Information Externalities, Free Riding, and Optimal Exploration in the UK Oil Industry

Charles Hodgson*

November 10, 2018

[Job Market Paper: Please Click HERE for the Latest Version]

Abstract

Information spillovers between firms can reduce the incentive to invest in R&D if property rights do not prevent firms from free riding on competitors' innovations. Conversely, strong property rights over innovations can impede cumulative research and lead to inefficient duplication of effort. These effects are particularly acute in natural resource exploration, where discoveries are spatially correlated and property rights over neighboring regions are allocated to competing firms. I use data from offshore oil exploration in the UK to quantify the effects of information externalities on the speed and efficiency of exploration by estimating a dynamic structural model of the firm's exploration problem. Firms drill exploration wells to learn about the spatial distribution of oil and face a trade-off between drilling now and delaying exploration to learn from other firms' wells. I show that removing the incentive to free ride brings exploration forward by about 1 year and increases industry surplus by 31%. Allowing perfect information flow between firms raises industry surplus by a further 38%. Counterfactual policy simulations highlight the trade off between discouraging free riding and encouraging cumulative research - 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.

Department of Economics, Stanford Univeristy, 579 Serra Mall, Stanford, CA 94305 Email: chodgson@stanford.edu

*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, Jackson Dorsey, Linda Welling, and participants at IIOC 2018 and WEAI 2018 for helpful comments on this project. Thanks also to Oonagh Werngren and Jen Brzozowska for help with sourcing the data.

1 Introduction

The incentive for a firm to invest in research and development depends on the extent to which it can benefit from the investments of its competitors. If the knowledge generated by R&D, such as new technologies, the results of experiments, or the discovery of mineral deposits, is publicly observable, firms may have an incentive to *free ride* on their competitors' innovations, for example by introducing similar products or mining in locations near their rivals' discoveries. When each firm would rather wait to observe the results of other firms' research than invest in R&D themselves, the equilibrium rate of innovation can fall below the socially optimal level (Bolton and Harris, 1999). On the other hand if information flow between firms is limited, for example by property rights on existing innovations, the progress of research may be slowed because of inefficient duplication and the inability of researchers to build on each other's discoveries (Williams, 2013).

The growth of knowledge and the generation of new ideas are the most important drivers of economic growth (Romer, 1990; Jones, 2002), and inefficiencies in the rate of innovation have potentially significant economic effects. 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, patent law assigns property rights over innovations so that firms who profit from an innovation must compensate the inventor for their research investment. 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).

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 a 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.[1] I use data covering the history of offshore drilling in the UK between 1964 and 1990 to 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 others' wells, and the spatial arrangement of drilling licenses assigned to different firms.

I start by measuring the spatial correlation of well outcomes. I fit a logistic Gaussian process model to data on the locations and outcomes of all exploration wells drilled before 1990. This model allows binary outcomes - wells are either successful or unsuccessful - to be correlated across space. The estimated Gaussian process can be used as a Bayesian prior that embeds spatial learning. When a successful or unsuccessful well is drilled, the implied posterior beliefs about the probability of finding oil are updated at all other locations, with the perceived probability at nearby locations updating more than at distant locations. The updating rule corresponds to a geostatistical technique for interpolating over space that is widely used in natural resource exploration.

The estimated spatial correlation indicates that the results of exploration wells should have a significant effect on beliefs about the probability of well success at distances of up to 50 km. To test whether firm behavior is consistent with this spatial correlation, I regress firm drilling decisions on past well results. I find that firms' probability of exploration at a location is significantly increasing in the number of successful past wells and significantly decreasing in the number of unsuccessful past wells. The response declines in distance in line with the measured spatial correlation. Firms' response to the results of their own past wells is 2 to 5 times as as large as their response to other firms' wells, suggesting imperfect information flow between firms.

Next, I measure how exploration probability varies with the spatial distribution of property rights. Drilling licenses are issued to firms on 22x18 km blocks. I find that the monthly probability of exploration on a block increases by 0.8 percentage points when the number of nearby blocks licenses to the same firm is doubled and decreases by 0.4 percentage points when the number of nearby blocks licensed to other firms is doubled. These effects are statistically and economically significant and consistent with the presence of a free riding incentive - firms are less likely to explore where there is a greater potential to learn from other firms' exploration.

Together, these descriptive findings suggest that information spillovers over space and between firms play an important role in firms' exploration decisions. To measure the effect of these externalities on equilibrium exploration rates and industry surplus I incorporate the model of spatial beliefs into a structural model of the firm's exploration problem. Firms face a dynamic discrete

[1] 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."

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 firm's wells with some probability, $\alpha \in [0, 1]$, and update their beliefs about the spatial distribution of oil.

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 needs to form beliefs about the outcomes of unobserved wells and about other firms' beliefs. To make estimation of the model and computation of equilibria feasible I adopt the simplifying assumption that firms believe blocks held by other firms are explored at a fixed rate which is equal to the true average probability of exploration in equilibrium. This removes certain strategic incentives - for example the incentive to signal to other firms through drilling - but leaves in tact the asymmetric information structure and the incentives I am interested in measuring. In particular, 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 estimated value of the spillover parameter, α, indicates that firms observe the results of other firms' wells with 37% probability. The presence of substantial but imperfect information spillovers means that equilibrium exploration behavior could be affected by both free riding - since firms observe each other's well results and have an incentive to delay exploration - and inefficient exploration - since spillovers are imperfect, each firm has less information on which to base its drilling decisions than the set of all firms combined.

I perform counterfactual simulations to quantify these two effects. First, I remove the incentive for firms to free ride and simulate counterfactual exploration and development behavior. I find that exploration and development is brought forward in time by about one year, increasing the number of exploration wells drilled between 1964 and 1990 by 7.4%. Removing free riding increases the 1964 present discounted value of 1964-1990 industry surplus by 31%. Next, I allow for perfect information sharing between firms, holding firms' incentive to free ride fixed at the baseline level. The number of exploration wells increases by 12.6% and the efficiency of exploration increases substantially - since firms can perfectly observe each other's well results, cumulative learning is faster. The number of exploration wells per block developed falls and exploration wells are more concentrated on productive blocks. Industry surplus is 70% higher than the baseline in this information sharing counterfactual.

I next ask to what extent these inefficiencies could be mitigated through alternative property rights. Under the current regulations in the UK, data from exploration wells is property of the firm for 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. When the confidentiality window is increased to 10 years, the increase in the exploration rate dominates the reduction in exploration efficiency and industry surplus increases by 11%. When the confidentiality window is reduced to 0, the increased the speed of learning and efficiency of exploration overcomes the free riding effect, and industry surplus increases by 57%. Although a marginal increase in window length would increase surplus, the free riding effect is sufficiently small such that it is optimal for well data to be released immediately.

Finally, I show how the spatial distribution of property rights affects exploration incentives. When each firm's drilling licenses neighbor fewer other-firm licenses the incentive for firms to delay exploration is reduced and the value to firms of the information generated by their own wells is greater. 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 the number of exploration wells drilled increases by 8% and the number of exploration wells per developed block falls from 22.45 to 18.9. I do not claim that this is the optimal arrangement of property rights, so these figures represent a lower bound on the possible effect of spatial reorganization.

The results highlight the tension between discouraging free riding and encouraging efficient cumulative research in the design of property rights over innovations. In this setting, there are ranges of the policy space in which strengthening property rights leads to a marginal improvement in surplus and ranges where weakening property rights 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. The quantitative results on the spatial assignment of licenses can be thought of as an example of decentralized research where a principal (here, the government) assigns research projects to independent agents (here, firms). The results suggest that there are significant gains from assignments of projects that minimize the potential for information spillovers across agents. This finding could be applied to, for example, publicly funded research efforts that coordinate the activity of many independent scientists.

This paper contributes to the large literature on firms' incentives to conduct R&D (Arrow, 1971; Dasgupta and Stiglitz, 1980; Spence, 1984). In particular, I build on recent papers that ask whether and to what extent intellectual property rights hinder subsequent innovation (Murray and Stern, 2007; Williams, 2013; Murray et al., 2016). Both Williams (2013) and Murray et al. (2016) address this issue in a similar spirit to this paper, by focusing on specific settings where the set of possible research projects and cumulative nature of research is well defined, rather than looking at research in general and using metrics such as patent citations to measure cumulative innovation (see for example, Jaffe, Trajtenberg, and Henderson, 1993). 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.

The results in this paper also contribute to an existing empirical literature on the effect of information externalities in oil exploration. Much of this literature, summarized by Porter (1995) and Haile, Hendricks, and Porter (2010), has focused on bidding incentives in license auctions using data from the Gulf of Mexico. 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), who finds no evidence that firms are more likely to drill exploration wells after neighboring tracts are explored. The descriptive results I present are closest to those of Levitt (2016), who shows how exploration decisions respond to past well outcomes using data from Alberta and finds evidence of limited information spillovers across firms operating within the same region. I show how these spillovers vary with distance and the spatial distribution of drilling licenses.

Existing papers on oil and gas exploration that estimate structural models of the firm's exploration problem include Levitt (2009), Lin (2013), and Agerton (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. This allows me to simulate exploration paths under counterfactual policies which change the dependence of each firm's beliefs on the results of other firms' exploration wells, for example under different spatial assignments of blocks to firms.

