NBER WORKING PAPER SERIES

INFORMATION EXTERNALITIES, FREE RIDING, AND OPTIMAL EXPLORATION IN THE UK OIL INDUSTRY

Charles Hodgson

Working Paper 33067
http://www.nber.org/papers/w33067

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2024

This research was supported by the Israel Dissertation Fellowship through a grant to the Stanford Institute for Economic Policy Research. I am grateful to Liran Einav, Matthew Gentzkow, Timothy Bresnahan, Paulo Somaini, Lanier Benkard, Brad Larsen, Nikhil Agarwal, Phil Haile, Steve Berry, Ariel Pakes, Ken Hendricks, Jackson Dorsey, Linda Welling, and participants at various seminars and conferences for helpful comments on this project. Thanks also to Oonagh Werngren and Jen Brzozowska for help with sourcing the data. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2024 by Charles Hodgson. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Information Externalities, Free Riding, and Optimal Exploration in the UK Oil Industry
Charles Hodgson
NBER Working Paper No. 33067
October 2024
JEL No. L0, L72, O30, O32

ABSTRACT

Information spillovers between firms can reduce R&D incentives if competitors can free ride on innovations. However, strong property rights may impede cumulative research and lead to inefficient duplication. These effects are particularly relevant in natural resource exploration, where discoveries are spatially correlated. Using UK offshore oil exploration data, I estimate a dynamic model that captures the trade-off between drilling now and waiting to learn from competitors. Removing free-riding incentives increases industry surplus by 52%, while perfect information flow raises it by 24%. Counterfactual policy simulations highlight a trade-off in property rights design: stronger property rights over exploration well data increase the rate of exploration, while weaker property rights increase the efficiency and speed of learning but reduce the rate of exploration. Spatial clustering of each firm's drilling licenses both reduces the incentive to free ride and increases the speed of learning.

Charles Hodgson
Department of Economics
Yale University
87 Trumbull Street
New Haven, CT 06511
and NBER
cbchodgson@gmail.com

1 Introduction

A firm's incentive to invest in R&D depends on how much it can benefit from competitors' investments. If R&D outcomes, such as new technologies or mineral discoveries, are public, firms may *free ride* by imitating products or exploring near rivals' discoveries.When each firm would rather wait to observe the results of other firms' research than invest in R&D itself, the equilibrium rate of innovation can fall below the socially optimal level (Bolton and Harris, 1999). On the other hand limiting information flow between firms, for example by property rights on existing innovations, can slow research by causing inefficient duplication and preventing cumulative innovation (Williams, 2013). [1]

In this paper, I quantify the effects of information externalities on R&D in the context of oil exploration. Several features of this industry make it an ideal setting for studying the general problem of information spillovers and the design of optimal property rights regulation. When an oil firm drills an *exploration well* it generates knowledge about the presence or absence of resources in a particular location. Exploration wells can therefore be thought of as experiments with observable outcomes located at points in geographic space. Since oil deposits are spatially correlated, the result of exploration in one location generates information about the likelihood of finding oil in nearby, unexplored locations. The spatial nature of research in this industry means that the extent to which different experiments are more or less closely related is well defined. Research is cumulative in the sense that the findings from exploration wells direct the location of future wells and the decision to *develop* fields and extract oil.

Since multiple firms operate in the same region, the results of rival firms' wells provide information that can determine the path of a firm's exploration. If firms can see the results of each other's exploration activity, then there is an incentive to free ride and delay investment in exploration until another firm has made discoveries that can direct subsequent drilling. However, if the results of exploration are confidential then firms are likely to engage in *wasteful exploration* of regions that are known by other firms to be unproductive.[2] Using

[1]Policy that defines property rights over innovations plays an important role in controlling the effects of information externalities and balancing the trade-off between discouraging free riding and encouraging cumulative research. For example, broader patents minimize the potential for free riding but increase the cost of research that builds on existing patents, and may therefore direct research investment away from socially efficient projects (Scotchmer, 1991).

[2]This trade-off between free riding and inefficient exploration has been identified as important for policy making in the industry literature. For example, in their survey of UK oil and gas regulation, Rowland and Hann (1987, p. 13) note that "if it is not possible to exclude other companies from the results of an exploration well... companies will wait for other companies' drilling results and exploration will be deferred," but if "information is treated highly confidentially... an unregulated market would be likely to generate repetitious exploration activity."

UK offshore drilling data from 1964-1990, I quantify these inefficiencies and the extent to which they can be mitigated by counterfactual property rights policies. The magnitude of these effects depends on the spatial correlation of well outcomes, the extent to which firms can observe the results of each other's wells, and the spatial arrangement of drilling licenses assigned to different firms.

I measure the effect of information externalities on equilibrium exploration rates and industry surplus by estimating a structural model of the firm's exploration problem. Firms face a dynamic discrete choice problem in which, each period, they can choose to drill exploration wells on the set of blocks over which they have property rights. At the end of each period firms observe the results of their exploration wells, observe the results of other firms' wells with some probability, $\alpha \in [0, 1]$, and update their beliefs about the spatial distribution of oil. Firms face a trade-off between drilling now and delaying exploration to learn from the results of other firms' wells that depends on the spatial arrangement of drilling licenses and the probability of observing the results of other firms' wells.

The model's asymmetric information structure complicates the firm's problem. Firms observe different sets of well outcomes, and in order to forecast other firms' drilling behavior each firm must form beliefs about the outcomes of unobserved wells and about other firms' beliefs. To make estimation of the model and computation of equilibrium tractable, I propose an equilibrium concept in which each firm has beliefs about the rate of exploration of blocks held by other firms that are correct in expectation, conditional on the firm's current state. This idea is similar to the experience based equilibrium approach of Fershtman and Pakes (2012). A further complication to estimating the model is that firms' state variables are not fully observable. In particular, the econometrician does not know which rival well outcomes a firm has observed at each point in time. I deal with this problem by presenting a novel identification argument leveraging the fact that choice probabilities are a mixture over potential information states.

The estimated value of the spillover parameter, α, indicates that firms observe the results of other firms' wells with 64% probability. The presence of substantial but imperfect information spillovers means that equilibrium exploration behavior could be affected by both free riding and inefficient exploration. I perform counterfactual simulations to quantify these two effects. I find that removing the free riding incentive brings exploration forward in time by about 1.5 years, substantially increasing total exploration and development and raising industry surplus by 52%. Next, I allow for perfect information sharing between firms, holding firms' incentive to free ride fixed at the baseline level. Industry surplus is increased by 24% due to an improvement in the efficiency of exploration - since firms can perfectly observe each other's well results, cumulative learning is faster and exploration wells are more concentrated on productive blocks.

I use the estimated model to evaluate counterfactual information policy. Under UK regulations, data from exploration wells is property of the firm for around five years before being made public. Weakening property rights by shortening the confidentiality window will increase the flow of information between firms, and is likely to increase the efficiency of exploration but may also increase the incentive to free ride. On the other hand, strengthening property rights by extending the confidentiality window will decrease the incentive to free ride but slow cumulative learning and reduce the efficiency of exploration. I simulate equilibrium behavior under different confidentiality window lengths and find that industry surplus is increased under both longer and shorter confidentiality windows. Among simulated policies, a 0-length confidentiality period is optimal, increasing surplus by 28% over the baseline, though a longer window of 7.5 years is also locally optimal.

Finally, I show how the spatial distribution of property rights affects exploration incentives. I construct a counterfactual spatial assignment of property rights that clusters each firm's licenses together, holding the total number of blocks assigned to each firm fixed. Under the clustered assignment industry surplus increases by 4%.

The results highlight the tension between discouraging free riding and encouraging efficient cumulative research in the design of property rights over innovations. Optimal policy depends sensitively on the interplay of these two effects, and the design of information policy can have large effects on industry surplus. In this setting, there are ranges of the policy space in which reducing information sharing leads to a marginal improvement in surplus and ranges where reducing information sharing is optimal. This trade-off applies in other settings, for example in defining the breadth of patents, regulations about the release of data from clinical trials, and the property rights conditions attached to public funding of research. In these settings, the role of information externalities in policy design may interact with the more commonly studied issue of property rights giving firms monopoly power over innovations.

This work contributes to the literature on firms' R&D incentives (Arrow, 1971; Dasgupta & Stiglitz, 1980; Spence, 1984), particularly building on studies of how intellectual property rights impact innovation (Murray & Stern, 2007; Williams, 2013). I contribute to this literature by quantifying the trade-off between this effect on cumulative research and the free riding incentive that has been discussed in the theory literature (Hendricks and Kovenock, 1989; Bolton and Farrell, 1990; Bolton and Harris, 1999). This paper differs from much of the innovation literature by using a structural model of the firm's sequential research (here, exploration) problem to quantify the effects of information externalities and alternative property rights policies.

This paper also contributes to the literature on the effect of information externalities in oil exploration, mostly focused on bidding in license auctions (Porter, 1995; Haile, Hendricks, and Porter 2010; Nguyen, 2021). Less attention has been given to the post-licensing exploration incentives induced by different property rights policies. Notable exceptions include Hendricks and Porter (1996), who show that the probability of exploration on tracts in the Gulf of Mexico increases sharply when firms' drilling licenses are close to expiry, and Lin (2009) and Levitt (2016), who document firms' drilling response to exploration on nearby tracts.

Existing papers on oil and gas exploration that estimate structural models of the firm's exploration problem include Levitt (2009), Lin (2013), Agerton (2018), and Steck (2018). The model I estimate in this paper differs from existing work by incorporating both Bayesian learning with spatially correlated beliefs and information leakage across firms. Steck (2018) uses a closely related dynamic model of the firm's decision of when to drill in the presence of social learning about the optimal inputs to hydraulic fracturing. Steck's finding of a significant free riding effect when there is uncertainty about the optimal *technology* is complementary to the findings of this paper, which measures the free riding effect in the presence of uncertainty about the *location* of oil deposits.

Other related papers in the economics of oil and gas exploration include Kellogg (2011) and Covert (2015). Covert's methodology is particularly close to mine, as he also uses a Gaussian process to model firms' beliefs about the effectiveness of different drilling technologies in different locations. The results I present in Section 4, which show that firms are more likely to drill exploration wells in locations where the outcome is more uncertain, contrast with the findings of Covert (2015), who shows that oil firms do not actively experiment with fracking technology when the optimal choice of inputs is uncertain.

Finally, the procedure used to estimate the structural model of the firm's exploration problem builds on the literature on estimation of dynamic games using conditional choice probability methods, following Hotz and Miller (1993), Hotz, Miller, Sanders, and Smith (1994), Bajari, Benkard, and Levin (2007), and Fershtman and Pakes (2012). In particular, I extend these methods to a setting with asymmetric information in which the econometrician is uninformed about each agent's information set. I propose a novel source of identification of conditional choice probabilities in the presence of this latent state variable that is different from the panel variation used by Kasahara and Shimotsu (2008).

The remainder of this paper proceeds as follows. Section 2 provides an overview of the setting and a summary of the data. Section 3 presents a model of spatial beliefs about the location of oil deposits. Section 4 presents reduced form results that provide evidence of spatial learning, information spillovers, and free riding. In Section 5 I develop a dynamic structural model of optimal exploration with information spillovers, and in Section 6 I discuss estimation of the model. Results and policy counterfactuals are presented in Sections 7 and 8. Section 9 concludes.

2 UK Oil Exploration: Setting and Data

I use data covering the history of oil drilling in the UK Continental Shelf (UKCS) from 1964 to 1990. Oil exploration and production on the UKCS is carried out by private companies who hold drilling licenses issued by the government. The first such licenses were issued in 1964, and the first successful (oil yielding) well was drilled in 1969. Discoveries of the large Forties and Brent oil fields followed in 1970 and 1971. Drilling activity took off after the oil price shock of 1973, and by the 1980s the North Sea was an important producer of oil and gas. I focus on the region of the UKCS north of $55°N$ and east of $2°W$, mapped in Figure 1, which is bordered on the north and east by the Norwegian and Faroese economic zones. This region contains the main oil producing areas of the North Sea and has few natural gas fields, which are mostly south of $55°N$.[3]

2.1 Technology

Offshore oil production involves two investment phases. First, *exploration wells* are drilled from mobile rigs or ships to locate oil reservoirs. It is important to note that the results of a single exploration well provide limited information about the size of an oil deposit, and many exploration wells must be drilled to estimate the volume of a reservoir. When a sufficiently large oil field has been located, the field is *developed*. This second phase of investment involves the construction of a production platform - a large static facility typically anchored to the seabed.

I observe the coordinates and operating firm of every exploration well drilled and development platform constructed from 1964 to 1990. The left panel of Figure 1 maps exploration wells in the relevant region. For each exploration well, I observe a binary outcome - whether or not it was successful. In industry terms, a successful exploration well is one that encounters an "oil column", and an unsuccessful well is a "dry hole".[4] In reality, although exploration wells yield more complex geological data, the success rate of wells based on a binary wet/dry classification is an important statistic in determining whether to develop, continue explor-

[3]In the estimation of the structural model, I also use equivalent data from the Norwegian sector.

[4]Well classification is based on well-level data that describes whether each well located oil. I classify as successful all wells coded as "oil well" or "oil shows". An underlying assumption is that the industry classification of wells did not change over time, for example with changing technology or prices.

ing, or abandon a region. See for example Lerche and MacKay (1995) and Bickel and Smith (2006) who present models of optimal sequential exploration decisions based on binary signals. I observe each development platform's monthly oil and gas production in m^3 up to the year 2000.

2.2 Regulation

The UKCS is divided into *blocks* measuring 12x10 nautical miles (approx. 22x18 km). These blocks are indicated by the grid squares on the maps in Figure 1. The UK government holds licensing *rounds* at irregular intervals (once every 1 to 2 years), during which licenses that grant drilling rights over blocks are issued to oil and gas companies. Unlike in many countries, drilling rights are not allocated by auctions. Instead, the government announces a set of blocks that are available, and firms submit applications which consist of a list of blocks, a portfolio of research on the geology and potential productivity of the areas requested, a proposed drilling program, and evidence of technical and financial capacity. Applications for each block are evaluated by government geoscientists. Although a formal scoring rubric allocates points for a large number of assessment criteria including financial competency, track record, use of new technology, and the extent and feasibility of the proposed drilling program, the assessment process allows government scientists and evaluators to exercise discretion in determining the allocation of blocks to firms. Although the evaluation criteria have changed over time, the discretionary system itself has remained relatively unchanged since 1964.[5]

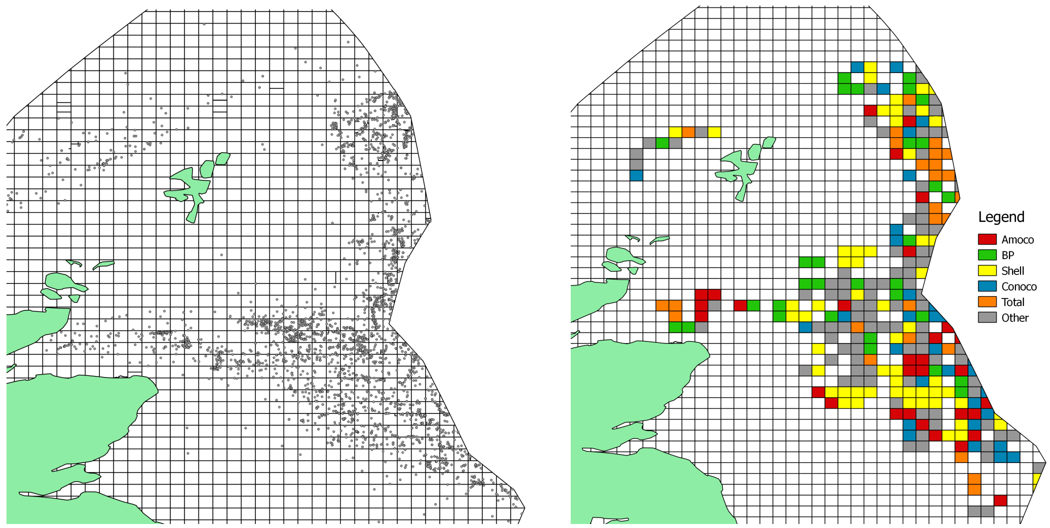
License holders pay an annual per-block fee, and are subject to 12.5% royalty payments on the gross value of all oil extracted. Licenses have an initial period of 4 or 6 years during which firms are required to carry out a minimum work requirement. I refer to the end of this period as the license's *work date*. Minimum work requirements are typically light, even in highly active areas. During the 1970s "3 exploration wells per... 7 blocks became the norm" in the main "contested" areas (Kemp, 2012a p. 58). Licenses in less contested "frontier" areas often did not require any drilling, only seismic analysis.

I observe the history of license allocations for all blocks. I perform all analysis on a region corresponding to the northern North Sea basin which contains almost all of the large oil deposits discovered on the UK continental shelf.[6] In assigning blocks to firms I make two

[5] A few blocks were offered at auction in the early 1970s, but this experiment was determined to be unsuccessful. According to a regulatory manager at the Oil and Gas Authority (OGA), the result of the auctions was that "the Treasury raised a lot of money but nobody drilled any wells." By contrast, the discretionary system has "stood the test of time". The belief among UK regulators is that auctions divert money away from firms' drilling budgets.

[6] Specifically, this region corresponds to the area north of $55N$, east of $0°W$, and south-west of the UK

Figure 1: Wells and License Blocks



Notes: Grid squares are license blocks. The left panel plots the location of all exploration wells drilled from 1964 to 1990. The right panel records license holders for each block in January 1975. Note that if multiple firms hold licenses on separate sections of a block, only one of those firms (chosen at random) is represented on this map.

important simplifying assumptions. First, I focus only on the "operator" firm for each block. Licenses are often issued to consortia of firms, each of which hold some share of equity on the block. The operator, typically the largest equity holder, is given responsibility for day to day operations and decision making. Non-operator equity holders are typically smaller oil companies that do not operate any blocks themselves, and are often banks or other financial institutions. Major oil companies do enter joint ventures, with one of the companies acting as operator, but these are typically long lasting alliances rather than block by block decisions.[7] In the main analysis below, I will be ignoring secondary equity holders and treating the operating firm as the sole decision maker, with all secondary equity holders being passive investors.[8] Second, licenses are sometimes issued over *parts* of blocks, splitting the original

economic zone's border. For some of the descriptive exercises I restrict the data to 1970-1980, since the location of early exploration before any oil discoveries is hard to rationalize.

[7]For example, 97% of blocks operated by Shell between 1964 and 1990 were actually licensed to Shell and Esso in a 50-50 split. Esso was at some point the operator of 16 unique blocks, compared to more than 740 blocks that were joint ventures with Shell. Only 8.6% of block-months operated by one of the top 5 firms (who together operate more than 50% of all block-months) have another top 5 firms as a secondary equity holder. This falls to 2.8% among the top 4 firms.

[8]Appendix Table A4 presents regressions of drilling probability on the distribution of surrounding licenses that suggest this is a reasonable assumption. The number of nearby licenses operated by the same firm as block j has a consistent, statistically significant positive effect on the probability of exploration on block j. The number of nearby licenses with the same secondary equity holders as block j, on which the operator of block j is a secondary equity holder, and on which one of the secondary equity holders on block j is the operator, all have no statistically significant effect on drilling probability.

blocks into smaller areas that can be held by different firms. All of the analysis below will take place at the block level. Therefore, if two firms have drilling rights on the two halves of block j, I will record them both as having independent drilling rights on block j. In practice, 88.2% of licensed block-months have only one license holder. 11.5% of block-months have two license holders and a negligible fraction have more than two. Subject to these simplifications, the right panel of Figure 1 maps the locations of licensed blocks operated by the 5 largest firms in January 1975. There are 73 unique operators between 1964 and 1990, but 90% of block-months are operated by one of the top 25 firms, and over 50% are operated by one of the top 5. Appendix Figure A2 illustrates the distribution of licenses at the block-month level across firms.[9]

A final set of regulations defines property rights over the information generated by wells. The production of development platforms is reported to the government and published on a monthly basis. Data from exploration wells, including whether or not the well was successful, is property of the firm for several years after a well is drilled. After this confidentiality period, well data is reported to the government and made publicly available. For most wells in the sample, the confidentiality period was five years. However, well data is typically released by the government in batches and is often delayed, so the average time from well completion to data release is closer to 6 years. Appendix Figure A4 records the distribution of realized confidentiality periods. In reality there is likely to be information flow between firms during this confidentiality period for a number of reasons: firms can exchange or sell well data, information can leak through shared employees, contractors, or common equity holders across blocks operated by different firms, and the activities associated with a successful exploration well might be visibly different than the activities associated with an unsuccessful exploration well. The extent to which information flows between firms during this confidentiality period is an object of interest in the empirical analysis that follows.

[9] One additional complication is the case of an oil reservoir which crosses multiple blocks operated by different firms. In these cases the oil reservoir is "unitized" by regulation, and profit is split proportionally between operators of the blocks (Weaver and Asmus, 2006). This provision removes the "common pool" incentive discussed by Lin (2013) and the incentive to develop an overlapping reservoir before a neighboring rival. Consistent with this, the model presented below abstracts from common pools. On the other hand, unitization means that a firm can obtain revenue from a block without being observed to develop it in the data, if a neighboring block on a common pool is developed. This channel is not present in the model. In reality, blocks are large relative to most fields and unitized shareholders are often small relative to the operating firm. In data on firm shares of fields, the median largest shareholder of a field has an 87% interest. This data includes interests held by passive investors in addition to unitized operators

2.3 Data

Table 1 contains summary statistics describing the data. Observations are at the firm-block level. That is, if a particular block is licensed multiple times to different firms, it appears in Table 1 as many times as it is licensed. There are a total of 628 blocks ever licensed and 1470 firm-block pairs between 1964 and 1990. I focus on two actions - the drilling of exploration wells and the development of blocks. I consider the development of a block as a one off decision to invest in a development platform. I record a block as being *developed* on the drill date of the first development well. In reality, this would come several months after construction of the development platform begins. I consider development to be a terminal action, so any additional exploration activity on firm-blocks that have been developed is dropped from the data.[10]

Table 1: Summary Statistics: Blocks & Wells

Firm-Blocks	All	Explored	Exp. & Devel-oped	Exp. & Not Dev.	Not Exp.
N	1470	721	160	561	749
Share Explored	.490	1.000	1.000	1.000	0.000
Share Developed	.120	.222	1.000	0.000	.021
First Exp. After Work Date	.	.227	.280	.215	.
Own Share of Nearby Blocks:					
Mean	.199	.178	.181	.177	.219
SD	.217	.199	.206	.197	.231
Exploration Wells per Block	2.002	4.082	10.138	2.355	0.000
Share Successful	.199	.199	.444	.129	.

Notes: Table records statistics on all license-block pairs active between 1964 and 1990. In particular, if a block is licensed to multiple firms it appears multiple times in this Table. Each column records statistics on subsets of license-blocks defined according to whether they are ever explored or developed. Own share of nearby blocks is defined as the share of license-blocks that are at most third degree neighbors that are licensed to the same firm.

The second column of Table 1 records statistics on the set of firm-blocks that are ever explored - that is, those firm-blocks where at least one exploration well was drilled - and the

[10] I associate output from a development platform with the block it is located on. In reality, fields may overlap multiple blocks although blocks are large relative to most fields in this setting. When development occurs on a field which overlaps multiple blocks, the overlapping sections of the blocks where the platform is not located are typically split off from their parent block. I treat the remainder of the parent block as a continuation of the original undeveloped block. Exploration on already developed firm-blocks accounts for only 5% of exploration wells in the raw data, and likely represents activity related to the operation of the existing development platform.

third column records statistics for those firm-blocks that are ever developed. 49% of firm-blocks are ever explored, and among these, 22% are developed. Note that the information generated by a single well is insufficient to establish the size of an oil reservoir, and firms must drill many exploration wells on a block before making the decision to develop. On average, 10 exploration wells are drilled before a block is developed, while 2.3 wells are drilled on blocks explored but not developed. Table 1 shows a 44% well success rate on developed blocks, and 13% on non-developed blocks. The success rate of exploration wells on a block is correlated with the size of any underlying oil reservoir. Thus, if an initial exploration well yields oil, but subsequent wells do not, the block is likely to only hold small oil deposits and is unlikely to be developed. Within blocks that are developed, there is a positive correlation between exploration well success and estimated reserves, illustrated in Appendix Figure A3.[11]

Note that the work requirement policy leaves significant scope for firms to delay exploration. The work requirement typically demands at most one exploration well be drilled per block, but it is clear that many more than one exploration well must be drilled before a block is developed. While the work requirement policy is therefore likely to hasten the drilling of the first exploration well on a block, there are no requirements on the speed with which the subsequent program of exploration must take place. The fourth row of Table 1 indicates that almost a quarter of blocks that are ever explored are first explored *after* the work requirement date. These findings corroborate claims from industry literature that indicate the terms of drilling licenses issued in the UK are considerably more generous than those issued, for example, in the Gulf of Mexico, and provide considerable room for firms to "stockpile" unexplored and undeveloped acreage for many years (Gordon, 2015).

