher in all

[59]Note that here I am using the intermediary as the comparison, so these figures are the change in welfare relative to the total surplus of an intermediary.

periods and the total improvement in match quality is 40.1%. The magnitude of the quality effect is slightly cyclical, with the gains in the boom 22.7%, which is higher than the gains in the bust at 17.4%. The quantity of matches increases in both booms and busts under an intermediary, however the magnitude of the quantity effect (10.7%) is much lower than the magnitude of the quality effect.

Given the gains from an intermediary, it might seem surprising that there is not currently one in the market. Attempting to explain the non-existence of an intermediary is arguably outside the scope of this paper. That said, participants in the industry are attempting to reduce search frictions. For example, as previously mentioned, recent advances in technology and e-procurement (using the internet to share information and find matching partners) are slowly being incorporated into the industry. This suggests that some of the gains to improving the search process may soon be realized.

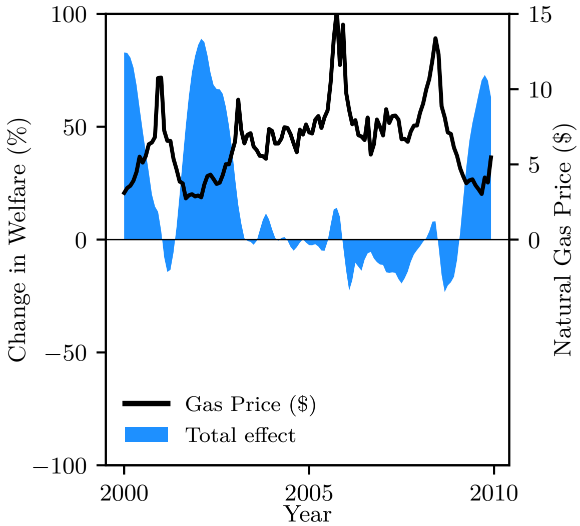
6.3 Effects of a demand smoothing policy

I now consider the effects of a demand smoothing policy. There is a long history in the oil and gas industry of policies designed to smooth out the disruptive effects of the boom-bust cycle. Between 1954 and 1978 natural gas producer prices were fixed in the United States for interstate trade. Today, 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.[60] 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.

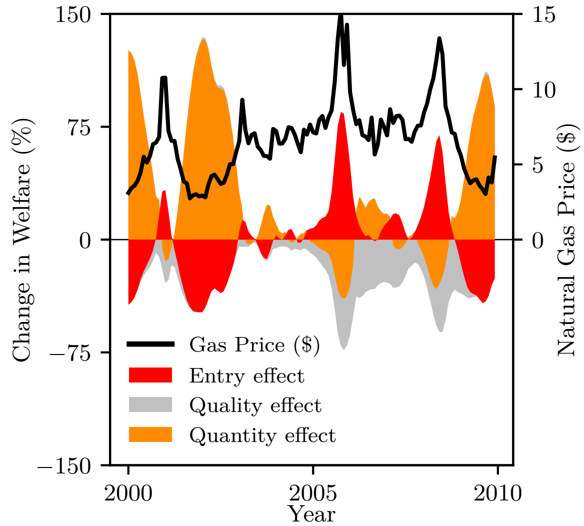
[60] Potter et al. (2017) summarizes the tax credits oil and gas producers receive in low-price environments.

Figure 10: Demand smoothing counterfactual results

(a) Total Change



(b) Decomposition



(c) Summary of changes

	Boom	Bust	Average
Quality Effect	-11.1%	-1.1%	-12.2%
Quantity Effect	-2.0%	31.5%	29.5%
Entry Effect	10.9%	-11.6%	-0.7%
Total	-2.2%	18.8%	16.6%

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

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.7.

The results are depicted in Figure 10. Panel (a) shows that the smoothing policy results in large shifts in drilling activity. The shift is somewhat cyclical, with a slight decrease in welfare in the boom (-2.2%) versus an increase in the bust (18.8%).

Panel (b) illustrates the determinants of the total change. First, consider the effects of demand smoothing on the quantity effect. The changes here are straightforward: demand smoothing increases both the quantity of matches in a bust and reduces the quantity in the boom, mainly due to changes in the number of potential well draws. Next, in terms of the quality effect, demand smoothing decreases the quality of matches in a boom by -11.1% since agents are less selective. Interestingly, match quality also decreases slightly by -1.1% in the *bust*: one reason is that demand smoothing results in different compositions of wells entering in these periods (usually those which result in lower-value matches), but these matches are accepted because the market is at the long-run average gas price and not in a boom.

The effects on the sign of the entry effect are 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 -11.6%) and less entry in a boom. Putting all of the effects together results in an increase in total surplus in a bust that is counteracted by a decrease in total surplus in the boom.

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 16.6%. 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, and provide an identification strategy to separate the sorting effect from changes in the composition of searching projects (demand). 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.

Baccara, M., Lee, S. and Yariv, L. (2020), 'Optimal dynamic matching', *Theoretical Economics* **15**, 1221–1278.

Barlevy, G. (2002), 'The Sullying Effect of Recessions', *The Review of Economic Studies* **69**(1), 65–69.

Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2013), 'Cross-Country Differences in Productivity: The Role of Allocation and Selection', *American Economic Review* **103**(1), 305–334.

Becker, G. S. (1973), 'A Theory of Marriage: Part I', *Journal of Political Economy* **81**(4), 813–846.

Brancaccio, G., Kalouptsidi, M. and Papageorgiou, T. (2020), 'Geography, Transportation, and Endogenous Trade Costs', *Econometrica* **88**, 657–691.

Brancaccio, G., Kalouptsidi, M., Papageorgiou, T. and Rosaia, N. (2022), 'Search Frictions and Efficiency in Decentralized Transport Markets', *Working Paper* .

Buchholz, N. (2022), 'Spatial Equilibrium, Search Frictions, and Dynamic Efficiency in the Taxi Industry', *Review of Economic Studies* **89**, 556–591.

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), 131–171.

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**.

Frechette, G., Lizzeri, A. and Salz, T. (2019), 'Frictions in a Competitive Regulated Market: Evidence from Taxis', *American Economic Review* **109**(8), 2954–2992.

Gavazza, A. (2016), 'An Empirical Equilibrium Model of a Decentralized Asset Market', *Econometrica* **84**(5), 1755–1798.

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. and Snyder, B. (2013), *The Offshore Drilling Industry and Rig Construction in the Gulf of Mexico*, Springer-Verlag, London.

Kalouptsidi, 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), 1104–1135.

Moscarini, G. and Postel-Vinay, F. (2018), 'The cyclical job ladder', *Annual Review of Economics* **10**(1), 165–188.

Ottonello, P. (2018), 'Capital Unemployment', *Working Paper* .

Petrongolo, B. and Pissarides, C. (2001), 'Looking into the Black Box: A Survey of the Matching Function', *Journal of Economic Literature* **39**, 390–431.

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* .

Rothgerber, U. (2002), 'E-Procurement Still Evolving For E and P Industry', *Oil and Gas Journal* **100**(16).

Salz, T. (2022), 'Intermediation and Competition in Search Markets: An Empirical Case Study', *Journal of Political Economy* **130**(2).

Villas-Boas, S. B. (2007), 'Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data', *The Review of Economic Studies* **74**(2), 625–652.

Vreugdenhil, N. (2021), 'Incomplete Environmental Regulation and Spatial Equilibrium in the Offshore Oil and Gas Industry', *Working Paper* .

Appendices

Contents

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
- The well activity report dataset

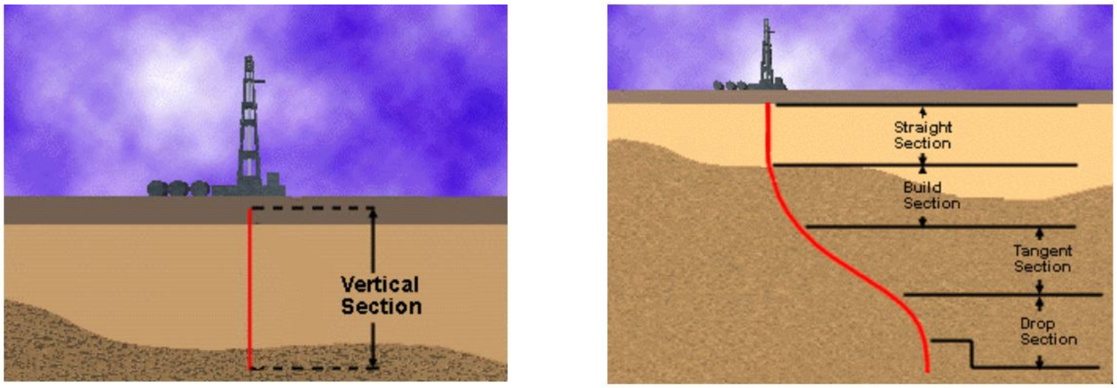
I begin by merging the borehole, permit, lease, and well activity report datasets on the unique API well number to obtain a single dataset with all well characteristics at the well level which I call the 'well database'.

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 3380 contracts for jackup rigs in total. I further remove 17 'workover' rigs, 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*, *marinetraffic.com*, and *vesselfinder.com*. I also use these websites plus the Rigzone order book to find the maximum drilling depth of each rig. I merge all but one of the rig names in the IHS contract dataset to the well dataset in this way. In total, removing the workover rigs and merging with the rig name crosswalk reduces the contract dataset to 3103 contracts.

Figure 11: Diagram of a simple vs complex well

(a) A simple well   (b) A complex well



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

In total there are 3103 contracts and 5129 wells in the datasets for the years 2000-2009. The procedure successfully results in matching 2355 contracts and 4937 wells. I further impute the characteristics of 300 contracts if the contract was an extension/renegotiation. In total 2655 contracts (86%) and 4937 wells (96%) 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. Finally, I compute the 'number of rigs' as the average number of rigs of each type per month in the sample.

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

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

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

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

Next 'key drilling factors' are computed. These are: $\psi_1 = 3$ if there is a horizontal sections; $\psi_2 = 3$ if there is a J-curve; $\psi_3 = 2$ if there is an S-curve; ψ_4 if there is a subsea well; $\psi_5 = 1$ if there is an H_2S/CO_2 environment; $\psi_6 = 1$ if there is a hydrate environment; $\psi_7 = 1$ if there is a depleted sand section; $\psi_8 = 1$ if there is a salt section; $\psi_9 = 1$ if there is a slimhole, $\psi_{10} = 1$ if there is a mudline suspension system installed; $\psi_{11} = 1$ if there is coring; $\psi_{12} = 1$ if there is shallow water flow potential; $\psi_{13} = 1$ if there is riserless mud to drill shallow water flows; $\psi_{14} = 1$ if there is a loop current.

The Mechanical Risk Index is then computed as:

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

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

Computing the 'key drilling factors' ψ_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 Microfoundation for the targeting weights

In this sub-section I provide a micro-foundation for how projects contact capital (i.e. Equation (4)).

Denote each unit of available capital by j and the corresponding type as y_j. Similarly, denote each searching project by i and its corresponding type by x_i. Using this notation, for a project of type x_i, the (expected) value of targeting capital j is $\pi_{y_j t}(x_i)$.