Other closely related papers in the economcis of oil and gas exploration include Kellogg (2011), who provides evidence of learning about drilling technology, showing that pairs of oil production companies and drilling contractors develop relationship-specific knowledge, and Covert (2015), who investigates firm learning about the optimal drilling technology at different locations in North Dakota's Bakken Shale. 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 in 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 probabily methods, following Hotz and Miller (1993), Hotz, Miller, Sanders, and Smith (1994), and Bajari, Benkard, and Levin (2007). In particular, I extend these methods to a setting in which the econometrician is uninformed about each agent's information set. The procedure I propose to deal with this latent state variable is less generally applicable but less computationally intensive than the Expectation-Maximization procedure proposed by Arcidiacono and Miller (2011).

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 $55°N$.

2.1 Technology

Offshore oil production can be divided into two phases of investment and two distinct technologies. First, oil reservoirs must be located through the drilling of *exploration wells*. These wells are typically drilled from mobile rigs or drill ships and generate information about the geology under the seabed at a particular point, including the presence or absence of oil in that location. 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 sea bed by stilts or concrete columns with the capacity to extract large volumes of oil.

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". 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 exploring, 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.

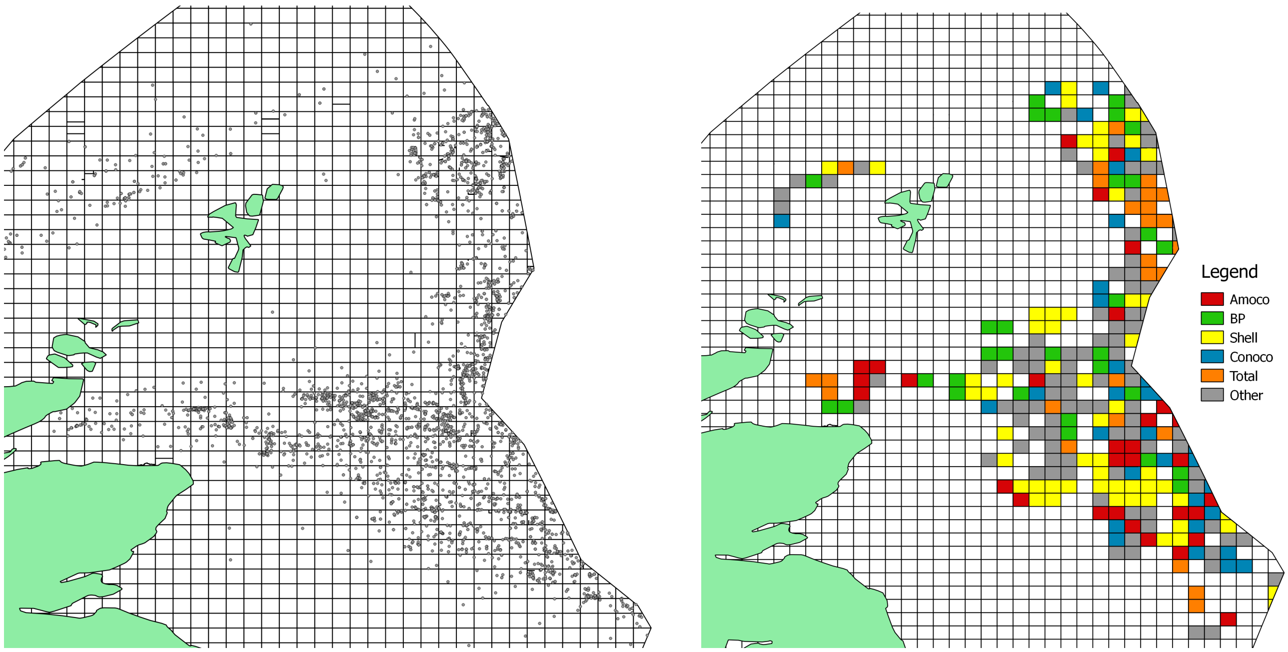
2.2 Regulation

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.[2]

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.

[2]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 got a whole bunch 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.

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.

I observe the history of license allocations for all blocks. In assigning blocks to firms I make two 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.[3] 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.[4] Second, licenses are sometimes issued over *parts* of blocks, splitting the original blocks into smaller areas that can be held by

[3]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.

[4]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.

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 A1 illustrates the distribution of licenses at the block-month level across firms.

A final set of regulations define 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 the first five years after a well is drilled. After this confidentiality period, well data is reported to the government and made publicly available. 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 investors, 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.

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. Once a block is developed, I drop it from the data.

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 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, over 10 exploration wells are drilled

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	.

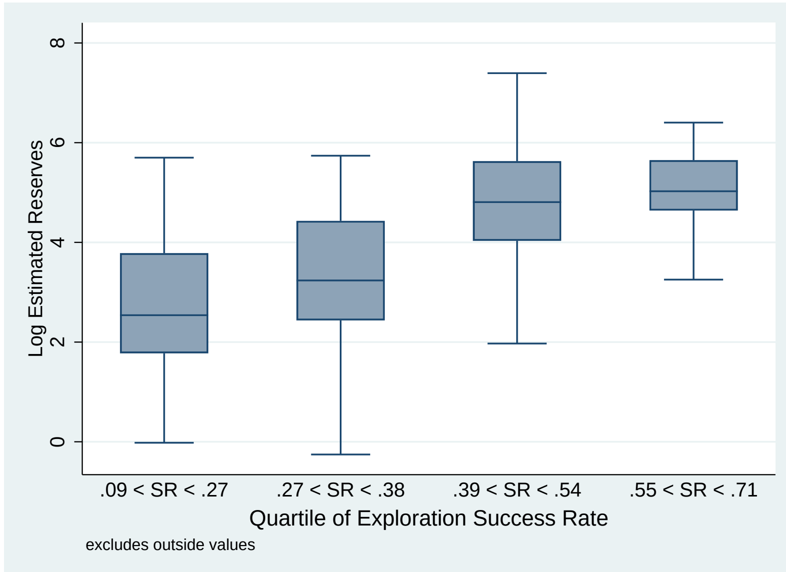
Notes: Table records statistics on all license-block pairs active between 1964 and 1990. In particular, if a block is licensed to multiple firms it appears multiple times in this Table. Each column records statistics on subsets of license-blocks defined according to whether they are ever explored or 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.

before a block is developed, and 2.3 exploration wells are drilled on blocks that are explored but not developed. The bottom row of Table 1 records the success rate of exploration wells across the different types of firm-block. 44% of exploration wells are successful on blocks that are eventually developed, while only 13% of wells are successful on blocks that are never developed. The success rate of exploration wells on a block is correlated withe 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. Figure 2 illustrates the distribution of estimated reserves in log millions of barrels over all developed blocks.[5] The distribution is plotted separately for four quartiles of the exploration success rate. There is a positive, approximately linear relationship between exploration success rate prior to development and log estimated reserves

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

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

Figure 2: 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 box plot markers record the lower adjacent value, 25th percentile, median, 75th percentile, and upper adjacent value. The distribution is plotted separately for four subsets of blocks defined by the quartiles of the pre-development exploration well success rate. A regression of log estimated reserves on success rate has a slope coefficient of 5.990 with a standard error of 0.964.

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 Appendix A 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 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 describing the true spatial correlation of oil deposits determined by underlying geology.

I then show how this statistical model can be used as a Bayesian prior about the distribution of oil.

If firms know the true parameter values, then the estimated model implies a Bayesian updating rule for firms with rational beliefs. In particular, firms' posterior beliefs about the probability of exploration well success at a given location are a function of past well outcomes at nearby locations. The true correlation of well outcomes informs the extent to which firms should make inferences over space when updating their beliefs after observing well outcomes. This model of spatial learning allows me to compute firms' posterior beliefs about the location of oil deposits after observing different sets of wells.

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 A2 for a one dimensional example). As I discuss further below, Gaussian processes are widely used in natural resource exploration to model the spatial distribution of geological features (see for example Hohn, 1999).

Formally, let $\rho(X) : \boldsymbol{X} \rightarrow [0,1]$ 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 $\boldsymbol{X}$.[6] 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 $\mathbb{R}$. Equation 1 is a logistic function that "squashes" $\lambda(X)$ so that $\rho(X) \in [0,1]$.[7]

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

[6] If well success rates were independent across locations j, a natural model would draw $\rho_j \in [0,1]$ from a beta distribution. However, it is likely that well outcomes are correlated across space. Indeed, the results presented below in Figure 6 indicate that firms' exploration decisions on block j respond to the results of exploration wells on nearby blocks. There is no natural multivariate analogue of the beta distribution that allows me to specify a covariance between ρ_j and ρ_k for $j \neq k$.

[7] If well success rates were independent across locations, a natural model would draw $\rho(X) \in [0,1]$ from a beta distribution. However, it is likely that well outcomes are correlated across space. There is no natural multivariate analogue of the beta distribution that allows me to specify a covariance between $\rho(X)$ and $\rho(X')$.

covariance matrix with (j,k) element $\kappa(X_j, X_k)$. The prior mean function $\mu : \boldsymbol{X} \to \mathbb{R}$ is assumed to be smooth and the covariance function $\kappa : \boldsymbol{X} \times \boldsymbol{X} \to \mathbb{R}$ 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:[8]

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

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 analogues of these statistics in the well outcome data. Most importantly, the estimated parameter

[8] This is a partial likelihood in the sense of Cox (1975). In Appendix B I provide a condition on the process that determines well locations X under which this is a valid likelihood function. See also chapter 13.8 of Wooldridge (2002). I use the hyperparameter estimation code provided by Rasmussen and Williams (2006) to implement the maximum likelihood estimation. The integral in equation 3 is approximated using Laplace's method. See section 5.5 of Rasmussen and Williams (2006) for details.