3 A Model of Spatially Correlated Beliefs

The effect of information externalities on firms' exploration decisions depends on the spatial arrangement of licenses, the extent to which firms can observe the results of each other's wells, and on the correlation of exploration results at different locations. In Hodgson (2024) I show that in a simple two-firm, two-block model, spatial correlation in well outcomes reduces the equilibrium rate of exploration below the social optimum. The magnitude of this free riding effect is determined by the extent to which well results are correlated over space. In particular, the more correlated are outcomes on neighboring blocks, the lower the equilibrium rate of exploration.

In this section, I measure this spatial correlation by estimating a statistical model of the distribution of oil that allows the results of exploration wells at different locations to

[11] The methodology used to estimate reserves is outlined in Appendix C.5.

be correlated. By fitting the model to data on the outcomes of all exploration wells drilled between 1964 and 1990, I obtain an estimate of the extent to which this covariance of well outcomes declines with distance. I interpret the estimated model as a Bayesian prior about the distribution of oil

3.1 Statistical Model of the Distribution of Oil

I start by describing a statistical model of the distribution of oil over space. I model the probability that an exploration well at a particular location is successful as a continuous function over space drawn from a Gaussian process. This model assumes that the location of oil is distributed randomly over space but allows spatial correlation - the outcomes of exploration wells close to each other are highly correlated and the degree of correlation declines with distance. A draw from this process is a continuous function that, depending on the parameters of the process, can have many local maxima corresponding to separate clusters of oil fields (see Appendix Figure A5 for a one dimensional example). As I discuss further below, such models are widely used in natural resource exploration to model the spatial distribution of geological features. The estimation of these models is sometimes known as Kriging in the geostatistics literature.[12].

Formally, let $\rho(X) : \boldsymbol{X} \to [0, 1]$ be a function that defines the probability of exploration well success at locations $X \in \boldsymbol{X}$. I model $\rho(X)$ as being drawn from a *logistic Gaussian process* $G(\rho)$ over the space $\boldsymbol{X}$.[13] In particular, for any location X,

$$\rho(X) \equiv \rho(\lambda(X)) = \frac{1}{1 + exp(-\lambda(X))}, \tag{1}$$

where $\lambda(X)$ is a continuous function from $\boldsymbol{X}$ to $\mathbb{R}$.

The function $\lambda(X)$ is drawn from a *Gaussian process* with mean function $\mu(X)$ and covariance function $\kappa(X, X')$. This means that for any finite collection of K locations $\{1, ..., K\}$, the vector $(\lambda(X_1), ..., \lambda(X_K))$ is a multivariate normal random variable with mean $(\mu(X_1), ..., \mu(X_K))$ and a covariance matrix with (j, k) element $\kappa(X_j, X_k)$. The prior mean

[12]See Chapter 7.2 of Rasmussen and Williams (2006) for a discussion of the conditions required for consistency of maximum likelihood estimates of a Gaussian Process with binary outcomes. For the case of Kriging with spatial data see Mardia and Marshall (1984) and Cressie (1991). In general, estimates generated using likelihoods of the form 3 are consistent under *increasing-domain asymptotics*. That is, fixing *any* set of locations $\boldsymbol{X}$, every $X \in \boldsymbol{X}$ is explored for a sufficiently large sample. This asymptotic framework is consistent with the assumption that the data is generated by the model of license issuing and exploration outlined in Section 5. Small sample bias due to the selection of explored blocks remains a concern, but such selection should be limited since 91% of licensed blocks have been explored at least once by 1990.

[13]If well success rates were independent across locations j, a natural model would draw $\rho_j \in [0, 1]$ from a beta distribution. However there is no natural multivariate analogue of the beta distribution that allows me to specify a covariance between ρ_j and ρ_k for $j \neq k$.

function $\mu : \boldsymbol{X} \rightarrow \mathbb{R}$ is assumed to be smooth and the covariance function $\kappa : \boldsymbol{X} \times \boldsymbol{X} \rightarrow \mathbb{R}$ must be such that the resulting covariance matrix for any K locations is symmetric and positive semi-definite. One covariance function that satisfies these assumptions is the square exponential covariance function (Rasmussen and Williams, 2006) given by

$$\kappa(X, X') = \omega^2 exp\left(\frac{-|X - X'|^2}{2\ell^2}\right). \tag{2}$$

The parameter ω controls the variance of the process. In particular, for any X, the marginal distribution of $\lambda(X)$ is given by $\lambda(X) \sim N(\mu(X), \omega)$. The parameter ℓ controls the covariance between $\lambda(X)$ and $\lambda(X')$ for $X \neq X'$. Notice that as the distance $|X - X'|$ between two locations increases, the covariance falls at a rate proportional to ℓ. As $|X - X'|$ goes to 0, the correlation of $\lambda(X)$ and $\lambda(X')$ goes to 1, so draws from this process are continuous functions.

I estimate the parameters, $(\mu(X), \omega, \ell)$, of the Gaussian process model using data on the binary outcomes of all well exploration wells drilled between 1964 and 1990. Let $s = (s_1, s_2, ..., s_W)$ be a vector of length W where W is the total number of exploration wells drilled by all firms and $s_w = 1$ if well s was successful, and otherwise $s_w = 0$. Let $X = (X_1, ..., X_W)$ be a matrix recording the block centroid coordinates of each well. Then the likelihood of well outcomes s conditional on well locations X is given by:

$$L(s|X, \mu, \omega, \ell) = \int \left(\prod_{w=1}^{W} \rho(X_w)^{1(s_w=1)}(1 - \rho(X_w))^{1(s_w=0)}\right) dG(\rho; \mu, \omega, \ell) \tag{3}$$

The integrand is the product of Bernoulli likelihoods for each well for a particular draw of ρ, which encodes success probabilities at every location X_w. The integral is over draws of ρ with respect to the distribution $G(\rho)$, which is a function of the parameters. Note that I assume a flat mean function, $\mu(X) = \mu(X') = \mu$.

Table 2 records maximum likelihood estimates. The first column records the estimated values of the three parameters of the Gaussian process, while the second column records implied statistics of the distribution of $\rho(X)$ at the estimated parameters - the expected success probability, the standard deviation of success probability, and the correlation of success probability between two locations one block (18 km) away from each other. The parameters are identified by the empirical analogs of these statistics in the well outcome data. Most importantly, the estimated parameter ℓ captures the true spatial correlation of exploration well outcomes.

Table 2: Oil Process Parameters

Parameter	Estimate	Implied Statistics	
μ	-1.728	$E(\rho(X))$	0.207
	(0.202)		
ω	1.2664	$SD(\rho(X))$	0.179
	(0.146)		
ℓ	0.862	$Corr(\rho(0), \rho(1))$	0.471
	(0.102)		

Notes: The first column records parameter estimates from fitting the likelihood function given by equation 3 to data on the outcome of all exploration wells drilled between 1964 and 1990 on the relevant area of the North Sea. Well locations are taken to be the centroids of corresponding blocks. Standard errors computed using the Hessian of the likelihood function in parentheses. The second column records the implied expected probability of success, the standard deviation of the prior beliefs about probability of success, and the correlation of success probability between two locations one block (18 km) away from each other.

3.2 Interpretation as a Bayesian Prior

The estimated parameters, (μ, ω, ℓ), can be thought of as describing primitive geological characteristics that determine the distribution of oil deposits over space. If these parameters are known by firms and the Gaussian process model is a good approximation to the geological process that generates the distribution of oil, then the estimated process $G(\rho|\mu, \omega, \ell)$ describes the rational beliefs that firms should hold about the probability of exploration well success at each location X prior to observing the outcome of any wells. The parameters of this prior also determine how beliefs are updated according to Bayes' rule after well results are observed.

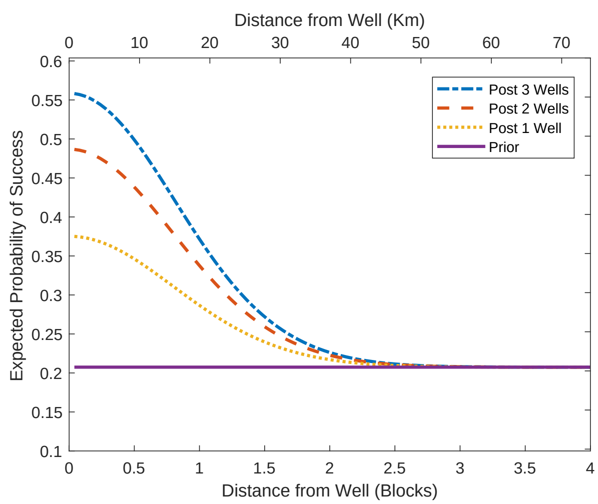
In particular, firms whose prior is described by $G(\rho)$ update their beliefs over the entire space $\boldsymbol{X}$ after observing a success or failure at a particular location X. Posterior beliefs at locations closer to X will be updated more than those at more distant locations. Figure 2 illustrates how posterior beliefs respond to well outcomes at different distances under the estimated parameters. The solid purple line illustrates the firm's constant prior expected probability of success of around 0.2.[14] The dotted yellow line represents the firm's posterior expected probability of success after observing one successful well at 0 on the x-axis. The dashed red and blue lines correspond to posteriors after observing two and three successful

[14]The assumption of a constant prior mean could be relaxed to allow μ to depend on, for example, prior knowledge of geological features. μ represents firms' mean beliefs in 1964, before any exploratory drilling took place. Brennand et al. (1998) emphasize that knowledge of subsea geology was extremely limited before exploration began. Using a modern map of actual geological features as inputs to the prior mean would therefore be inappropriate. For this reason, I believe it is not unreasonable to adopt a constant prior mean. It may be possible to identify a location varying prior probability of success from firms' initial exploration choices, as recorded in the maps in Appendix Figure A12. However, this would mean that the two step estimation approach described below, which uses beliefs as an input to policy function estimation, would be infeasible. This would instead require a more computationally costly full solution estimation method.

wells at the same location. Notice that the expected probability of success increases most at the well location, and decreases smoothly at more distant locations.

The true spatial correlation of well outcomes, captured by the parameter ℓ, determines the rate at which belief updating declines with distance. In particular, the estimated value of ℓ implies that firms should update their beliefs about the probability of success in response to well outcomes on neighboring blocks and those two blocks away, but not in response to well outcomes three or more blocks away. At these distances, the correlation in well outcomes dies out and thus so does the implied response of beliefs to well outcomes.[15]

Figure 2: Response of Beliefs to Well Outcomes



Notes: Figure depicts prior and posterior expected value of $\rho(X)$ in a one dimensional space for posteriors computed after observing one, two, and three successful wells at $X = 0$. The parameters (μ, ω, ℓ) of the logistic Gaussian process prior are set to the estimated values from Table 2.

Formally, let $w \in W$ index wells, let $s(w) \in \{0, 1\}$ be the outcome of well w, and let X_w denote the location of well w. If prior beliefs are given by the logistic Gaussian Process $G(\rho)$ then the posterior beliefs $G'(\rho)$ after observing $\{(s(w), X_w)\}_{w \in W}$ are given by

$$G'(\rho) = B(G(\rho), \{(s(w), X_w)\}_{w \in W}), \tag{4}$$

where $B(\cdot)$ is a Bayesian updating operator. Since the signals that firms receive are binary, there is no analytical expression for the posterior beliefs given the Gaussian prior and the observed signals. In particular, $G'(\rho)$ is non-Gaussian. I compute posterior distributions using the Laplace approximation technique of Rasmussen and Williams (2006) which provides a Gaussian approximation to the non-Gaussian posterior $G'(\rho)$. I discuss the procedure used

[15] In Appendix Figure A5 I illustrate belief updating under different values of ℓ in a numerical example.

to compute $B(\cdot)$ in more detail in Appendix B. Using the Bayesian updating rule it is possible to generate posterior beliefs for any set of observed well realizations. Appendix Figure A7 is a map of posterior beliefs for a firm that observed the outcome of *all* exploration wells drilled from 1964-1990, illustrating regions with different posterior expected probability of success and uncertainty. In general, the standard deviation of posterior beliefs is lower in regions where more exploration wells have been drilled.

The Gaussian process model is a parsimonious approximation to more complex inferences about nearby geology made by geologists based on exploration well results. The method of spatial interpolation between observed wells that is achieved by computing the Gaussian process posterior (known as Kriging) is widely applied in predicting the distribution of geological features over space. The model of beliefs employed here corresponds to "trans-Gaussian Kriging", so called because of the use of a transformed Gaussian distribution (Diggle, Tawn, and Moyeed, 1998). Whether or not we think these beliefs are a correct representation of how oil deposits are distributed, the model of learning described above *is representative* of how geologists (and presumably oil companies) *think*.

3.3 Beliefs and Development Payoffs

In what follows, I adopt the additional simplifying assumption that firms have beliefs about the probability of success at the *block level*. In particular, let $\rho_j = \rho(X_j)$ where X_j are the coordinates of the centroid of block $j \in \{1, ..., J\}$. When an exploration well is drilled *anywhere* on block j, firms update their beliefs as if the success of that well is drawn with probability ρ_j. One way to rationalize this assumption is to assume that the locations of exploration wells *within* blocks are drawn uniformly at random (see Appendix Figure A6 for an illustration). The probability of success, ρ_j, then has a natural interpretation as the share of block j that contains oil, and the observed success rate is an estimate of this probability which becomes more precise as the number of wells on the block increases.

Formally, I assume that the potential oil revenue yielded by block j, π_j, is drawn from a distribution $\Gamma(\pi|\rho_j, P)$ where P is the oil price and $\frac{\partial E(\pi_j)}{\partial \rho_j} > 0$. A higher exploration success probability ρ_j corresponds to higher expected oil revenue. Beliefs about exploration well success $G(\rho)$ then imply beliefs about the potential oil revenue on block j given by:

$$\tilde{\Gamma}_j(\pi|G, P) = \int \Gamma(\pi|\rho_j, P)dG(\rho). \tag{5}$$

This interpretation of block-level success rates is supported by a positive relationship between the realized exploration success rate and estimated oil reserves on developed blocks, illustrated by Appendix Figure A3. Note that the assumption that probability of success is a primitive feature of *a block* and within-block location choice is random implies that the realized success rate on a block should be constant over time. This abstracts from the reality of within block exploration - for instance success rates would not be constant if firms continue to drill near previous successful wells *within* the block. In Appendix Table A5 I present the results of regressions that show that within blocks, the success rate is not significantly higher or lower for later wells than for earlier wells. This suggests that the assumption of constant block-level success rates is a reasonable approximation for the purpose of studying the firm's exploration decision across blocks.

Let J_d be a set of blocks that have been developed and for which the oil revenue, π_j, has therefore been observed. Since the distribution of oil revenue is a function of ρ_j, it is possible to define a new Bayesian updating operator,

$$G'(\rho) = B(G(\rho), \{(s(w), X_w)\}_{w \in W}, \{\pi_j, P\}_{j \in J_d}). \tag{6}$$

That is, realizations of development payoffs affect firms' posterior beliefs about ρ.

4 Descriptive Evidence

The estimated model of beliefs suggests that there is high degree of correlation between well outcomes on neighboring blocks. This spatial correlation is estimated from data on well *outcomes* at different locations. In this section, I use data on firms' drilling *decisions* to test whether firm behavior is consistent with the estimated model of rational beliefs.

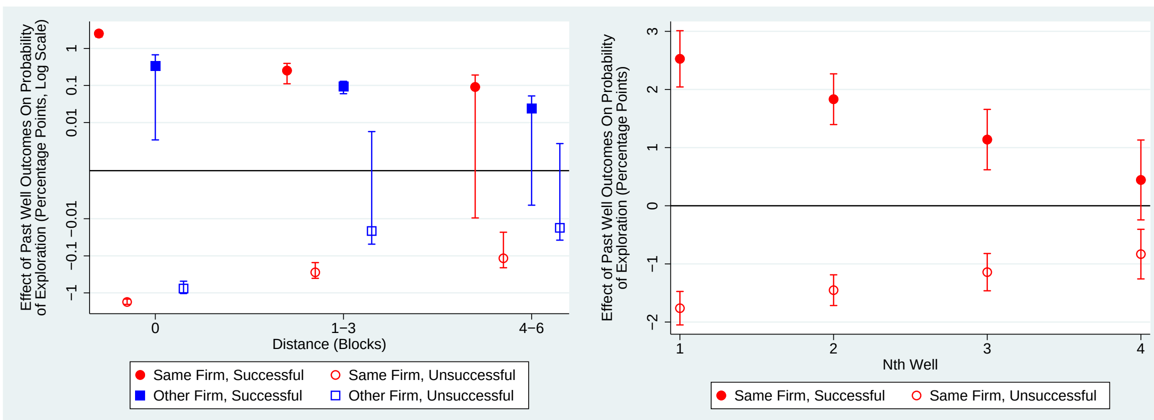
4.1 Exploration Drilling Patterns

The estimated spatial correlation illustrated by Figure 2 suggests that firms should make inferences across space based on past well results. I test this prediction using data on firm behavior. Let Suc_{jdot} be the cumulative number of successful wells drilled on blocks distance d from block j before date t by firms $o \in \{f, -f\}$, where $-f$ indicates all firms other than firm f (including exploration wells drilled on now-developed blocks). $Fail_{jdot}$ is analogously defined as the cumulative number of past unsuccessful wells. I estimate the following regression specification using OLS:

$$Explore_{fjt} = \alpha_{ft} + \beta_j + \sum_d \sum_{o \in \{f, -f\}} g_{do}\left(Suc_{jdot}, Fail_{jdot})\right) + \epsilon_{fjt}. \tag{7}$$

Where g_{do} is a flexible function of cumulative successful and successful well counts for wells of type (d, o). $Explore_{fjt}$ is an indicator for whether or not firm f drilled an exploration well on block j in month t. Notice that the specification includes firm-month and block fixed effects. This means that the effects of past wells are identified by within-block changes in the set of well results over time, and not by the fact that some blocks have higher average success rates than others and these blocks tend to be explored more.

Figure 3: Response of Drilling Probability to Cumulative Past Results



Notes: Points are the estimated marginal effect of each type of past well on $Explore_{fjt}$ from the specification given by equation 7 where $g_{do}(\cdot)$ is quadratic in each of the arguments. In the left panel, marginal effects are computed for the first well of each type. In the right panel, the marginal effects of same-firm same-block wells are reported for the nth well of each type. Error bars are 95% confidence intervals computed using standard errors clustered at the firm-month level. Sample includes block-months in the relevant region before 1991. I drop observations from highly explored regions where the number of nearby own wells (those on 1st and 2nd degree neighboring blocks) is above the 80th percentile of the distribution in the data.

The left panel of Figure 3 records the estimated marginal effect of the first well of each type on the probability of subsequent exploration. I include three distance bands in the regression - wells on the same block, those 1-3 blocks away, and those 4-6 blocks away. Solid red circles indicate the effect on the probability of firm f drilling an exploration well on block j of an additional past successful well drilled by firm f at each distance. Hollow red circles record this effect for unsuccessful past wells drilled by firm f. The results indicate that additional successful wells on the same block and 1-3 blocks away significantly increase the probability of subsequent exploration, and additional unsuccessful wells significantly decrease the probability of subsequent exploration. These results suggest that firms make inferences across space at distances consistent with the spatial correlation of well results illustrated by Figure 2, with the size of the drilling response declining with distance. These effects also diminish with the total number of wells drilled - the right panel shows that the marginal effect of a successful same-firm well on exploration probability falls from around

2.5 percentage points for the first well to 0 for the fourth well. This is consistent with the diminishing marginal effect of exploration wells on beliefs illustrated by comparing the curves in Figure 2.

Blue squares in the left panel indicate the effect of past wells drilled by other firms on firm f's probability of exploration. The effects are of the same sign but have magnitudes between 20% and 50% of the same-firm well effects. As with the same-firm effects, the other-firm effects diminish with distance and lose statistical significance at distances of 4-6 blocks.[16] This suggests that information flow across firms is imperfect - consistent with the confidentiality regulations limiting observability of other firm wells. Hypothesis tests reject the equality of the same- and other-firm effects for all pairs of coefficients except those on successful wells 4-6 blocks away which are both indistinguishable from zero.[17]

To test directly whether firm behavior responds to changes in beliefs, I regress firm exploration decisions on model-implied posteriors. Since exploration wells generate *information*, and their value is in informing firms' future drilling decisions, a natural hypothesis is that the probability of drilling an exploration well should be increasing in the expected information generated by that well. For instance, the first exploration well drilled on a block should be more valuable than the tenth because its marginal effect on beliefs is greater.

I compute the model-implied posterior beliefs for each block j, each month t, under the assumption that firms observe all wells drilled before that month according to the Bayesian updating rule (4).[18] I obtain $E_t(\rho_j)$, the posterior mean, and $Var_t(\rho_j)$, the posterior variance of beliefs about the probability of success on block j, ρ_j. To measure the expected information gain of an additional well I obtain the expected Kullback-Leibler divergence, $KL_{j,t}$, between the prior and posterior distributions following an additional exploration well for each (j, t) (Kullback, 1997).

Column 1 of Table 3 records the coefficients from a regression of $KL_{j,t}$ on the computed posterior variance and a quadratic in posterior mean at (j, t). There is an inverse u-shaped relationship between expected KL divergence and $E_t(\rho_j)$ that is maximized when

[16]Since the regression includes block fixed effects, the effect of other firm wells on the same block comes from variation in the number of wells over time when multiple firms hold licenses on the same block. See Section 2.2 for discussion of how I assign blocks to firms.

[17]I report a number of additional specifications in the Appendix. Appendix Table A6 reports analogous results for different sub-periods of the data. These results indicate that the ratio of the effect of wells 1-3 blocks away to the effect of wells on the same block is relatively constant over time. This is consistent with the assumption that the firms are learning about the location of oil, not about the true value of the spatial covariance parameter ℓ which I assume is known to firms ex-ante. Appendix Figure A1 shows that firm's response to the outcome of past wells that are no longer confidential is significantly greater than firms' response to wells in the confidentiality period.

[18]In this section, I compute beliefs as if all firms observe the results of all other firms' exploration wells. This assumption is relaxed in the structural model developed in Section 5.

$E_t(\rho_j) = 0.48$. This reflects the classic result in information theory (see for example MacKay, 2003) that the information generated by a Bernoulli random variable is maximized when the probability of success is 0.5. There is a positive relationship between $Var_t(\rho_j)$ and KL_{jt}, consistent with the information gain from an additional well increasing in variance.

The second column of Table 3 presents estimated coefficients from a regression of $Explore_{fjt}$ on $Var_t(\rho_j)$, a quadratic in $E_t(\rho_j)$, and (f, j) level fixed effects. Note that the coefficients follow the same pattern as those in the first column: firms are less likely to drill exploration wells on blocks with very high or very low expected probability of success, and are more likely to drill exploration wells on blocks with higher variance in beliefs. Firm *behavior* aligns closely with the *theoretical* relationship between moments of the posterior beliefs and the expected information generated by exploration wells. The third column shows that development is more likely on blocks with high mean and low variance beliefs, consistent with the correlation between exploration well success and oil reserves illustrated by Figure A3.