In the special case where search is perfectly directed then each potential project i will choose j to solve $\max_j \pi_{y_j t}(x_i)$. In my setting, I allow for a more flexible search technology by instead modeling potential projects targeting capital based on a *perceived value* $\hat{\pi}_{y_j t}(x_i)$ which is defined as:

$$\hat{\pi}_{y_j t}(x_i) = \pi_{y_j t}(x_i) - \gamma_1 1\left[S_t(x_i, y_j) < 0\right] + \epsilon_{ijt}^{target} \tag{16}$$

I assume that ϵ_{ijt}^{target} are drawn from an i.i.d. type-1 extreme value distribution with scale parameter $1/\gamma_0$. The interpretation of γ_0 and γ_1 is that they are 'targeting parameters' that index how precisely a project can target capital. I allow for targeting to be responsive to both whether the match will be rejected (the parameter γ_1) as well as the overall quality of the match: γ_0.[61]

The conditional choice probabilities that result from $\max_j \hat{\pi}_{y_j t}(x_i)$ are given as follows:

$$P_{ijt} = \frac{\exp\Big(\gamma_0\left[\pi_{y_j t}(x_i) - \gamma_1 1\left[S_t(x_i, y_j) < 0\right]\right]\Big)}{\sum_k \exp\Big(\gamma_0\left[\pi_{y_k t}(x_i) - \gamma_1 1\left[S_t(x_i, y_k) < 0\right]\right]\Big)} \tag{17}$$

Finally, noting that $P_{ijt} = P_{ij't}$ for all capital of the same type $y_j = y_{j'}$, and that there are n_{yt} rigs of type y, Equation (17) can be aggregated to form the probability of a project of type x targeting a rig of type y:

$$\omega_{yt}(x) = \frac{n_{yt}\exp\Big(\gamma_0\Big[\pi_{yt}(x) - \gamma_1 1\left[S_t(x, y) < 0\right]\Big]\Big)}{\sum_{k \in Y} n_{kt}\exp\Big(\gamma_0\Big[\pi_{kt}(x) - \gamma_1 1\left[S_t(x, k) < 0\right]\Big]\Big)} \tag{18}$$

B.2 Constructing $U_t(y)$ from the data

The aim is to show that the value of search $U_t(y)$ can be computed from the data.

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

$$\begin{aligned}
U_t(y) &= \int_z \max\Big\{V_t(z, y), \beta\mathbb{E}_t U_{t+1}(y)\Big\} h_{yt}(z)dz + h_{yt}(\emptyset)\beta\mathbb{E}_t U_{t+1}(y) && (19)\\
&= \int_{z \in A_{yt}} h_{yt}(z)V_t(z, y)dz + \left(\int_{z \notin A_{yt}} h_{yt}(z)dz + h_{yt}(\emptyset)\right) \cdot \beta\mathbb{E}_t U_{t+1}(y) && (20)\\
&= \sum_{n \in \{2,3,4\}} \mathbb{P}_{\tau=n,t}(y) \cdot \mathbb{E}_t\mathbb{E}_{x_{-\tau}|\tau=n,t} V_t(x, y) + \mathbb{P}_{\tau=0,t}(y) \cdot \beta\mathbb{E}_t U_{t+1}(y) && (21)
\end{aligned}$$

[61] Whether a match could be rejected might be more salient to capital owners than other features of the match - I allow the data to determine whether this is the case.

where, in the third equality, n indexes an element in the set of possible contract lengths and $\mathbb{E}_{x_{-\tau}|\tau=n,t}$ is an expectation taken over all the covariates in a potential project (except τ) conditional on $\tau = n$ at time t (i.e. the expectation is over the covariates x_{quantity} and $x_{\text{complexity}}$). Recall that value of a match to the capital owner is:

$$V_t(x,y) = \sum_{k=0}^{\tau-1} \beta^k p_t(x,y) + \beta^\tau \mathbb{E}_t \Big[\eta_{t+\tau}(x,y) V_{t+\tau}(x,y) + (1-\eta_{t+\tau}(x,y)) U_{t+\tau}(y)\Big] \tag{22}$$

Taking conditional expectations of both sides of Equation (22) with respect to the covariates (except τ):

$$\begin{aligned}
\mathbb{E}_{x_{-\tau}|\tau=n,t} V_t(x,y) &= \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[\sum_{k=0}^{n-1} \beta^k p_t(x,y)\Big] \\
&\quad + \beta^n \mathbb{E}_t \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[\eta_{t+n}(x,y) V_{t+n}(x,y) + (1-\eta_{t+n}(x,y)) U_{t+n}(y)\Big] \qquad (23) \\
&= \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[\sum_{k=0}^{n-1} \beta^k p_t(x,y)\Big] \\
&\quad + \beta^n \mathbb{E}_t \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[\eta_{t+n}(x,y) \sum_{k=0}^{n-1} \beta^k p_{t+n}(x,y) + (1-\eta_{t+n}(x,y)) U_{t+n}(y)\Big] \\
&\quad + \beta^{2n} \mathbb{E}_t \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[\\
&\qquad \eta_{t+n}(x,y)\eta_{t+2n}(x,y) \sum_{k=0}^{n-1} \beta^k p_{t+2n}(x,y) + \eta_{t+n}(x,y)(1-\eta_{t+2n}(x,y)) U_{t+2n}(y)\Big] \\
&\quad + \ldots \qquad (24) \\
&= \sum_{k=0}^{n-1} \beta^k \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[p_t(x,y)\Big] \\
&\quad + \beta^n \mathbb{E}_t \sum_{k=0}^{n-1} \beta^k \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[\eta_{t+n}(x,y) p_{t+n}(x,y)\Big] \\
&\quad + \beta^n \mathbb{E}_t \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[(1-\eta_{t+n}(x,y)) U_{t+n}(y)\Big] \\
&\quad + \beta^{2n} \mathbb{E}_t \sum_{k=0}^{n-1} \beta^k \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[\eta_{t+n}(x,y)\eta_{t+2n}(x,y) p_{t+2n}(x,y)\Big] \\
&\quad + \beta^{2n} \mathbb{E}_t \mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[\eta_{t+2n}(x,y)(1-\eta_{t+n}(x,y)) U_{t+2n}(y)\Big] \\
&\quad + \ldots \qquad (25)
\end{aligned}$$

Note that the per-period payoff $\mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[p_t(x,y)\Big] = \bar{p}_{\tau=n,t}(y)$, which is the average price (conditional on the rig type, state, and contract length), and is directly observed in the data. Furthermore, the values of extending the contract $\mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[\eta_{t+n}(x,y) p_{t+n}(x,y)\Big]$, and $\mathbb{E}_{x_{-\tau}|\tau=n,t}\Big[(1-\eta_{t+n}(x,y)) U_{t+n}(y)\Big]$, etc, can be constructed, since they are just a function of the observed probabilities of extension, observed prices of extensions, or future values of U.

Although the probabilities of extension and the prices of extensions are observed directly in the data and so could be computed in principle, I only have limited data on extensions and a large dimensional space over which these components need to be computed. Therefore, I need to make some further assumptions about how to aggregate these components. The main assumption that I make is that these components are well-approximated by combinations of the mean empirical probabilities of extensions and mean empirical probabilities of prices computed only at the state where the extension is occurring. Therefore, for example, I approximate the second component of Equation (25) by:

$$\mathbb{E}_{x_{-\tau}|\tau=n,t}\big[\eta_{t+n}(x,y)p_{t+n}(x,y)\big] = \bar{\eta}_{\tau=n,t+n}(y)\cdot\bar{p}_{\tau=n,t+n}(y) \tag{26}$$

In the above equation $\bar{\eta}_{\tau=n,t+n}(y)$ denotes the probability of extending a contract of length n for a rig of type y at the state at time $t+n$. Similar, $\bar{p}_{\tau=n,t+n}(y)$ denotes the average price. The assumption that the average probability of extending a contract, and the average price if an extension occurs, are well-approximated by these components computed only at the state where the extension is occurring, is restrictive. For example, it rules out the dependence of the probability of extending a contract for a second time at $t+2n$ on the state at the first extension at $t+n$. This implicitly ignores that different states at $t+n$ might cause different types of matches to survive at $t+n$, which might lead to different probabilities of extension at $t+2n$. However, allowing the average probabilities of extension (and also the extension prices) to be dependent on the entire sequence of states from when the contract was initially signed is not feasible in this setting due to limited data. I emphasize that in practice this assumption is probably not too strong in this application. For example, contracts typically do not survive after one or two extensions, and so most of the variation in the value functions is driven by the state affecting the probabilities of being matched in a new contract (and the corresponding price).

Overall:

$$\begin{aligned} U_t(y) = \sum_{n\in\{2,3,4\}} \mathbb{P}_{\tau=n,t}(y)\cdot \Bigg\{ & \\ & \sum_{k=0}^{n-1}\beta^k\bar{p}_{\tau=n,t}(y) + \beta^n\mathbb{E}_t\Bigg[\bar{\eta}_{\tau=n,t+n}(y)\Big(\sum_{k=0}^{n-1}\beta^k\bar{p}_{\tau=n,t+n}(y)+\ldots\Big) + (1-\bar{\eta}_{\tau=n,t+n}(y))U_{t+n}(y)\Bigg] \\ \Bigg\} & + \mathbb{P}_{\tau=0,t}(y)\cdot\beta\mathbb{E}_tU_{t+1}(y) \end{aligned} \tag{27}$$

Notice that the above expression implies that $U_t(y)$ can be written in terms of the average price of extensions, the average price of new contracts, the probability of matching different length contracts, the extension probability, and future values of U, which proves the result.

B.3 Proof that prices can be written as in Section 4.3

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

$$V_t(x,y) - \beta\mathbb{E}_tU_{t+1}(y) = \delta S_t(x,y) \tag{28}$$

Substituting in Equation (7) into $V_t(x,y)$:

$$\sum_{k=0}^{\tau-1}\beta^k p_t(x,y)+\beta^\tau\mathbb{E}_t\Big[\eta_{t+\tau}(x,y)V_{t+\tau}(x,y)+(1-\eta_{t+\tau}(x,y))U_{t+\tau}(y)\Big]-\beta\mathbb{E}_tU_{t+1}(y)$$
$$=\delta\Big[W_t(x,y)+V_t(x,y)-\beta\mathbb{E}_tU_{t+1}(y)\Big] \tag{29}$$

Further substituting in for $V_t(x,y)$ and $W_t(x,y)$:

$$\sum_{k=0}^{\tau-1}\beta^k p_t(x,y)+\beta^\tau\mathbb{E}_t\Big[\eta_{t+\tau}(x,y)V_{t+\tau}(x,y)+(1-\eta_{t+\tau}(x,y))U_{t+\tau}(y)\Big]-\beta\mathbb{E}_tU_{t+1}(y)$$
$$=\delta\Big[\sum_{k=0}^{\tau-1}\beta^k v_{t,k}(x,y)+\beta^\tau\mathbb{E}_t\Big[\eta_{t+\tau}(x,y)V_{t+\tau}(x,y)+\eta_{t+\tau}(x,y)W_{t+\tau}(x,y)+(1-\eta_{t+\tau}(x,y))U_{t+\tau}(y)\Big]-\beta\mathbb{E}_tU_{t+1}(y)\Big] \tag{30}$$

Substituting in for surplus in the above equation:

$$\sum_{k=0}^{\tau-1}\beta^k p_t(x,y)+\beta^\tau\mathbb{E}_t\Big[\eta_{t+\tau}(x,y)\Big[\delta S_{t+\tau}(x,y)+\beta U_{t+\tau+1}(y)\Big]+(1-\eta_{t+\tau}(x,y))U_{t+\tau}(y)\Big]-\beta\mathbb{E}_tU_{t+1}(y)$$
$$=\delta\Big[\sum_{k=0}^{\tau-1}\beta^k v_{t,k}(x,y)+\beta^\tau\mathbb{E}_t\Big[\eta_{t+\tau}(x,y)\Big[S_{t+\tau}(x,y)+\beta U_{t+\tau+1}(y)\Big]+(1-\eta_{t+\tau}(x,y))U_{t+\tau}(y)\Big]-\beta\mathbb{E}_tU_{t+1}(y)\Big] \tag{31}$$