ℓ captures the true spatial correlation of exploration well outcomes.

Table 2: Oil Process Parameters

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

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. 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.

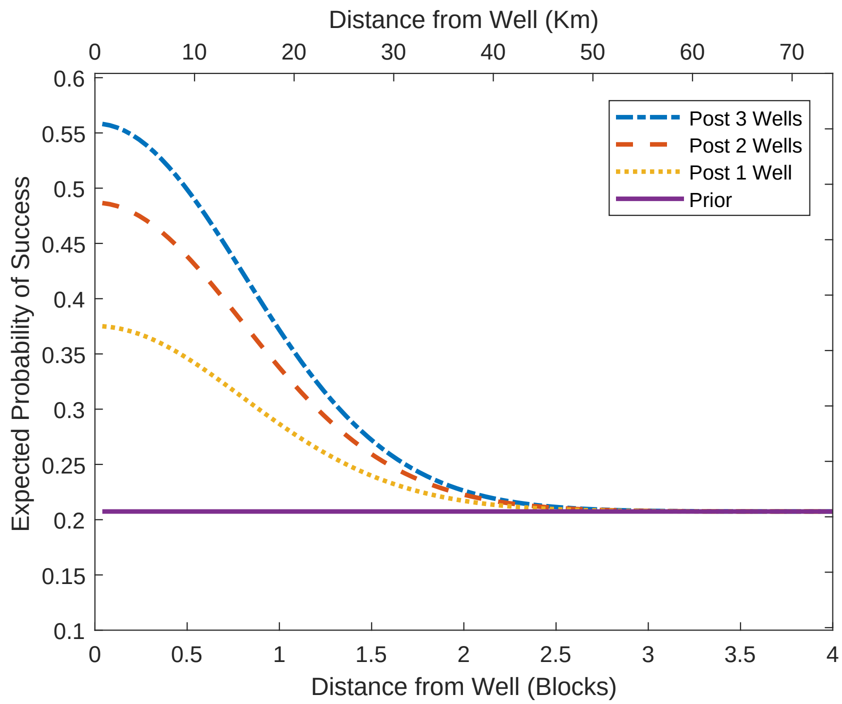
In particular, firms whose prior is described by $G(\rho)$ update their beliefs over the entire space $\boldsymbol{X}$ after observing a success or failure at a particular location X. Posterior beliefs at locations closer to X will be updated more than those at more distant locations. Figure 3 illustrates 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.[9] 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

[9] 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. In addition, as the maps in Appendix Figure A8 indicate, exploration did not begin in a particularly productive area, and the geographic focus of exploration shifted dramatically after the first early discoveries. For these reasons, I believe it is not unreasonable to adopt a constant prior mean.

observing two and three successful 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.[10]

Figure 3: 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

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

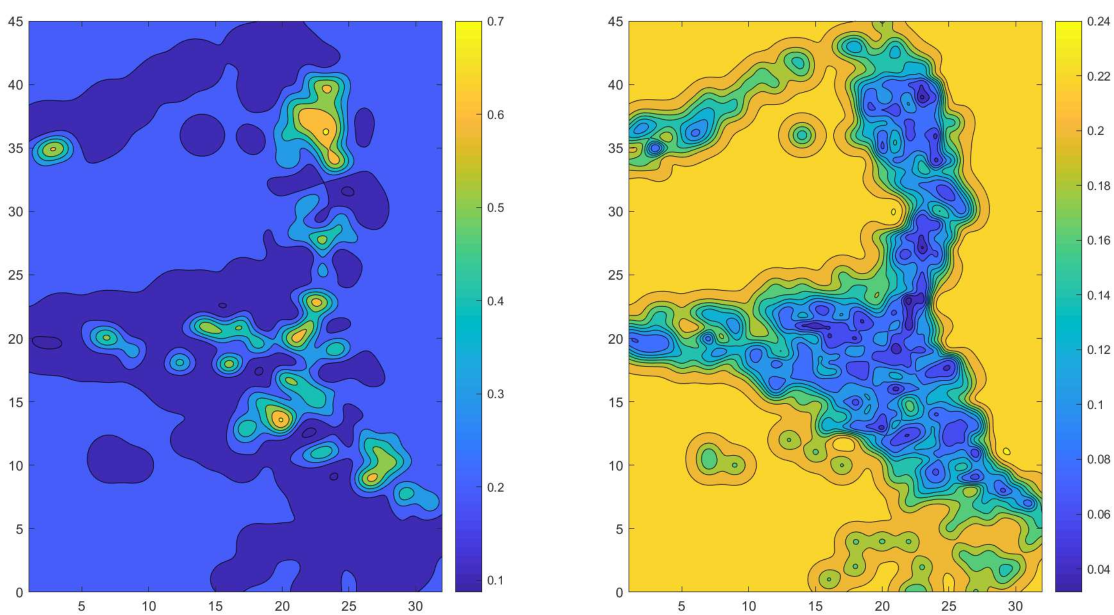
where $B(\cdot)$ is a Bayesian updating operator. Since the signals that firms receive are binary, there is no analytical expression for the posterior beliefs given the Gaussian prior and the observed signals.

[10]In Appendix Figure A3 I illustrate belief updating under different values of ℓ in a numerical example.

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 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. Figure 4 is a map of posterior beliefs for a firm that observed the outcome of *all* exploration wells drilled from 1964-1990. 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. In general, the standard deviation of posterior beliefs is lower in regions where more exploration wells have been drilled.[11]

Figure 4: 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.

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 is known in the geostatistics literature as Kriging (see for example standard geostatistics textbooks

[11] This is not necessarily the case everywhere. In particular, if the realized outcome of a well at location X is unlikely given prior beliefs, posterior variance around X can increase.

such as Hohn, 1999). Kriging is a widely applied statistical technique for making predictions about the distribution of geological features, including oil deposits, over space. Standard Kriging of a continuous variable corresponds exactly to Bayesian updating of a Gaussian process with continuous, normally distributed signals. 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*.

In addition to being representative of industry techniques, the model of spatial beliefs is closely linked to the literature on Gaussian processes in machine learning, as summarized by Rasmussen and Williams (2006). In this literature, optimal Bayesian learning based on Gaussian process priors is used to construct algorithms for efficiently maximizing unknown functions. In a close analogue to the machine learning problem studied by, for example, Osborne et al. (2009), exploration wells can be thought of as costly evaluations of a function mapping geographical locations to the presence of oil, with the firm's problem being to locate the largest oil deposits at minimum cost. The logistic Gaussian Process model of beliefs is a flexible (in terms of covariance and mean function specification) and computationally tractable model of spatial updating of beliefs with binary signals that is applicable to settings beyond oil exploration. See for example Hodgson and Lewis (2018) on learning in consumer search.

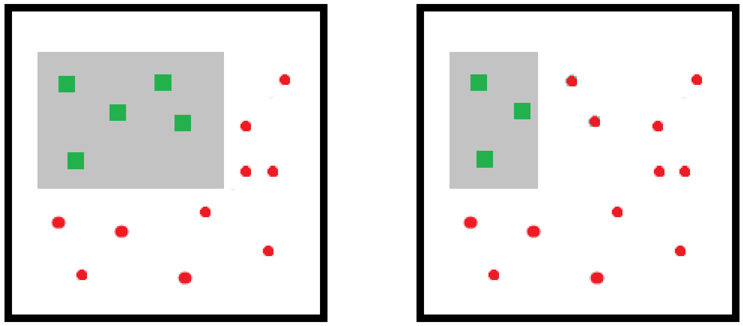
3.3 Beliefs and Development Payoffs

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 random.[12] 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. For example, Figure 5 illustrates a stylized example in which wells have been drilled at random locations within two blocks. In the left block, the oil field occupies one-third of the area, and in the right block, the oil field occupies one-fifth of the area. The success rates, indicated by the ratio of green wells to all wells, are equal to the sizes of the oil fields - with one third of wells successful on the left block and one fifth successful on the right block.

Formally, I assume that the potential oil revenue yielded by block j, π_j, is drawn from a distribution

[12] In particular, that well coordinates are drawn from a uniform distribution over the area of the block.

Figure 5: 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.

$\Gamma(\pi|\rho_j, P)$ 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 positive relationship between the realized exploration success rate and estimated oil reserves on developed blocks, illustrated by Figure 2. 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 might not be true if, for example, firms continue to drill near previous successful wells *within* the block. I test this implication 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. That is, the effect of the well sequence number on success probability is not statistically significant. This is consistent with a model in which within-block well locations are drawn at random.

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.

I provide evidence that firms respond to the results of past wells, both their own wells and those of other firms, in a way that is consistent with the estimated spatial correlation of well results. I then use the estimated model of beliefs to quantify the free riding incentive faced by firms operating in the North Sea. I provide direct evidence of free riding by showing how drilling behavior changes when the spatial arrangement of licenses changes.