Table 3: Response of Drilling Probability to Posterior Beliefs and License Distribution

Dependent Variable:	KL Divergence	Exploration Well	Develop Block
Posterior Mean	0.414***	0.267***	0.004
	(0.001)	(0.070)	(0.003)
Posterior Mean^2	-0.450***	-0.206**	.
	(0.001)	(0.100)	.
Posterior Variance	0.082***	0.0199***	-0.002***
	(0.000)	(0.006)	(0.001)
Own Share Nearby Blocks	.	0.017**	0.000
	.	(0.008)	(0.001)
R^2	0.970	0.106	0.077
N	86,362	86,362	86,362
Firm-Month and Block FE	No	Yes	Yes

Notes: Standard errors clustered at the firm-month level. Mean, variance, and KL divergence of posterior beliefs computed for each (f, j, t) as if all wells drilled by all firms up to month $t-1$ are observed. Sample is all *undeveloped* firm-block-months in the relevant region,. *** indicates significance at the 99% level. ** indicates significance at the 95% level. * indicates significance at the 90% level. Standard errors are two-way clustered at the firm-month and block levels. Own share of nearby blocks is the number of nearby own blocks as a share of all nearby licensed and undeveloped blocks, where nearby means at most third degree neighbors. Sample includes block-months in the relevant region between 1970 and 1990.

The final row of Table 3 reports coefficients on the share of nearby undeveloped blocks belonging to the same firm as block j. The results indicate that a firm is more likely to explore when a larger share of the surrounding blocks are owned by that firm. This is consistent with information spillovers across blocks driving firms' exploration decisions. Since the information generated by an exploration well on block j is informative about nearby blocks, a firm learns more about the distribution of oil on its own blocks when it holds licenses on more of the block surrounding block j. On the other hand, when other firms hold more of the blocks surrounding block j, then a firm may have a greater incentive to delay exploration and learn from other firms' results. [19]

5 An Econometric Model of Optimal Exploration

To measure the extent to which information externalities affect industry surplus, I estimate a structural econometric model of the firm's exploration problem in which I assume that firm beliefs follow the logistic Gaussian process model of Section 3.2. I set up the firm's problem by specifying a full information game in which firms observe the results of all wells. I then extend the model to one of asymmetric information in which firms do not observe the results of other firms' wells with certainty.

5.1 Full Information

I start by specifying a full information game played by a set of firms F. Firms are indexed by f, discrete time periods are indexed by t, and blocks are indexed by j. J is the set of all blocks. $J_{ft} \subset J$ is the set of undeveloped blocks on which firm f holds drilling rights at the beginning of period t. $J_{0t} \subset J$ is the set of undeveloped blocks on which no firm holds drilling rights at the beginning of period t. P_t is the oil price.

Exploration wells are indexed by w, and each well is associated with an outcome $s(w) \in \{0,1\}$, a block $j(w)$, a firm $f(w)$, and a drill date $t(w)$. The set of all locations and realizations of exploration wells drilled on date t is given by $W_t = \{(j(w), s(w)) : t(w) = t\}$. The set of all blocks that are developed on date t is J_{dt}.

The firm's prior beliefs about the probability of exploration well success on each block are given by the logistic Gaussian process G_0 defined in Section 3.1. G_{ft} is firm f's posterior at the beginning of period t. Under the assumption of full information firms observe the results of all wells, so $G_{ft} = G_t$ for all firms $f \in F$.

The industry state at date t is described by

$$\mathcal{S}_t = \{G_t, \{J_{ft}\}_{f \in F \cup \{0\}}, P_t\}. \tag{8}$$

[19] An alternative explanation for this pattern is that firms may obtain more leases in areas that they expect to be productive, generating a correlation within each firm between density of leases and exploration activity. To provide additional support to the free riding hypothesis, Appendix Table A7 presents results indicating that firms' exploration rates fall significantly after licensing rounds in which more new licenses are issued to other firms on nearby blocks, controlling for own licenses, supporting the hypothesis of strategic delay. An ideal experiment would control for each firm's license applications, but this data is not available.

Each period, the firm makes two decisions sequentially. First, in the *exploration stage*, it selects at most one block on which to drill an exploration well. Then, in the *development stage*, it selects at most one block to develop.

Drilling an exploration well on block j incurs a cost, $c(j, \mathcal{S}_t) - \epsilon_{ftj}$. Developing block j incurs a cost $\kappa(j, \mathcal{S}_t) - \nu_{ftj}$. ϵ_{ftj} and ν_{fjt} are private information cost shocks drawn iid from Type I extreme value distributions with variance parameters σ_ϵ and σ_ν. Developing block j at date t yields a random payoff π_{jt}. Firms' beliefs about the distribution of payoffs on block j are $\tilde{\Gamma}_j(\pi|G_t, P_t)$, defined in equation 5.

The timing of the game is as follows: *Exploration Stage*

1. Given state $\mathcal{S}_t$, each firm f observes a vector of private cost shocks $\boldsymbol{\epsilon}_{ft}$.

2. Firm f chooses an exploration action, $a^E_{ft} \in J_{ft} \cup \{0\}$. If $a^E_{ft} \neq 0$, then firm f incurs an exploration cost.

Development Stage

1. Given state $\mathcal{S}'_{ft} = (\mathcal{S}_t, a^E_{ft})$, each firm f observes a vector of private cost shocks $\boldsymbol{\nu}_{ft}$.

2. Firm f chooses a development action, $a^D_{ft} \in J_{ft} \cup \{0\}$. If $a^D_{ft} \neq 0$, then firm f incurs a development cost.

Well Outcomes Realized

1. Exploration well results W_t are realized.

2. If $a^D_{ft} = j$ then $j \in J_{dt}$, and firm f draws oil revenue π_{jt}.

3. The industry state evolves to $\mathcal{S}_{t+1} = \{G_{t+1}, \{J_{ft+1}\}_{f \in F \cup \{0\}}, P_{t+1}\}$.Exploration well outcomes and development payoffs are not realized until the end of the development phase. Beliefs G_t are then updated based on exploration well results W_t and realized revenues $\{\pi_{jt}\}_{j \in J_{dt}}$according to $G_{t+1} = B(G'_t, W_t, \{\pi_j, P_t\}_{j \in J_{dt}})$, where $B(\cdot)$ is defined in equation 6.

Other state variables evolve at the end of the development stage as follows. I assume that oil price follows an exogenous AR(1) process, so $P_{t+1} = \delta_0 + \delta_1 P_t + \zeta_t)$ where $\zeta_t \sim N(0, \sigma_\zeta)$. I assume that firm licenses on undeveloped blocks are issued and surrendered according to an exogenous stochastic process defined by probabilities $P(j \in J_{ft+1}|\{J_{gt}\}_{g \in F \cup \{0\}}, a^D_{ft})$. Developed blocks are removed from firms' choice sets, so $P(j \in J_{ft+1}|a^D_{ft} = j) = 0$ and $P(j \in J_{ft+1}|j \notin \cup\{J_{gt}\}_{g \in F \cup \{0\}}) = 0$. This assumption eliminates any strategic consideration in the timing of drilling with respect to regulatory deadlines, the announcement of new licensing rounds, and the firm's decision to surrender a block.

The firm's continuation values at the beginning of the exploration and development stages (before private cost shocks are realized) are described by the following two Bellman equations:

$$V_f^E(\mathcal{S}_t) = E_{\boldsymbol{\epsilon}_{ft}}\left[\max_{a_{ft}^E \in J_{ft}\cup\{0\}} \left\{V_f^D(\mathcal{S}'_{ft}) - c(a_{ft}^E, \mathcal{S}_t) + \epsilon_{ftj}\right\}\right] \tag{9}$$
$$V_f^D(\mathcal{S}'_{ft}) = E_{\boldsymbol{\nu}_{ft}}\left[\max_{a_{ft}^D \in J_{ft}\cup\{0\}} \left\{E_{\pi_{a_{ft}^D},\mathcal{S}_{t+1}}\left[\beta V_f^E(\mathcal{S}_{t+1}) + \pi_{a_{ft}^D} | a_{ft}^D, \mathcal{S}'_{ft}\right] - \kappa(a_{ft}^D, \mathcal{S}'_{ft}) + \nu_{ftj}\right\}\right].$$

Where β is the one period discount rate.The inner expectation in the development Bellman equation is taken over realizations of other firms' exploration and development actions, development revenues $\{\pi_{jt}\}_{j\in J_{dt}}$ and exploration results W_t which affect next period's s beliefs G_{t+1}, as well as the evolution of P_{t+1} and J_{t+1}. There is no expectation inside the exploration Bellman because beliefs are not updated until the end of the development phase.

Define choice-specific ex-ante (before private cost shocks are realized) value functions as,

$$\begin{aligned} v_f^E(a_t^E, \mathcal{S}_t) =& E_{\mathcal{S}'_t}\left[V_f^D(\mathcal{S}'_t) | a_t^E, \mathcal{S}_t\right] - c(a_t^E, \mathcal{S}_t) \\ v_f^D(a_t^D, \mathcal{S}'_t) =& E_{\pi_{a_t^D},\mathcal{S}_{t+1}}\left[\beta V_f^E(\mathcal{S}_{t+1}) + \pi_{a_t^D} | a_t^D, \mathcal{S}'_t\right] - \kappa(a_t^D, \mathcal{S}'_t). \end{aligned} \tag{10}$$

I assume that ϵ_{ftj} and ν_{ftj} are distributed type I extreme value with standard deviation parameters σ_ϵ and σ_ν, yielding conditional choice probabilities (CCPs):

$$P(a_f^E = j|\mathcal{S}_t) = \frac{exp\left(\frac{1}{\sigma_\epsilon} v_f^E(j, \mathcal{S}_t)\right)}{\sum_{k\in J_{ft}\cup\{0\}} exp\left(\frac{1}{\sigma_\epsilon} v_f^E(k, \mathcal{S}_t))\right)}. \tag{11}$$

With a similar expression for the CCP of development action j, $P(a_f^D = j|\mathcal{S}'_t)$.

A Markov perfect equilibrium of this game is then defined by strategies $a_f^E(\mathcal{S}, \boldsymbol{\epsilon})$ and $a_f^D(\mathcal{S}, \boldsymbol{\nu})$ that maximize the firm's continuation value, conditional on the state variable and the privately observed cost shocks,

$$\begin{aligned} a_f^E(\mathcal{S}, \boldsymbol{\epsilon}) &= \arg\max_{a^E\in J_f\cup\{0\}} \left\{v_f^E(a^E, \mathcal{S}) + \epsilon_{ta^E}\right\} \\ a_f^D(\mathcal{S}', \boldsymbol{\nu}) &= \arg\max_{a^D\in J_f\cup\{0\}} \left\{v_f^D(a^D, \mathcal{S}') + \nu_{ta^D}\right\}, \end{aligned} \tag{12}$$

where the firm forecasts all firms' actions conditional on the industry state using the true conditional choice probabilities is equation 11.

5.2 Asymmetric Information

The model described above assumes that firms can perfectly observe the results of each other's exploration wells as soon as they are drilled. In reality, industry regulation allows for confidentiality of well data for several years after an exploration well is drilled, and the empirical evidence presented in Section 3 suggests imperfect spillover of information between firms. The extent to which information flows between firms before the end of the well data confidentiality period is a potentially important determinant of firms' incentive to delay exploration.

To allow for imperfect spillovers of information in the model, I make an alternative assumption about when firms observe the results of exploration wells. In particular, when a well w is drilled by firm f, I let each firm $g \neq f$ observe the outcome, $s(w)$, with probability α. $s(w)$ is revealed to all firms τ_w periods after the well is drilled, on expiry of the confidentiality window. The length of the confidentiality window is drawn i.i.d. from the distribution $F_\tau(\tau_w)$.[20]

Formally, let $o_f(w) \in \{0, 1\}$ be a random variable drawn independently across firms after the exploration stage of period $t(w)$ where $P(o_f(w) = 1|f(w) \neq f) = \alpha$ and $P(o_f(w) = 1|f(w) = f) = 1$. The cumulative set of well results observed by firm f in period t is

$$W^o_{ft} = \{(j(w), s(w)) : (o_f(w) = 1 \text{ and } t(w) \leq t) \text{ or } (o_f(w) = 0 \text{ and } t(w) \leq t - \tau(w))\} . \tag{13}$$

In general, $G_{ft} \neq G_{gt}$ since firms observe different sets of well outcomes, so firm f's state variable can now be written:

$$\mathcal{S}_{ft} = \{G_{ft}, \{J_{ft}\}_{f \in F \cup \{0\}}, P_t, W^u_{ft}\}, \tag{14}$$

where W^u_{ft} is the set of locations and dates of all wells whose results are *not* observed by firm f at date t,

$$W^u_{ft} = \{(j(w), t(w)) : o_f(w) = 0 \text{ and } t(w) \geq t - \tau(w)\} . \tag{15}$$

Note that I assume firms know the distribution of end confidentiality window lengths, $F_\tau(\tau_w)$, but not the confidentiality end date, $\tau(w)$ for each well.

The introduction of this asymmetric information structure complicates the firm's problem. In Markov perfect Bayesian equilibrium, each firm f must form beliefs about every

[20] Note that unobserved exploration well outcomes are revealed after τ_w periods, even if the block is developed before this date.

other firm g's beliefs, G_{gt}, in order to forecast the next period's state. The history of firm g's actions is informative about G_{gt} and about well outcomes unobserved by firm f. For instance, if firm g drilled many exploration wells on block j, this should signal to firm f something about g's beliefs about the success probability on that block, even if firm f did not observe the outcome of any of those wells directly. In contrast to the full information game, this means that the entire history of drilling and license allocations should enter the firm's state.

As discussed by Fershtman and Pakes (2012), the addition of these "informationally relevant" but not "payoff relevant" state variables makes estimating the asymmetric information game and finding equilibria computationally infeasible. Because of these difficulties, dynamic games with asymmetric information are rarely used in empirical work. Rather than include all potentially informationally relevant observables in the firm's state, I adopt an equilibrium concept that conditions beliefs only on $\mathcal{S}_{ft}$. First, I define belief functions that map a firm's current information, $\mathcal{S}_f$, to perceived probabilities of other firms' actions and well outcomes.

Assumption 1. *Firm f believes that at every period t the probability of exploration by a firm $g \neq f$ on block $j \in J_{gt}$ is given by $Q^E(g, j; \mathcal{S}_{ft}) \in [0, 1]$. Likewise, beliefs over the probability of development are $Q^D(g, j; \mathcal{S}_{ft}) \in [0, 1]$.*

Assumption 1 says that firms' beliefs are specified by functions Q^E and Q^D of their current state $\mathcal{S}_{ft}$. For example, these functions could be the expected action probabilities based on Bayesian posteriors about other firms' states. Although this assumption allows for Bayesian beliefs, it restricts the set of information that firms can use to make forecasts about other firms' actions to the payoff-relevant state variables in equation 14.

Assumption 2. *Firm f forecasts the outcomes of exploration and development wells using beliefs $Q^G(\rho; G(\rho), W^u_{ft})$.*

Assumption 2 says that firms adjust their beliefs about this distribution of oil based on the number of wells with unobserved results on each block, W^u_{ft}. Thus, beliefs depend on observed actions, not just the observed outcomes which define $G(\rho)$. For instance, it might be that $E(\rho_j|G(\rho)) < E(\rho_j|\tilde{G}(\rho))$ if there are many exploration wells on block j whose results are unobserved. As with Assumption A1, I restrict the observable variables that enter Q^G. For instance I do not allow beliefs to depend on the timing of past exploration.

The following equilibrium definition requires the belief objects Q^E, Q^D, and Q^G to be consistent with firm behavior.

Definition 1. An equilibrium is a set of policies, $a^E(\mathcal{S}_f, \boldsymbol{\epsilon})$, $a^D(\mathcal{S}_f, \boldsymbol{\epsilon})$, beliefs about other firms' actions, Q^E, Q^D, and beliefs about the distribution of oil, Q^G such that:

1. $a^E(\mathcal{S}_f, \boldsymbol{\epsilon})$ and $a^D(\mathcal{S}_f, \boldsymbol{\epsilon})$ solve the dynamic program described by equation 9, where expectations over other firms actions and well outcomes are described by Assumptions A1 and A2.

2. $Q^E(g, j; \mathcal{S}_{ft})$, $Q^D(g, j; \mathcal{S}_{ft})$ are consistent with equilibrium policies. Let $P(a^E_{f,t} = j|\mathcal{S}_{f,t})$ and $P(a^D_{f,t} = j|\mathcal{S}'_{f,t})$ be firms' equilibrium CCPs.

$$Q^E(\mathcal{S}_{ft}, g, j) = E\left[P(a^E_{gt} = j|\mathcal{S}_{gt})|\mathcal{S}_{ft}\right] \quad (16)$$
$$Q^D(S_{ft}, g, j) = E\left[P(a^D_{gt} = j|\mathcal{S}_{gt})|\mathcal{S}_{ft}\right].$$

   Where the expectations are taken over states with respect to the equilibrium distribution of $\mathcal{S}_{gt}$ conditional on $\mathcal{S}_{ft}$.

3. $Q^G(\rho; G(\rho), W^u_{ft})$ is the posterior distribution of ρ conditional on $(G(\rho), W^u_{ft})$ consistent with the equilibrium distribution of $(G(\rho), W^u_{ft})$ conditional on ρ and the prior distribution of ρ, $G_0(\rho)$.

Condition 2 means that in equilibrium, firm f's beliefs about the probability of exploration and development by other firms are correct *in expectation*, conditional on S_{ft}. This assumption is close in spirit to that made by Fershtman and Pakes (2012), whose Experience Based Equilibrium (EBE) imposes consistency between firm's beliefs about state transitions conditional on its information set and the behavior of other firms. The crucial difference between this approach and Markov perfect Bayesian equilibrium is that beliefs are specified over actions or own-state transitions, not over other firms' states.[21] Similar assumptions that require agents' beliefs about outcomes to be correct *on average* conditional on some information set have been made in static models, for example by Doraszelski, Lewis, and Pakes (2018), and Agarwal and Somiani (2018).

Condition 3 imposes a similar consistency requirement on firms' beliefs about exploration success rates. Firms beliefs $G(\rho)$ are formed using well outcomes, as described in Section 3, and then updated according to Bayes' rule using information on the set of unobserved wells W^u_{ft} and the equilibrium distribution of $(G(\rho), W^u_{ft}, \rho)$.

[21] One difference between EBE and the equilibrium concept used in this paper is that EBE imposes consistency of beliefs on the set of recurrent states - those that are repeatedly visited in stationary equilibrium. This means that EBE can be estimated by simulating a stationary environment for a long time using reinforcement learning. In this paper's environment, any state will not be visited infinitely often because, for example, a developed block will never be returned to the set of undeveloped blocks. However, there still exists an equilibrium distribution of states which can be computed by simulating the model many times (rather than once for a long time). I use this distribution to define consistency of beliefs. Computation of beliefs is discussed further in Section 6 below.

This equilibrium concept greatly simplifies estimation and computation of equilibria relative to a Markov perfect Bayesian equilibrium. As discussed below, the functions $Q^E(\mathcal{S}, g, j)$ and $Q^D(\mathcal{S}, , g, j)$ are CCP-like objects that can be estimated from simulated data in a straightforward way. Notice that, like the firm's choice probabilities, $Q^E(\mathcal{S}, g, j)$ and $Q^D(\mathcal{S}, g, j)$ are equilibrium objects and will change in counterfactuals.

These assumptions preserve important behavioral implications of the model. The model allows for free riding because firms have expectations about the future exploration of other firms. The model also allows for the possibility of an "encouragement effect" discussed by Dong (2017) because firms anticipate that their actions may affect the probability of other firms drilling in future through the dependence of Q^E on S_{ft}. This approach to modeling beliefs is restrictive insofar as the variables that enter S_f do not include all informationally relevant state variables that are potentially observed by the firm. This can be thought of as a bounded rationality assumption that restricts the set of objects on which firms can condition their beliefs.

6 Estimation & Identification

6.1 Parameterization

I measure the oil price using the monthly West Texas Intermediate oil price inflated to 2011 dollars using the UK Producer Price Index converted to dollars using the UK/US exchange rate. The WTI price is highly correlated with the Brent crude price over the time period studied, and data on Brent crude prices is not available before 1980. For years before 1980 where the Brent price is unavailable I use projected values from a regression of Brent on the West Texas Intermediate price. I let a period be one month.[22] I set the one month discount rate to $\beta = 0.992$, which corresponds to a 10% annual discount.

I specify exploration and development costs according to equation 17.

$$\begin{aligned} c(j, \mathcal{S}_{ft}) &= c_0 \\ \kappa(j, \mathcal{S}_{ft}) &= \kappa_0 + \kappa_1 1(Var_{ft}(\rho_j) < \underline{v}). \end{aligned} \tag{17}$$

Development cost on block j is allowed to depend on the variance of firm's beliefs about ρ_j. In particular, I allow for lower (or higher) development costs on blocks where the firm's

[22]The choice of a one month period imposes an implicit capacity constraint - each firm can choose at most one block to explore and one block to develop each month. In practice, in 94% of (f, t) observations where exploration takes place, only one exploration well is drilled. I never observe more than one block developed by the same firm in the same month. In my detailed discussion of the estimation routine in Appendix C, I describe how I deal with observations where there are multiple exploration wells in a month.

posterior variance is below a threshold $\underline{v}$. This captures the fact that development cannot take place until the firm has extensively studied the target field through exploratory drilling, which provides additional technical information that I do not model directly, beyond the presence of oil. Empirically, this specification helps rationalize the concentration of development on low-variance blocks. I set $\underline{v} = 0.5$, which is greater than the posterior variance for 86% of developed blocks.[23]. I do not adopt a more flexible specification because of the small number of development actions in the data (50).

The model parameters are therefore $\{\theta_1, \theta_2, \alpha, \delta, \sigma_\zeta\}$, where $\theta_1 = \{\mu, \omega, \ell\}$ are the parameters of the firm's beliefs defined in Section 3, $\theta_2 = \{c_0, \kappa_0, \kappa_1, \sigma_c, \sigma_\kappa\}$ are the cost parameters, α is the probability of observing another firm's well outcome before it is made public, and (δ, σ_ζ) are the parameters of the oil price process. Other objects to be estimated are the transition probabilities of the license issuing process $P(j \in J_{ft+1}|J_t, \{J_{gt}\}_{\forall g \in F})$, the distribution of development profits, $\Gamma(\pi; \rho_j, P_t)$, and firm beliefs about other firms' actions, Q^E, Q^D, and Q^{Past}.

6.2 Estimation

Parameters θ_1 are taken from the estimation procedure described in Section 4.1. I estimate σ_ζ with the variance of monthly changes in the log oil price. I estimate $\Gamma(\cdot)$ using data on realized oil flows from all developed wells. I detail this part of estimation in Appendix C.4. Probabilities $P(j \in J_{f,t+1}|J_t, \{J_{g,t}\}_{\forall g \in F})$ that are used by firms to forecast the evolution of license assignments are estimated using a logit regression of the licenses issued in period $t+1$ on the period t license assignment. The distribution of confidentiality deadlines, $F_\tau(\tau_w)$, is estimated using the empirical CDF. I detail this part of estimation in Appendix B.5.

The remaining parameters are estimated using a two step conditional choice probability method related to those described by Hotz, Miller, Sanders and Smith (1994) and Bajari, Benkard and Levin (2007). In the first step, I obtain estimates of the conditional choice probabilities (CCPs) given by equation 11 and the parameter α. First step estimates of the CCPs are then used to simulate data which I use to estimate the belief functions Q^E, Q^D, and Q^{Past}.

Using the estimated CCPs and belief functions, I compute the firm's state-specific continuation values (9) as functions of the remaining parameters θ_2 by forward simulation. I then find the value of θ_2 that minimizes the distance between the first step estimates of the CCPs and the choice probabilities implied by the simulated continuation values. I describe this two step procedure in detail in Appendix C.