Rearranging the equation further:

$$\sum_{k=0}^{\tau-1}\beta^k p_t(x,y)=\delta\Big[\sum_{k=0}^{\tau-1}\beta^k v_{t,k}(x,y)\Big]+(1-\delta)\mathbb{E}_t\Big[\beta U_{t+1}(y)-\beta^\tau(1-\eta_{t+\tau}(x,y))U_{t+\tau}(y)-\beta^{\tau+1}\eta_{t+\tau}(x,y)U_{t+\tau+1}(y)\Big] \tag{32}$$

Finally, rearranging the above equation produces the result:

$$p_t(x,y)=(1-\delta)a_t(x,y)+\delta\left[\frac{\sum_{k=0}^{\tau-1}\beta^k v_{t,k}(x,y)}{\sum_{k=0}^{\tau-1}\beta^k}\right] \tag{33}$$
$$=(1-\delta)a_t(x,y)+\delta m_{0,y}+\delta m_{1,y}x_{\text{complexity}}+\delta\left[\frac{\sum_{k=0}^{\tau-1}\beta^k\mathbb{E}_t[g_{t+k}]}{\sum_{k=0}^{\tau-1}\beta^k}\right]x_{\text{quantity}} \tag{34}$$

where $a_t(x,y)=\frac{1}{\sum_{s=0}^{\tau-1}\beta^s}\mathbb{E}_t\left[\beta U_{t+1}(y)-\beta^\tau(1-\eta_{t+\tau}(x,y))U_{t+\tau}(y)-\beta^{\tau+1}\eta_{t+\tau}(x,y)U_{t+\tau+1}(y)\right]$.

B.4 Bargaining parameter equation in footnote B.4

I show that the bargaining equation in footnote B.4 can be derived from the price equation (Equation (34)). Note that the bargaining parameter is calibrated from data from the year 2005 which is a period where the industry is approximately at its long-run steady-state. Therefore, I first drop time subscripts and work with the steady-state equivalent of Equation (34). Taking averages (denoted with a bar) on both sides of this equation results in: $\bar{p}=(1-\delta)\bar{a}+\delta\bar{v}$. Finally, substituting in for the average match value using $\bar{v}=\bar{p}/(1-margin)$ and rearranging leads to the result that $\delta=1-\frac{margin\cdot\bar{p}}{\bar{p}-(1-margin)\cdot\bar{a}}$.

B.5 Details about identification in Section 4.3

The aim is to show that the targeting parameters γ_0, γ_1, the matching efficiency parameters a_y, the distribution of potential projects $f(x)$, and the potential project draw parameters k_0, k_1, are identified. The observed distribution of projects that type y capital matches with is:

$$\tilde{f}_t(x|y) = \begin{cases} q_y^{capital}(\theta_{yt}) \frac{\omega_{yt}(x)e_t(x)f(x)}{\int_z \omega_{yt}(z)e_t(z)f(z)dz} & \text{if } x \in A_{yt} \\ 1 - q_y^{capital}(\theta_{yt}) \frac{\int_{z \in A_{yt}} \omega_{yt}(z)e_t(z)f(z)dz}{\int_z \omega_{yt}(z)e_t(z)f(z)dz} & \text{if } x = \emptyset \end{cases} \tag{35}$$

where $x = \emptyset$ corresponds to the capital being unmatched. At this stage the probability of projects meeting a match $q_y^{project}(\theta_{yt})$, as well as the weights $\omega_{yt}(x)$, are not known.

Note that I previously showed that the value function for capital can be constructed from the data. Similarly, as discussed in the main text, the match value parameters can be identified from prices. Therefore I assume that the value of a match to a project, $W_t(x, y)$, as well as acceptance sets, are known for this proof. Furthermore, I assume that the entry cost c is calibrated from external data.

I complete the proof in several steps:

1. Identify the targeting weights $\omega_{yt}(x)$
2. Identify the targeting parameter γ_0
3. Identify the matching efficiency parameters a_y
4. Identify the distribution of potential projects $f(x)$ and the parameters that underlie the potential project draws k_0, k_1 where $K_t = k_0 + k_1 g_t$
5. Identify the targeting parameter γ_1

Part 1 I show how the targeting weights $\omega_{yt}(x)$ can be identified. Rewriting the equation for the targeting weights:

$$\omega_{yt}(x) = \frac{n_{yt} \exp\left(\gamma_0 \left[\pi_{yt}(x) - \gamma_1 1\left[S_t(x, y) < 0\right]\right]\right)}{\sum\limits_{k \in Y} n_{kt} \exp\left(\gamma_0 \left[\pi_{kt}(x) - \gamma_1 1\left[S_t(x, k) < 0\right]\right]\right)} \tag{36}$$

Recall that the acceptance sets are known at this stage. Consider one project type x that will accept a match with any capital type (i.e. $x \in A_{yt}$ for each $y \in \{\text{low}, \text{mid}, \text{high}\}$ and therefore $1\left[S_t(x, y) < 0\right] = 0$). Comparing the probability of the same type of project x matching conditional on different capital y, y'

at t:

$$\begin{aligned}
\ln\Big(\tilde{f}_t(x|y)\Big) - \ln\Big(\tilde{f}_t(x|y')\Big) &= \ln\Big(\omega_{yt}(x)/\omega_{y't}(x)\Big) \\
&\quad + \ln\Big(q_y^{capital}(\theta_{yt})/q_{y'}^{capital}(\theta_{y't})\Big) \\
&\quad + \ln\Big(\int_z \omega_{y't}(z)e_t(z)f(z)dz/\int_z \omega_{yt}(z)e_t(z)f(z)dz\Big) \qquad (37)\\
&= \gamma_0\Big(\pi_{yt}(x) - \pi_{y't}(x)\Big) \\
&\quad + \ln\Big(n_{yt}/n_{y't}\Big) \\
&\quad + \ln\Big(q_y^{capital}(\theta_{yt})/q_{y'}^{capital}(\theta_{y't})\Big) \\
&\quad + \ln\Big(\int_z \omega_{y't}(z)e_t(z)f(z)dz/\int_z \omega_{yt}(z)e_t(z)f(z)dz\Big) \qquad (38)
\end{aligned}$$

Differencing Equation (38) over two points x and x':

$$\begin{aligned}
&\left(\ln\Big(\tilde{f}_t(x|y)\Big) - \ln\Big(\tilde{f}_t(x|y')\Big)\right) - \left(\ln\Big(\tilde{f}_t(x'|y)\Big) - \ln\Big(\tilde{f}_t(x'|y')\Big)\right) \\
&\qquad = \gamma_0\Big(\pi_{yt}(x) - \pi_{y't}(x) - \big(\pi_{yt}(x') - \pi_{y't}(x')\big)\Big) \qquad (39)\\
&\qquad = \gamma_0 q_y^{project}(\theta_{yt}) \cdot \Big(W_t(x,y) - W_t(x',y)\Big) \\
&\qquad\quad - \gamma_0 q_{y'}^{project}(\theta_{y't}) \cdot \Big(W_t(x,y') - W_t(x',y')\Big) \qquad (40)
\end{aligned}$$

Here the second equality follows from substituting the expression:

$$\pi_{yt}(x) = q_y^{project}(\theta_{yt})W_t(x,y) \qquad (41)$$

In Equation (40) the left-hand-side is data and the $W_t(x,y)$ terms are known. Therefore, I can identify $\gamma_0 q_y^{project}(\theta_{yt})$ for each $y \in \{\text{low}, \text{mid}, \text{high}\}$. Hence, $\gamma_0\pi_{yt}(x)$ can be constructed for any match (x,y) using Equation (41) and $\gamma_0 q_y^{project}(\theta_{yt})$. Finally, I can recover the weights $\omega_{yt}(x)$ because they are just a function of $\gamma_0\pi_{yt}(x)$ (for the case here where all rigs accept a match with a well of type x).

Part 2: Next I show how γ_0 can be identified from the data. Denote $\widehat{q_y}^{project}(\theta_{yt}) = \gamma_0 q_y^{project}(\theta_{yt})$. As in Part 1, continue to consider one project type x that will accept a match with any capital type. Denote the expected value from entering as $(1/\gamma_0)g_t(x) = (1/\gamma_0)\sum_{k\in Y}\omega_{kt}(x)W_t(x,k)\widehat{q_k}^{project}(\theta_{kt})$. Rearranging the entry condition:

$$e_t(x) = \frac{\exp\Big(\sum_{k\in Y}\omega_{kt}(x)\pi_{kt}(x) - c\Big)}{1+\exp\Big(\sum_{k\in Y}\omega_{kt}(x)\pi_{kt}(x) - c\Big)} = \frac{\exp\Big((1/\gamma_0)g_t(x) - c\Big)}{1+\exp\Big((1/\gamma_0)g_t(x) - c\Big)} \qquad (42)$$

Note that $g_t(x)$ can be constructed from the data and also note that I calibrate the entry cost c using external data. Therefore, all the components of Equation (42) are known with the exception of the parameter γ_0.

Next, consider the relative probability of two projects x and x' entering at different time periods t and t'. Differencing Equation (35) over these two points x and x' and over the two time periods, and rearranging:

$$\begin{aligned}
&\left(\ln\left(\tilde{f}_t(x|y)/\omega_{yt}(x)\right) - \ln\left(\tilde{f}_t(x'|y)/\omega_{yt}(x')\right)\right) \\
&- \left(\ln\left(\tilde{f}_{t'}(x|y)/\omega_{yt'}(x)\right) - \ln\left(\tilde{f}_{t'}(x'|y)/\omega_{yt'}(x')\right)\right) \\
&= \ln\left(\frac{e_t(x)/e_t(x')}{e_{t'}(x)/e_{t'}(x')}\right) && (43) \\
&= (1/\gamma_0)\big((g_t(x) - g_t(x')) - (g_{t'}(x) - g_{t'}(x'))\big) \\
&\quad + \ln\left(\frac{1+\exp((1/\gamma_0)g_t(x') - c)}{1+\exp((1/\gamma_0)g_t(x) - c)} \cdot \frac{1+\exp((1/\gamma_0)g_{t'}(x) - c)}{1+\exp((1/\gamma_0)g_{t'}(x') - c)}\right) && (44)
\end{aligned}$$

Here, the left-hand-side can be constructed from the data: the $\tilde{f}_t(x|y)$ are directly observed and the weight terms $\omega_{yt}(x)$ are identified using Part 1. Furthermore, since the right-hand-side is known up to the parameter γ_0, the parameter γ_0 is identified. Therefore, the $q_y^{project}(\theta_{yt})$ are also identified.

Part 3: Suppose that there is a period in the data where the acceptance sets for capital type y is all wells x (that is, suppose no matches are rejected). Then, $q_y^{capital}(\theta_{yt})$ is directly observed in the data since it is the probability of capital type y matching. Since $\theta_{yt} = q_y^{project}(\theta_{yt})/q_y^{capital}(\theta_{yt})$, and $q_y^{capital}(\theta_{yt})$ is directly observed in the data, the market tightness terms θ_{yt} are also identified. Therefore, the matching efficiency terms a_y are also identified, through the effect of variation in θ_{yt} on the (observed) probability of capital matching $q_y^{capital}(\theta_{yt})$.