4.1 Exploration Drilling Patterns

The estimated spatial correlation illustrated by Figure 3 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. $Fail_{jdot}$ is analogously defined as the cumulative number of past unsuccessful wells. To provide suggestive evidence of the extent to which firms' exploratory drilling decisions are correlated with the results of past wells drilled by different firms at different locations, I estimate the following regression specification using OLS:

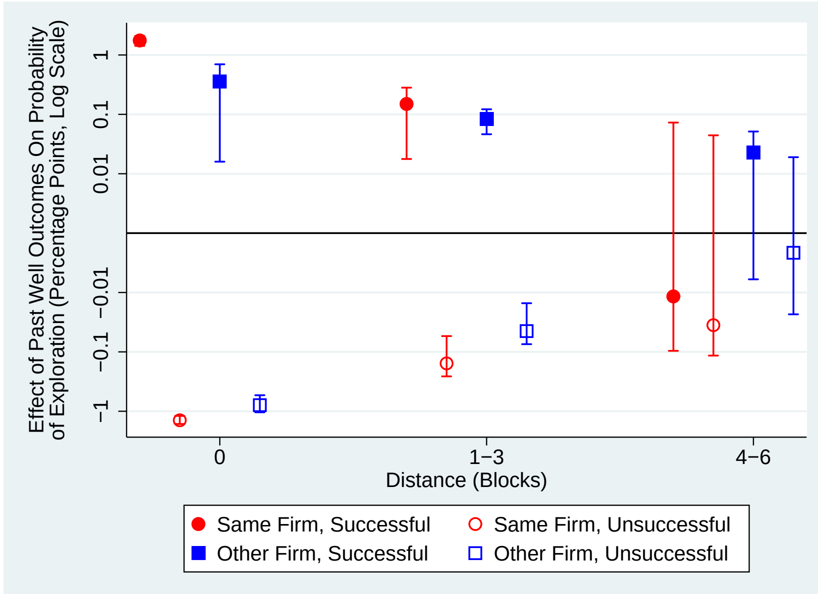
$$Explore_{fjt} = \alpha_f + \beta_j + \gamma_t + \sum_d \sum_{o \in \{f, -f\}} g_{do} \left(Suc_{jdot}, Fail_{jdot} \right)) + \epsilon_{fjt}. \tag{6}$$

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, block, and month 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 6 records the estimated marginal effect of an the first past well of each type on the probability of 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 an additional unsuccessful wells significantly decrease the probability of subsequent exploration.

The effect of an additional same firm, same block well is approximately 120% of the mean of the dependent variable, $Explore_{fjt}$, which is 0.0161, and the size of the effect is roughly equal for successful and unsuccessful wells. The magnitude of the effect decreases with distance. Notice that the y-axis of Figure 6 is on a log scale. The effect of past wells at a distance of 1-3 blocks is

Figure 6: 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 6 where $g_{do}(\cdot)$ is quadratic in each of the arguments. Marginal effects are computed for the first well of each type. The y-axis is scaled by multiplying the effect by 10^4 and taking the log. Error bars are 95% confidence intervals computed using robust standard errors. All estimates are from one regression which includes quadratics in each of the 8 types of past well. The mean of the dependent variable is 0.0161. Sample includes block-months in the relevant region up to December 1990. 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.

about 10% of the effect of past same-block wells. The effect at distances of 4-6 blocks is on the order of 1% of the same-block effect and is not statistically significant.

Blue squares 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.[13]

These results suggest that firm's decisions about where to drill depend on the results of nearby past wells, both their own wells and those of their rivals. The probability of drilling on block j responds both to the results of past wells on block j as well as to the results of wells on nearby blocks, suggesting that firms make inferences across space at distances consistent with the spatial correlation of well results illustrated by Figure 3, with the size of the drilling response declining with distance. Exploration probability is also more responsive to own-firm exploration results than to other-firm exploration results, suggesting that information flow across firms is imperfect.[14]

[13] 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.

[14] One potential concern is that these results could be explained by the arrival over time of public information

In Appendix Table A6 I report 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. Firms do not appear to have been systematically over- or under-extrapolating across space during early exploration. This finding 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.

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.[15] 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, based on all wells drilled before that month according to the Bayesian updating rule (4).[16] 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 Kulback-Leibler divergence, $KL_{j,t}$, between the prior and posterior distributions following an additional exploration well for each (j, t).[17]

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 $E_t(\rho_j) = 0.48$. 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}. It is clear that as variance goes to 0, the change in beliefs from an additional well will also go to 0.

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

that is independent of drilling results and is correlated over space. To test of whether the information generated by past wells is driving these results, I use the fact that the confidentiality period on exploration data expires 5 years after a well is drilled. In Appendix Figure A4 I show that moving an successful other-firm well *back* in time by more than 6 months has a positive and significant effect on the probability of exploration. The effect is greatest for wells close to the confidentiality cutoff, drilled between 4.5 and 5 years ago. For wells that are older than 5 years, there is no significant effect, consistent with the outcomes of these wells already being public knowledge.

[15] This prediction is true in the simple model presented in Appendix A. In more general settings, it is not necessarily the case that more informative wells are always more valuable. Note that the value of an exploration well is not just the amount of information it generates, but its effect on the firm's future behavior and payoffs.

[16] 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.

[17] The KL divergence is a measure of the difference between two distributions. It can be interpreted as the information gain when moving from one distribution to another (see Kullback and Leibler, 1951, and Kullback, 1997). See Appendix B for details.

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. This is confirmed by the results in column 3, which presents the estimated positive and significant coefficient from a regression of $Explore_{fjt}$ on KL_{jt}.

Table 3: Response of Drilling Probability to Posterior Beliefs

Dependent Variable:	KL Divergence	Exploration Well		Develop Block
Posterior Mean	.547***	.275***	.	.011***
	(.001)	(.062)	.	(.003)
Posterior Mean^2	-.570***	-.188**	.	.
	(.002)	(.089)	.	.
Posterior Variance	.092***	.029***	.	.001
	(.000)	(.008)	.	(.001)
KL Divergence	.	.	.190***	-.039***
	.	.	(.070)	(.010)
R^2	.914	.045	.043	.077
N	95690	95330	95330	93569
Firm-Block and Month FE	No	Yes	Yes	No
Firm-Month FE	No	No	No	Yes

Notes: Standard errors clustered at the firm-block 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.

The last column of Table 3 present the results of a regression with $Develop_{fjt}$, an indicator for whether firm f developed block j in month t, as the dependent variable. As illustrated in Figure 2, a block's exploration well success rate is positively correlated with size of the oil field located on that block. Consistent with this, the results indicate that probability of development is increasing in $E(\rho_j)$. In contrast to the exploration results there is a negative effect of KL_{jt} on development - the more information could be generated by an additional exploration well on a block, the less likely is a firm to develop that block.[18]

4.2 The Value of Information and the Incentive to Free Ride

The results presented in Section 4.1 suggest that information spillovers across space and firms have a significant effect on drilling behavior. To what extent do these externalities provide an incentive

[18] The development regression includes a firm-month fixed effect rather than a firm-block fixed effect because development happens at most once within each (f, j), at the end of that firm-block's time series. Results with firm-block fixed effects would therefore capture the fact that variance and KL_{jt} tend to decline over time.

for firms to delay exploration and free ride off the information generated by other firms' wells? Using the estimated model of beliefs, it is possible to perform a back of the envelope quantification of the incentive to delay exploration without invoking a further structural model of firm behavior.

I consider a firm f's decision to delay drilling the first exploration well on block j by one year. I suppose that the firm's beliefs are given by the estimated prior process and that, each month, each block held by another firm is drilled with a fixed probability Q^E, which I set equal to the empirical mean exploration rate of 0.0219. I further assume that firm f observes the results of each well drilled by another firm with probability α. For a given arrangement of licenses, I run twelve month simulations of other firms' drilling behavior and update the beliefs of firm f. For each simulation, I calculate the information gained about block j by firm f from observing the results of other firms' wells, and compare the mean information gain across simulations (in particular, the expected Kullback-Leibler divergence between the firm's prior beliefs and the posterior after 12 months) to the expected information gain from firm f drilling its own exploration well on block j.

Table 4: Information Gain from Delay of Exploration

	Other Firm Neighbors			One Year Delay at $\alpha = 0.4$	
Percentile	Same Block	First Degree	Second Degree	Info. Generated	Net Gain
1	0	0	0	0	-43.02
25	0	3	5	0.080	-15.42
50	0	5	9	0.120	-1.51
75	0	7	12	0.174	17.23
90	1	8	13	0.335	72.67
99	2	14	22	0.603	165.45

Notes: The first three columns report percentiles of the distribution of other firm neighbors across all (f, j, t) observations in the relevant area from 1964-1990. First and second degree neighbors are those one or two blocks away (including diagonal neighbors). Columns 4 reports the mean information generated from 1000 12 month simulations, as described in the text. Column 5 presents the implied net gain in millions of dollars from delaying exploration for 12 months, as described in the text.