[23] See Appendix Figure A8

Estimation of Conditional Choice Probabilities If the state variable were observable in the data, then CCPs $\hat{P}(a_f^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$ could be estimated directly using the empirical choice probability conditional on the state. However, the asymmetric information structure of the model means that the true state is not observed by the econometrician. In particular, the econometrician knows the outcome of every well, but does not know *which* outcomes were observed by each firm. Formally, the data does not include the vector $\boldsymbol{o}_f$ that records which other-firm well outcomes were observed by firm f. Different realizations of $\boldsymbol{o}_f$ imply different states through the effect of observed well outcomes on G_{ft}. The data is therefore consistent with a *set* of possible states $\tilde{\mathcal{S}}_f$ for each firm.[24]

To recover CCP estimates, observe that different values of the parameter α define distributions $P(\mathcal{S}_f|\tilde{\mathcal{S}}_f, \alpha)$ over the elements of $\tilde{\mathcal{S}}_f$. For example, suppose at date t there was one other-firm well w that may have been observed by firm f. The data is consistent with two possible states: let $\mathcal{S}_{ft}^1$ be the state if $o_f(w) = 1$ and $\mathcal{S}_{ft}^0$ be the state if $o_f(w) = 0$. From the econometrician's perspective, $P(\mathcal{S}_{ft}^1|\{\mathcal{S}_{ft}^1, \mathcal{S}_{ft}^0\}, \alpha) = \alpha$. I provide a formal definition of the distribution $P(\mathcal{S}_f|\tilde{\mathcal{S}}_f, \alpha)$ in Appendix C.

Given this distribution over states, the likelihood of a sequence of exploration and development choice observations is:

$$\mathcal{L}_f = \sum_{\mathcal{S}_f \in \tilde{\mathcal{S}}_f} \left[\left(\prod_{t=1}^{T} \prod_{j \in J_{ft} \cup \{0\}} P(a^E = j|\mathcal{S}_f)^{1(a_{ft}^E = j)} P(a^D = j|\mathcal{S}_f)^{1(a_{ft}^D = j)} \right) P(\mathcal{S}_f|\tilde{\mathcal{S}}_f, \alpha) \right]. \quad (18)$$

I maximize this likelihood to obtain estimates of the conditional choice probabilities $\hat{P}(a_f^E = j|\mathcal{S}_{ft})$ and the information spillover parameter, $\hat{\alpha}$, which controls the probability weight placed on each of the different states $\mathcal{S}_f \in \tilde{\mathcal{S}}_f$ that could have obtained given the data.

Since the state variable is high dimensional, I use the logit structure of $\hat{P}(a_f^E = j|\mathcal{S}_{ft})$ implied by equation 11 and approximate the choice-specific value functions $v_f^E(j, \mathcal{S}_{ft})$ for each alternative j with a linear equation in elements of the state variable. For the exploration choice, this approximation to the state variable includes a third order polynomial in the mean and variance of beliefs G_{ft} on each block, interactions with the oil price, counts of nearby own- and other-firm licenses for each block, and the number of nearby exploration wells with unobserved outcomes. Denote these lower dimensional statistics of the state variable s_{ft}.

This approach involves both dimension reduction - projecting value functions onto lower dimensional statistics of the state variable - and approximation - estimating the value func-

[24]More precisely, and element of $\tilde{\mathcal{S}}_f$ is a particular sequence of firm-f states $\mathcal{S}_f = \{\mathcal{S}_{ft}\}_{t=1}^T$. See Appendix B for a formal definition of $\tilde{\mathcal{S}}_f$.

tion using a flexible polynomial. The use of flexible functional forms to estimate value functions is common in the applied literature that estimates dynamic discrete choice models with conditional choice probability methods.[25] Dimension reduction requires the additional assumption that firms make decisions and form beliefs based on lower dimensional statistics s_{ft}, and that value functions are averaged over the equilibrium distribution of the full state variable S_f conditional on s_f. That is, firms' choice-specific value functions take the form

$$v_f^E(j, s_{ft}) = E\left[v_f^E(j, \mathcal{S}_{ft})|s_{ft}\right]. \tag{19}$$

This approach is similar to the dimension reduction techniques introduced by Ifrach and Weintraub (2017) and Gowrisankaran et al. (2024). Full details are provided in Appendix B.

Estimation of Belief Functions The estimated choice probabilities, prior distribution of oil, $G_0(\rho)$, and spillover parameter α can be used to simulate drilling histories for the entire market. Under the assumption that CCPs reflect equilibrium policy functions, the data generated by these simulations is drawn from the equilibrium distribution of states. These simulations can be used to estimate empirical analogs $\hat{Q}^E(\mathcal{S}_{ft}, g, j)$ and $\hat{Q}^D(\mathcal{S}_{ft}, g, j)$ of equation 16, since in the simulated data each firm's state is fully observed. Similarly, since the true productivity of each block, ρ_j, is observed in the simulations, it is possible to estimate adjusted beliefs about the oil distribution, $\hat{Q}^G(\rho; G(\rho), W_{ft}^u)$, that are consistent with the equilibrium distribution of states.

I parameterize $\hat{Q}^E(\mathcal{S}_{ft}, g, j)$ using a linear equation in s_{ft}, similar to that used to estimate the CCPs. I run a logit regression of other firms' exploration choices on firm f's state, pooling data across all firms. $\hat{Q}^D(\mathcal{S}_{ft}, g, j)$, and $\hat{Q}^{Past}(\mathcal{S}_{ft}, g, j)$ are estimated analogously. I also use this simulated data to estimate $\hat{Q}^G(\rho; G(\rho), W_{ft}^u)$ by regressing the true exploration well success, ρ_j, on the mean and variance of $G(\rho_j)$ and fixed effects for the number of nearby unobserved wells in W_{ft}^u. Details on these procedures are provided in Appendix C.2.

Firm f's estimated beliefs about the actions of other firms are therefore set equal to expected actions under equilibrium play, conditional on state s_{ft}. This approach to estimating equilibrium beliefs extends the logic of the CCP approach of Hotz and Miller (1993). If equilibrium policy functions can be recovered directly from the data, then equilibrium beliefs can be recovered by simulating data using these policy functions. This approach does not require solving for optimal policy functions or equilibrium beliefs using value function iteration and is an alternative to the reinforcement learning algorithm of Fershtman and Pakes (2012).

[25]For example, see Ryan and Tucker (2011), Collard-Wexler (2013), and Agarwal et al. (2020). See Aguirregabiria et al. (2021) for a discussion.

Norwegian Drilling The region covered by the data is adjacent to the Norwegian economic zone where several significant oil discoveries were made between 1970 and 1990. To account for the fact the firms operating in the UK sector might be able to learn from the results of Norwegian drilling, I allow firms' beliefs $G(\rho)$ to depend on the results of exploration wells drilled in the Norwegian zone. In particular, I estimate an average block-month-level drilling rate in Norway using analogous data to the main UK sample, and assume that blocks in Norway are drilled at this exogenous rate. Firms observe the results of Norwegian wells immediately with probability $\alpha + (1-\alpha)\alpha^n$, where α^n is the empirical probability that a Norwegian well is drilled by a firm with a presence in the UK. All Norwegian wells are made public after 2 years of confidentiality.

6.3 Identification

Identification of CCPs The first step of the estimation procedure recovers the parameter α and conditional choice probabilities $\hat{P}(a=j|\mathcal{S})$ at each state $\mathcal{S}$ from data in which each observation is consistent with a set of states $\tilde{\mathcal{S}}$. The model's information structure means these objects are separately identified despite the fact that the econometrician does not observe the full state. In particular, I claim that the list of choice probabilities $P(a=j|\tilde{\mathcal{S}})$ for each set of states $\tilde{\mathcal{S}}$ that it is possible to *observe* in the data can be inverted to uniquely identify choice probabilities conditioned on the *unobserved* states $P(a=j|\mathcal{S})$ *and* the information spillover parameter α.

To illustrate identification, consider the following simplified example. Suppose that a state is described by a triple, $\mathcal{S} = (suc, fail, unobs)$, where suc is the number of successful wells observed, $fail$ is the number of unsuccessful wells observed, and $unobs$ is the number of wells with unobserved outcomes. Consider data that contains observations consistent with the following sets of states:

$$\tilde{\mathcal{S}_A} = \{(1,0,0)\},\ \tilde{\mathcal{S}_B} = \{(0,1,0)\},\ \tilde{\mathcal{S}_C} = \{(1,0,0),(0,0,1)\},\ \tilde{\mathcal{S}_D} = \{(0,1,0),(0,0,1)\} \quad (20)$$

$\tilde{\mathcal{S}_A}$ and $\tilde{\mathcal{S}_B}$ are observed by the econometrician when there is one own-firm well outcome. The econometrician then knows the state with certainty since the firm always observes their own well outcome. $\tilde{\mathcal{S}_C}$ and $\tilde{\mathcal{S}_D}$ are observed by the econometrician when there is one other-firm well outcome. In this case, the econometrician knows whether the well was successful or unsuccessful, but not whether the firm observed the outcome or not. Given a value of the parameter α, choice probabilities conditional on the observed set of states can be written as:

$$\begin{aligned} P(a=j|\tilde{\mathcal{S}_A}) &= P(a=j|\mathcal{S}=(1,0,0)) \\ P(a=j|\tilde{\mathcal{S}_B}) &= P(a=j|\mathcal{S}=(0,1,0)) \\ P(a=j|\tilde{\mathcal{S}_C}) &= \alpha P(a=j|\mathcal{S}=(1,0,0)) + (1-\alpha)P(a=j|\mathcal{S}=(0,0,1)) \\ P(a=j|\tilde{\mathcal{S}_D}) &= \alpha P(a=j|\mathcal{S}=(0,1,0)) + (1-\alpha)P(a=j|\mathcal{S}=(0,0,1)). \end{aligned} \tag{21}$$

The left hand side of each equation is a probability that is observable in the data. Notice that there are four equations and four unknowns - three conditional choice probabilities and the parameter α. The first two equations yield estimates of $P(a = j|\mathcal{S} = (1, 0, 0))$ and $P(a = j|\mathcal{S} = (0, 1, 0))$ directly. Rearranging the third and fourth equations yields:

$$\alpha = \frac{P(a=j|\tilde{\mathcal{S}_C}) - P(a=j|\tilde{\mathcal{S}_D})}{P(a=j|\tilde{\mathcal{S}_A}) - P(a=j|\tilde{\mathcal{S}_B})}. \tag{22}$$

This says that α is identified by the difference between how much the firm responds to other firm wells (the numerator) and how much the firm responds to its own wells (the denominator). As documented in Figure 3, firms' exploration choices respond more to the results of their own wells than to those of other firm wells, implying $0 < \alpha < 1$.[26] $P(a = j|\mathcal{S} = (0, 0, 1))$ is then identified by the level of $P(a = j|\tilde{\mathcal{S}_C})$ or $P(a = j|\tilde{\mathcal{S}_D})$.

This identification argument relies on two features of the model's information structure. First, the belief updating rule (4) treats own-firm and other-firm well results identically. This means that we can use the firm's response to their own wells to infer how they would have responded *if* they had observed another firm's well. Second, if firm f does not observe the outcome $s(w)$ of well w at date t, then the $s(w)$ does not enter $\mathcal{S}_{ft}$. This means that if a well was not observed, then the firm's actions should not depend on the well's outcome. Relaxing either assumption would break identification by introducing an extra free parameter.

This argument extends to states with multiple well results and well results at different distances and dates. In Appendix D I provide a proof that shows, in general, how $\hat{P}(a = j|\mathcal{S})$ can be identified from observable quantities for any $\mathcal{S}$.

This approach to identification exploits a different source of variation than existing literature on the identification of unobserved heterogeneity in dynamic discrete choice (Berry and Compiani, 2021). Kasahara and Shimotsu (2008) show that mixture models with per-

[26] This argument could be extended to identifying firm pair specific spillover parameters, α_{fg}. For instance, if firm f responds more to firm g's well results than to firm h's well results, this would suggest $\alpha_{fg} > \alpha_{fh}$. One might allow these terms to depend on covariates, such as the overlap in equity holders between blocks operated by f and g. I opt not to estimate these firm-specific parameters because this would increase the computational cost of estimation considerably and my counterfactuals focus on aggregate outcomes.

sistent firm types are identified by time series variation within each firm. Estimation of a model with this type of persistent unobservable heterogeneity can be achieved using the iterative expectation-maximization procedure proposed by Arcidiacono and Miller (2011). In contrast, my approach to identification relies on cross-sectional variation in the econometrician's information about the firm's state and the mixture distribution over possible information states, and estimation of the CCPs and mixture parameter α is performed in a single step.[27]

Identification of Cost Parameters The cost parameters are estimated in the second step of estimation, which minimizes the distance between the first step CCPs and model-predicted choice probabilities, detailed in Appendix C. Intuitively, cost parameters c_0, κ_0, and κ_1 are identified by the average probability of exploration and development. Additional identifying variation comes from the difference in the response of drilling probability to nearby own-firm and other-firm licenses. Higher exploration drilling costs, c_0, imply that firms have more of an incentive to free ride and should have a lower exploration probability when the surrounding blocks are owned by other firms than when they are owned by the same firm. The variance parameters σ_ϵ and σ_ν are identified by the extent to which firms are more likely to explore blocks for which the expected future revenue stream conditional on exploration is higher.[28]

7 Results

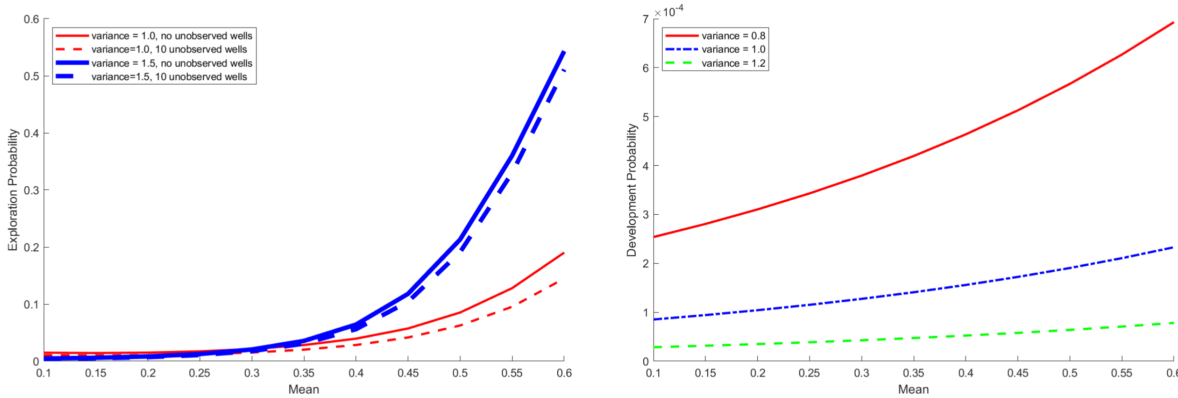
7.1 Estimates

Appendix Table A2 reports coefficients from the estimated conditional choice probabilities (CCPs) $\hat{P}(a^E = j|\mathcal{S}_{ft})$ and $\hat{P}(a^D = j|\mathcal{S}_{ft})$. Figure 4 illustrates the predicted probability of exploration and development for different values of the state variables. The left panel shows that the probability of exploration is greatest on blocks with high posterior mean and high posterior variance (Appendix Figure A9 shows this pattern for other variance levels). Exploration probability is also lower when there are more unobserved wells nearby, consistent with the free riding incentive. The right panel shows that development probability is increasing in posterior mean and decreasing in posterior variance.

[27] This suggests that it may be possible to identity and estimate a model with both unobserved asymmetric information and persistent unobserved firm types by combining these two sources of variation.

[28] As discussed by Bajari, Benkard, and Levin (2007), the two step procedure obtains consistent estimates of the model parameters if the data is generated by a single equilibrium. I assume this here since I cannot guarantee that there is a unique equilibrium of the asymmetric information game.

Figure 4: Estimated CCPs



Notes: Simulated exploration and development probabilities using the estimated CCPs. Probabilities are simulated for a firm with a single block and one neighboring block held by another firm. Oil price is set to the average value in the data.

Table 4 reports estimated model parameters. The spillover parameter α which is estimated simultaneously with the CCPs indicates that firms behave as if they observe the results of 64% of other firm wells before they are made public. This finding is in line with the descriptive results reported in Figure 3, which suggest imperfect information spillover across firms. A likelihood ratio test comparing the estimated first stage CCPs to those estimated under the restriction $\alpha = 1$ strongly rejects the full information model, suggesting asymmetry in observability of well outcomes between own-firm and other-firm wells.[29]

Cost parameters are reported in billions of 2015 dollars. As expected, exploration costs are lower after the first exploration well on a block, and increasing in the variance of firms' beliefs. Similarly, development costs are increasing in the variance of firms' beliefs. The last row of Table 4 records average realized costs including the cost shocks ϵ and ν computed from simulations. The average marginal cost of an exploration well is \$36 million. The average marginal cost of development is \$870 million. These average realized costs are close to estimates of the capital costs of exploration and development from auxiliary data on capital expenditure provided by the regulator that was not used in estimation. The average capital expenditure per exploration well from this auxiliary data is \$20.9 million. The capital expenditure per development platform is \$1.26 billion.

[29]As discussed in Section 6 above, α is identified if firms interpret information from same- and other-firm wells identically and firms either observe well outcomes perfectly or not at all. An alternative model might have firms observe noisy signals of the results of other firms' wells before the confidentiality window ends. Under such a model, the signal to noise ratio would be identified in a similar way to α. The same externalities are present in both models, and the counterfactual policy analysis presented in Section 8 would likely be similar under this alternative model. A model with noisy signals and some probability of non-observation (i.e. $\alpha < 1$) would not be identified.

Table 4: Parameter Estimates

Parameter	Estimate	SE	Parameter	Estimate	SE
α	0.64	0.017	Oil Price Process		
Exploration Cost	0.141	0.054	Intercept	0.302	0.208
σ_c	0.036	0.017	P_{t-1}	0.988	0.008
Development Cost			Variance of Shock	3.323	1.147
High Variance Locations	2.022	0.651			
Low Variance Locations	1.043	0.4148			
σ_κ	0.252	0.111			
Average Realized Costs					
$E(c(j, \mathcal{S}_{ft})\|a_t^E = j)$	0.036		$E(\kappa(j, \mathcal{S}_{ft})\|a_t^E = j)$	0.870	
Comparison to $\alpha = 1$ Model					
λ_{LR}	53.054		$C_{0.95}$	3.841	

Notes: Cost parameters are in billions of 2015 dollars. Standard error of α is computed using the inverse Hessian of the likelihood function given by equation 18 at the estimated parameter values. Standard errors for the remaining (cost) parameters are computed using 200 bootstrap draws from the first step CCP estimates. Average realized costs are computed by simulating exploration and development behavior using the estimated first step CCPs and taking the average realized costs, including the cost shocks, ϵ and ν, over 40 simulations. The LR test is a comparison of the 1st step likelihoods (equation 29) under the restricted and baseline models. The test statistic λ_{LR} has a χ_1^2 distribution with 95% critical value $C_{0.95}$.

The estimates of the price process imply a stationary mean price of \$26.16. The variance of monthly cost shocks is \$3.32. Combined with a high persistence parameter, this process generates large swings in oil prices, consistent with the trajectory of realized prices in this period (see Appendix Figure A10). Appendix Table A8 shows a comparison of equilibrium simulations to the data. The simulated number of exploration wells is a close match to the data. The simulations predict fewer development wells than in the data, likely due to estimation error from the small number of development observations and the restrictive functional form used for the first step development CCPs.

7.2 Quantifying the Effects of Information Spillovers

To illustrate how information spillovers affect the equilibrium speed and efficiency of exploration, I simulate counterfactual exploration and development decisions. I separately quantify the effects of free riding and wasteful exploration on the equilibrium rates of exploration and development and on industry surplus by removing these sources of inefficiency from the model.

First, I remove the free riding incentive by computing firm's optimal policy functions assuming that firms believe $Q^E = 0$ and that wells remain confidential forever. That is, I ask how firms would behave if, at each period, they believed that no new exploration wells would be drilled by other firms at any period in the future and no past wells with unobserved outcomes would ever be revealed. Under this assumption there is no incentive to strategically delay exploration.[30] This counterfactual is not an equilibrium since firms beliefs about the average exploration probability are inconsistent with the actual probability of exploration. Simulation of firm behavior under these non-equilibrium beliefs *isolates* the direct effect of free riding on firm behavior since I allow firms to learn the results of past wells as in the baseline, but I remove the forward-looking incentive to delay.

The effect of eliminating the incentive to free ride on industry outcomes is illustrated by comparing the simulation results in the first and second columns of Panel A in Table 5. Removing the free riding incentive brings exploration and development forward in time. The average number of exploration wells drilled up to 1990 increases by 7.7% and the number of blocks developed increases by almost 70%. The large increase in development is both due to faster learning from exploration and less strategic delay of development. The fourth and fifth rows record the 1964 present total surplus, computed as oil revenue less exploration and production costs. Eliminating free riding increases total surplus by $22.11 billion, or about 52%, and discounted revenue by about 73% by increasing the number of developed blocks and bringing development forward in time.

These effects are also illustrated by comparing the solid and dotted lines in the left panel of Figure 5, which records the average number of exploration wells and blocks developed each month from 1975 to 1990. Removing the free riding incentive shifts exploration back in time by around 1.5 years.

The second exercise removes wasteful exploration due to imperfect information spillovers. I simulate the model at the baseline equilibrium choice probabilities but allow firms to observe the results of each other's wells with certainty. That is, I set $\alpha = 1$. I hold firms' choice probabilities fixed at the baseline level. This means that firms behave *as if* they expect the results of other firms' wells to be revealed with probability equal to the estimated value of α. This isolates the direct effect of increased flow of information from the equilibrium effects on firms' drilling decisions.

The third column of Table 5 records drilling, revenue, and profit statistics for this information sharing simulation. Allowing for perfect information flow without changing firms' policy functions increases the number of exploration wells drilled before 1990 by 1.3% relative to the baseline, but increases the number of blocks developed by 16.4%. The efficiency of exploration improves substantially, reflecting a reduction in duplicative exploration - the

[30] One only additional effect of setting $Q_E = 0$ is that it changes firm's expectations about future development (since if other firms do not explore they are unlikely to develop), and therefore also changes expectations about the future distribution of licenses. Because development is infrequent, so any change in the distribution of future licenses due to this effect is likely to be small.

Table 5: Outcomes of Counterfactual Simulations

	Baseline	No Free Riding	Info Sharing	No Info Sharing	Info Sharing (Equilibrium)	No Price Shocks	Monopoly	Clustered
Exp. Wells	1554.35	1674.65	1574.45	1539.85	1556.60	1553.95	2202.30	1553.95
	[6.46]	[14.68]	[7.98]	[6.92]	[9.65]	[6.50]	[10.17]	[6.50]
Blocks Dev.	33.70	56.15	39.25	26.50	36.85	38.55	89.40	38.55
	[1.19]	[2.27]	[1.99]	[1.69]	[1.74]	[2.09]	[2.89]	[2.09]
Exp. Wells/Dev.	47.42	30.81	42.12	63.89	44.73	42.94	25.26	42.94
	[1.91]	[1.28]	[2.06]	[4.95]	[2.82]	[2.57]	[0.99]	[2.57]
Total Surplus								
Discounted	5.35	9.27	6.96	3.23	4.72	5.57	13.02	5.57
	[0.37]	[0.77]	[0.59]	[0.35]	[0.44]	[0.55]	[0.86]	[0.55]
Not Discounted	42.41	64.52	52.27	30.16	39.79	44.87	81.56	44.87
	[2.44]	[4.66]	[3.75]	[2.60]	[2.90]	[3.86]	[4.70]	[3.86]

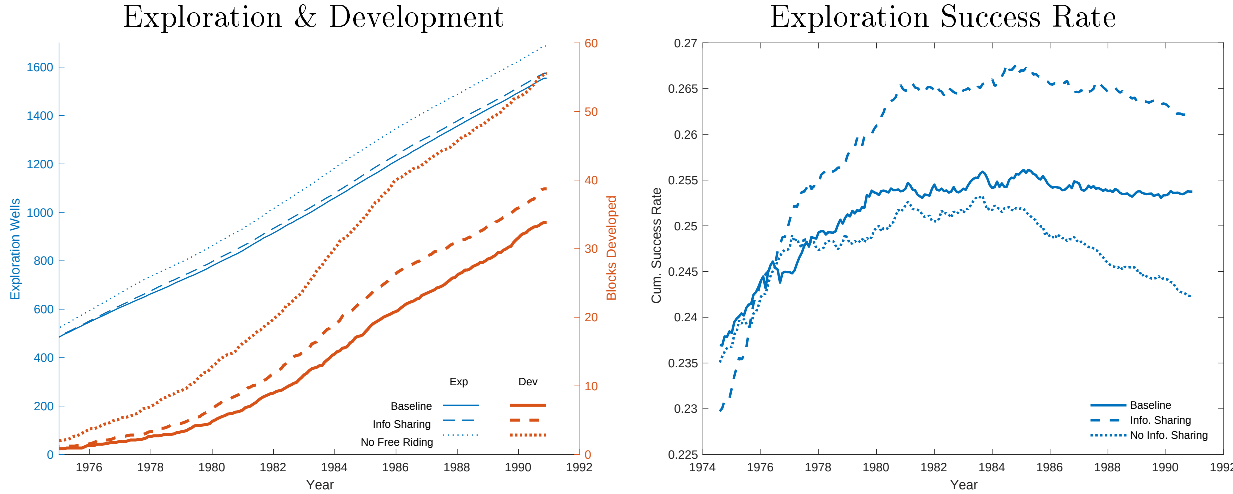
Notes: Results are averages over 40 simulations that cover 1964-1990. The assignment of blocks to firms and the oil price are set at their realized values. Well outcomes and development revenue are drawn from the posterior of the estimated Gaussian process using all observed wells. Revenue and profits are in billions of 2015 dollars. Total surplus is firm and government revenue less costs, including the value of cost shocks for every period. PDV revenue and profit are 1964 values where the annual discount factor is 0.9. Simulation standard errors are in square brackets.

number of exploration wells drilled per block developed is reduced to 42.12 from 47.42 in the baseline.