Part 4: The distribution of potential projects $f(x)$ can be identified using the time periods in Part 3 where no matches are rejected. Concretely, using these time periods, $f(x)$ can be identified by inverting the distribution of observed matches $\tilde{f}_t(x|y)$ through the targeting weights and the entry condition. Similarly, in these periods the number of potential well draws K_t can be identified because the market tightness θ_{yt} is known and $\theta_{yt} = \frac{n_{yt}}{K_t \cdot \int_z \omega_{yt}(z)e_t(z)f(z)dz}$, where the number of available capital n_{yt} is known from the data and $\int_z \omega_{yt}(z)e_t(z)f(z)dz$ is identified from the previous steps. The potential well draw parameters (k_0 and k_1) can then be recovered by exploiting variation of K_t across time.

Part 5: Finally, there is one parameter remaining to identify: the targeting parameter γ_1. This parameter is identified by ensuring that capital's probability of matching is not too low in periods where the acceptance sets are narrow (that is, periods where many potential matches are rejected).

C Computation

C.1 Computing capital's value of searching $U_t(y)$

I estimate the empirical objects used to construct the value functions in Section 4.2 as follows.

Prices To estimate $\bar{p}_{\tau=n,t}(y)$ 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 I estimate $\bar{\eta}_{\tau=n,t}(y)$ (the average probability of extending a τ length contract given the state at time t for capital type y) as follows. For each capital type $y \in \{low, mid, high\}$ I estimate a logit model for whether a contract is extended on first and second order polynomial combinations of the state vector plus an extra term for the contract length.

The probability of matching To estimate $\mathbb{P}_{\tau=n,t}(y)$ 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.

Algorithm The algorithm that I use is based on the results where I show how $U_t(y)$ can be constructed from the data in Appendix B.2.

Overall, I use forward simulation to compute the value function over a set of points (nodes). I linearly interpolate the value of searching over the state space using a grid with 10 nodes for the gas price dimension and 5 nodes for each available capital state dimension (so there are $5^3 \times 10 = 1250$ grid nodes in total). I perform the algorithm separately for each $y \in \{\text{low}, \text{mid}, \text{high}\}$. For each node s in the state space grid, repeat the following forward simulation algorithm:[62]

1. Initialize the state as the state corresponding to the node.
2. Given the state, draw a contract from the estimated multinomial logit for the probability of matching. (Note that this incorporates the possibility that the rig will be unmatched i.e. draw a 0 length contract.)

[62] I end the algorithm after 1200 periods have elapsed. Since a time that is k periods in the future is discounted by β^k, periods in the later values of the forward simulation are given very low weight.

3. If a contract is drawn:

   (a) Get the corresponding price using the empirical object estimated for prices.

   (b) Update the state; count down the number of periods remaining on the contract by 1.

   (c) If the number of periods left on the contract is 0, simulate the probability of extension using the empirical object computed above. If the contract is extended, begin again from Part (a).[63]

   (d) If an extension does not occur: return to 2.

4. If a contract is not drawn: update the state and return to 2.

I repeat the above forward simulation algorithm 200 times. I then take the average value of these 200 simulations to compute $U_t(y)$.

C.2 Computing the match surplus

When simulating the model, I compute the match surplus $S_t(x, y)$ using forward simulation. I explain the details in this section.

Extension probability I estimate $\eta_t(x, y)$ as follows.[64] For each capital type $y \in \{low, mid, high\}$ I estimate a logit model for whether a contract is extended on first and second order polynomial combinations of the state vector, as well as two extra terms: one for the contract length and one for well complexity.

Algorithm: overview Overall, I use forward simulation to compute the surplus over a set of points (nodes). I linearly interpolate the surplus using a grid with 10 nodes for the gas price dimension, 5 nodes for each available capital state dimension, and 5 nodes for the well complexity dimension (so there are

[63] Since this step occurs after the state is updated, it is consistent with the timing outlined in the model section that the extensions occur first. Also recall that the extension will be of the same match (including the contract duration) as the original contract.

[64] When computing the rig's value of searching recall that I computed a different object for extensions based on the average probability of extending a contract (where the average is taken over the distribution of potential matches at a future state in time). In computing the match surplus the exact (x,y) match is fixed. Therefore, for computational simplicity, I estimate the probability of extending a contract conditional on a particular match.

$5^3 \times 10 \times 5 = 6250$ grid nodes in total).[65] I perform the algorithm separately for each $y \in \{\text{low}, \text{mid}, \text{high}\}$ and for each contract length $\tau \in \{2, 3, 4\}$.[66]

When searching over the objective function in the simulated method of moments, I need to quickly compute the match surplus for different candidate parameter values of $m_{0,y}, m_{1,y}, m_2, \rho_0, \rho_1, \rho_2, \rho_3$. The algorithm I set out below exploits that the match surplus is a linear function of these parameters and that the nodes of the surplus interpolation grid are fixed. Specifically, the algorithm computes the coefficients on these parameters in the match surplus for each node. The computational benefit is that at a given node in the interpolation grid, I only need to perform the forward simulation algorithm once.

The coefficients on the parameters that I simulate can be derived in the following way:

$$\begin{aligned} S_t(x,y) &= W_t(x,y) + V_t(x,y) - \beta \mathbb{E}_t U_{t+1}(y) \\ &= \sum_{k=0}^{\tau-1} \beta^k v_{t,k}(x,y) + \beta^\tau \mathbb{E}_t \Big[\eta_{t+\tau}(x,y) \left(\sum_{k=0}^{\tau-1} \beta^k v_{t+\tau,k}(x,y) + ... \right) + (1 - \eta_{t+\tau}(x,y)) U_{t+\tau}(y) \Big] - \beta \mathbb{E}_t U_{t+1}(y) \end{aligned}$$

and therefore:

$$S_t(x,y) = m_{0,y} b_0 + m_{1,y} b_1 + m_2 \cdot (\rho_0 b_2 + \rho_1 b_3 + \rho_2 b_4 + \rho_3 b_5) + b_6 - \beta \cdot b_7$$

---

[65] By assumption, in the model the quantity of hydrocarbons $x_{quantity}$ is assumed to by a polynomial function of the well complexity variable $x_{complexity}$. Therefore, given this polynomial function, I can compute $x_{quantity}$ from a particular $x_{complexity}$. That is, I do not need to include $x_{quantity}$ as an additional dimension to be interpolated over.

[66] Therefore, there are 9 different components that I compute for surplus. Later, when I need to compute the surplus at a particular match, the code first chooses the component relating to the capital type and contract length of the match, and then linear interpolates surplus over the continuous variables i.e. the state and well complexity.

where the coefficients are:

$$b_0 = \mathbb{E}_t \left[\sum_{k=0}^{\tau-1} \beta^k + \eta_{t+\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{\tau+k} + \eta_{t+\tau}(x,y)\eta_{t+2\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{2\tau+k} + ... \right]$$

$$b_1 = \mathbb{E}_t \left[\sum_{k=0}^{\tau-1} \beta^k + \eta_{t+\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{\tau+k} + \eta_{t+\tau}(x,y)\eta_{t+2\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{2\tau+k} + ... \right] \cdot x_{complexity}$$

$$b_2 = \mathbb{E}_t \left[\sum_{k=0}^{\tau-1} \beta^k g_{t+k} + \eta_{t+\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{\tau+k} g_{t+\tau+k} + \eta_{t+\tau}(x,y)\eta_{t+2\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{2\tau+k} g_{t+2\tau+k} + ... \right]$$

$$b_3 = \mathbb{E}_t \left[\sum_{k=0}^{\tau-1} \beta^k g_{t+k} + \eta_{t+\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{\tau+k} g_{t+\tau+k} + \eta_{t+\tau}(x,y)\eta_{t+2\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{2\tau+k} g_{t+2\tau+k} + ... \right] \cdot x_{complexity}$$

$$b_4 = \mathbb{E}_t \left[\sum_{k=0}^{\tau-1} \beta^k g_{t+k} + \eta_{t+\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{\tau+k} g_{t+\tau+k} + \eta_{t+\tau}(x,y)\eta_{t+2\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{2\tau+k} g_{t+2\tau+k} + ... \right] \cdot x^2_{complexity}$$

$$b_5 = \mathbb{E}_t \left[\sum_{k=0}^{\tau-1} \beta^k g_{t+k} + \eta_{t+\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{\tau+k} g_{t+\tau+k} + \eta_{t+\tau}(x,y)\eta_{t+2\tau}(x,y) \sum_{k=0}^{\tau-1} \beta^{2\tau+k} g_{t+2\tau+k} + ... \right] \cdot x^3_{complexity}$$

$$b_6 = \mathbb{E}_t \Big[(1-\eta_{t+\tau}(x,y))U_{t+\tau}(y) + \eta_{t+\tau}(x,y)(1-\eta_{t+2\tau}(x,y))U_{t+2\tau}(y) + ... \Big]$$

$$b_7 = \mathbb{E}_t U_{t+1}(y)$$

Algorithm: computing the surplus at each node For each node s in the state space grid, I compute b_j for $j \in \{0, 1, ..., 7\}$ by forward simulation. I take the average value of each b_j over 200 simulations. For each simulation, I stop the algorithm after 120 periods have elapsed. Since the values of later periods in the surplus equation are weighted by the probability that the contract has been extended until that period, the simulation error from ending the algorithm after 120 periods is extremely low.

C.3 Computing prices

Algorithm: overview Similar to the surplus computation above, when estimating the model I need to quickly compute prices for different values of the parameters. In order to do this, I use Equation (34) which shows how prices can be written as a function of the bargaining parameter, capital's value of searching, and the value of a match. For every contract in the data, I simulate the following components:

$$
\begin{aligned}
a_t(x,y) &= \frac{1}{\sum_{s=0}^{\tau-1}\beta^s}\mathbb{E}_t\left[\beta U_{t+1}(y) - \beta^{\tau}(1-\eta_{t+\tau}(x,y))U_{t+\tau}(y) - \beta^{\tau+1}\eta_{t+\tau}(x,y)U_{t+\tau+1}(y)\right] \\
b_0 &= \frac{1}{\sum_{s=0}^{\tau-1}\beta^s}\cdot\sum_{k=0}^{\tau-1}\beta^k \\
b_1 &= \frac{1}{\sum_{s=0}^{\tau-1}\beta^s}\cdot\sum_{k=0}^{\tau-1}\beta^k\cdot x_{complexity} \\
b_2 &= \frac{1}{\sum_{s=0}^{\tau-1}\beta^s}\cdot\sum_{k=0}^{\tau-1}\beta^k\cdot\mathbb{E}_t[g_{t+k}]x_{quantity}
\end{aligned}
$$

and, therefore, prices can be constructed in the following way (recalling that the bargaining parameter δ is calibrated):

$$p_t(x,y) = (1-\delta)a_t(x,y) + \delta(m_{0,y}b_0 + m_{1,y}b_1 + m_2b_2) \tag{45}$$

Therefore, when estimating the model, the algorithm requires simulating the components $a_t(x,y), b_0, b_1, b_2$ only once for a given match. I use 200 simulations to compute these components. Prices can then be quickly computed using Equation (45) for different candidate parameters $m_{0,y}, m_{1,y}, m_2$.

C.4 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.4.1 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. 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 $share^i_{yt}$ where the share of wells that do not enter is $1 - \sum_{k\in\{low,mid,high\}} share^i_{kt}$. The variable i denotes the iteration (so the guess initializes at $i = 0$).