Table 4 presents the expected information generated from 12 month delay as a fraction of the information generated by drilling an exploration well for six different arrangements of licenses. Each row corresponds to a license arrangement where the numbers of other firms holding licenses at different distances from block j are drawn from percentiles of the empirical distribution. The fourth column records the information generated from one year of delay when $\alpha = 0.4$, as a fraction of the information generated by drilling one exploration well. The information gain from delay is increasing in the density of other firm neighbors. For the 25th percentile arrangement, delaying exploration by one year generates 8% of the information of an exploration well. For the 99th percentile arrangement, delay achieves 60% of the information generation of an exploration well. The fifth column records an approximation of the net gain in millions of dollars from delaying exploration by one year, suggesting that firms with an arrangement of neighboring licenses in the 1st, 25th, and 50th percentiles would not benefit from delay, while firms above the 75th percentile would gain on net.[19] To illustrate how these incentives change with the flow of information between firms, Appendix Figure A7 records the net gain from delay for different license arrangement percentiles and for values of $\alpha \in [0, 1]$. The gain from delay is increasing in α.

These results suggest that, if there is sufficient flow of information between firms, variation in spatial arrangement of licenses in the data should result in changes in the incentive to free ride by delaying exploration. To provide direct empirical evidence that such free riding incentives matter, I run regressions exploiting the variation in the spatial arrangement of licenses.

The number of licensed blocks in a region is likely to be correlated with, for example, the arrival of information that is not captured by well outcomes or changes in region specific drilling costs. To isolate the causal effect of changes in license distribution on the incentive to explore, I focus on quasi-experimental variation by selecting (f, j, t) observations before and after discrete jumps in the number of licenses issued, corresponding to the months before and after the government announces the results of licensing rounds. In particular, I identify (f, j, t) observations for which the total number of licensed blocks neighoring block j increases from the previous month. I select nine month windows centered on these licensing events and index these windows with γ. For observations in a licensing window, I define $\Delta(f, j, t) \in \{-4, -3, ..., 4\}$ as the number of months before or after the relevant licensing event. I estimate the following specification on the set of observations in licensing windows:

$$Explore_{fjt} = \alpha_\gamma + \alpha_{\Delta(f,j,t)} + \beta_1 BlocksOwn_{fjt} + \beta_2 BlocksOther_{fjt} + X_{fjt}\delta + \epsilon_{fjt}. \quad (7)$$

Where X_{fjt} contains all the regressors in equation 6. $BlocksOwn_{fjt}$ is the number of neighboring blocks licensed to firm f and $BlocksOther_{fjt}$ is the number of neighboring blocks licensed to other firms. The change in the number of licensed blocks near block j *within* a window is unlikely to reflect the arrival of new information about the productivity of block j, since issued licenses are the result of applications that are made *before* the beginning of the window. Any changes in drilling costs or arrival of information within each window is therefore likely uncorrelated with changes in $BlocksOwn_{fjt}$ and $BlocksOther_{fjt}$.

The first column of Table 5 reports the coefficients on $BlocksOwn_{fjt}$ and $BlocksOther_{fjt}$. Within-window increases in the number of own-firm blocks are correlated with increased exploration probability, and within-window increases in the number of other-firm blocks are correlated with decreased exploration probability. The second column reports results using the log of $BlocksOwn_{fjt}$

[19] Suppose the information generated from delay as a share of one well is s. If the cost of drilling an exploration well is c, then delaying the first exploration well reduces the expected cost of exploration by sc. The cost of delay is the resulting discounting of future profits, V. If the annual discount rate is β, then I compute the net gain from delay as $sc - (1 - \beta)V$. I set $\beta = 0.9$. I set $V = 43.02$ based Hunter's (2015) account of the per-block auction revenue generated by one-off auction licensing round held by the UK regulator in 1971, inflated to 2015 dollars. I set $c = 34.55$ based on the average per-well capital expenditure between 1970 and 2000 reported by the regulator, inflated to 2015 dollars.

Table 5: Regressions of Drilling Probability on Nearby Licenses

	Exploration Well			Develop Block
$BlocksOwn_{fjt}$	4.739	.	3.300***	-.101
	(5.800)	.	(.961)	(.256)
$BlocksOther_{fjt}$	-1.446	.	.915***	-.059
	(1.330)	.	(.267)	(.064)
$log(BlocksOwn_{fjt})$	.	.028**	.	.
	.	(.014)	.	.
$log(BlocksOther_{fjt})$	.	-.013***	.	.
	.	(.004)	.	.
N	21971	21618	136430	136430
Firm-Block, and Month FE	No	No	Yes	Yes
Experiment Fixed Effects	Yes	Yes	No	No
Coefficients Scaled by 10^3	Yes	No	Yes	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 event, for which the firm f has held a license on block j for at least 6 months. Block counts are of all licenses on block j and neighboring blocks on date t. *** indicates significance at the 99% level. ** indicates significance at the 95% level. * indicates significance at the 90% level.

and $BlocksOther_{fjt}$, with both coefficients significant and of the same sign as in the first column. These results suggest that doubling the number of neighboring blocks licensed to firm f will increase the probability of exploration by firm f on block j by 0.8 percentage points, and doubling the number of blocks licensed to other firms will reduce the probability of exploration by 0.4 percentage points. Notice that these effects are large relative to the mean of the dependent variable, which is 0.016 in this sample. This finding is suggestive of a significant incentive to delay investment in exploration when the probability that another firm will explore nearby increases. In particular, changes in the number of blocks licensed to other firms should not change the value to firm f of the results of exploration on block j, but can increase the value of *delaying* exploration.

The third and fourth columns of Table 7 presents regressions of $Explore_{fjt}$ and $Develop_{fjt}$ on $BlocksOwn_{fjt}$ and $BlocksOther_{fjt}$ that do not restrict the sample to licensing windows. Notice that the probability of exploration is increasing in both measures of nearby licenses, but the effect of $BlocksOwn_{fjt}$ is substantially larger. The distribution of licenses neighboring block j is not significantly correlated with the probability that block j is developed. It seems reasonable that a firm would not delay development on a block known to hold large reserves because of expected exploration by rivals on nearby blocks, and the revenue produced by a development well is not a function of the number of surrounding blocks owned by the same firm.[20]

[20] One exception to this 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 revenue is split proportionally between operators of the blocks. This provision removes the "common pool" incentive discussed by Lin (2013) and the incentive to develop an overlapping reservoir before a neighboring rival.

In Appendix Figures A5 and A6 I present further evidence that is suggestive of free riding. In particular, I reproduce a result from Hendricks and Porter (1996), who showed that the probability of drilling an exploration well on unexplored tracts in the Gulf of Mexico increased near the drilling deadline imposed by the tract lease. The authors argue that this delay until the end of the lease term is evidence of a free riding incentive. I show that the same pattern obtains on North Sea blocks when the drilling deadline (which, as discussed in Section 2, is not as strict as the deadline imposed in the Gulf) approaches. I also show that this pattern obtains for license blocks with a large number of other firm license nearby, but is not present for blocks that are far from other firm licenses, consistent with the predictions presented in Table 4.

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. Motivated by the empirical findings described in Section 3, I then extend the model to one of asymmetric information in which firms do not observe the results of other firms' wells with certainty. I describe a simplifying assumption on firm beliefs and specify an equilibrium concept that makes estimation of the asymmetric information game feasible.

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 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 equation X. 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 $G_{ft+1} = B(G_{ft}, W_t)$ and $G_{ft} = G_t$ for all firms $f \in F$, where $B(\cdot)$ is defined in equation 4.

The industry state at date t is described by

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

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 which I allow to depend on the state, $c(j, \mathcal{S}_t) - \epsilon_{ftj}$. Developing block j incurs a cost $\kappa - \nu_{ftj}$. ϵ_{ftj} and ν_{fjt} are private information cost shocks drawn iid from logistic 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.
3. Exploration well results W_t are realized.
4. The industry state evolves to $\mathcal{S}'_t = \{G_{t+1}, \{J_{ft}\}_{f \in F \cup \{0\}}, P_t\}$.

Development Stage

1. Given state $\mathcal{S}'_t$, 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.
3. If $a^D_{ft} = j$ then the firm f draws oil revenue π_{jt}.
4. The industry state evolves to $\mathcal{S}_{t+1} = \{G_{t+1}, \{J_{ft+1}\}_{f \in F \cup \{0\}}, P_{t+1}\}$.[21]

[21] Note that I have assumed that firms do not update their beliefs based on the outcomes of *development* decisions. Formally, this assumption means that although firms *obtain* revenues π_j after making development decisions, they do not *observe* π_j. The assumption that firms do not update their beliefs based on this realization is likely not unreasonable. In reality oil flow is obtained from a reservoir over many years, and additional information about the true size of the field is gradually obtained. Furthermore, since development platforms are very expensive, the information value of development is unlikely to be pivotal to the development decision, and the marginal effect of information revealed by the development outcome is likely to be small since development takes place only after extensive exploration.

State variables evolve at the end of the development stage as follows. I assume that log oil price follows an exogenous random walk, so $P_{t+1} = exp(log(P_t) + \zeta_t)$ 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:

$$V^E_f(\mathcal{S}_t) = E_{\boldsymbol{\epsilon}_{ft}} \left[\max_{a^E_t \in J_{ft} \cup \{0\}} \left\{ E_{\mathcal{S}'_t} \left[V^D_f(\mathcal{S}'_t) | a^E_t, \mathcal{S}_t \right] - c(a^E_t, \mathcal{S}_t) + \epsilon_{ftj} \right\} \right] \tag{9}$$

$$V^D_f(\mathcal{S}'_t) = E_{\boldsymbol{\nu}_{ft}} \left[\max_{a^D_t \in J_{ft} \cup \{0\}} \left\{ E_{\pi_{a^D_t}, \mathcal{S}_{t+1}} \left[\beta V^E_f(\mathcal{S}_{t+1}) + \pi_{a^D_t} | a^D_t, \mathcal{S}'_t \right] - \kappa(a^D_t | \mathcal{S}'_t) + \nu_{ftj} \right\} \right].$$

Where β is the one period discount rate. The inner expectation in the exploration Bellman equation is taken over realizations of the intermediate state $\mathcal{S}'_t$, with respect to the firm's beliefs G_t and beliefs about other firms' exploration actions. The inner expectation in the development Bellman equation is taken over realizations of development revenues π_{a^D} and realizations of next period's state variable $\mathcal{S}_t$, with respect to the firm's beliefs G_{t+1} and beliefs about other firms' actions.