Perfect information flow increases discounted industry surplus by 30%. This effect is about 40% of the the effect of removing free riding. This change in industry surplus is the result of two effects. First, perfect information flow increases industry surplus by reducing wasteful exploration of unproductive areas and per-development costs, thereby reducing expenditure on exploration wells. Second, increased information flow allows firms to identify productive areas faster, bringing development forward in time. The dashed lines in the left panel of Figure 5 show that perfect information flow brings development forward in time modestly but substantially increases the speed of development, reflecting the increased efficiency of exploration.

The improved efficiency of exploration from information sharing can also be observed in the right panel of Figure 5, which records the cumulative exploration success rate between 1975 and 1990 in the baseline simulation, the information sharing simulation, and a no information sharing counterfactual in which well information is never shared and the firms follow their baseline equilibrium policy functions. In all three simulations, the success rate increases over time until around 1983, after which it starts falling. This is a result of firms identifying productive blocks and then those productive blocks becoming developed, reducing

Figure 5: Exploration Paths



Notes: The left panel plots the cumulative number of exploration wells drilled and blocks developed for each month from 1975 to 1990 for three simulations. Thick red lines plot the number of blocks developed and correspond to the right axis. Thin blue lines plot the number of exploration wells and correspond to the left axis. The solid lines are the average of 40 simulations using the baseline equilibrium choice probabilities. The dashed lines are the average of 40 simulations under the information sharing counterfactual. The dotted lines are the average of 40 simulations under the no free riding counterfactual. The right panel plots the cumulative success rate of exploration wells under the baseline,information sharing, and no information sharing counterfactuals.

the success rate of the remaining undeveloped blocks. The success rate under information sharing is around 1 percentage point higher than the baseline success rate at its maximum.[31]

The fifth column of of Table 5 records the results of an information sharing counterfactual that uses equilibrium choice probabilities and belief functions. Comparing the third and fifth columns illustrates how the surplus gains from information sharing are diminished by increased free riding. When information is shared perfectly, firms have more incentive to delay exploration in equilibrium. This countervailing effect reduces discounted surplus by around 12.8%.

The sixth and seventh columns of Table 5 record outcomes from benchmark simulations. In the sixth column, I report the results of a simulation in which firms believe that prices will remain constant at all periods in the future. This removes the incentive to delay drilling due to the option value of waiting for higher prices, which is potentially important in this era because of the volatility of prices (for instance, see Kellogg (2014)). Removing strategic delay due to oil price volatility brings development forward in time and increases discounted

[31]Statistics from the no information sharing counterfactual are also recorded in the fourth column of Table 5, which indicate that stopping all information flow between firms reduces the efficiency of exploration substantially, increasing the number of exploration wells per developed block by around 20% relative to the baseline.

surplus. Despite the high volatility of prices, the surplus loss from delays due to oil price volatility are significantly smaller than the surplus losses from free riding or imperfect information flow.[32] The seventh column reports outcomes from a collusive simulation in which information is shared between firms and firms optimize total industry profit, approximating a monopoly's (or social planner's) problem. I use a collusive scenario rather than a monopoly to facilitate comparison with the other simulations: each firm's drilling capacity and block assignment are as in the baseline simulation. Collusion allows firms to fully internalize the information externalities generated by information sharing, increasing discounted profit by $7.6 billion from the baseline. The simulations which eliminating free riding and allow perfect information sharing achieve 50% and 21% respectively of this total potential increase in surplus.[33]

8 Counterfactual Property Rights Policy

The gains in industry surplus from removing free riding and increasing information flow are substantial, but are not necessarily achievable in equilibrium. In this section I ask how much of industry surplus could be increased in equilibrium through alternative design of property rights that minimize the inefficiencies resulting from information spillovers.

Confidentiality Window Well outcome data is property of the firm that drilled the well until the confidentiality deadline, after which it becomes public knowledge. By changing the confidentiality deadline, the government can increase or decrease the speed with which information flows between firms and manipulate firms' incentive to delay exploration. Changing the length of the well data confidentiality period has two potential effects on firms' equilibrium drilling behavior. First, increasing the confidentiality period decreases the incentive to free ride. If the release of well data is pushed further into the future, then the cost of delaying exploration is increased due to the discounting of future profits, and the equilibrium

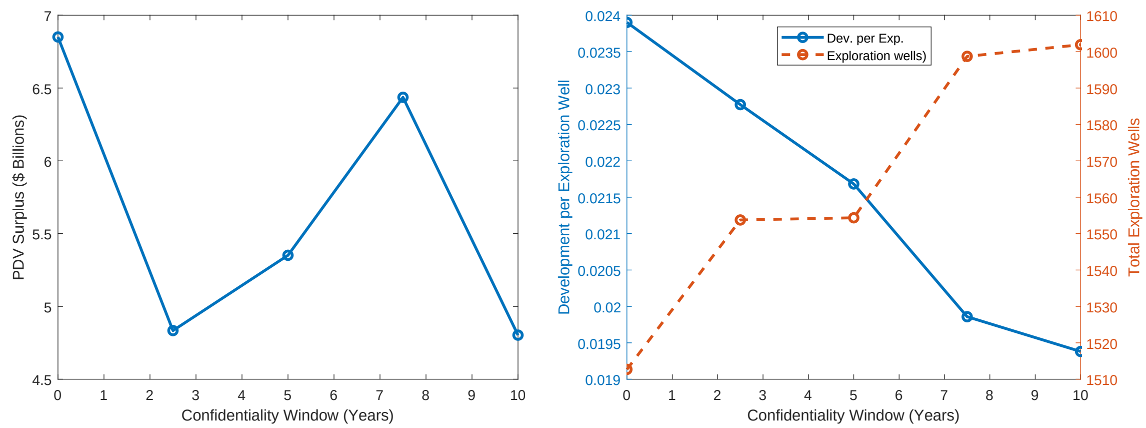
[32] To examine the sensitivity of the results in Table 5 to the particular path of oil prices realized in reality, Appendix Table A9 records the results of simulations where prices are held fixed at the long run average of the estimated AR process. The qualitative patterns across columns are similar, but the level of surplus is lower because prices do not reach the high level realized in the early 1980s.

[33] The large gains from information sharing raise the question of why firms do not engage in more exchange of information before the confidentiality windows expires (per the Coase theorem). The empirical evidence indicates that this efficient exchange of information does not take place in reality, and anecdotal evidence (Moreton, 1995) describes a culture of secrecy around exploration outcomes. There are several potential sources of transaction costs that might limit efficient trade. First, sharing well data is not costless to the firm because it may be valuable in future competitive license applications. Second, firms have asymmetric information about the value of additional well data, preventing efficient trade (Myerson and Satterthwaite, 1983; Farrell, 1987; Bessen, 2004).There is an additional set of barriers to efficient trade when the object being traded *is information* (Anton and Yao, 2002, Ali, Chen-Zion, and Lillethun, 2017).

probability of exploratory drilling should increase. On the other hand, lengthening the confidentiality window will reduce the efficiency of exploration by increasing wasteful drilling. The regulatory problem of setting the optimal confidentiality window is therefore a case of trading off these two effects.

To determine the effect of changing the confidentiality window on surplus, I run counterfactual simulations of the model under different window lengths. For each window length, I first compute the equilibrium choice probabilities implied by the estimated model parameters. I then simulate the model using these choice probabilities, imposing the relevant confidentiality window lengths. The left panel of Figure 6 records the average present discounted revenue under confidentiality windows of 0, 2.5, 5 (the baseline), 7.5, and 10 years.

Figure 6: Confidentiality Window



Notes: The left panel records the 1964 present discounted value of 1964-1990 industry surplus in counterfactual simulations with different confidentiality window lengths. In the right panel, the blue line, corresponding to the left y-axis, records the average number of developed blocks per exploration well evaluated using baseline choice probabilities under different confidentiality window lengths. The dashed red line, corresponding to the right y-axis, records the average number of exploration wells evaluated at counterfactual equilibrium choice probabilities under baseline information flow. Surplus is in billions of 2015 dollars. All figures are average over 40 simulations.

Industry surplus would be increased both by shortening the confidentiality window to 0, or increasing the confidentiality window to 7.5 years. Among the policies studied reducing the confidentiality window to 0 and enforcing full information achieves the maximum surplus of \$6.85 billion, 28% higher than the baseline policy of 5 years.

The non-monotonicity of surplus in window length results from the interaction of the effect of limiting information flow on the free riding incentive and the effect on the speed of learning and efficiency of exploration. The right panel of Figure 6 illustrates these two effects separately. The dashed red line records the average number of exploration wells in simulations that hold information flow fixed but use counterfactual choice probabilities, and the blue line records exploration efficiency from simulations that hold choice probabilities fixed but vary information flow.

As expected, the rate of exploration decreases and the efficiency of exploration increases with shorter window lengths. For instance reducing the confidentiality window from 5 to 0 years reduces the number of exploration wells by about 2.5%. Countervailing this free riding effect, exploration efficiency increases by 11.6%.

Increasing the confidentiality window to 7.5 year leads to a significant reduction in free riding and an increase in industry surplus. However, for further increases in window length the loss in efficiency dominates, resulting in a local optimum around 7.5 years. The marginal effect of increasing the confidentiality window on the exploration rate diminishes with window length, and the size of the free riding effect is small relative to the full elimination of free riding recorded in Table 5. This is explained by the fact that $\alpha = 0.64$, and there is a significant degree of information sharing that does not respond to changes in the confidentiality window.[34]

Spatial Arrangement of Licenses In addition to manipulating the flow of information between firms, the regulator can change the spatial arrangement of property rights. If, as suggested by the results in Table 5, the potential to learn from the results of other firms' wells reduces the exploration rate in equilibrium, then the regulator should take this effect into account when assigning blocks to firms. In particular, spatial arrangements of property rights in which each firm's blocks are clustered together should minimize the free riding problem and improve the speed at which each firm learns about its blocks.

To quantify the effect of spatial reallocation of licenses, I construct an alternative license allocation for each month in the data using an algorithm that maximizes the spatial clustering of firms' licenses. The new assignment holds fixed the number of blocks assigned to each firm in each year. The drilling capacity of the industry (one well per firm per month in the model) is therefore held fixed relative to the baseline, and only the location of each firm's licenses changes. Details of the license clustering algorithm are available on request. Appendix Figure A11 illustrates the true and counterfactual license assignments in January 1975, note that each block is more likely to neighbor blocks held by the same firm in the clustered allocation. In particular, the average number of own-firm neighbors increases from 1.64 in the baseline assignment to 2.01 in the clustered assignment.

The final column of Table 5 records the results of the clustered simulation. Clustering

[34] Kemp's (2012a) account of the process by which the regulations were designed indicates that the 5-year window was arrived at through negotiations between the government and the major oil companies, who were resistant to any regulation that diminished their property rights over well data. Although not optimal, the settlement the parties arrived at improved industry surplus over full confidentiality.

firms' licenses increases the number of blocks developed by 14%. The discounted value of profit increases by 4% and the efficiency of exploration improves, with the number of exploration wells per developed block falling from 47.42 to 42.94.

The results suggest that the government could substantially increase revenue and industry surplus through a simple rearrangement of the spatial allocation of blocks to firms. Note that there is no sense in which this particular allocation is optimal, and it may be that other allocations would result in faster learning and a higher surplus. These results therefore provide a lower bound on the potential gain from spatial reassignment of licenses.[35]

9 Conclusion

In many industries the creation of new knowledge through R&D is carried out in a decentralized manner by competing firms. The growth of the industry-wide stock of knowledge depends on the extent to which firms can observe and build on each other's innovations. Allowing information spillovers between firms can improve the speed of cumulative research and reduce duplicative or socially inefficient investments. On the other hand, information spillovers can diminish firms' individual incentives to innovate by enabling free riding on the innovations of other firms. The design of property rights over innovations plays an important role in balancing these effects.

I study the effects of information spillovers on R&D in the context of oil exploration, using historical data from the UK North Sea. Oil exploration by individual firms can be thought of as a process of cumulative learning about the location of oil deposits. Exploration wells are experiments located in geographical space with observable outcomes. If firms can learn from the results of other firms' wells they face an incentive to delay exploration. However, if other firms' well outcomes are unobserved, firms are likely to make inefficient drilling decisions, for example exploring regions that are known by other firms to be unproductive.

To quantify the effects of information spillovers, I build and estimate a model of the firm's dynamic exploration problem with spatial learning and information spillovers across firms. The estimated model indicates that there is imperfect information flow between firms. In counterfactual simulations, I show that removing the incentive to free ride brings

[35]As with the confidentiality window, it is worth asking why the actual allocation of licenses to firms does not appear to fully take into account information externalities. One reason that firms may not apply for a large number of licenses close together is that this type of clustered allocation increases the risk borne by each individual firm. Because of the spatial correlation of oil deposits, a risk averse firm with a constant prior mean would prefer to be allocated licenses that are spread over a wide area. One alternative policy that could achieve some of the gain from license clustering would be to require firms to apply for licenses at a regional rather than block level, with the government determining the exact allocation of blocks to firms within the region.

exploration and development forward in time, increasing industry surplus in the same time period by 52%. Holding the free riding incentive fixed and allowing perfect information flow between firms increases surplus by 23% by increasing the speed of learning, increasing the cost efficiency of exploration by reducing the number of development wells drilled per developed block, and increasing the concentration of development on productive blocks.

Equilibrium simulations under counterfactual property rights policies highlight the trade-off between free riding and efficient cumulative research. Strengthening property rights by extending the well data confidentiality period increases industry surplus by increasing the rate of exploration, while weakening property rights by limiting the confidentiality period increases industry surplus by increasing the speed of learning and efficiency of exploration. Over the range of policies I examine, reducing the confidentiality window to 0 achieves the highest industry surplus.

Notice that the gains from strengthening property rights here are due to the effect of limiting inter-firm information flow on the incentive to free ride on other firms' discoveries. This differs from the more commonly discussed motive of allowing firms to capture the surplus from their innovations.There is a substantial body of recent work quantifying the extent to which property rights limit follow-on research in a number of settings (Murray and Stern, 2007; Williams, 2013; Murray et al., 2016), but little empirical work on the potential for weaker property rights to encourage free riding. The policy results in this paper suggest that the question of the optimal generosity of property rights is subtle, even in the absence of an effect of stronger property rights on firms' ability to extract rent from their discoveries. In some settings it may be optimal to strengthen property rights to reduce the free riding incentive even though stronger property rights hinder cumulative research.

Methodologically, this paper makes two contributions that are applicable to other settings. First, the model of beliefs and learning can be used to study other industries where research takes place in a well defined space. For example, measures of molecular similarity are important metrics in the exploratory phase of pharmaceutical development, and measures of the distance between molecular structures are increasingly used in the economics literature on pharmaceutical R&D (Krieger, Li, and Papanikolau, 2017; Cunningham, Ederer, and Ma, 2018). An application of this model to research in chemical space might be able to inform the design of property rights, for example the disclosure of clinical trial results, in that industry. Second, the estimation approach developed in this paper is potentially applicable to other settings in which agents have asymmetric information and the econometrician is not fully informed about each agent's information set.

References

Agerton, M. 2018. "Learning Where to Drill: Drilling Decisions and Geological Quality in the Haynesville Shale." *Working Paper.*Agarwal, N. and P. Somaini. 2018. "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism." *Econometrica.* 86 (2) 391-444.

Ali, S. N., A. Chen-Zion, and E. Lillethun. 2017. "Reselling Information." *Working Paper.*

Anton, J. J., and D. A. Yao. 2002. "The Sale of Ideas: Strategic Disclosure, Property Rights, and Contracting." *Review of Economic Studies.* 69, 513-531.

Arcidiacono, P., and R. A. Miller. 2011. "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity." *Econometrica.* 79 (6), 1823-1867.

Arrow, K. J. 1971. "Economic Welfare and the Allocation of Resources for Invention." in *Essays in the Theory of Risk Bearing.* Chicago: Markham.

Bajari, P., C. L. Benkard, and J. Levin. 2007. "Estimating Dynamic Models of Imperfect Competition." *Econometrica.* 75 (5), 1331-1370.

Berry, S. and G. Compiani. 2021. "Empirical Models of Industry Dynamics with Endogenous Market Structure." *Annual Review of Economics.* 13 11.1-11.26.

Bessen, J. 2004. "Holdup and Licensing of Cumulative Innovations with Private Information." *Economics Letters.* 82 (3), 321-326.

Bickel, J. E., and J. E. Smith. 2006. "Optimal Sequential Exploration: A Binary Learning Model." *Decision Analysis.* 3 (1), 1-59.

Bolton, P., and C. Harris. 1999. "Strategic Experimentation." *Econometrica.* 67 (2), 349-174.

Bolton, P., and J. Farrell. 1990. Decentralization, Duplication, and Delay." *Journal of Political Economy.* 98 (4), 803-826.

Brennand, T. P., B. Van Hoorn, K. H. James, and K. W. Glennie. 1998. "Historical Review of North Sea Exploration." in *Petroleum Geology of the North Sea.* Oxford: Blackwell.

Collard-Wexler, A. 2013. "Demand Fluctuations in the Ready-Mix Concrete Industry." *Econometrica.* 81 (3), 1003-1037.

Covert, T. R. 2015. "Experimental and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale." *Working Paper.*

Cressie, N. 1991. "Statistics for Spatial Data." New York: Wiley.

Cunningham, C., F. Ederer, and S. Ma. 2018. "Killer Acquisitions." *Working Paper.*

Dasgupta, P., and J. Stiglitz. 1977. "Industrial Structure and the Nature of Innovative Activity." *Economic Journal.* 90 (358), 266-293.

Diggle, P. J., J. A. Tawn, and R. A. Moyeed. 1998. "Model-based Geostatistics." *Applied Statistics.* 47 (3) 299-350.

Dong, M. 2018. "Strategic Experimentation with Asymmetric Information." *Working Paper.*

Doraszelski, U, Lewis, G., and A. Pakes. 2018. "Just Starting Out: Learning and Equilibrium in a New Market." American Economic Review. 108 (3) 565-615.

Farrell, J. 1987. "Information and the Coase Theorem." *Journal of Economic Perspectives.* 1 (2) 113-129.

Fershtman., C. and A. Pakes. 2012. "Dynamic Games with Asymmetric Information: A Framework for Empirical Work." *Quarterly Journal of Economics* 127 (4) 1611-1661.

Gordon, G. W. 2015. "Production Licensing on UK Continental Shelf: Ministerial Powers and Controls." *LSU Journal of Energy Law and Resources.* 4 (1), 75-95.

Haile, P., K. Hendricks, and R. H. Porter. 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 D. Kovenock. 1989. "Asymmetric Information, Information Externalities, and Efficiency: The Case of Oil Exploration." *RAND Journal of Economics.* 20 (2), 164-182.

Hendricks, K. and R. H. Porter. 1996. "The Timing and Incidence of Exploratory Drilling on Offshore Wildcat Tracts." *American Economic Review.* 86 (3), 388-407.

Hodgson, C. 2024. "Information Externalities in Oil Exploration - Theoretical Note." *Working Paper*

Hohn, M. E. 1999. "Geostatistics and Petroleum Geology." Dordrecht: Kluwer Academic Publishers.

Hotz, V. J., and R. A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The Review of Economic Studies.* 60 (3), 497-529.

Hotz, V. J., R. A. Miller, S. Sanders, and J. Smith. 1994. "A Simulation Estimator for Dynamic Models of Discrete Choice." *The Review of Economic Studies*. 61 (2) 265-289.

Hunter, T. 2015. "Regulation of the Upstream Petroleum Sector." Cheltenham: Edward Elgar.

Jaffe, A, M. Trajtenberg, and R. Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics*. 113 (4), 1137-1167.

Jahn, F., M. Cook, and M. Graham. 1998. "Hydrocarbon Exploration and Production." Amsterdam: Elsevier.

Kasahara, H., and K. Shimotsu. 2008. "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices." *Econometrica*. 77 (1) 135-175.

Kellogg, R. 2011. "Learning by Drilling: Interfirm Learning and Relationship Persistence in the Texas Oilpatch." *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.

Kemp, A. 2012a. "The Official History of North Sea Oil and Gas Vol. I: The Growing Dominance of the State." New York: Routledge.

Kemp, A. 2012b. "The Official History of North Sea Oil and Gas: Vol. II: Moderating the State's Role." New York: Routledge .

Krieger, J., D. Li, and D. Papanikolaou. 2017. " Developing Novel Drugs." *Working Paper.*

Kullback, S. 1997. "Information Theory and Statistics." New York: Dover.

Laffont, J. J., H. Ossard, and Q. Vuong. 1995. "Econometrics of First-Price Auctions." *Econometrica*. 63 (4), 953-980.

Lerche, I, and J. A. MacKay. 1995. "Economic Risk in Hydrocarbon Exploration." San Diego: Academic Press.

Levitt, C. J. 2009. "Learning Through Oil and Gas Exploration." *Working Paper.*

Levitt, C. J. 2016. "Information Spillovers in Onshore Oil and Gas Exploration." *Resource and Energy Economics.* 45, 80-98.

Lin, C.-Y. C, 2009. "Estimating Strategic Interactions in Petroleum Exploration." *Energy Economics.* 31 (4), 586-594.

Lin, C.-Y. C., 2013. "Strategic Decision Making with Information Externalities: A Structural Model of the Multistage Investment Game in Offshore Petroleum Production." *The Review of Economics and Statistics.* 95 (5), 1601-1621.

MacKay, D. J. C. 2003. "Information Theory, Inference, and Learning Algorithms." Cambridge: Cambridge University Press.

Mardia, K. V., and R. J. Marshall. 1984. "Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression." *Biometrika*. 71 (1), 135-146.

Murray, F., P. Aghion, M. Dewatripont, J. Kolev, and S. Stern. 2016 "Of Mice and Academics: Examining the Effect of Openness on Innovation." *American Economic Journal: Economic Policy*. 8 (1) 212-252.

Murray, F. and S. Stern. 2007. "Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the Anti-commons Hypothesis." *Journal of Economic Behavior and Organization*. 63 (4), 648-687.

Moreton, R. 1995. "Tales from Early UK Oil Exploration: 1960-1979." Croydon: Petroleum Exploration Society of Great Britain.

Myerson, R. B, and M. A. Satterthwaite. 1981. "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory*. 29, 265-281.

Porter, R. H. 1995 "The Role of Information in U.S. Offshore Oil and Gas Lease Auctions." *Econometrica.* 63 (1). 1-27.

Rasmussen, C. E., and C. K. I. Williams. 2005. "Gaussian Processes for Machine Learning." Cambridge: The MIT Press.