2. Get the submarket tightness θ^i_{yt} using:

$$\theta^i_{yt} = \frac{n_{yt}}{share^i_{yt} \cdot K_t} \tag{46}$$

and note that this pins down the expected surplus of project type x to searching in the type y capital submarket:

$$\pi^i_{yt}(x) = q^{project}_y(\theta^i_{yt}) W_t(x, y) \tag{47}$$

3. Update the shares using:

$$share^{i+1}_y = \int_z w^i_{yt}(z) e^i_t(z) f(z) dz \tag{48}$$

where the weights are defined using Equation (4):

$$\omega^i_{yt}(x) = \frac{n_{yt} \exp\left(\gamma_0 \left[\pi^i_{yt}(x) - \gamma_1 1\left[S_t(x, y) < 0\right]\right]\right)}{\sum_{k \in Y} n_{kt} \exp\left(\gamma_0 \left[\pi^i_{kt}(x) - \gamma_1 1\left[S_t(x, k) < 0\right]\right]\right)} \tag{49}$$

and the entry condition is defined using Equation (5):

$$e^i_t(x) = \frac{\exp\left(\sum_{k \in Y} \omega^i_{kt}(x) \pi^i_{kt}(x) - c\right)}{1 + \exp\left(\sum_{k \in Y} \omega^i_{kt}(x) \pi^i_{kt}(x) - c\right)} \tag{50}$$

4. Repeat steps 2-3 until the targeting shares converge. Denote the equilibrium shares as $share^*_{yt}$, the equilibrium weights as $w^*_{yt}(x)$, and the equilibrium probability of entry as $e^*_t(x)$.

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. In addition, since some components of the empirical state can be noisy (e.g. the number of available rigs of each type), I smooth the empirical state using a local polynomial regression.

C.4.2 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, for each capital type y, a distribution of current matches.

For the distribution of current matches, I discretize $x_{complexity}$ into 30 bins. In addition, each match also has a corresponding contract length τ, and a number of remaining periods on the contract. So, for example, the number of possible detailed states is divided into $30 \times 2 = 60$ bins for $\tau = 2$ contracts, 90 bins for $\tau = 3$ contracts, and 120 bins for $\tau = 4$ contracts.

Note that agents' expectations about future states (and their corresponding value functions) are computed over the state $s_t = (g_t, n_{low,t}, n_{mid,t}, n_{high,t})$ rather than the detailed state. Therefore, the value functions are not subject to a curse of dimensionality.

I compute the market as follows (starting from a guess of the detailed state for January 2000 - which includes the empirical natural gas price - where I first burn-in the simulation).

1. Compute the probability that each match is extended and update the 'detailed state' with these contract extensions.
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. Use these equilibrium values to update the 'detailed state'.
4. Repeat Steps 2-4 over the monthly natural gas price evolution from January 2000 - December 2009.

C.4.3 Other implementation details for the simulated method of moments

I stack the simulated moments: $m_s(\boldsymbol{\lambda})$, where I denote $\boldsymbol{\lambda}$ as the vector of parameters to be estimated. I fit the simulated moments m_s to the empirical moments m_d by minimizing the following objective function (denoting Ω as the weighting matrix): $(m_d - m_s(\boldsymbol{\lambda}))'\Omega(m_d - m_s(\boldsymbol{\lambda}))$.

I set the weighting matrix Ω to be a diagonal matrix. I set the weights for the medium-specification rig matches, as well as the low- and mid-specification mean utilization moments, to 0.1. I set the 2006 high-specification utilization moment weight to 0.01. I set the utilization variance for the high-specification rig to 10, and the moment on the coefficient β_3 to 100. I set all the remaining weights to 1.[67]

C.5 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

[67]These weights are chosen to ensure that the model closely replicates moments such as the sorting patterns, which are of primary interest, as well as to ensure the moments have a similar scale.

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 parameters are set so that $\gamma_0 = \gamma_1 = 0$ 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 $E_t U_{t+k}(y) = 0$ for $k \in \{1, 2, ...\}$ (that is, agents do not reject matches due to high future outside options).

C.6 Algorithm: intermediary

In this section I describe the algorithm I use to compute the equilibrium with an intermediary. Given that the match value is supermodular, positive assortative matching is optimal in a static model. Motivated by this idea, I look for cutoff solutions in the following form: one cutoff $\bar{x}_{complexity}$ where all wells where $x_{complexity} > \bar{x}_{complexity}$ are assigned to target high-specification rigs, and another cutoff $\underline{x}_{complexity}$ where if $\bar{x}_{complexity} \geq x_{complexity} > \underline{x}_{complexity}$ then the well is assigned to target a mid-specification rig, and if $x_{complexity} \leq \underline{x}_{complexity}$ then the well is assigned to target a low-specification rig. Finally, I assume that if a rig and well meet under the intermediary's protocol then the match will be accepted, and also that matches will be extended randomly (i.e. with probability η).[68] Note that since I consider a myopic algorithm, as well as also keeping the composition of entered wells the same as in the baseline model and assuming that matches will be accepted, I do not need to re-solve for the value functions or agents' beliefs over the state evolution in this counterfactual.

An alternative algorithm would be an intermediary which also internalizes dynamic considerations. I choose to not implement this alternative for several reasons. First, in the myopic algorithm, rigs are already strongly assortatively matched. This blunts the incentives behind the sorting effect as there is less of a benefit to waiting for a better match. Second, the greedy matching algorithm does not require re-solving for value functions and beliefs that are changing across time and so is not computationally intensive. Third the matching protocol is relatively simple and so could be feasibly implemented by a real-world intermediary. Finally, to the extent that there are additional benefits from incorporating dynamic considerations, this simple greedy algorithm can be interpreted as a lower bound on the potential benefits of an intermediary.

The algorithm proceeds as follows *within* each period:

1. Given a set of entered wells and available rigs, and potential cutoffs $(\underline{x}_{complexity}, \bar{x}_{complexity})$, allocate wells to rig submarkets.

[68] The rationale for this assumption - as well as computational tractability - is that since the intermediary's allocation results in high quality matches, there is little incentive to reject a match (or to not extend a match) in order to wait for a better match.

2. Given the above allocations, compute the total expected value of the (static) matches.
3. Repeat Steps 1.-2. and recompute over all potential cutoffs.
4. Given the above steps, choose the solution which maximizes the total expected value of the (static) matches.

Then, using the above solution, I update the state of the market and move to the next period where I run the above algorithm again.

C.7 Algorithm: demand smoothing

In this section I detail the algorithm I use to compute the demand smoothing counterfactual. 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). Since by assumption in this counterfactual the state evolution is in a steady state at the mean natural gas price over 2000-2009 ($\bar{g}$), the transitions are in the following form:

$$R_0 = \begin{bmatrix} \bar{g} \\ \bar{n}_{low} \\ \bar{n}_{mid} \\ \bar{n}_{high} \end{bmatrix}, \quad R_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \sigma_\epsilon = 0 \tag{51}$$

In the above equations the mean number of available rigs of each type $\bar{n}_{low}, \bar{n}_{mid}, \bar{n}_{high}$ will change. Therefore, these components need to be recomputed in the algorithm, amongst other things.

Overall, the algorithm can be viewed as featuring an inner loop and outer loop. In the inner loop, I recompute value functions and re-simulate the model using the same algorithm as the original model. In the outer loop I treat the results of the re-simulated model as 'data'. From these 'data' I recompute the state evolution beliefs and other empirical objects used to construct value functions, and continue to iterate until convergence.

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.
2. Initialize a guess of the objects used to compute the value of searching: the average prices for each capital type, the average probability of matching for each capital type, and the average probability of extending a contract for each capital type.[69]

[69]When implementing this procedure for computational simplicity I fix the probability of extending the contract

3. Denote the value of searching in the j-th iteration as $U^j(y)$. Construct this value of searching for each capital type y using the current guess of the objects used to compute the value of searching using the procedure in Section C.1.

4. Using $U^j(y)$, simulate the model using the algorithm in Section C.4.

5. From the simulated model, recompute the objects used to compute the value of searching (the average prices for each capital type, the average probability of matching for each capital type, and the average probability of extending a contract for each capital type) and update the state evolution beliefs R_0^j, R_1^j.

6. Check for convergence.[70] If the algorithm has not converged, repeat from Step 3.

as η. Note that this restriction is made without loss of generality. Since the market is in steady state the surplus of a match is fixed over time. Therefore, if a match is initially accepted, the match surplus is positive and so will also be positive at the time when the contract is extended. Hence, each initial match is extended with probability $\eta_t(x,y) = \eta 1[S_t(x,y) > 0] = \eta$.

[70] My metric for convergence is the maximum absolute difference in the average prices. I define convergence as whether this metric is < 0.001.

D Additional Graphs, Tables, and Results

D.1 Synergies between high efficiency rigs and high complexity wells during boom periods

I show that there are larger synergies between high efficiency rigs and high complexity wells during boom periods, as reflected in prices. To do so I run the following regression at the contract level (for a match between well i and rig j at time t):

$$\begin{aligned} price_{ijt} = \beta_{0,y_j} + \beta_{1,y_j} 1[\text{High-complexity}_i] + \beta_{2,y_j} g_t + \beta_{3,low} 1[\text{High-complexity}_i] g_t \\ + \beta_{3,y_j} 1[y_j \in \{\text{mid}, \text{high}\}] 1[\text{High-complexity}_i] g_t + \epsilon_{ijt} \end{aligned} \tag{52}$$

In Equation (52), $1[\text{High-complexity}_i]$ is an indicator for whether the well complexity for well i is in the top 10% of complexity. The component g_t is the natural gas price. The coefficients are dependent on the type of rig $y_j \in \{low, mid, high\}$. The component $\beta_{3,low}$ is included in all regressions, while if $y_j \in \{mid, high\}$ then the components $\beta_{3,mid}$ or $\beta_{3,high}$ are additionally included.

Equation (52) captures how high-specification rig and complex well synergies change in busts vs booms through the coefficent $\beta_{3,high}$. I report regressions based on Equation (52) in Table 7. In column (1) I report the value of $\beta_{3,high}$ for an initial specification without controls. In column (2) I include project controls (a control for the value of hydrocarbons) and in column (3) I additionally include contract-type controls (contract duration and whether the contract is an extension). Overall, the results show that the price premium for a 'well-matched' high-specification rig is higher in a boom than a bust. Concretely, moving from a bust (e.g. a natural gas price of \$3) to a boom (e.g. a natural gas price of \$9) results in a price premium of 'well-matched' high-specification rigs of around \$30 thousand dollars per day. For comparison, the average price per day for a high-specification rig over the cycle is around \$70 thousand dollars per day, so the \$30 thousand dollars per day number is substantial.

D.2 Why do I assume the extension is the same length as the initial contract?

I assume that an extended contract is the same length as the original contract because this is typically the case in the data. Concretely, using the aggregated measures of duration that I use in the model ($\tau \in \{2, 3, 4\}$), I test how often the duration changes between subsequent extended contracts. I find that the probability that the duration remains the same across extended contracts is 64.8% . Furthermore, since the majority of contracts are new contracts, extended contracts with a different duration from the original contract represent only 12.2% of all contracts.