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

$$\begin{aligned} v^E_f(a^E_t, \mathcal{S}_t) =& E_{\mathcal{S}'_t} \left[V^D_f(\mathcal{S}'_t) | a^E_t, \mathcal{S}_t \right] - c(a^E_t, \mathcal{S}_t) \\ v^D_f(a^D_t, \mathcal{S}'_t) =& E_{\pi_{a^D_t}, \mathcal{S}_{t+1}} \left[\beta V^E_f(\mathcal{S}_{t+1}) + \pi_{a^D_t} | a^D_t, \mathcal{S}'_t \right] - \kappa(a^D_t, \mathcal{S}'_t). \end{aligned} \tag{10}$$

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

$$a^E_f(\mathcal{S}, \boldsymbol{\epsilon}) = \arg \max_{a^E \in J_f \cup \{0\}} \left\{ v^E_f(a^E, \mathcal{S}) + \epsilon_{ta^E} \right\} \tag{11}$$

$$a^D_f(\mathcal{S}', \boldsymbol{\nu}) = \arg \max_{a^D \in J_f \cup \{0\}} \left\{ v^D_f(a^D, \mathcal{S}') + \nu_{ta^D} \right\},$$

where the firm forecasts all firms' actions conditional on the industry state using the true conditional choice probabilities (CCPs) given by:

$$P(a_f^E = j|\mathcal{S}_t) = \frac{exp\left(\frac{1}{\sigma_\epsilon} v_f^E(j, \mathcal{S}_t)\right)}{\sum_{k \in J_{ft} \cup \{0\}} exp\left(\frac{1}{\sigma_\epsilon} v_f^E(k, \mathcal{S}_t))\right)}. \tag{12}$$

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

5.2 Asymmetric Information

A key assumption made in the model described above is 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 the first five 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 periods after the well is drilled, on expiry of the confidentiality window.

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 set of well results observed by firm f in period t is

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

Firms observe the location, $j(w)$, and the drill date, $t(w)$, for all wells. This assumption reflects the fact that the regulator makes this data public immediately after a well is drilled. Firms f's information about past wells with unobserved outcomes is

$$W_{ft}^U = \{(j(w), t(w)) : o_f(w) = 0 \text{ and } t(w) > t - \tau\}. \tag{14}$$

The introduction of this asymmetric information structure complicates the firm's problem. In general, $G_{ft} \neq G_{gt}$ since firms observe different sets of well outcomes. To forecast next period's state in equilibrium, firm f must form beliefs about every other firm g's beliefs, G_{gt}. The history of firm g's actions is informative about G_{gt} and about well outcomes unobserved by firm f. Firm f should therefore update its beliefs based not only on observed outcomes, but on the past *behavior* of other firms. For instance, if firm g drilled many exploration wells on block j, this should signal to firm f something 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.

These complexities make estimating the asymmetric information game and finding equilibria computationally infeasible. To make progress, I impose the following simplifying assumption on firms' beliefs about other firms' actions.

- Assumption A1: Firm f believes that at every period t the probability of a new exploration well being drilled by a firm $g \neq f$ on block $j \in J_{gt}$ is given by $Q_t^E \in [0,1]$. Likewise firm f believes that at every period t the probability of firm $g \neq f$ developing block $j \in J_{gt}$ is $Q_t^D \in [0,1]$.

Assumption A1 says that firms believe that blocks held by other firms are explored at a fixed rate Q^E and developed at a fixed rate Q^D. Under this assumption I can redefine the state variable as:

$$\mathcal{S}_{ft} = \{G_{ft}, J_{ft}, \cup\{J_{gt}\}_{g\neq f}, J_{0t}, P_t, W_{ft}^U\}. \tag{15}$$

This firm-specific state is sufficient for firm f's date t decision under asymmetric information. Note that firm f only needs to know which blocks it holds and which are held by *some other* firm ($\cup\{J_{gt}\}_{g\neq f}$), not the identity of the license holding firm for each block, since the identity of the block owner does not affect drilling probability under firms' beliefs.[22] Further, $G_{ft+1} = B(G_{ft}, W_{ft})$ as before. In particular, G_{ft+1} does not depend on W_{ft}^U since firms believe past wells were drilled at an exogenous rate and drilling history does not contain information about other firms' beliefs. The state variable includes W_{ft}^U since firms anticipate the release of well outcome data at the end of each well's confidentiality period.

Fixing Q^E and Q^D, the firm's problem becomes a single agent problem where other wells are drilled at an exogenous rate. The firm's optimal strategy is given by equation 11 and CCPs are given by 12, where firm's expectations about the future actions of other firms are now given by (Q^E, Q^D), not the true CCPs. Fixing the initial conditions, defined by J_0 and P_0, and a value of (Q^E, Q^D), firms' optimal strategies imply probability distributions over realized states for each (f,t). I use these distributions to define equilibrium in the asymmetric information model as follows.

- Assumption A2: Let $P(a_{f,t}^E = j|\mathcal{S}_{f,t})$ and $P(a_{f,t}^D = j|\mathcal{S}'_{f,t})$ be firms' equilibrium CCPs. Fix a time horizon T. In equilibrium, firms have beliefs about other firms' exploration and

[22] Formally this requires additional assumptions on the stochastic process that governs the issuing and surrender of licenses. In particular, $P(j \in J_{ft+1}|\{J_{gt}\}_{g\in F\cup\{0\}}, a_{ft}^D) = P(j \in J_{ft+1}|J_{ft}, \cup\{J_{gt}\}_{g\neq f}, J_{0t}, a_{ft}^D)$, and $P(j \in \cup\{J_{gt+1}\}_{g\neq f}|\{J_{gt}\}_{g\in F\cup\{0\}}, \{a_{gt}^D\}_{g\in F}) = P(j \in \cup\{J_{gt+1}\}_{g\neq f}|J_{ft}, \cup\{J_{gt}\}_{g\neq f}, J_{0t}, \{a_{gt}^D\}_{g\in F})$. I also assume $J_{f0} = \{\}$ for all $f \in F$.

development rates given by:

$$Q^E = E\left[\frac{1}{TF}\sum_{t=1}^{T}\sum_{f=1}^{F}\frac{1}{|J_{ft}|}\sum_{j\in J_{ft}} P(a_{ft}^E = j|\mathcal{S}_{ft})\right] \tag{16}$$
$$Q^D = E\left[\frac{1}{TF}\sum_{t=1}^{T}\sum_{f=1}^{F}\frac{1}{|J_{ft}|}\sum_{j\in J_{ft}} P(a_{ft}^D = j|\mathcal{S}'_{ft})\right].$$

Where the expectations are taken over states with respect to equilibrium state distributions.

This assumption means that in equilibrium, a firm's beliefs about the probability of exploration and development by other firms are *on average* correct. Q^E is equal to the average over firms, periods, and blocks of the *expected* equilibrium probability of exploration. This means that Q^E is an equilibrium object, and, for example, policy changes that change firms' incentive to explore will change Q^E in equilibrium.

Assumptions A1 and A2 retain the asymmetric information structure but greatly simplify estimation and computation of equilibria. These assumptions also simplify the behavioral implications of the model in three significant ways. First, firms' beliefs about the actions of other firms are identical at all locations and times. This means that free riding incentives only vary with the number of other firms' *blocks* near a block j, not with, for example, the number of unique *firms* that hold drilling licenses nearby. Secondly, the model does not allow firms to reason about how their actions affect other firms' future behavior. For example, Assumption A1 precludes the "encouragement effect" discussed by Dong (2017), which mitigates the free riding incentive because firms have an added incentive to explore if doing so encourages other firms to explore. Third, this assumption shuts down any signaling incentives, since firms to not update their beliefs based on the presence of wells, only well results.

6 Estimation & Identification

6.1 Sample & Parameterization

I estimate the model using the subsample of the data that records activity on a 270 block region corresponding to the northern North Sea basin. This region contains many of the large oil deposits discovered on the UK continental shelf.[23] I restrict the estimation sample to this region in order to reduce computational time. I use the monthly Brent crude price inflated to 2015 dollars using the UK GDP deflator to measure the oil price. For years before 1980 where the Brent price is

[23]Specifically, this region corresponds to the area north of $59°N$, south of $62°N$, east of $1°W$, and west of the UK-Norway border.

unavailable I use projected values from a regression of Brent on the West Texas Intermediate price. I let a period be one month, and set the number to periods after which well outcomes are made public to $\tau = 60$.[24] This corresponds to the 5 year confidentiality period imposed by the regulator. I set the one month discount rate to $\beta = 0.992$, which corresponds to a 10% annual discount.