Rowland, C., and D. Hann. 1987. "The Economics of North Sea Oil Taxation." London: Palgrave Macmillan.

Ryan, S. P., and C. Tucker. 2011. "Heterogeneity and the Dynamics of Technology Adoption." *Quantitative Marketing and Economics*. 10, 63-103.

Scotchmer, S. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law." *Journal of Economic Perspectives.* 5 (1), 29-41.

Spence, M. 1984. "Cost Reduction, Competition, and Industry Performance." *Econometrica*. 52 (1), 101-122.

Steck, A. 2018. "Industry Dynamics with Social Learning: Evidence from Hydraulic Fracturing." *Working Paper.*

Weaver, J. L., and D. F. Asmus. "Unitizing Oil and Gas Fields Around the World." *Houston Journal of International Law.* 28 (1), 1-197.

Williams, H. 2013. "Intellectual Property Rights and Innovation: Evidence from the Human Genome." *Journal of Political Economy*. 121 (1), 1-27.

Appendix

A Details of Logistic Gaussian Process Model

This section describes the Bayesian updating rule for the logistic Gaussian process model and relies heavily on Section 3 of Rasmussen and Williams (2006). The code that I use to implement the numerical Bayesian updating rule is a modified version of the Matlab package made available by Rasmussen and Williams.[36]

The latent variable, $\lambda(X)$ is assumed to be distributed according to a Gaussian process. That is, $\lambda(X)$ is a continuous function, and any finite collection of K locations $\{1, ..., K\}$, the vector $(\lambda(X_1), ..., \lambda(X_K))$ is a multivariate normal random variable with mean $(\mu(X_1), ..., \mu(X_K))$ and a covariance matrix with (j, k) element $\kappa(X_j, X_k)$ where $\kappa(X_j, X_k) \to \kappa(X_j, X_j)$ as $|X_j - X_k| \to 0$.

I assume a constant prior mean and a covariance specification given by equation 2. The prior distribution is therefore defined by three parameters, (μ, ω, ℓ). Denote the density function of prior distribution of λ by $p_0(\lambda)$. Observed data is described by $y = \{(s(w), X_w)\}_{w \in W}$ for a set of wells, W. The Bayesian posterior distribution of λ conditional on y is given by:

$$p_1(\lambda|y) = \frac{p_0(\lambda)p(y|\lambda)}{p(y)} \tag{23}$$

$$p(y|\lambda) = \prod_{w \in W} \left(1(s(w) = 1)\rho(\lambda(X_w)) + 1(s(w) = 0)\left(1 - \rho(\lambda(X_w))\right)\right)$$

$$p(y) = \prod_{w \in W} \left(1(s(w) = 1) \int \rho(\lambda(X_w))p_0(\lambda)d\lambda + 1(s(w) = 0)\left(1 - \int \rho(\lambda(X_w))p_0(\lambda)d\lambda\right)\right)$$

Where $\rho(\lambda(X))$ is defined by equation 1. This posterior distribution is difficult to work with. In particular, in order to compute the posterior $E(\rho(X)|y)$ for some location X I must first compute the marginal distribution of $\lambda(X)$, which is given by:

$$p(\lambda(X) = \tilde{\lambda}|y) = \int 1(\lambda(X) = \tilde{\lambda})p_1(\lambda|y)d\lambda \tag{24}$$

Then the expected value of $\rho(X)$ is given by:

$$E(\rho(X)|y) = \int \rho(\tilde{\lambda})p(\lambda(X) = \tilde{\lambda}|y)d\tilde{\lambda} \tag{25}$$

The posterior marginal distribution of $\lambda(X)$ given by equation 24 is non-Gaussian and has

[36] Available at http://www.gaussianprocess.org/.

no analytical expression. This means that it is computationally costly to compute $E(\rho(X)|y)$.

To solve this problem I use a Gaussian approximation to the posterior $p_1(\lambda|y)$ computed using the Laplace approximation technique detailed in Section 3.4 of Rasmussen and Williams (2006), based on Williams and Barber (1998). This method is widely used for Bayesian classification problems in computer science and in geostatistics (Diggle, Tawn, and Moyeed, 1998)..

Denote the Gaussian approximation to $p_1(\lambda|y)$ by $q_1(\lambda|y)$. Since $q_1(\lambda|y)$ is Gaussian, the posterior distribution over any finite collection of K locations can be written as a $N(\mu^1, \Sigma^1)$ where μ^1 is $K \times 1$ and Σ^1 is $K \times K$. In particular, the marginal distribution given by equation 24 is a Normal distribution.

Notice that, since $q_1(\lambda|y)$ is itself a Gaussian process, it is straightforward to update beliefs *again* given a new set of data, y', following the same procedure. This updating procedure defines the operator $B(\cdot)$ in equation 4, where $G(\rho)$ is the distribution of ρ implied by the prior Gaussian distribution of λ and the logistic squashing function 1, and $G'(\rho)$ is the distribution over ρ defined by the Gaussian approximation to the posterior distribution of λ.

B Estimation Details

B.1 Estimation of Development Payoffs

Firms decide to develop blocks based on the expected payoff from the block, π_j and the fixed cost of developing the block, κ_j. π_j is drawn from a distribution $\Gamma(\pi; \lambda_j, P)$. Recall that the relationship between ρ_j and λ_j is defined in equation 1.

I assume that $\pi_j = \pi(R_j, P)(1 - \mu)$, where R_j are the reserves on block j and P is the oil price at development. I do not observe R_j directly in the data, but I do observe the realized flow of oil from all production wells drilled from a development platform up to 2000. I cannot use the total oil produced from each block to measure R_j because most fields were still producing in January 2000, the last month in my data, and the sum of all oil produced is therefore less than the total reserves.

A classic production profile involves a pre-specified number of wells being drilled, over which time the production flow of the field ramps up. Once the total number of wells is reached, production peaks and then begins to fall off (Lerche and MacKay, 1999). To estimate the volume of reserves, I use data on the set of wells that were drilled *before* production peaked on each block, and extrapolate into the future using an estimate of the rate of post-peak decline in production. Let $t_0(j)$ be the month that production began on block j and let $t^*(j)$ be the month of peak production. Let $r_j(t)$ be the observed flow of oil from block j in month t. I estimate a parameter b_j that measures the rate of post-peak decline in production separately for each block j by applying non-linear least squares to the following specification:

$$r_j(t) = r_j(t^*(j))exp(-b_j(t - t^*(j))) + \epsilon_{jt} \tag{26}$$

Where the estimation sample includes all months after $t^*(j)$ for all developed blocks, j. Estimated initial reserves are then given by:

$$R_j = \sum_{t=t_0(j)}^{t^*(j)} r_j(t) + \sum_{t=0}^{\infty} r_j(t^*(j))exp(-\hat{b}_j t) \tag{27}$$

Where the first term is the realized pre-peak production, and the second term is the extrapolated post-peak production. Note that the oil flow used to estimate the block-specific rate of decline b_j includes data from future redevelopments (so-called "enhanced oil recovery," see Jahn, Cook, and Graham, 1998). Thus, I assume that firms anticipate future improvements in technology that extend the life of the field.

Figure A3 illustrates the relationship between exploration success rate and log estimated reserves. Notice that the expected size of the reserves is increasing in the success rate of exploration wells on the same block. I assume that reserves are drawn from a log normal distribution: $R_j \sim logN(\alpha_R + \mu_R \lambda_j, \sigma_R)$ and estimate via MLE where I integrate out the posterior distribution of λ_j conditional on all wells drilled up to 1990.The estimated parameters are reported in Table A1.

Finally, note that $\pi_j = \pi(R_j, P)(1 - \mu)$ This function converts the total reserves in barrels to the expected present discounted value of revenue taking into account the future evolution of prices, less the 12.5% royalty paid to the government, where oil is assumed to flow at a constant rate for 30 years at which point the reserves, R_j are exhausted. $\mu \in [0, 1]$ is a profit margin parameter that represents the share of revenues after the royalty tax that are captured by operating costs. This parameter is not identified by the data, and I normalize it to 50%. Changing this parameter would change the estimated cost parameters proportionally and have no effect on firm choices.

Table A1: Distribution of Development Payoffs

Parameter	Estimate	SE
α_R	3.903	0.167
μ_R	0.753	0.226
σ^2_R	1.493	0.109
N	111	

Notes: Reported coefficients are from OLS estimation of regression specification given by equation ??. Sample includes one observation for each of the 111 blocks developed before 2000 in the area north of $55°N$ and east of $2°W$. Left hand side variable is the log of the predicted oil reserves on block j, measured in millions of barrels. Right hand side variable is the observed exploration well success rate for block j calculated using all exploration wells drilled on block j before development.

B.2 Estimating Conditional Choice Probabilities

In the first step, I estimate CCPs $\hat{P}(a^E = j|\mathcal{S}_{ft})$ and $\hat{P}(a^D = j|\mathcal{S}_{ft})$ - the probabilities that a firm takes an action j in the exploration and development stages of the game conditional on its state $\mathcal{S}_{ft}$ and the information spillover parameter α. It is not possible to obtain a consistent estimator of the CCPs conditional on the full state, $\mathcal{S}_{ft}$, because states are not recurrent. For instance, once a block is developed it is removed from the set of blocks for which licenses can be issued. As discussed in Section 6, I instead assume that firms condition their choices on a set of lower dimensional statistics, s_{ft}. In particular, I impose the following additional structure on firms' CCPs. Consider first the exploration decision. Notice that equation 11 can be rewritten as

$$P(a_f^E = j|\mathcal{S}_{ft}) = \frac{exp\left(\tilde{v}_f^E(j, s_{ft})\right)}{1 + \sum_{k \in J_{ft}} exp\left(\tilde{v}_f^E(k, s_{ft})\right))} \tag{28}$$

where $\tilde{v}_f^E(j, s_{ft}) = \frac{1}{\sigma_\epsilon} v_f^E(j, s_{ft}) - \frac{1}{\sigma_\epsilon} v_f^E(0, s_{ft})$.

I approximate $\tilde{v_f}^E(j, s)$ with a linear equation with the following terms (components of s):

- A cubic function of the mean and variance of the firm's beliefs: $E(\rho_j|G_{ft}), Var(\rho_j|G_{ft})$.
- The number of licenses held near block j by firm f and by other firms: $|\{k : k \in J_{ft} \text{ and } d(j,k) \leq 3\}|$, $|\{k : k \in \cup\{J_{gt}\}_{g \neq f} \text{ and } d(j,k) \leq 3\}|$, and $|\{k : k \in \cup\{J_{gt}\}_{g \neq f} \text{ and } d(j,k) = 0\}|$, where $d(j,k) = 1$ if j and k are neighbors, $d(j,k) = 2$ if j and k are second degree neighbors etc.
- The number of nearby unobserved well outcomes$|\{w \in W_{ft}^u : d(j, j_w) \leq 3\}|$ wherej_w is the location of well j,

- The number of firms operating near block j: $|\{g : \exists k \in J_{gt} : d(j,k) \leq 3\}|$.
- A quadratic in the price level: P_t and P_t^2.
- Interactions of these variables (listed in Table A2).

The approximation to $\tilde{v_f}^D(j, s_{ft})$ contains fewer terms because of the limited number of development actions observed in the data.

Estimating $\hat{P}(a_f^E = j|s_{ft})$ is then a case of estimating the parameters of this approximation to $\tilde{v_f}^E(j, s_{ft})$. Notice that I assume that s depends on the distribution of licenses and beliefs "near" block j, rather than on the entire set of licenses $\cup\{J_{gt}\}$ and the entire distribution G_{ft}. Intuitively, the *difference* between the value of drilling on block j and taking no action should depend more on the local distribution of licenses and wells than the distribution at distant locations. This "local state" is recurrent and $\tilde{v_f}^E(j, s_{ft})$ can be estimated consistently.

If the state variable were observable in the data, then $\hat{P}(a_f^E = j|s_{ft})$ could be estimated using the likelihood function implied by equation 28. However, the asymmetric information structure of the model means that the true state is not observed by the econometrician. The data does not include the vector $\boldsymbol{o}_f$ that records which other-firm well outcomes were observed by firm f. Different realizations of $\boldsymbol{o}_f$ imply different states through the effect of observed well outcomes on G_{ft}. The data is therefore consistent with a *set* of possible states $\tilde{\mathcal{S}}_f$ for each firm.[37]

To recover CCP estimates, observe that different values of the parameter α define distributions $P(\mathcal{S}_f|\tilde{\mathcal{S}}_f, \alpha)$ over the elements of $\tilde{\mathcal{S}}_f$. For example, suppose at date t there was one other-firm well w that may have been observed by firm f. Let $\mathcal{S}_{ft}^1$ be the state if $o_f(w) = 1$ and $\mathcal{S}_{ft}^0$ be the state if $o_f(w) = 0$. From the econometrician's perspective, $P(\mathcal{S}_{ft}^1|\{\mathcal{S}_{ft}^1, \mathcal{S}_{ft}^0\}, \alpha) = \alpha$. I provide a formal definition of the distribution $P(\mathcal{S}_f|\tilde{\mathcal{S}}_f, \alpha)$ in subsection C.4 below.

Given this distribution over states, the likelihood of a sequence of exploration choice observations is:

$$\mathcal{L}_f^E = \sum_{\mathcal{S}_f \in \tilde{\mathcal{S}}_f} \left[\left(\prod_{t=1}^{T} \prod_{j \in J_{ft} \cup \{0\}} P(a_{ft}^E = j|\mathcal{S}_f)^{1(a_{ft}^E = j)} P(a_{ft}^D = j|\mathcal{S}_f)^{1(a_{ft}^D = j)} \right) P(\mathcal{S}_f|\tilde{\mathcal{S}}_f, \alpha) \right]. \quad (29)$$

I maximize this likelihood to jointly estimate the coefficients of the approximations $\tilde{v}_f^E(j, s_{ft})$ and $\tilde{v}_f^D(j, s_{ft})$, and the parameter α. Since I sometimes observe multiple exploration wells for the same (f,t) I treat these as separate observations inside the brackets

[37] More precisely, and element of $\tilde{\mathcal{S}}_f$ is a particular sequence of firm-f states $\mathcal{S}_f = \{\mathcal{S}_{ft}\}_{t=1}^T$. See the subsection below for a formal definition of $\tilde{\mathcal{S}}_f$.

in equation 29. The estimated coefficients imply conditional choice probability estimates, $\hat{P}(a^E = j|\mathcal{S}_{ft})$ and $\hat{P}(a^D = j|\mathcal{S}_{ft})$.

Logit coefficients for the estimated CCPs are recorded in Table A2.

Table A2: Conditional Choice Probabilities: Logit Coefficients

	Exploration		Development	
	Coefficient	SE	Coefficient	SE
$E(\rho_j)$	-16.840	6.061	2.010	1.205
$E(\rho_j)^2$	25.862	11.113		
$E(\rho_j)^3$	-13.721	8.722		
$Var(\rho_j)$	-12.073	2.859	-5.454	0.852
$Var(\rho_j)^2$	7.057	1.948		
$Var(\rho_j)^3$	-1.845	0.582		
$E(\rho_j)Var(\rho_j)$	11.097	2.080		
Oil Price	0.107	0.025	0.309	0.119
Oil Price^2	-0.001	0.000	-0.005	0.002
Oil Price * $Var(\rho_j)$	-0.004	0.008		
Oil Price * $E(\rho_j)$	-0.029	0.022		
Own Blocks Nearby	0.188	0.043		
Other Blocks Nearby	-0.024	0.018		
Unobserved Wells	-0.077	0.028		
Count of Other Firms Nearby	0.087	0.015		
Unobserved Wells $*$ Own Blocks Nearby	-0.001	0.001		
Unobserved Wells $*$ Other Blocks Nearby	0.001	0.000		
$E(\rho_j)$ $*$ Own Blocks Nearby	-0.195	0.054		
$E(\rho_j)$ $*$ Other Blocks Nearby	0.004	0.025		
$Var(\rho_j)$ $*$ Own Blocks Nearby	-0.129	0.026		
$Var(\rho_j)$ $*$ Other Blocks Nearby	0.002	0.010		
$E(\rho_j)$ $*$ Unobserved Wells	0.007	0.030		
$Var(\rho_j)$ $*$ Unobserved Wells	0.040	0.015		
Intercept	1.067		-8.910	1.781
LR Test of $\alpha = 1$	$\lambda_{LR} = 3924.8$		$C_{0.99}$ =6.635	

Notes: Table records logit coefficients on state variable summary statistics the enter the approximation to the state for the firm's exploration and development decisions. Standard errors are computed using the inverse hessian of the likelihood function. The lower panel reports the results of an LR test is a comparing of the maximized 1st step likelihoods (equation 29) of the baseline model to a restricted model with $\alpha = 1$. The test statistic λ_{LR} has a χ^2_1 distribution with 95% critical value $C_{0.95}$.

B.3 Estimating Belief Functions

Beliefs about Other Firms Actions I use $\hat{P}(a^E = j|\mathcal{S}_{ft})$ and $\hat{P}(a^D = j|\mathcal{S}_{ft})$ to generate simulated data sets that can be used to estimate firms' equilibrium beliefs about other firms' actions.

Starting at the date when the first licenses were issued, I draw exploration and development decisions for each firm using $\hat{P}(a^E = j|\mathcal{S}_{ft})$ and $\hat{P}(a^D = j|\mathcal{S}_{ft})$. Well outcomes on block j are generated using "true" success probabilities ρ_j that are drawn from the estimated Gaussian process, as discussed in Section 3.

Every period, for every well that is drilled by firm f in a period, I allow each other firm to observe the outcome with probability $\hat{\alpha}$, with unobserved wells revealed at the end of the confidentiality period drawn from from $F_\tau(\tau_w)$. At the end of each period, I update each firm's state variable according to the new observed well outcomes and allow the price to evolve according to the estimated process. Each period, firm licenses are set to the observed licenses in the data, unless a firm-block has been developed, in which case the license is removed.

This simulation procedure generates a data set which records each firm's state $\mathcal{S}_{ft}$ and actions (a^E_{ft}, a^D_{ft}) at each date. I simulate this data 20 times using different seeds for the random number generator. Each simulation is run for 316 months, corresponding to the period of the data from the issuing of the first license up to the end of 1990. This means that firms' beliefs are consistent with the equilibrium distribution of states and actions *over this time period*. Let r index simulations. I estimate the following equations on the simulated data using logistic regression:

$$\begin{aligned} 1(a^E_{rgt} = j) &= q^E(s_{rft}, g, j) + \varepsilon_{ftrj} \\ 1(a^D_{rgt} = j) &= q^D(s_{rft}, g, j) + \varepsilon_{ftrj} \end{aligned} \tag{30}$$

Where ε_{ftrj} is i.i.d. logistic. These specifications are regressions of firm g's action on firm f's state. Data is pooled across simulations r, dates t, pairs of firms (f, g), and locations $j \in J_{gt}$. In practice, I estimate these specification on a random 20% sub-sample of this pooled data set which has around 20 million (f, g, j, t) observations.

The functions $q^E(s_{rft}, g, j)$, $q^{Past}(s_{rft}, g, j)$, and $q^D(s_{rft}, g, j)$ are linear in elements of s_{rft}. In particular, $q^{Past}(\mathcal{S}_{rft}, g, j)$, and $q^D(\mathcal{S}_{rft}, g, j)$ have the same specification as $\tilde{v_f}^E(j, s)$ outlined in Appendix B.1 above, where the counts of "own" and "other" blocks now refer to firm g. That is, the functions summarize firm f's beliefs about the well success probability on block j and counts of firm g and other fim's licenses near block j. Likewise, $q^D(s_{rft}, g, j)$

has the same specification as $\tilde{v_f}^D(j,s)$.

Estimates of firms' equilibrium beliefs about other firms actions, defined in equations16 are then given by

$$\hat{Q}^E(\mathcal{S}_{rft},g,j) = \frac{exp(\hat{q}^E(s_{rft},g,j))}{1+exp(\hat{q}^E(s_{rft},g,j))}, \tag{31}$$

and analogously for $\hat{Q}^D$.

Adjusted Beliefs about the Distribution of Oil As discussed in Section 5.2, equilibrium condition 3 requires firms to forecast well outcomes using the distribution $Q^G(\rho; G(\rho), W^u_{ft})$ which is the posterior distribution of ρ conditional on $G(\rho)$, the which depends on observed well outcomes, and W^u_{ft}, the set of wells with unobserved results. Recall from Appendix A that $G(\rho)$ is a transformation of the Gaussian posterior of λ, denoted by $p_1(\lambda|y)$ (equation 23). Denote $p_Q(\lambda|y, W^u_{ft})$ the equivalent transformation of $Q^G(\rho; G(\rho), W^u_{ft})$. Under the assumption that $p_Q(\lambda|y, W^u_{ft})$ is Gaussian, I need to recover an expectation $E_{p_Q}(\lambda_j)$ and a variance $Var_{p_Q}(\lambda_j)$ in order to forecast binary exploration well outcomes and development payoffs (equation **??**) I need to

To compute these adjusted beliefs I use the simulated data described above and run the following regression:

$$\lambda_j = \beta_G E_{p_1}(\rho_j) + \sum_{n=0}^{5} \beta_n 1(|W^u_{ft}| \geq n)| + \sum_{n=0}^{5} \beta_{Gn} 1(|W^u_{ft}| \geq n)|) E_{p_1}(\rho_j) + \epsilon_{jft}. \tag{32}$$

Where an observation is (j,f,t). The fitted values from this regression are then used as $E_{Q^G}(\rho)$. Intuitively, equation 32 estimated the expected value of λ_j conditional on $p_1(\lambda|y)$ and $|W^u_{ft}|$, restricted to a particular linear model. I then obtain residuals $\hat{\epsilon}_{jft}$ and run another regression,

$$\hat{\epsilon}^2_{jft} = \delta_G Var_{p_1}(\rho_j) + \sum_{n=0}^{5} \delta_n 1(|W^u_{ft}| \geq n)| + \sum_{n=0}^{5} \delta_{Gn} 1(|W^u_{ft}| \geq n)|) Var_{p_1}(\rho_j) + \eta_{jft}. \tag{33}$$

The fitted values from this equation are estimated of the expected value of $\hat{\epsilon}^2_{jft}$; i.e. the estimated of the variance of λ_j conditional on $p_1(\lambda|y)$ and $|W^u_{ft}|$.

B.4 Estimating Dynamic Parameters

In the second step, I use the estimated conditional choice probabilities $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$, estimated beliefs $\hat{Q}^E(\mathcal{S}_{ft},g,j)$ and $\hat{Q}^D(\mathcal{S}_{ft},g,j)$, and estimated spillover parameter, $\hat{\alpha}$, to estimate the cost parameters θ_2. The firm's value functions (9) can be written in terms of the expected sum of future payoffs and costs as

$$V^E_{ft_0}(\mathcal{S},\theta_2) = E\left[\sum_{t=t_0}^{\infty}\beta^t \sum_{j=J_{ft}}\left(1(a^D_{ft}=j)\left(\pi_j - (\kappa(j,\mathcal{S}_{ft})-\nu_{ftj})\right) - 1(a^E_{ft}=j)\left(c(j,\mathcal{S}_{ft})-\epsilon_{ftj}\right)\right)\right]. \tag{34}$$

Where the expectations are taken over all future cost shocks, firm actions, and realizations of $s(w)$, $o_f(w)$, and π_j with respect to the firm's beliefs at state $\mathcal{S}$. To estimate this expectation, I forward simulate the model from initial state $\mathcal{S}$.[38] Simulation proceeds as follows:

1. Draw an exploration action using probabilities $\hat{P}(a^E_{ft} = j|\mathcal{S}_{ft})$. Compute expected cost shock ϵ_{fta^E}, given realized action. If a well is drilled, let it be successful with probability corresponding to firm f's *beliefs at state* $\mathcal{S}_{ft}$.

2. Draw other firms' exploration actions using $\hat{Q}_E$. Let wells be successful with probability corresponding to firm *f's beliefs at state* $\mathcal{S}_{ft}$.