Table 7: Regressions of synergies

	(1) Price/day	(2) Price/day	(3) Price/day
Δ synergies in boom: complex wells/high-spec rigs	5.42***	5.41***	4.95***
	(1.22)	(1.22)	(1.26)
Project controls	No	Yes	Yes
Contract-type controls	No	No	Yes
N	2655	2655	2655

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This regression reports the estimated $\beta_{3,high}$ from the equation: $price_{ijt} = \beta_{0,y_j} + \beta_{1,y_j} 1[\text{High-complexity}_i] + \beta_{2,y_j} g_t + \beta_{3,low} 1[\text{High-complexity}_i] g_t + \beta_{3,y_j} 1[y_j \in \{\text{mid}, \text{high}\}] 1[\text{High-complexity}_i] g_t + \epsilon_{ijt}$. (The component $\beta_{3,low}$ is included in all regressions, while if $y_j \in \{mid, high\}$ then the components $\beta_{3,mid}$ and $\beta_{3,high}$ are additionally included.) Prices are in thousands of USD per day. The Δ synergies in boom term corresponds to $\beta_{3,high}$ i.e. the change in a price per day for a 'well-matched' high-specification rig for a \$1 increase in the natural gas price. Project controls indicates a control for the value of hydrocarbons (i.e. the quantity of hydrocarbons multiplied by the current natural gas price). Contract-type controls corresponds to controls for contract duration and also whether the contract is an extension or not. Robust standard errors are in brackets.

D.3 Do rigs change ranking between contracts?

In order to answer this question I split the data up into two five-year periods: January 2000 to December 2004 and January 2005 to December 2009. I begin by de-meaning the price by the average in each month (to remove price cyclicality). Using the de-meaned average prices I then rank rigs into three categories (low, medium, and high) in each of the five-year periods for the rigs that are present in both periods. I then investigate the extent to which rig rankings stay constant between the two periods.

I find that rig ranking remains remarkably constant throughout the two time periods, with 84.6% of rigs maintaining their ranking across periods.[71] In terms of the small fraction of rigs that change their category under this metric, note that average prices are a noisy measure of rig type since they are also dependent on the well that the rig is matched to. In addition, the average price for each rig in each period might only be computed from a few contracts.

D.4 Additional results

Table 8: Utilization by rig specification and whether the market is in a boom or bust

	Bust	Boom
Utilization: low-specification rigs	0.549	0.657
Utilization: mid-specification rigs	0.661	0.804
Utilization: high-specification rigs	0.894	0.976

Note: The variable ‘boom’ is an indicator for whether the market is in a boom or not (defined as the gas price being above or below average).

[71] I also experimented with computing the rankings by splitting the data up into the boom versus the bust. This results in a similarly high number of rigs that maintain their price ranking, with 85.5% of rigs keeping the same rank.

Table 9: Regressions of sorting patterns

	(1) Complexity	(2) Complexity	(3) Complexity
1[Low-spec]	0.787***	0.795***	0.692***
	(0.02)	(0.02)	(0.038)
1[Low-spec] × 1[Boom]	-0.063**	-0.064**	-0.061**
	(0.027)	(0.027)	(0.027)
1[Mid-spec]	0.844***	0.851***	0.751***
	(0.02)	(0.02)	(0.037)
1[Mid-spec] × 1[Boom]	0.016	0.016	0.012
	(0.028)	(0.028)	(0.028)
1[High-spec]	0.9***	0.91***	0.805***
	(0.023)	(0.024)	(0.042)
1[High-spec] × 1[Boom]	0.083**	0.084**	0.089**
	(0.038)	(0.038)	(0.038)
Project controls	No	Yes	Yes
Contract-type controls	No	No	Yes
N	2655	2655	2655

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The variable 'boom' is an indicator for whether the market is in a boom or not (defined as the gas price being above or below average). Project controls indicates a control for the quantity of hydrocarbons. Contract-type controls corresponds to controls for contract duration. Robust standard errors are in brackets.

Table 10: Hedonic regressions of price on characteristics

	(1) Price/day	(2) Price/day	(3) Price/day
1[Low-spec]	46.2***	45.8***	38.4***
	(0.5)	(0.5)	(1.4)
1[Low-spec] × Complexity	-4.8***	-4.7***	-5.5***
	(1.7)	(1.7)	(1.7)
1[Mid-spec]	52.7***	52.3***	45.0***
	(0.5)	(0.5)	(1.4)
1[Mid-spec] × Complexity	2.1*	2.0*	1.8
	(1.1)	(1.1)	(1.1)
1[High-spec]	71.3***	70.7***	63.3***
	(0.8)	(0.8)	(1.6)
1[High-spec] × Complexity	10.6***	10.7***	10.6***
	(1.6)	(1.6)	(1.6)
Value of hydrocarbons		2.0***	1.9***
		(0.7)	(0.7)
State and rig-type interactions	Yes	Yes	Yes
Contract-type controls	No	No	Yes
N	2655	2655	2655

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Prices are in thousands of US dollars (per day). Contract duration is in months. Value is the quantity of hydrocarbons in a well multiplied by the current natural gas price. State and rig type interactions correspond to including controls for four state variables: gas price, and the number of available rigs of each type, as well as the interactions of these variables with an indicator for each of the three rig types. I de-mean the complexity index variable, as well as the state variables, so that the indicators 1[Low-spec], 1[Mid-spec], 1[High-spec], correspond to the average price per day of each rig type. Robust standard errors are in brackets.

Table 11: Regressions of well drilling duration on whether the market is in a boom or bust

	(1)	(2)	(3)
Dependent Variable	Duration	Duration	Duration
1[Boom]	0.14	0.08	-0.03
	(0.79)	(0.8)	(0.8)
Third order polynomial of complexity	Yes	Yes	Yes
Rig type FEs	No	Yes	Yes
Rig type FEs interacted with complexity polynomial	No	No	Yes
N	5652	5652	5652

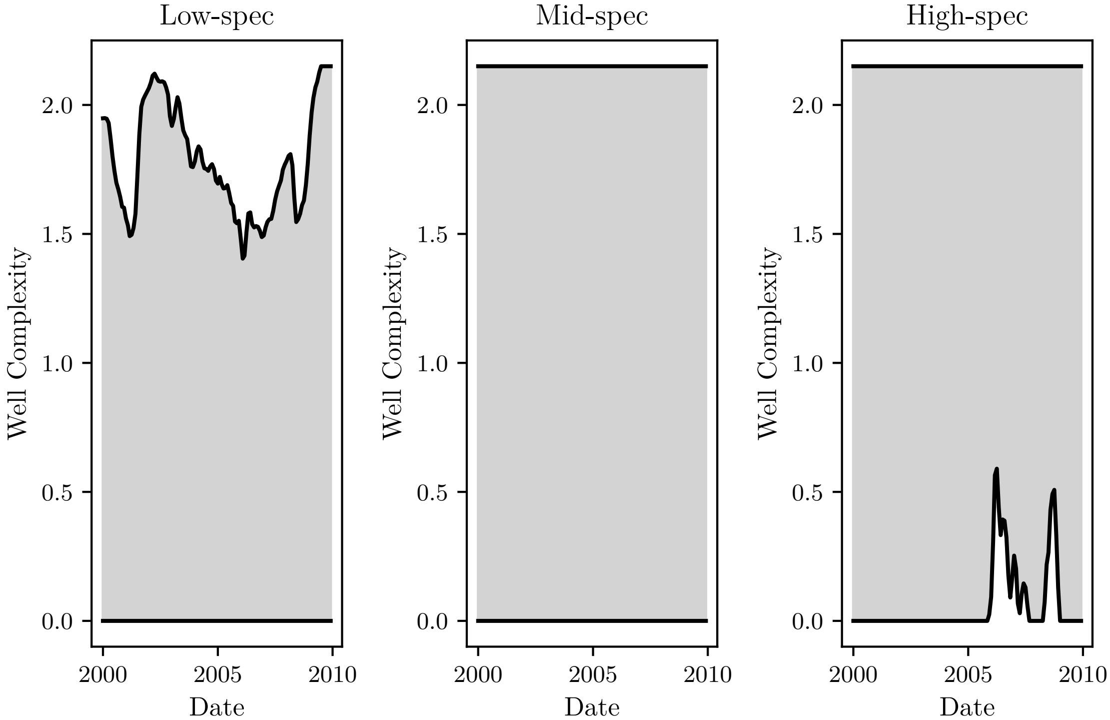
Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Drilling duration is measured in days. The variable 'boom' is an indicator for whether the market is in a boom or not (defined as the gas price being above or below average). The results illustrates that drilling speed (after controlling for the complexity of a well) does not change in booms vs busts. Robust standard errors are in brackets.

Table 12: Fit of the simulation to the moments used in the estimation

		Data	Simulated		Data	Simulated
Matches: Complexity				**Matches: Duration**		
Mean, Bust:	Low	0.79	0.79	Prob. τ_2	0.7	0.7
	Mid	0.84	0.87	Prob. τ_3	0.19	0.19
	High	0.9	0.9	**Price**		
Mean, Boom:	Low	0.72	0.72	$\hat{\beta}_0$	0.4	0.3
	Mid	0.86	0.87	$\hat{\beta}_{1,low}$	-0.09	-0.1
	High	0.98	0.98	$\hat{\beta}_{0,mid}$	-0.06	-0.05
Variance		0.26	0.26	$\hat{\beta}_{1,mid}$	0.02	0.002
Utilization				$\hat{\beta}_{0,high}$	-0.1	-0.1
Mean:	Low	0.6	0.61	$\hat{\beta}_{1,high}$	0.1	0.1
	Mid	0.73	0.71	$\hat{\beta}_2$	0.04	0.04
	High	0.94	0.93	Difference: High-Mid	0.1	0.2
Variance:	Low	0.032	0.0076	Difference: Mid-Low	0.07	0.1
	Mid	0.024	0.0096	**Extensions**		
	High	0.0086	0.0051	Mean	0.35	0.35
Covariance:	Low	0.18	0.19			
	Mid	0.22	0.21			
	High	0.12	0.12			
Level in 2006:	High	0.97	0.92			

Note: This table contains the moments used in the simulated method of moments step. In this table I present the 'price' moments in hundreds of thousands of dollars for readability (in the model and throughout the paper I measure prices in millions of dollars). I report both the value observed in the data and the simulated moments at the optimal parameters.

Figure 12: Acceptance sets for 2 month contracts over time



Note: This table contains the acceptance sets computed for a contract of $\tau = 2$ 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 three rig types computed at the empirical state value and with the minimum bound 0 and the maximum bound 2.15. 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.

D.5 Simple version of the model

I now set out a model with two types of rigs and wells, to show the intuition for the result. The main takeaway is that whether stronger sorting in a boom is optimal - and whether the model will predict stronger sorting in a boom - is ultimately an empirical question and could go either way in the theoretical model. This is because drilling a well is more valuable during boom periods because the value of the oil and gas that is extracted from the well is higher. Therefore, in theory, agents might be *less* selective in booms if the value of the project increases more than the dynamic benefits of waiting for a better match.

Simple model setup Suppose that capital can take two types $y \in \{low, high\}$ and wells can take two types $x \in \{simple, complex\}$. Assume that capital takes on each type with probability 0.5, and similarly wells take on each type with probability 0.5. Assume that contract length is $\tau = 2$, the market is in a steady state (that is, if the market is in a boom then agents expect it to remain in a boom in the following period), and that agents meet under random search.

Denote the value of a match as $\bar{\text{m}}$ if it is between a simple well and a low-type capital, or between a complex well and high-type capital. Alternatively, denote the value of a match as $\underline{\text{m}} < \bar{\text{m}}$ if it is between a complex well and a low-type rig, or a simple well and a high-type rig. Therefore, given one rig of each type and one well of each type, without any search frictions it is optimal to allocate the complex well to the high-efficiency capital and the simple well to the low-efficiency capital. In addition to this match value agents also receive the value of hydrocarbons in the well which in this simple model I set to the gas price g (that is, $x_{quantity} = 1$).