I impose the following parametric restrictions on exploration costs:

$$c(j, \mathcal{S}_{ft}) = c_0 + c_1 \ln(Nearby_{jt}). \tag{17}$$

Where $Nearby_{jt}$ be the number of licensed blocks "near" block j at date t, counting same-block licenses, first and second degree neighbors. This specification allows for information and technology spillovers in exploration drilling that are not explicitly modeled. For example, more heavily licensed areas are likely to be better understood in terms of geology and optimal drilling technology (see for example Covert (2015) on inter-firm learning about location-specific drilling technology).

The model parameters are therefore $\{\theta_1, \theta_2, \alpha, \sigma_\zeta\}$, where $\theta_1 = \{\mu, \omega, \ell\}$ are the parameters of the firm's beliefs defined in Section 3, $\theta_2 = \{c_0, c_1, \kappa_0, \sigma_c, \sigma_\kappa\}$ are the cost parameters, α is the probability of observing another firm's well outcome before it is made public, and σ_ζ is the variance of innovations to the oil price random walk. 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 and Q^D.

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 two probit regressions. First, I estimate the probability of a block j being licensed to *any* firm in period $t+1$ as a function of whether it was licensed to any firm in period t and the number of neighboring blocks licensed in period t. I then estimate the probability of block j being licensed to firm f in period $t+1$ *conditional* on it being licensed to some firm as a function of whether it was licensed to firm f in period t, whether it was licensed to any firm in period t, and the number of neighboring blocks licensed to firm f in period t. I detail this part of estimation in Appendix C.5.

[24] 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 destailed discussion of the estimation routine in Appendix C, I describe how I deal with observations where there are multiple exploration wells in a month.

The remaining parameters, θ_2 and α, 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 12 and the parameter α. Using these estimates, 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. First step estimates of the CCPs are also used to estimate the average exploration and development rates Q^E and Q^D which correspond to firms' beliefs. I describe this two step procedure in detail in Appendix C.

6.2.1 Estimation of Conditional Choice Probabilities

The most important difference between the procedure I implement and the existing literature is in the first step estimation of CCPs $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$ - the probabilities that a firm takes an action j in the exploration and development stages of the game conditional on its state $\mathcal{S}$.

If the state variable were observable in the data, then $\hat{P}(a_f^E = j|\mathcal{S})$ 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} and W_{ft}^U. The data is therefore consistent with a *set* of possible states $\tilde{\mathcal{S}}_f$ for each firm.[25]

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 choice observations is:

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

[25]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 C for a formal definition of $\tilde{\mathcal{S}}_f$.

I maximize this likelihood to obtain estimates of the conditional choice probabilities $\hat{P}(a_f^E = j|\mathcal{S})$ and the information spillover parameter, $\hat{\alpha}$, which controls the proabbility 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 $\hat{P}(a_f^E = j|\mathcal{S})$ implied by equation 12 and approximate the choice specific value function for each alternative with a linear equation in summary statistics of the state variable. Full details are provided in Appendix C. In approximating a high dimensional state variable with lower dimensional statistics I follow much of the applied literature that estimates dynamic discrete choice models with conditional choice probability methods. For example, see Ryan and Tucker (2011) and Collard-Wexler (2013).

6.3 Identification

6.3.1 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:

$$\begin{aligned} \tilde{\mathcal{S}_A} &= \{(1,0,0)\} \\ \tilde{\mathcal{S}_B} &= \{(0,1,0)\} \\ \tilde{\mathcal{S}_C} &= \{(1,0,0),(0,0,1)\} \\ \tilde{\mathcal{S}_D} &= \{(0,1,0),(0,0,1)\}. \end{aligned} \tag{19}$$

$\tilde{\mathcal{S}_A}$ 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:

$$
\begin{aligned}
P(a = j|\tilde{\mathcal{S}_A}) &= P(a = j|\mathcal{S} = (1,0,0)) \\
P(a = j|\tilde{\mathcal{S}_B}) &= P(a = j|\mathcal{S} = (0,1,0)) \\
P(a = j|\tilde{\mathcal{S}_C}) &= \alpha P(a = j|\mathcal{S} = (1,0,0)) + (1-\alpha)P(a = j|\mathcal{S} = (0,0,1)) \\
P(a = j|\tilde{\mathcal{S}_D}) &= \alpha P(a = j|\mathcal{S} = (0,1,0)) + (1-\alpha)P(a = j|\mathcal{S} = (0,0,1)).
\end{aligned} \tag{20}
$$

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:

$$
\alpha = \frac{P(a = j|\tilde{\mathcal{S}_C}) - P(a = j|\tilde{\mathcal{S}_D})}{P(a = j|\tilde{\mathcal{S}_A}) - P(a = j|\tilde{\mathcal{S}_B})}. \tag{21}
$$

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 6, firms' exploration choices respond more to the results of their own wells than to those of other firm wells, implying $0 < \alpha < 1$. $P(a = j|\mathcal{S} = (0,0,1))$ 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. For example, $P(a = j|\mathcal{S} = (1,0,0))$ enters both the first and third equation in (20). 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. That is, the second terms of the third and fourth equation in (20) are identical. 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 particular for states with n wells there are always at least as many equations as unknowns in the n well analogue of (20). This means that the number of observable sets of states $\tilde{\mathcal{S}}$, which correspond to equations, is always at least one greater than the number of true states $\mathcal{S}$. 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 $\mathcal{S}$. In practice, additional identification comes from the approximation of the state variable which smooths choice probabilities across states and allow extrapolation to states not observed in the data.

This procedure, which estimates the conditional choice probabilities and α in one step, is significantly less computationally intensive than alternatives such as the Expectation-Maximization procedure proposed by Arcidiacono and Miller (2011), which requires iteration of the two step estimator. Although calculation of the sum in equation 18 for different values of α is computationally expensive, this first estimation step only has to be performed once.

6.3.2 Identification of Cost Parameters

The cost parameters are estimated in the nonlinear regression given by equation 33. Intuitively, cost parameters c_0 and κ are identified by the average probability of exploration and development. Lower average probability of drilling is rationalized by higher costs. Cost parameter c_1 is identified by the extent to which the probability of drilling is higher on blocks with more licensed blocks nearby. 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 exploration variance parameter σ_ϵ is 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. The development variance parameter σ_ν is similarly identified. To see this, notice that $\frac{1}{\sigma_\epsilon}$ multiplies the choice specific continuation value $v_f^E(j, \mathcal{S}_t)$ in equation 12, and the sum of future revenue enters linearly in the firm's continuation value.[26] As the variance of cost shocks becomes large, the probability of any choice $j \in J_{ft} \cup \{0\}$ tends to $\frac{1}{|J_{ft}|+1}$.

Finally note that, 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.

7 Results

7.1 Estimates

Detailed results for each part of the estimation procedure are presented in Appendix C. Appendix Table A1 reports descriptive statistics on the estimated conditional choice probabilities (CCPs) $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$. In particular, I report the marginal effects of varying different

[26] See equation 30 in Appendix C.

elements of the approximation to the state variable on the estimated choice probabilities. The patterns are broadly as expected. The probability of exploration is increasing in the expected probability of success and in the variance of beliefs, in line with the descriptive results recorded in Table 3. Development probability is increasing in expected probability of success and decreasing in variance, also consistent with the descriptive results. Exploration probability is also increasing in both the number of neighboring own-firm licenses and other-firm licenses. However, the effect of own firm licenses of the probability of exploration is almost twice the effect of other firm licenses. The level of these effects is rationalized in the model by the parameter c_1, which allows exploration costs to be lower in regions with a high number of licenses. The difference between these two effects is then explained by the free riding incentive induced by additional other-firm licenses and the increased value of information when there are more same-firm licenses nearby.

Table 6 reports estimated model parameters and the average exploration and development probabilities, Q^E and Q^D. The parameter α, which is estimated simultaneously with the CCPs indicates that firms behave as if they observe the results of 36.6% of other firm wells before they are made public. This finding is in line with the descriptive results reported in Figure 6, which indicated that the marginal effect of an additional other-firm well on the probability of exploration was between 20% and 50% of the effect of an own-firm well. Recall that the exploration cost is given by $c(j, \mathcal{S}_{ft}) = c_0 + c_1 \ln(Nearby_{jt})$. The estimated value of c_1 indicates that the cost of exploration is, as expected, decreasing in the number of nearby licenses. Exploration cost at the average value of $Nearby_{jt}$, reported as $\bar{c}$ in Table 6, is about 25% of the development cost κ.

Table 6: Parameter Estimates

Parameter	Estimate	SE	Parameter	Estimate	SE
α	0.3661	0.0412	κ_0	16.3400	0.2431
c_0	10.3514	0.1861	σ_c	1.4484	0.0354
c_1	-1.9910	0.0464	σ_κ	2.0523	0.0720
$\bar{c}$	4.0571	0.1002	σ_ξ^2	0.0048	0.0004
		Average Choice Probabilities			
Q^E	0.0223		Q^D	0.0017	

Notes: Cost parameters are in billions of 2015 dollars. $\bar{c}$ is computed as the value of the expression given by equation 17 at the average value of $Nearby_{jt}$. Standard error of α is computed using the Jacobian of the likelihood function given by equation 18 at the estimated parameter values. Standard error of σ_ξ^2 is computed using the fourth centered moment of month to month changes in log price. Standard errors for the remaining (cost) parameters are computed using the Hessian of the second step nonlinear least squares specification given by equation 33 in Appendix C. Note that the standard error for the cost parameters does not take into account the first step error, and is therefore likely to be biased down.