3. Draw $o_f(w)$ for wells drilled by other firms using $\hat{\alpha}$. If $o_f(w) = 0$ the draw deadline $\tau(w)$ from $F_\tau(\tau_w)$

4. Reveal wells at the end of their confidentiality period: if $t - t(w) = \tau(w)$ set $o_f(w) = 1$.

5. Update state to $\mathcal{S}'_{ft} = (\mathcal{S}_{ft}, a^E_{ft})$.

6. Draw a development action using $\hat{P}(a^D_{ft} = j|\mathcal{S}'_{ft})$. Compute expected cost shock ν_{fta^E}, given realized action. If block j is developed draw development revenue π_j from the distribution corresponding to firm f's *beliefs* $Q^G(\rho; G(\rho), W^u_{ft})$.

7. Draw other firms' development actions using $\hat{Q}_D$. Draw development revenue π_j from $Q^G(\rho; G(\rho), W^u_{ft})$.

8. Update state to $\mathcal{S}_{ft+1}$. Beliefs G_t are updated based on exploration well results W_t and realized revenues $\{\pi_{jt}\}_{j\in J_{dt}}$. Price and licenses evolve according to estimated processes.[39]

[38] Hotz and Miller (1993) obtain estimates of the firm's value function using finite dependence by normalizing one state to have a continuation value of 0. This approach is complicated here since the "absorbing state" of developing all blocks is the result of a series of choices, rather than a single choice that is available at every state (for example exit in a standard dynamic oligopoly model).

[39] Notice that since cost parameters θ_2 enter equation 34 linearly, I only need to perform the simulation step once. Simulated continuation values can be obtained under different parameter vectors θ_2 by multiplying the simulated costs and revenues by the relevant elements of the parameter vector (Bajari, Benkard, and Levin, 2007).

Let r index simulation runs and $V_{fr}^E(\mathcal{S}_{ft}, \theta_2)$ be the present discounted sum of firm f's payoffs and costs from run r. This sum includes incurred costs and oil revenue as well as the sum of future cost shocks, ϵ_{ftj} and ν_{ftj}. I normalize the means of ϵ_{ftj} and ν_{ftj} so that the expected cost of no action is 0. That is, $E(\epsilon_{ft0}|a_{ft}^E = 0, S_{ft}) = 0$ and $E(\nu_{ft0}|a_{ft}^E = 0, S_{ft}) = 0$. Given R simulations from state $\mathcal{S}_{ft}$, estimates of the value functions given by equation 34 are:

$$\hat{V}_f^E(\mathcal{S}_{ft}, \theta_2) = \frac{1}{R} \sum_{r=1}^{R} \left[V_{fr}^E(\mathcal{S}_{ft}, \theta_2) \right]. \tag{35}$$

A similar procedure is used to compute estimates of development stage value functions $\hat{V}_f^D(\mathcal{S}_{ft}, \theta_2)$ where the simulation algorithm is started at step 5. Simulated choice-specific value functions, defined relative to the value of choosing not to drill, are:

$$\hat{v}_f^E(j, \mathcal{S}_{ft}, \theta_2) = \frac{1}{R} \sum_{r=1}^{R} \left[V_{fr}^E(\mathcal{S}_{ft}, \theta_2) | a_{ft_0}^E = j \right] - \frac{1}{R} \sum_{r=1}^{R} \left[V_{fr}^E(\mathcal{S}_{ft}, \theta_2) | a_{ft_0}^E = 0 \right] \tag{36}$$

I simulate these choice-specific value differences for 500 firm-date-block observations drawn from the data. For each observation I run $R = 100$ simulations with $a_{f0}^E = j$ and 100 simulations with $a_{f0}^E = 0$, and then compute the difference in expected continuation values as a function of θ_2. Each simulation is run for 360 periods (30 years). A similar procedure generates $\hat{v}_f^D(j, \mathcal{S}, \theta_2)$.

I then find the cost parameters, θ_2, that minimize the difference between the fist step estimated choice-specific value functions, $\tilde{v}_f^E(j, s_{ft})$ and $\tilde{v}_f^D(j, s_{ft})$, and the model-implied choice-specific value functions, $\hat{v}_f^E(j, \mathcal{S}_{ft}, \theta_2)$ and $\hat{v}_f^D(j, \mathcal{S}_{ft}, \theta_2)$, using the objective function given by equation 37.

$$\theta_2^* = \arg\min_{\theta_2} \sum_f \sum_t \sum_{k \in \{E,D\}} \left(\tilde{v}_f^k(j, s_{ft}) - \hat{v}_f^k(j, \mathcal{S}_{ft}, \theta_2) \right)^2 \tag{37}$$

Where the sum is over the set of 500 firm-date-blocks drawn from the data. Notice that $\tilde{v}_f^k(j, s_{ft})$ is a function of the lower dimensional state variable s_{ft}. Since the sampled states are drawn from the equilibrium distribution (i.e. the data), this regression imposes the assumption in equation 19 that firms' choice-specific values are averages over the distribution of $\mathcal{S}_{ft}$ conditional on s_{ft} within the time period of the data.

Standard errors for θ_2 are obtained by repeating the second step estimation procedure for 200 bootstrap draws from the first step CCP estimated in Table A2 and taking the standard deviation of the estimated parameters. Because the simulation of value functions is computationally intensive, I only use 50 simulations for each bootstrap draw.

B.5 Technical Details on Distribution of States

Define a period t observation as

$$X_t = \{\{(j(w), s(w), f(w)) : t(w) < t\}, \{J_{ft}\}_{f \in F \cup \{0\}}, P_t\}, \tag{38}$$

the set of wells, the set of licenses, and the oil price. The data consists of T such observations, $X = \{X_t\}_{t=1}^T$. If the states $\{\mathcal{S}_{ft}\}_{f \in F}$ were uniquely identified by X_t, then $\hat{P}(a_f^E = j|\mathcal{S})$ could be estimated using a straightforward logit. This is not possible since the econometrician does not observe the vector $\boldsymbol{o}_f$. That is, the econometrician does not know *which* well outcomes each firm observed in reality. Different realizations of $\boldsymbol{o}_f$ imply different states through the effect of observed well outcomes on G_{ft}. The state variable $\mathcal{S}_{ft}$ is therefore not directly observed in the data, and for every (f, t), the data is consistent with a *set* of states.

Formally, denote a sequence of firm f states as $\mathcal{S}_f = \{\mathcal{S}_{ft}\}_{t=1}^T$. There exists a function $s(\cdot)$ such that $\mathcal{S}_f = s(\boldsymbol{o}_f|X)$. Define $\tilde{\mathcal{S}}_f(X)$ as the range of this function. That is, $\tilde{\mathcal{S}}_f$ is the set of firm f states that are consistent with the data. There also exists an inverse correspondence $s^{-1}(\mathcal{S}_f|X)$ that maps states to (possibly multiple) vectors $\boldsymbol{o}_f$ that imply those states.

To recover CCP estimates, observe that different values of α define distributions over the elements of $\tilde{\mathcal{S}}_f$. In particular, the probability of sequence of states $\mathcal{S}_f \in \tilde{\mathcal{S}}_f$, conditional on the data is:

$$P(\mathcal{S}_f|X, \alpha) = \sum_{o \in s^{-1}(\mathcal{S}_f|X)} \left(\alpha^{\sum_w o(w)}(1-\alpha)^{\sum_w (1-o(w))}\right). \tag{39}$$

Given this distribution over true states, the likelihood of a sequence of exploration choice observations conditional on (X, α) is given by:

$$\mathcal{L}_f^E = \sum_{\mathcal{S}_f \in \tilde{\mathcal{S}}_f(X)} \left[\left(\prod_{t=1}^T \prod_{j \in J_{ft} \cup \{0\}} P(a_{ft}^E = j|\mathcal{S}_f)^{1(a_{ft}^E = j)} P(a_{ft}^D = j|\mathcal{S}_f)^{1(a_{ft}^D = j)}\right) P(\mathcal{S}_f|X, \alpha)\right]. \tag{40}$$

Note that the summation in equation 40 is an expectation. In practice, it is computationally infeasible to compute the action probabilities at every possible state sequence $\mathcal{S}_f \in \tilde{\mathcal{S}}_f$. I approximate this expectation for different values of α using importance sampling.

B.6 Estimation of License Issuing Process

New drilling licenses are issued by the government in rounds that occur irregularly every two to three years. To model this process, I assume that licensing rounds happen each month with a probability 1/24, so that every two years there is one round in expectation. When a round takes place, new licenses are issued and existing licenses may be revoked. Firm f has correct beliefs about this process

In a given round, one license can be issued on each block j. The probability that a new license on block j is issued to a firm f is given by:

$$P(j \in J_{ft}) = \frac{exp(\beta Z_{fjt})}{1 + \sum_{g:j \notin F_{jt}} exp(\beta Z_{gjt})}$$

That is, a new license can be issued to any firm that does not already have a license on that block. The probability of receiving the license depends on covariates Z_{fjt} which include whether any firm has a license on the block at date, whether any neighboring blocks are licensed at date, and whether any neighboring blocks are licensed to firm f.

After new licenses are issued, existing licenses are revoked at random with probability ω.

I estimate this licensing process by running a regression of licenses at date t on covariates Z_{fjt-24}. That is, I estimate the probability that a new license is issued in today as a function of the license distribution two years ago. Similarly, I estimate ω as the probability that a license at date t is revoked by date $t + 24$. The estimated parameters recorded in Table A3.

The average probability that a new license is issued on a block in a given round is 4.8%. The probability that a license on a particular block is issued to a particular firm is 0.06%. Marginal effects of the covariates Z_{fjt} on this probability, evaluated at the mean values of the covariates, are recorded in Table A3.

Table A3: License Issuing Process

Probability of New License			Probability of Revoked License	
	Parameters	Marginal Effect (percentage points)		
Constant	-7.808		ω	0.129
	(0.010)			(0.000)
Licensed in $t-24$	0.660	0.025		
	(0.018)			
Neighbors Licensed				
to $g \neq f$ in $t-24$	-0.579	-0.022		
	(0.072)			
to f in $t-24$	2.120	0.080		
	(0.024)			

Notes: Marginal effects on probability are computed at the mean value of the independent variableS.

C Identification Details

In this section I provide a proof of identification of the exploration conditional choice probabilities (CCPs) $P(a_f^E = j|\mathcal{S})$ and the information spillover parameter, α. Identical reasoning applies to development choice probabilities. Let $\boldsymbol{X}$ be the space of possible data points, where $X \in \boldsymbol{X}$ is an observation as defined by equation 38.

Proposition 1. *Suppose $P(a_f^E = j|\tilde{\mathcal{S}}_f(X))$ is observed for all f and all $X \in \boldsymbol{X}$. These observed probabilities are consistent with a unique value of α and a unique value of $P(a_f^E = j|\mathcal{S}_f)$ for every possible state $\mathcal{S}_f$.*

Proof. First, suppose that α is known.

Let w_t be a vector of length $W = |\{w : t(w) < t\}|$ indexed by $i \in [1, ..., W]$ is an index which contains the identity w of each well $w \in \{w : t(w) < t\}$ in some order such that we can refer to well identities by, $w_t(i)$. Let $\boldsymbol{\gamma}_{ft}$ be a vector of length W with ith element $\boldsymbol{\gamma}_{ft}(i) = 1(f(w_t(i)) = f)$. $\boldsymbol{\gamma}_{ft}$ is a vector of indicators for whether each well w was drilled by firm f.

We can then rewrite the observable data X_t as $X_t = \{x_t, \{\boldsymbol{\gamma}_{ft}\}_{f\in F}\}$. Where

$$x_t = \{\{(j(w), s(w)) : t(w) < t\}, \{J_{ft}\}_{f\in F\cup\{0\}}, P_t\}.$$

x_t describes the location and outcome of all wells drilled up to date t, the date t distribution of licenses, and the oil price.

Define $\boldsymbol{o}_{ft}$ as a vector of length W with ith element given by $\boldsymbol{o}_{ft}(i) = \boldsymbol{o}_f(w_t(i))$. $\boldsymbol{o}_{ft}$ is just an ordered vector of containing indicators for whether firm f observed each well $w \in \{w : t(w) < t\}$ (a subset of the elements of $\boldsymbol{o}_f$).

Suppose for simplicity that all wells w, $t - t(w) < \tau$, so no wells are older than the confidentiality period τ. This assumption simplifies notation, and the following argument easily generalizes. I now drop the t subscript for simplicity.

Firm f's state is uniquely defined by the pair $(\boldsymbol{o}_f, x)$. That is, there exists a *function* $\mathcal{S}_f = s(f, \boldsymbol{o}_f, x)$. The set of states that are consistent with the objects observed in the data is defined by a *correspondence* $\tilde{\mathcal{S}}_f = \tilde{s}(f, \boldsymbol{\gamma}_f, X)$. In particular:

$$\tilde{s}(f, \boldsymbol{\gamma}_f, x) = \{s(f, \boldsymbol{o}_f, x) : \boldsymbol{\gamma}_f(i) = 1 \Rightarrow \boldsymbol{o}_f(i) = 1 \forall i \in [1, ..., W]\}.$$

So $\tilde{s}(f, \boldsymbol{\gamma}_f, x)$ contains states implied by all possible values of $\boldsymbol{o}_f$. In particular, each well drilled by a firm other than f may or may not have been observed.

Now fix a value of x. There are 2^W possible values of $\boldsymbol{\gamma}_f$ and therefore of $\tilde{\mathcal{S}}_f = \tilde{s}(f, \boldsymbol{\gamma}_f, x)$. There are also 2^W possible values of $\boldsymbol{o}_f$ and therefore of $\mathcal{S}_f = s(f, \boldsymbol{o}_f, x)$. Let $\boldsymbol{S}_f(x)$ be the set of possible values of $\mathcal{S}_f$ and $\tilde{\boldsymbol{S}}_f(x)$ be the set of possible values of $\tilde{\mathcal{S}}_f$. For any action choice $j \in J_f$ and any $\tilde{\mathcal{S}}_f \in \tilde{\boldsymbol{S}}_f(x)$ we can write:

$$P(a_f^E = j|\tilde{\mathcal{S}}_f) = \sum_{\mathcal{S}_f \in \mathbf{S}_f(x)} P(a_f^E = j|\mathcal{S})P(\mathcal{S}_f|\tilde{\mathcal{S}}_f).$$

Where $P(\mathcal{S}_f|\tilde{\mathcal{S}}_f)$ is a function of α given by equation 39 if $\mathcal{S}_f \in \tilde{\mathcal{S}}_f$ and $P(\mathcal{S}|\tilde{\mathcal{S}}_f) = 0$ if $\mathcal{S}_f \notin \tilde{\mathcal{S}}_f$.

There are 2^W such equations which define a linear system $\tilde{\boldsymbol{P}} = \boldsymbol{AP}$ where $\tilde{\boldsymbol{P}}$ is a $2^W \times 1$ vector which stacks the probabilities $P(a_f^E = j|\tilde{\mathcal{S}}_f)$, $\boldsymbol{P}$ is a $2^W \times 1$ vector which stacks the probabilities $P(a_f^E = j|\mathcal{S})$, and $\boldsymbol{A}$ is a $2^W \times 2^W$ matrix containing the probabilities $P(\mathcal{S}_f|\tilde{\mathcal{S}}_f)$ which are known functions of α. $\tilde{\boldsymbol{P}}$ is observed in the data. $\boldsymbol{A}$ is a known function of the single parameter α. $\boldsymbol{P}$ is an unknown vector for which we would like to solve.

The vector of true CCPs $\boldsymbol{P}$ can be recovered from the observed probabilities, $\tilde{\boldsymbol{P}}$ when $\boldsymbol{A}$ has full rank. This is the case here because the system of equations can be written such that $\boldsymbol{A}$ is lower triangular with non-zero diagonal elements. I show this by providing an algorithm to solve the system by forward substitution, which is only possible in a triangular system of equations. The algorithm proceeds as follows:

1. Denote the vector with all entries equal to 1 by $\mathbf{1}$ Start with $\boldsymbol{\gamma}_f^1 = \mathbf{1}$. Let $\tilde{\mathcal{S}}_f^1 = \tilde{s}(f, \mathbf{1}, x)$ and $\mathcal{S}_f^1 = s(f, \mathbf{1}, x)$. Notice $\tilde{\mathcal{S}}_f^1 = \mathcal{S}_f^1$. If all wells were drilled by firm f, then they are all observed. Therefore

$$P(a_f^E = j|\tilde{\mathcal{S}}_f^1) = P(a_f^E = j|\mathcal{S}_f^1).$$

   $P(a_f^E = j|\mathcal{S}_f^1)$ is uniquely identified.

2. Denote the vector with all entries except the ith equal to 1 and the ith equal to 0 by $\mathbf{1}^{\{i\}}$. Let $\boldsymbol{\gamma}_f^2 = \mathbf{1}^{\{i\}}$. Let $\tilde{\mathcal{S}}_f^2 = \tilde{s}(f, \mathbf{1}^{\{i\}}, x)$ and $\mathcal{S}_f^2 = s(f, \mathbf{1}^{\{i\}}, x)$. Notice that $\tilde{\mathcal{S}}_f^2 = \{\mathcal{S}_f^1, \mathcal{S}_f^2\}$. The firm either did or did not observe the ith well. Therefore

$$P(a_f^E = j|\tilde{\mathcal{S}}_f^2) = \alpha P(a_f^E = j|\mathcal{S}_f^1) + (1-\alpha)P(a_f^E = j|\mathcal{S}_f^2).$$

   Since the other terms are already known, $P(a_f^E = j|\mathcal{S}_f^2)$ is uniquely identified.

3. Repeat step 2 for each index $\forall i \in [1, ..., W]$.

4. Proceed to vectors $\boldsymbol{\gamma}_f$ with two entries equal to 0 and repeat step 2.

5. Continue iterating through vectors with increasingly more entries equal to 0 until $P(a_f^E = j|\mathcal{S}_f)$ has been solved for for all $\mathcal{S}_f \in \mathbf{S}_f(x)$.

This algorithm generates the unique solution $\boldsymbol{P}$ of the system of equations $\tilde{\boldsymbol{P}} = \boldsymbol{AP}$. This can be repeated for any value of x.

Now I argue that α is uniquely identified. Fix a pair (x, x') where x and $x'$are identical except for the outcome of the ith well. The following four equations hold:

$$\begin{aligned}
P(a_f^E = j|\tilde{s}(f, \mathbf{1}, x)) &= P(a_f^E = j|s(f, \mathbf{1}, x)) \\
P(a_f^E = j|\tilde{s}(f, \mathbf{1}, x')) &= P(a_f^E = j|s(f, \mathbf{1}, x')) \\
P(a_f^E = j|\tilde{s}(f, \mathbf{1}^{\{i\}}, x)) &= \alpha P(a_f^E = j|s(f, \mathbf{1}, x)) + (1-\alpha) P(a_f^E = j|s(f, \mathbf{1}^{\{i\}}, x)) \\
P(a_f^E = j|\tilde{s}(f, \mathbf{1},^{\{i\}} x')) &= \alpha P(a_f^E = j|s(f, \mathbf{1}, x')) + (1-\alpha) P(a_f^E = j|s(f, \mathbf{1}^{\{i\}}, x'))
\end{aligned}$$

The left hand side of each equation is observed. Notice that $P(a_f^E = j|s(f, \mathbf{1}^{\{i\}}, x)) = P(a_f^E = j|s(f, \mathbf{1}^{\{i\}}, x'))$ since when the ith well is unobserved the two states are identical to the firm. There are therefore three unknown choice probabilities and the parameter α on the right hand side. α can be solved for in terms of observed quantities. □

D Simulation Details

In this section, I describe the simulation algorithm used to compute counterfactual equilibria of the estimated model. Inputs to the simulation are a vector of model parameters, θ, a confidentiality window, τ, a license assignment$\{J_{ft}\}_{f\in F}$ for each period, first step conditional choice probability (CCP) estimates, $\hat{P}(a^E = j|\mathcal{S}_{ft})$ and $\hat{P}(a^D = j|\mathcal{S}_{ft})$, and belief functions $\hat{Q}^E$, $\hat{Q}^D$, and $\hat{Q}^G$. The output of the simulation are equilibrium CCPs, $P^*(a^E = j|\mathcal{S}_{ft})$ and $P^*(a^D = j|\mathcal{S}_{ft})$.

The algorithm works by taking a set of CCPs as input and forward simulating those probabilities to obtain model-implied choice-specific continuation values (given by equation 35). This is the same procedure as used in the second step of estimation, as described in Appendix C.3.

These value functions imply new choice probabilities that can be used to generate simulated data, from which new belief functions can be estimated (using the procedure described in Appendix C.2). The procedure can then be iterated until choice probabilities converge to a fixed point.

The algorithm proceeds as follows:

1. Draw a set of states $\boldsymbol{S}^1$ from the data.

2. Use first step CCPs $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$ and first step estimates of $\hat{Q}^E$, $\hat{Q}^D$, and $\hat{Q}^G$ to perform the forward simulation described in Appendix C.3 for each $S \in \boldsymbol{S}^1$. This procedure generates model implied exploration and development probabilities, $\hat{P}^1(a_f^E = j|\mathcal{S}, \theta_2)$ and $\hat{P}^1(a_f^D = j|\mathcal{S}, \theta_2)$.

3. Simulate data using $\hat{P}^1(a_f^E = j|\mathcal{S}, \theta_2)$ and $\hat{P}^1(a_f^D = j|\mathcal{S}, \theta_2)$, and use this data to estimate belief functions, $\hat{Q}^{E1}$, $\hat{Q}^{D1}$, and $\hat{Q}^{G1}$.

4. Draw a new set of states $\boldsymbol{S}^2$ from the simulated data.

5. Go back to step 2 and repeat using new exploration CCPs and belief functions. Repeat for k iterations until

$$\sum_{S \in \boldsymbol{S}} \left(\hat{P}^k(a_f^E = j|\mathcal{S}, \theta_2) - \hat{P}^{k+1}(a_f^E = j|\mathcal{S}, \theta_2) \right)^2 \approx 0$$

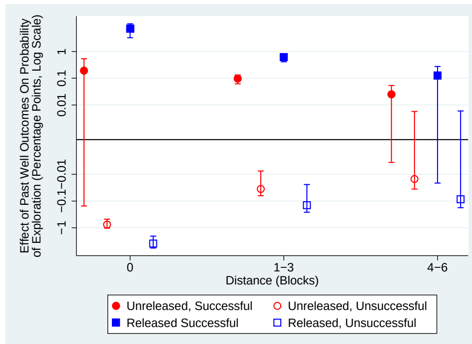
and

$$\sum_{S \in \boldsymbol{S}} \left(\hat{P}^k(a_f^D = j|\mathcal{S}, \theta_2) - \hat{P}^{k+1}(a_f^D = j|\mathcal{S}, \theta_2) \right)^2 \approx 0.$$

Note that by drawing a new set of states from the simulated data on each iteration I ensure that the equilibrium values functions are consistent with equation 19. In practice I modify this algorithm by terminating the forward simulation in iteration k after 100 months and assigning terminal continuation values $\hat{V}^{k-1}(\mathcal{S}, \theta_2)$ from the $k-1$th iteration, extrapolated across states using a linear regression in elements of $\mathcal{S}$. This greatly reduces the time it takes to perform one iteration (from > 10 hours to ~ 1 hours). This method is intermediate between full simulation of value function and iteration of the bellman equation. Shorter simulations are faster, but may be sensitive to the functional form used to approximate continuation values and may take more iterations to converge. Longer simulations are slower but likely converge in fewer iterations. The approach I take here uses the longest forward simulations that are practical given the computational constraints, and a flexible functional form for $\hat{V}^{k-1}(\mathcal{S}, \theta_2)$.

E Additional Tables and Figures

Figure A1: Response of Drilling Probability to Exploration Well Data Release



Notes: Points are the estimated marginal effect of each type of past well on $Explore_{fjt}$ from a specification similar to equation 7, that also includes the number of wells for which data has been released (i.e. the confidentiality period has expired) of each type as independent variables. Error bars are 95% confidence intervals computed using standard errors clustered at the firm-month level. Sample includes block-months in the relevant region before 1991. I drop observations from highly explored regions where the number of nearby own wells (those on 1st and 2nd degree neighboring blocks) is above the 80th percentile of the distribution in the data.