Solution The aim is to determine under what primitives rejections of bad matches ($\underline{\text{m}}$), and hence stronger sorting patterns, will occur. I begin by assuming that bad matches are rejected and find conditions on the parameters that support this equilibrium. The rig's value of searching is:[72]

$$U = 0.5qV_{\bar{\text{m}}} + \Big(1 - 0.5q\Big)\beta U \tag{53}$$

In Equation 53, a rig is matched with probability q and if matched then encounters a match it will accept with probability 0.5. Since only matches with value $\bar{\text{m}}$ are accepted and prices are determined by Nash bargaining the payoff to matching is: $V_{\bar{\text{m}}} = \delta S_{\bar{\text{m}}} + \beta U$. Match surplus is $S_{\bar{\text{m}}} = (1+\beta)(\bar{\text{m}} + g) + \beta^2 U - \beta U$. Substituting these objects into Equation 53 and rearranging results in:

$$U = \left(\frac{1+\beta}{1-\beta}\right)\left(\frac{0.5q\delta}{1+0.5q\delta}\right)(\bar{m} + g) \tag{54}$$

It is optimal to reject bad matches if $S_{\underline{\text{m}}} < 0$. Substituting Equation 54 into this surplus condition and rearranging:

$$q > \frac{\underline{\text{m}} + g}{0.5\delta\beta(\bar{\text{m}} - \underline{\text{m}})} \tag{55}$$

[72] Note that I avoid notation about rig type.

Similarly, it can be show that if the inequality in Equation 55 does not hold then it is optimal to accept the match $\underline{\mathrm{m}}$.

<u>Comparative statics</u> Equation 55 allows me to perform some simple comparative statics that illustrate the main forces at work in the model. First, as the bargaining parameter δ increases or the difference between a good vs bad match $\bar{\mathrm{m}} - \underline{\mathrm{m}}$ increases, it becomes more beneficial to reject bad matches (for fixed values of q and g).

Second, to illustrate the sorting effect, consider how changing the probability of matching q and the gas price g affects the decision to reject bad matches in Equation 55. As the market moves from a bust to a boom both q and g increase. Since both sides of Equation 55 will increase in a boom it is an empirical question as to whether rejections of bad matches (and therefore the sorting effect) will be pro-cyclical or counter-cyclical.

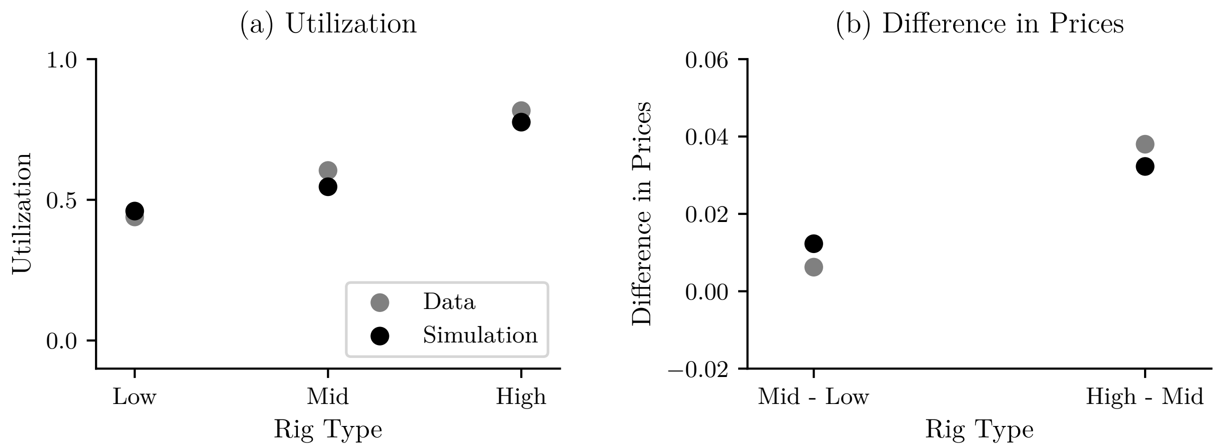
D.6 Out-of-sample fit

I evaluate out-of-sample fit in Figure 13. To do so I run the model for the period 2010-2013, removing the dates from April to December 2010 when the shallow-water market was either in an official moratorium after the 2010 Deepwater Horizon oil spill, or a 'defacto moratorium' where no new permits were awarded.[73] This period is out-of-sample in terms of the date range. Also, the average natural gas price in this period is low (\$3.65) when compared to the sample data used for estimation (where the average natural gas price is around \$6.66). Therefore, it presents an arguably onerous out-of-sample test.

In Figure 13 I compare two sets of simulated moments to their empirical counterparts. The first set of moments are related to mean rig utilization. These are most sensitive to parameters such as the potential project draws, and the meeting technology/targeting parameters. Even though the market is in a bust and the mean utilization is relatively low for all rig types, the model appears to fit rig utilization well in Panel (a). The second set of moments relate to the differences/orderings in prices between high and mid-specification rigs and mid and low-specification rigs. These moments are quite sensitive to the estimated match values, as well as the rig's value of searching (which enters as an outside option in the Nash bargaining equation). Again, the fit (documented in Panel (b)) appears to be quite good comparing the model to the data.

[73]For a longer discussion of the effects of the moratorium see Vreugdenhil (2021).

Figure 13: Out-of-sample fit



Note: The out-of-sample fit is performed for the period 2010-2013, removing the dates from April to December 2010 when the shallow-water market was either in an official moratorium after the 2010 Deepwater Horizon oil spill, or a 'defacto moratorium' where no new permits were awarded.

E Robustness Exercises

E.1 Robustness to instances of well deepening

I investigate how widespread instances of well deepening are, and also whether it is affecting the sorting patterns in Figure 4(a). Overall, I find that instances of well deepening are at most infrequent in the data and including them or removing them does not make a substantive difference to the results. I now explain in more detail about how I identify instances of well deepening, and illustrate the effects of deleting versus keeping these wells in the data.

In order to identify instances of well deepening I first investigate the 'proposal to drill' option in the 'Applications for Permit to Drill' database. Amongst the options here is to apply to the regulator to 'deepen' a well (as opposed to e.g. drill a 'new well'). In the data the 'proposal to drill' code is not recorded for every permit. However, for the permits in which I do observe this code I do not find any permits that select the 'deepen' option, which suggests that this is not a widespread behavior. Since deepening may be still be occurring for the subset of permits which do not directly contain this information, I alternatively try to conservatively identify any instances of deepening by looking at if any additional wellbores were drilled more than 365 days after the original hole.[74] Removing these wells from the dataset results in

[74]To implement this, I look at the 'wellbore code' which is contained in the last two digits of the well API number. A value of 00 for the wellbore code indicates the original wellbore. I look at if any additional wellbores

around 318 (12.0%) fewer contracts, but the sorting patterns in Figure 4(a) do not substantially change. Since this procedure is conservative and results in deleting many wells that are not being deepened, I retain these wells in the dataset.

E.2 Robustness to only using the natural gas price as a state variable

I now explore the implications of lumping both oil and gas together in Figure 14 and Table 13. I provide a more detailed discussion below, but overall I argue that because wells in the shallow water of the Gulf of Mexico are mainly producing natural gas it does not seem an unreasonable assumption to abstract away from differences in oil and gas prices in this context.[75]

In Figure 14(a) I plot the proportion of hydrocarbons produced by the wells drilled under each contract that are natural gas. To compute this measure I use the realized production of oil and natural gas from each well for 5 years after the date of first production (where these data are available). I then aggregate total production of oil and natural gas up to the contract level (since some contracts are for multiple wells). Finally, since oil and natural gas are measured in different units, I convert natural gas production into a 'barrel of oil equivalent' by dividing each cubic foot of natural gas production by 1/6000.[76]

Figure 14(a) documents that the wells in the shallow water of the US Gulf of Mexico are primarily producing natural gas. This is the motivation of simply using the natural gas price as the state variable rather than both oil and natural gas prices.

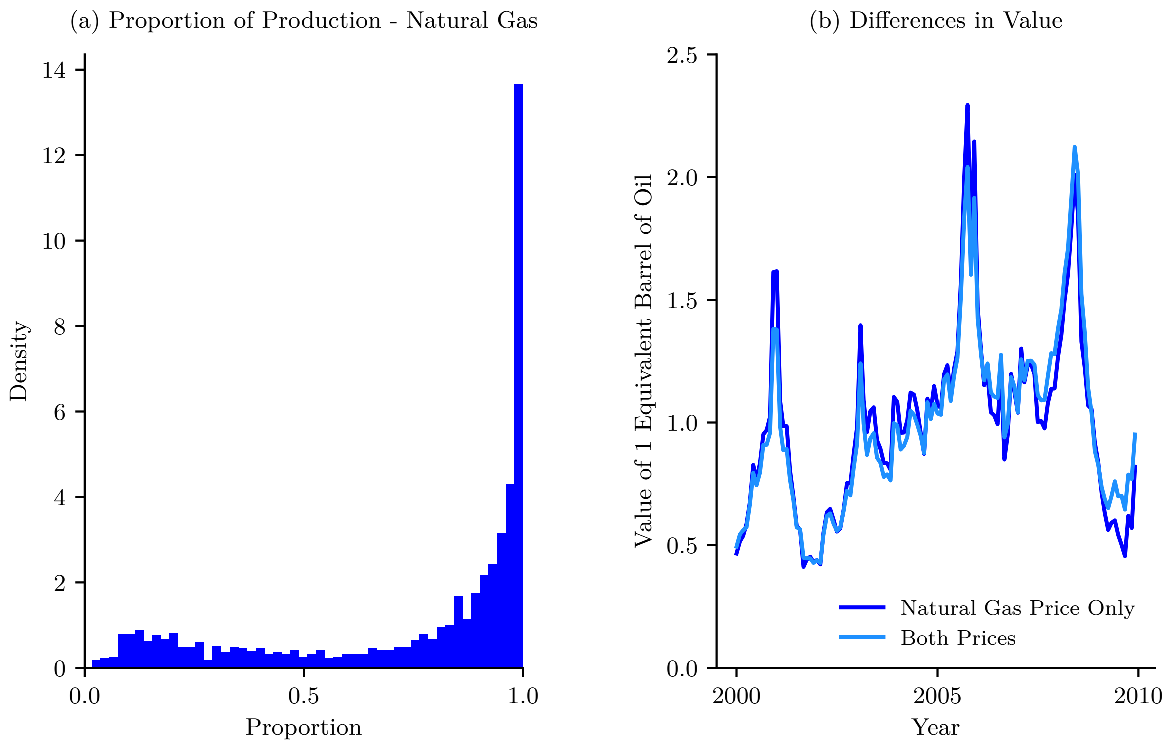
Figure 14(b) elaborates on the implications of Figure 14(a) through a thought experiment of how the value of the 'average well' (a well where natural gas is 75% of the total production of hydrocarbons) changes if I use only the natural gas price versus using a weighted average of both the oil and natural gas prices. Concretely, I consider a hypothetical well where the total value of hydrocarbons in the well = 1.0 at the average natural gas price. I then consider how the value of hydrocarbons changes across the cycle in 2000-2009 if I just use the natural gas price versus if I use a weighted average of both prices. The results Figure 14(b) show that the series are relatively similar, suggesting that using only the natural gas

were spudded more than 365 days after the initial wellbore depth date.

[75] Note that the way that oil and gas enter into the model is through lease bids which proxy for ex-ante expectations over the 'total quantity of hydrocarbons' in the well, rather than the amount of oil and gas produced by a well after it has been drilled. Therefore the primary way that not accounting for differences in oil and gas prices would affect the results is through how this value of hydrocarbons (computed as the quantity of hydrocarbons multiplied by the current natural gas price) differs over the cycle.

[76] This is the value used by the US Geological Survey to compute total hydrocarbon reserves (https://certmapper.cr.usgs.gov/data/PubArchives/WEcont/world/woutsum.pdf). Oil and gas companies also use similar 'barrel of oil equivalent' measures in financial statements when disclosing their reserves to investors.

Figure 14: Heterogeneity of oil and gas output



Note: This figure documents the degree of heterogeneity in oil/gas output. Panel (a) shows the distribution of the proportion of total production that is natural gas for each contract. Panel (b) show how the value of a hypothetical 'average well' with 75% natural gas would change across the sample period depending on whether the natural gas price is used versus both the oil price and the natural gas price. The hypothetical 'average well' has a value that is normalized to = 1.0 at the average gas price.

price is not too strong an assumption.

Finally, in Table 13 I present two regressions that test whether oil and gas heterogeneity makes a difference for the way wells are drilled. These regressions test whether the complexity of the design of the well is a function of whether the hydrocarbons that the well produces are predominately oil or natural gas. Overall, I find that the relationship between the proportion of production that is natural gas and the complexity of the well is not statistically significant. Therefore, I do not find any evidence that heterogeneity of oil/gas output affects the complexity of a well.

Table 13: Regressions of well complexity on the type of hydrocarbons the well produces

	(1) Complexity	(2) Complexity
Proportion of production - natural gas	0.028	0.025
	(0.031)	(0.031)
Time FEs	No	Yes
N	1796	1796

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors are in brackets. Time FEs incorporates fixed effects for the year and the month that the contract is signed. These regressions are performed at the contract level for contracts where production data are available.

E.3 Robustness to non-myopic wells

Robustness exercise: overview Overall, the main finding is that allowing projects to have an outside option of negotiation failure - together with re-estimating the parameters of the model to ensure that the model makes sensible in-sample predictions - does not substantially change the results or exacerbate the implied welfare effects from entry. I now provide more details about these robustness exercises, discuss the benefits and limitations, and the results.

Implementation I include an outside option for projects in the following form: $\beta(1-\mathbb{P}_{exit})q_y^{project}(\theta_{yt})W_t(x,y)$. Here, if there is a negotiation failure, projects wait for one period. They are then hit by an exogenous exit shock (so that $(1-\mathbb{P}_{exit})$ is the probability that the project survives and is still available to match). Finally, if the project survives, it can rematch with probability $q_y^{project}(\theta_{yt})$ resulting in the value $W_t(x,y)$. For simplicity, I assume that if the well rematches then it will match with the same capital type as the original match.[77]

The above specification allows for the project's outside option to be dependent on the state of the market at time t through the probability of (re-)matching $q_y^{project}(\theta_{yt})$ as well as the value $W_t(x,y)$. Concretely, in a boom the project's probability of re-matching $q_y^{project}(\theta_{yt})$ will decrease but the value of the project $W_t(x,y)$ will increase, which lead to overall theoretically ambiguous effects on the well's outside option.

There are at least two implicit assumptions in the above specification of the project's outside option,

[77] A related setup is used for exporters in Brancaccio et al. (2020).

Table 14: Robustness exercises with non-myopic projects

$\mathbb{P}_{exit}$	**1.0** *(Baseline)*	**0.95**	**0.75**
Value of sorting effect	9.9%	18.0%	13.9%

Note: This table provides an overview of how the results change when projects have an outside option of $\beta(1-\mathbb{P}_{exit})q_y^{project}(\theta_{yt})W_t(x,y)$, for different assumptions about $\mathbb{P}_{exit}$. Note that I re-estimate all the parameters in each of the robustness exercise scenarios.

which I now detail. First, I assume that the well rematches with the same capital type; I make this assumption to keep the robustness check as parsimonious as possible. Second, since I do not observe searching projects (recall I only observe matches and searching capital), I need to make an assumption on the probability of project exit $\mathbb{P}_{exit}$ and the probability of project re-matching $q_y^{project}(\theta_{yt})$. I perform the robustness exercise for several different values of $\mathbb{P}_{exit}$. For $q_y^{project}(\theta_{yt})$, I treat the equilibrium values from the benchmark model as 'data' and estimate a logit model of the probability of matching each capital type conditional on a second-order polynomial of the state. I then use this model to compute $q_y^{project}(\theta_{yt})$ in the robustness exercise, for different states of the market.

Finally, in all of the robustness exercises I also re-estimate the parameters. This is important to allow the data to discipline the model and ensure that - for example - the model in the robustness exercises continues to replicate the empirical sorting patterns.

<u>Results from the robustness exercises</u> I detail the results of the robustness exercises in Table 14. Overall, allowing for non-myopic projects - as well as re-estimating the model - appears to increase the value of the sorting effect. The exact change does not appear to be monotone in $\mathbb{P}_{exit}$. The ultimate conclusion that the sorting effect is economically significant still remains. If the parameters are not re-estimated then allowing for non-myopic projects increases the probability of rejecting a match in both booms and busts. This is because the match surplus now becomes:

$$S_t(x,y) = W_t(x,y) - \beta(1-\mathbb{P}_{exit})q_y^{project}(\theta_{yt})W_t(x,y) + V_t(x,y) - \beta\mathbb{E}_tU_{t+1}(y)$$

and the project's outside option $\beta(1-\mathbb{P}_{exit})q_y^{project}(\theta_{yt})W_t(x,y) \geq 0$. However, after re-estimating the parameters, it moderates this effect.[78]

[78] There are many components to the model that are re-estimated, and so it is not clear before performing the exercise if the net effect of allowing for non-myopic projects will increase or decrease the value of the sorting effect, or if the effects will be monotone in the value of $\mathbb{P}_{exit}$.

E.4 Robustness to the assumption that projects target capital

Implication of the assumption that project owners direct their search The implications of this assumption enter into the model through the targeting weights expression (Equation 4). In Appendix B.1 I show how Equation 4 can be micro-founded from the individual search decisions of well owners. I now derive the equivalent expression for the targeting weights when rig owners make the decision about whom to match with.

To begin this alternative derivation, denote each unit of available capital by j and the corresponding type as y_j. Similarly, denote each searching project by i and its corresponding type by x_i. Using this notation, for a rig of type y_j, the (expected) value of targeting well i is $\pi_{x_i t}(y_j) = q^{capital}_{y_j}(\theta_{y_j t})(\delta S_t(x_i, y_j) + \beta U_{t+1}(y_j))$. Here I am using the Nash bargaining result that the surplus of a match will be split so that the rig receives $\delta S_t(x_i, y_j) + \beta U_{t+1}(y_j)$ and the well receives $(1-\delta)S_t(x_i, y_j)$.

Next, as in the micro-foundation of the targeting weights in Appendix B.1, I setup the rig's problem as choosing which well to target based on a *perceived value* $\max_i \hat{\pi}_{x_i t}(y_j)$. This value is given by:

$$\hat{\pi}_{x_i t}(y_j) = \pi_{x_i t}(y_j) - \gamma_1 1\left[S_t(x_i, y_j) < 0\right] + \epsilon^{target}_{ijt} \tag{56}$$

I assume that ϵ^{target}_{ijt} are drawn from an i.i.d. type-1 extreme value distribution with scale parameter $1/\gamma_0$. The conditional choice probability that a rig j targets well i at time t is then:

$$P^{capital}_{ijt} = \frac{\exp\left(\gamma_0 \left[\delta q^{capital}_{y_j}(\theta_{y_j t}) S_t(x_i, y_j) - \gamma_1 1\left[S_t(x_i, y_j) < 0\right]\right]\right)}{\sum\limits_k \exp\left(\gamma_0 \left[\delta q^{capital}_{y_k}(\theta_{y_k t}) S_t(x_i, y_k) - \gamma_1 1\left[S_t(x_i, y_k) < 0\right]\right]\right)} \tag{57}$$

Here the $q^{capital}_{y_j}(\theta_{y_j t})\beta U_{t+1}(y_j)$ terms in $\pi_{x_i t}(y_j)$ enter additively and are the same for all alternatives and so they cancel. Finally, aggregating $P^{capital}_{ijt}$ over all j, the targeting weight is:

$$\omega'_{yt}(x) = \frac{n_{yt} \exp\left(\gamma_0 \left[\delta q^{capital}_{y}(\theta_{yt}) S_t(x, y) - \gamma_1 1\left[S_t(x, y) < 0\right]\right]\right)}{\sum\limits_{k \in Y} n_{kt} \exp\left(\gamma_0 \left[\delta q^{capital}_{k}(\theta_{kt}) S_t(x, k) - \gamma_1 1\left[S_t(x, k) < 0\right]\right]\right)} \tag{58}$$

This can be compared to the targeting weight in the model which - writing in terms of the total surplus of a match - is:

$$\omega_{yt}(x) = \frac{n_{yt} \exp\left(\gamma_0 \left[(1-\delta) q^{project}_{y}(\theta_{yt}) S_t(x, y) - \gamma_1 1\left[S_t(x, y) < 0\right]\right]\right)}{\sum\limits_{k \in Y} n_{kt} \exp\left(\gamma_0 \left[(1-\delta) q^{project}_{k}(\theta_{kt}) S_t(x, k) - \gamma_1 1\left[S_t(x, k) < 0\right]\right]\right)} \tag{59}$$

I now discuss under what conditions the targeting weights are the same, regardless of which side of the market is assumed to choose who to match with. That is, I discuss under what conditions $\omega'_{yt}(x) = \omega_{yt}(x)$. If search is random ($\gamma_0 = 0$) then the targeting weights are the same. Similarly, if both sides of the market have approximately equal bargaining power ($\delta = 1/2$) and the probability of matching is the same for

Table 15: Robustness test of rigs vs project owners (wells) targeting their search.

	Wells target rigs	Rigs target wells
Sorting Effect	9.9%	9.0%

capital and projects ($q_y^{project}(\theta_{yt}) = q_y^{capital}(\theta_{yt})$), then the targeting weights are then same. Therefore, the extent to which the assumption matters hinges on whether these conditions hold in the estimated model.

Robustness exercise Finally, I re-run the model using the targeting weights derived in Equation 58. (That is, I re-run the model assuming the other extreme that it is rigs - rather than project owners - who target their search). I report the results for the sorting effect counterfactual in Table 15. Overall, Table 15 shows that the results do not hinge on which side of the market is assumed to direct search, with qualitatively similar results under either assumption.

E.5 Robustness to changes in the period length

Table 16: Results from the period length robustness check

	One-month period	Two-week period
Welfare (billions USD)	\$4.91	\$5.39

Note: This table reports robustness of the model to different period lengths. Specifically, I compare the results for the sorting effect with a one-month period length (which is assumed in the model) versus a fortnightly period length. In order to compute the fortnightly period length model I take the following steps. First, I recompute Step 1 of the estimation with a two-week period length; this includes re-estimating the empirical objects that underpin the value functions, and re-simulating the value functions. Next, I re-compute the model at the baseline parameters (i.e. the parameters corresponding to a one-month period length) except for the potential well draw parameters d_0, d_1, which I halve. Therefore, each two-week period has approximately half the number of potential wells as the corresponding one-month period. I also halve the match value terms, $m_{0,y}, m_{1,y}$ for $y \in \{low, mid, high\}$ and m_2, since they now correspond to the match value for half a month.