Table A4: Regressions of Exploration Probability on Equity Holders' Nearby Licenses

	Exploration Well			
$BlocksOwn_{fjt}$	2.467***	2.479***	2.505***	2.401***
	(.875)	(0.858)	(.851)	(.868)
$BlocksOpEquity_{fjt}$	-.514	.	.	-1.026
	(1.277)	.	.	(1.304)
$BlocksEquityOp_{fjt}$	.	1.351	.	1.220
	.	(0.824)	.	(.816)
$BlocksEquityEquity_{fjt}$	.	.	.846	.902
	.	.	(.617)	(.623)
N	80562	80562	80562	80562
Firm-Block, and Month FE	Yes	Yes	Yes	Yes
Coefficients Scaled by 10^3	Yes	Yes	Yes	Yes

Notes: Each column records OLS estimates of the coefficients from a regression of $Explore_{fjt}$ on counts on of nearby licenses (1st and 2nd degree neighbors). $BlocksOpEquity_{fjt}$ is the number of blocks nearby block j at month t on which firm f, the operator of block j, is an equity holder but not an operator. $BlocksEquityOp_{fjt}$ is the count of blocks nearby block j at date t for which one of the non-operator firms with equity on block j is the operator. $BlocksEquityEquity_{fjt}$ is the count of blocks nearby block j at date t for which one of the non-operator firms with equity on block j is a non-operator equity holder. Regressions also include controls for past well results as in equation 7 Standard errors clustered at the firm-block level. *** indicates significance at the 99% level. ** indicates significance at the 95% level. * indicates significance at the 90% level.

Table A5: Block Level Success Rates Over Time

Dependent Variable: Well Success							
Well Sequence Number	.025***	-.001	.003	Well 2	-0.020	Well 7	-0.043
	(.002)	(.003)	(.003)		(0.029)		(0.047)
				Well 3	0.017	Well 8	0.026
Year	-.005***	.005**	.		(0.032)		(0.052)
	(.001)	(#3#)	.	Well 4	-0.014	Well 9	0.077
					(0.035)		(0.058)
				Well 5	-0.038	Well 10	0.017
					(0.039)		(0.042)
				Well 6	0.026		
					(0.043)		
N	2105	2105	2105				2105
Block FE	No	Yes	Yes				Yes

Notes: Sample includes all exploration wells drilled before 1991 on the region north of $55°N$ and east of $2°W$. Left hand side variable is an indicator for whether the well was successful. Well sequence number records the order in which wells were drilled on a block. The first well on block j has well sequence number 1, the second well has well sequence number 2, etc. *** indicates significance at the 99% level. ** indicates significance at the 95% level. * indicates significance at the 90% level.

Table A6: Ratio of Response to Nearby Wells to Response to Same-Block Wells

	Successful Wells		Unsuccessful Wells	
Years	Ratio	SE	Ratio	SE
1966-1980	0.160	0.118	0.090	0.030
1971-1985	0.103	0.066	0.048	0.036
1976-1990	0.124	0.057	0.078	0.045
1981-1995	0.090	0.067	0.082	0.040
1986-2000	0.131	0.168	0.049	0.029

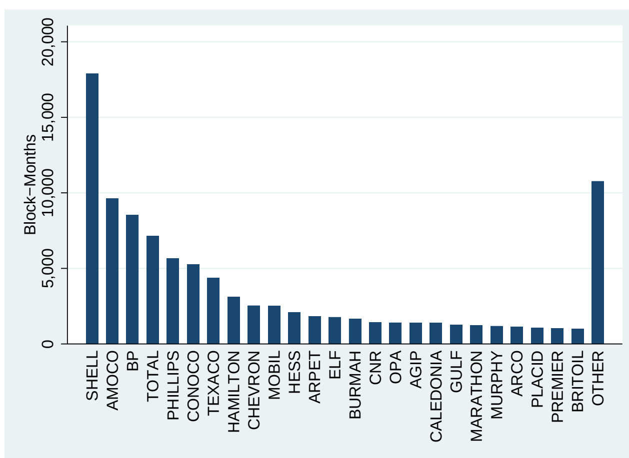
Notes: Table reports the ratio of the estimated marginal effect of past wells on nearby blocks (1-3 blocks away) to past wells on the same block on $Explore_{fjt}$ from the specification given by equation 7 where $g_{do}(\cdot)$ is quadratic in each of the arguments. Marginal effect is computed for the first well of each type. Sample includes block-months in the relevant region up for the time period indicated in the first column. An observation, (f, j, t) is in the sample if firm f had drilling rights on block j in month t, and block j had not yet been developed. I drop observations from highly explored regions where the number of nearby own wells (those on 1st and 2nd degree neighboring blocks) is above the 95th percentile of the distribution in the data. Robust standard errors are reported.

Table A7: Regressions of Drilling Probability on Nearby Licenses

	Exploration Well		
$log(BlocksOwn_{fjt})$	0.002	0.009	0.028**
	(0.002)	(0.014)	(0.014)
$log(BlocksOther_{fjt})$	0.015***	-0.008*	-0.013***
	(0.002)	(0.005)	(0.004)
N	21,618	21,618	21,618
Licensing Round Fixed Effects	No	Yes	Yes
Well Outcome Controls	No	No	Yes

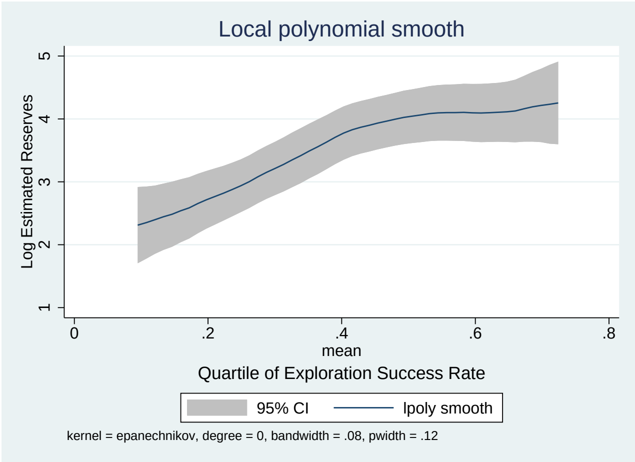
Notes: Standard errors clustered at the firm-block level. Observations are at the (f, j, t) level. Sample includes all (f, j, t) observations that are within 4 months of a licensing round, for which the firm f has held a license on block j for at least 6 months. Licensing rounds are identified as (f, j, t) observations for which the total number of licensed blocks neighboring block j increases from the previous month. Block counts are of all licenses on block j and neighboring blocks on date t. Well outcome controls are the same as in specification 7. *** indicates significance at the 99% level. ** indicates significance at the 95% level. * indicates significance at the 90% level.

Figure A2: Top 25 Firms



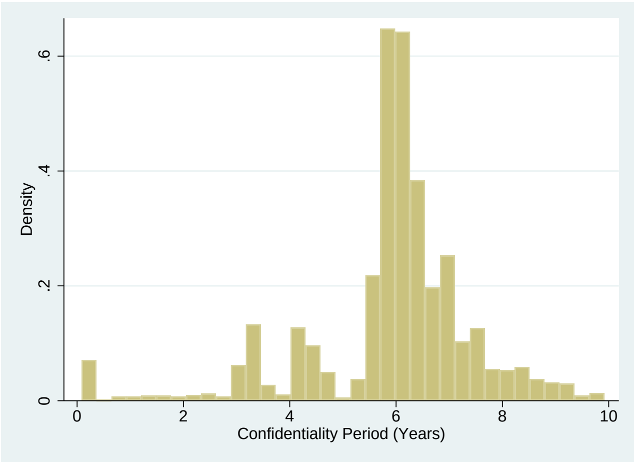
Notes: Figure plots the number of block-month pairs for 1964-1990 licensed to each of the top 25 firms, and the set of all other firms.

Figure A3: Estimated Reserves



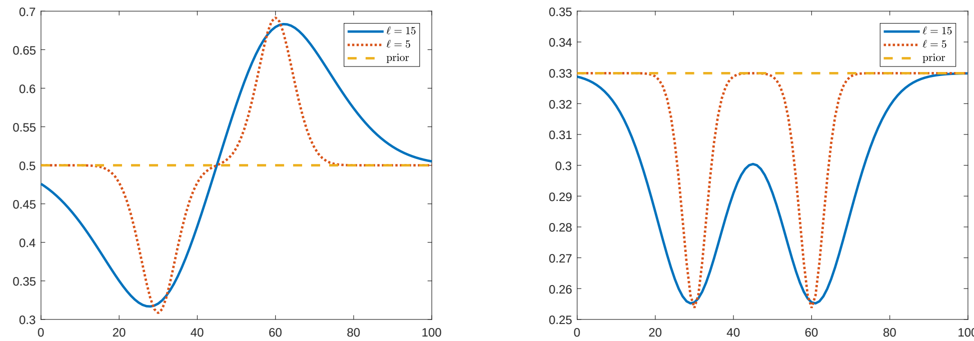
Notes: Figure records the distribution of estimated oil reserve volume, measured in log millions of barrels, across all developed blocks in the relevant area. The line is a local polynomial regression with a 95% confidence band. A regression of log estimated reserves on success rate has a slope coefficient of 5.990 with a standard error of 0.964.

Figure A4: Confidentiality Periods



Notes: Histogram of the time from well completion to data release for all exploration wells. 27 observations with confidentiality periods less than or equal to 0 or greater than 10 years are dropped.

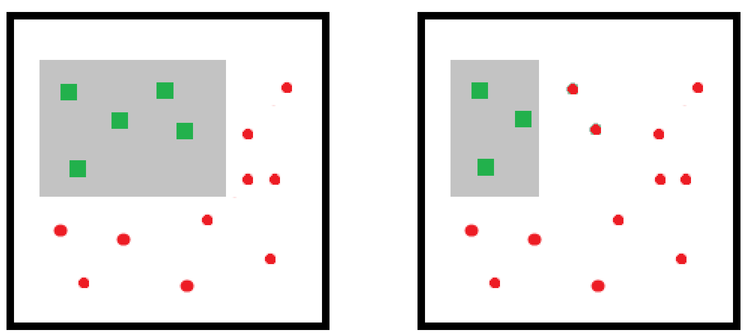
Figure A5: Gaussian Process Learning



Expected Probability   Standard Deviation of Probability

Notes: The x-axis of both panels represents the one dimensional space $[0, 1]$ on which the Gaussian process is defined. The dashed yellow line in the left panel plots the expected value of $\rho(X)$ for $X \in [0, 1]$ under prior beliefs represented by a logistic Gaussian process defined according to equations 1 - 2 with $\mu(X) = 1$ and $\omega = 5$. The solid blue line in the left panel represents the posterior expectation of $\rho(X)$ after observing a successful well at $X = 60$ and an unsuccessful well at $X = 30$ when $\ell = 15$. The dotted red line represnte the posterior expectation when $\ell = 5$. The right panel plots the standard deviation of $\rho(X)$ under the same prior (red dashed line) and posterior (solid blue line) beliefs.

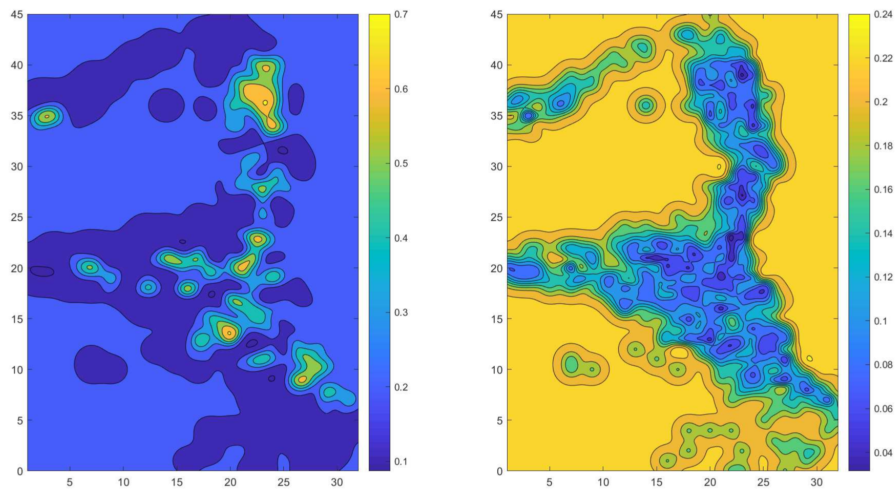
Figure A6: Success Rate and Reserve Size



$\rho_j = 0.333$ $\rho_j = 0.2$

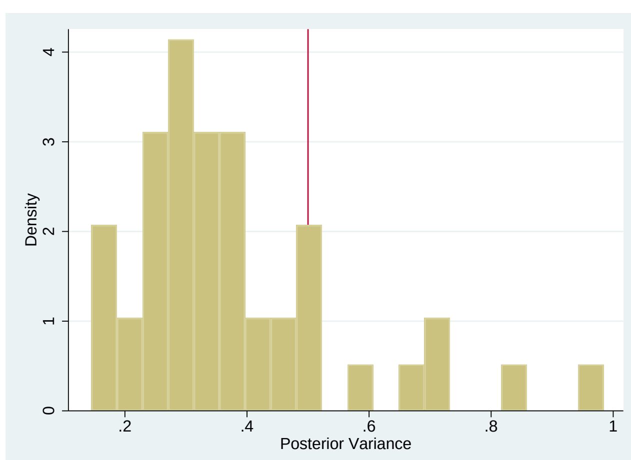
Notes: Stylized example. Each panel represents a block. The points are oil wells and the shaded area is the oil field. Green wells are "successful" (that is, they encountered an oil column), and red wells are "unsuccessful". The probability of exploration well success, ρ_j,on each block corresponds to the share of that block occupied by the oil field.

Figure A7: Posterior Oil Well Probabilities



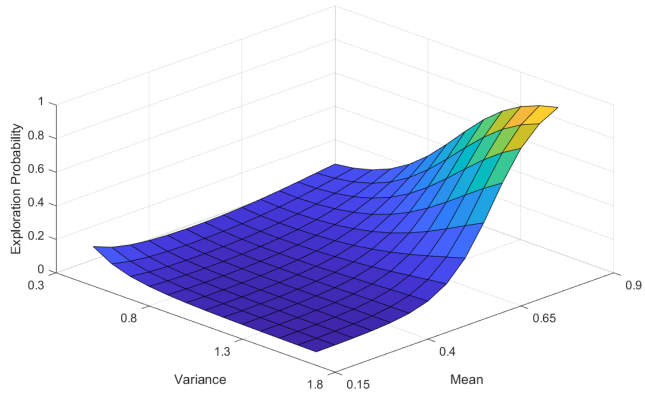
Notes: The left panel is a map of the posterior expected probability of success of a firm with prior beliefs given by the parameters in Table 2 that observes every well drilled between 1964 and 1990. The right panel is a map of the posterior standard deviation of beliefs for the same firm. In the left panel, lighter regions have a higher posterior expected probability of success, and correspond to areas where more successful wells were drilled. Darker regions indicate lower posterior expected probability of success, and correspond to areas where more unsuccessful wells were drilled. The right panel records the posterior standard deviation of beliefs, with darker regions indicating less uncertainty.

Figure A8: Distribution of Posterior Variance on Developed Blocks



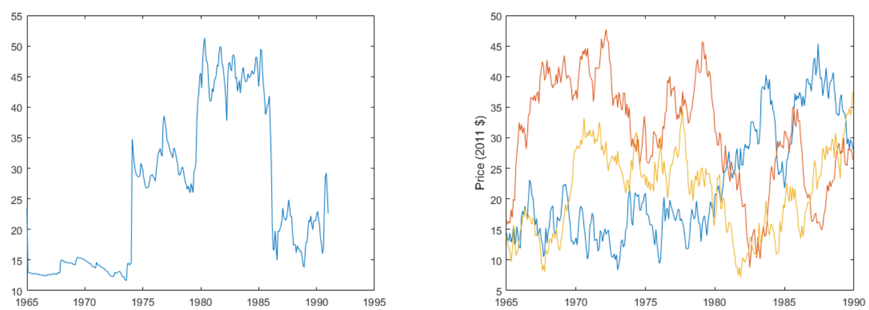
Notes: Figure plots the distribution of the posterior variance (assuming full information) on block-dates for which a development action is observed. There are 50 such observations in the data. 86% of development actions take place at posterior variances below 0.5, indicated by the vertical line.

Figure A9: Estimated CCPs



Notes: Simulated exploration probabilities using the estimated CCPs. Probabilities are simulated for a firm with a single block and one neighboring block held by another firm. Oil price is set to the average value in the data.

Figure A10: True and Simulate Prices



True Prices Three Simulated Paths

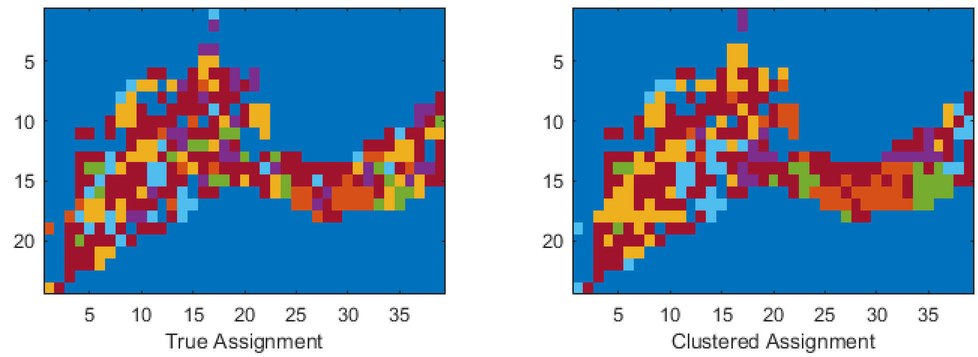
Notes: Price is WTI price converted to 2011 $ using the UK production price index for manufacturing and the UK/US exchange rate.

Table A8: Model Fit

	Data	Simulation	
		First Step Probabilities	Equilibrium Probabilities
Baseline Model			
Exploration Wells	1608	1774.4	1544.4
Blocks Developed	45	54.7	33.7
Exp. Wells/Dev.	35.7	30.8	46.1

Notes: Data column excludes multiple wills on same block on same date. Column 1 records statistics from the data covering 1964-1990 for the relevant region. Columns 2 and 3 are averages over 50 simulations that cover 1964-1990. For each month the assignment of blocks to firms and the oil price in the simulations are set at their realized values. Simulations in column 2 draw firm actions using the first step estimates of the conditional choice probabilities. Well outcomes and development revenue are drawn from the estimated Gaussian process. Simulations in column 3 use equilibrium conditional choice probabilities at the estimated parameter values.

Figure A11: Clustered Licenses



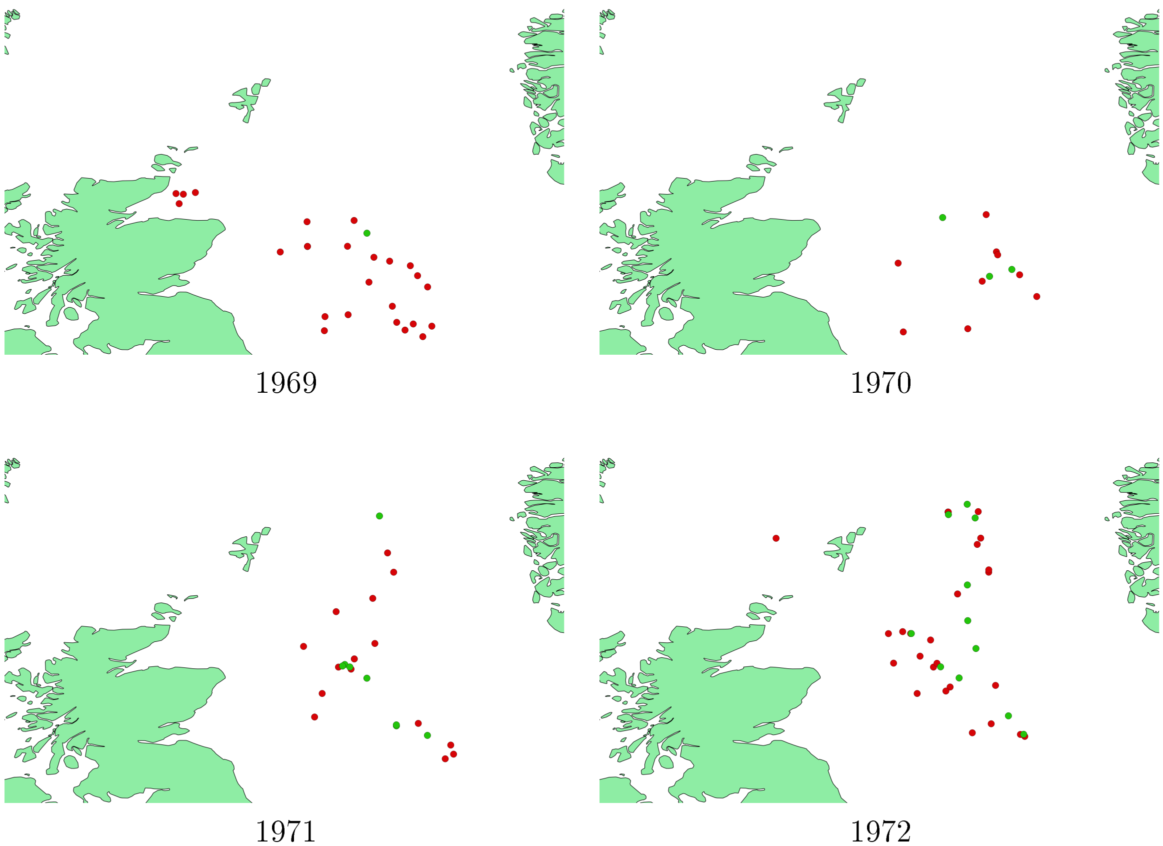
Notes: Left panel illustrates the location of drilling licenses for the five largest firms in January 1975 on the region of the North Sea used for structural estimation. Orange corresponds to Total, green to Conoco, yellow to Shell, purple to BP, and light blue to Amoco. Red blocks are licensed to other firms, and dark blue blocks are unlicensed. The right panel illustrates the counterfactual license assignment constructed using the clustering algorithm discussed in Appendix F. Table records statistics averaged over 40 simulations.

Table A9: Simulations: No Price Volatility

	Baseline	No Free Riding	Info Sharing	No Info Sharing	Info Sharing (Equilibrium)	No Price Shocks	Monopoly	Clustered
Exp. Wells	1556.35	1660.00	1575.90	1540.75	1566.35	1549.40	2030.25	1555.15
	[7.74]	[10.42]	[7.88]	[10.89]	[7.28]	[8.74]	[9.87]	[7.39]
Blocks Dev.	27.30	48.80	35.10	23.55	36.45	31.30	68.85	34.40
	[1.21]	[2.96]	[7.88]	[1.19]	[1.53]	[1.24]	[3.48]	[1.82]
Exp. Wells/Dev.	59.47	37.43	46.65	69.97	44.56	51.23	31.06	47.83
	[2.90]	[3.23]	[46.65]	[4.66]	[1.96]	[2.27]	[1.64]	[2.65]
Total Surplus								
Discounted	2.97	6.58	4.61	1.66	37.48	3.05	10.02	3.52
	[0.35]	[0.73]	[4.61]	[0.19]	[3.39]	[0.29]	[1.14]	[0.44]
Not Discounted	27.63	48.62	46.65	20.09	4.27	28.98	60.78	31.95
	[2.20]	[4.59]	[46.65]	[1.56]	[0.45]	[2.00]	[5.74]	[2.94]

Notes: Results are averages over 40 simulations that cover 1964-1990. The assignment of blocks to firms are set at their realized values. Price is held fixed at the long run average of the AR process. Well outcomes and development revenue are drawn from the posterior of estimated Gaussian process using all observed wells. Revenue and profits are in billions of 2015 dollars. Total surplus is firm and government revenue less costs, including the value of cost shocks for every period. PDV revenue and profit are 1964 values where the annual discount factor is 0.9. Simulation standard errors are in square brackets.

Figure A12: Maps of Early Exploration



Notes: Each map plots the location of exploration wells drilled that year. Red points are unsuccessful wells and green points are successful wells.