Cost parameters are reported in billions of 2015 dollars. The estimated cost parameters are substantially larger than estimates of the capital costs of exploration and development from data on expenditure provided by the regulator. The average capital expenditure per exploration well is \$34.6 million and per development platform is \$1.9 billion. To understand the discrepancy, notice that the estimated cost parameters likely include frictions such as the cost of relocating capital equipment, redeploying labor, and other capacity constraints. For example, I model exploration as a monthly decision. If, in reality, drilling an exploration well ties up capital equipment for several months, this would inflate estimated costs. Furthermore, since the model is estimated on a small region of the North Sea, the cost parameters implicitly contain the opportunity cost of drilling in this region rather than elsewhere. Realized costs also include the random terms ϵ and ν, which I have interpreted as cost shocks but could also capture shocks to information. One can think of the estimated costs as being equal to the sum of engineering costs and the additional frictions due to capacity constraints, opportunity costs, and information shocks. Although these frictions are relevant to the firm, it is not clear that they should be included in the calculation of industry surplus used by the policy maker. In what follows, I use the *estimated* parameters to compute counterfactual firm *actions*. However, when I add up revenues and expenditures to compute *industry profit* for a given sequence of actions I will use the engineering costs obtained from average capital expenditure rather than the model-implied costs.

To examine the fit of the model to the data, I simulate the model from 1964 to 1990. Simulations are generated by drawing an action for each firm, each month, and updating firms beliefs based on the observed results. For each month, I set the distribution of licenses $\{J_{ft}\}_{f \in F}$ and the oil price P_t equal to the truth. I use mean values of the posterior success probability recorded in Figure 4, which is estimated using the true outcomes of all wells drilled before 1990, to draw exploration well outcomes and development revenue.

Table 7 records statistics on firm activity from the data and two simulations. The first column records the total the number of exploration wells, blocks developed, blocks explored, and the average number of exploration wells drilled on developed and undeveloped blocks from the data. The second column records the average of these statistics over 40 simulations of the model using the first step CCPs, $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$, to draw firm actions. Since the CCPs are estimated directly from the data, it is not surprising that the total number of exploration wells drilled and blocks developed in these simulations match the data closely. The estimated choice probabilities slightly overstate the number of exploration wells drilled on blocks that are eventually developed, although the qualitative pattern that more wells are drilled on blocks that are developed is preserved. This slight mismatch is likely due to the approximation to the state variable used in the first step of the estimation procedure.

The third column records the average of these statistics over 40 simulations of the model using approximate *equilibrium* choice probabilities. Equilibrium choice probabilities are computed by forward simulating the model-implied choice probabilities, $P(a^E = j|\mathcal{S}, \hat{\theta}_2)$ using estimated parameters $\hat{\theta}_2$ to obtain new estimates of the value function given by equation 30. These new value function estimates are then used to compute new choice probabilities. The process is iterated until the estimated choice probabilities converge. On each iteration, the average exploration probabil-

Table 7: Model Fit

	Data	Simulation	
		First Step Probabilities	Equilibrium Probabilities
Exploration Wells	476	473.90	503.65
Blocks Explored	99	95.55	97.25
Blocks Developed	20	22.95	22.43
Exp. Wells on Dev. Blocks.	8.75	12.45	13.19
Exp. Wells on Undev. Blocks.	3.81	3.79	3.91

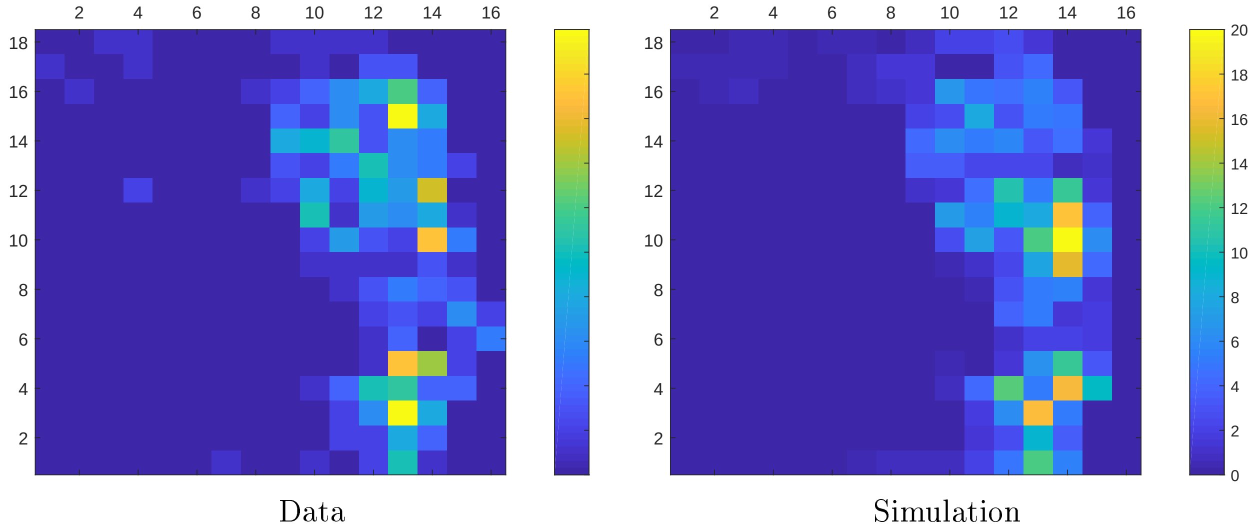
Notes: Column 1 records statistics from the data covering 1964-1990 for the relevant region. Columns 2 and 3 are averages over 40 simulations that cover 1964-1990. . For each month the assignment of blocks to firms and the oil price in the simulations are set at their realized values. Simulations in column 2 draw firm actions using the first step estimates of the conditional choice probabilities. Simulations in column 3 use approximate equilibrium conditional choice probabilities at the estimated parameter values.

ity $\hat{Q}^E$ is also updated. These equilibrium choice probabilities are approximate because I place restrictions on how the probabilities can change on each iteration to improve stability and reduce computational time. Details on this procedure are provided in Appendix E.

The difference between the second and third columns of Table 7 therefore reflects the difference between the first step choice probabilities estimated directly from the data, and the equilibrium choice probabilities implied by the model given the estimated cost parameters, $\hat{\theta}_2$. Equilibirium choice probabilities overstate the number of exploration by about 6% wells and the number of blocks developed by about 2% relative to the first step probabilities. When I examine the predictions of the model under counterfactual scenarios, I use these equilibrium simulations as a baseline.

As an additional test of the fit of the model, I compare the spatial distribution of exploration wells in the data to simulations using the equilibrium choice probabilities. The left panel of Figure 7 is a heat map that records the number of exploration wells drilled between 1964 and 1990 on each block in the data. Lighter colored blocks were drilled more often than darker blocks. The large dark region on the left side of the map was never licensed. Notice that there are three regions of concentrated drilling activity - in the south, centered on coordinate $(13, 3)$, in the middle of the map, centered on coordinate $(14, 10)$, and in the north, centered on $(13, 15)$. The right panel records equivalent well counts from the average of 40 simulations using the equilibrium action probabilities. Drilling is concentrated around the same points in the south and middle of the map, but not at the point $(13, 15)$ in the north. Many wells were drilled on this block despite it having been licensed for a relatively short period of 134 months (compared to 290 and 434 month-firm observations for $(13, 3)$ and $(14, 10)$ respectively). The observed monthly drilling rate on this block is an outlier that is difficult for the model to rationalize.

Figure 7: Model Fit: Well Locations



Notes: The left panel is a heat map recording the number of exploration wells drilled on each block of the region used for structural estimation from 1964 to 1990. More exploration wells were drilled on lighter blocks. The right panel is an analogous heat map of the average number of wells drilled on each block over 40 simulations using the baseline equilibrium choice probabilities. In both panels, the number of wells per block is truncated at 20 to better illustrate the cross-block variance.

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 effect 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 one at a time and then jointly.

First, I remove the free riding incentive by computing firm's optimal policy functions under the assumption that $Q^E = 0$. That is, I ask how firms would behave if, at each period, they believed that no new wells would be drilled by other firms at any period in the future. Under this assumption there is no incentive to strategically delay exploration. This counterfactual is not an equilibrium as defined in Section 5.2, 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 first and second columns of Table 8. The first column records statistics on exploration wells drilled, blocks developed, and industry revenue and profit for the baseline simulation. The second column records the same statistics for the no free riding counterfactual.

The first five rows record statistics on exploration well and development counts. Removing the free riding incentive brings exploration and development forward in time. The average number of exploration wells drilled up to 1990 increases by 7.4% from 503.65 to 541.15. The number of blocks developed before 1990 increases by 28% from 23.37 to 27.38. The efficiency of exploration, which I measure using the number of exploration wells drilled per development well, and the distribution of exploration wells between developed and undeveloped wells remain relatively constant. The sixth and seventh rows record the 1964 present discounted value of industry revenue and profit. Moving from the baseline to the no free riding counterfactual increases discounted revenue by \$6.21 billion or about 26% by bringing development forward in time. 45% of this increase in revenue comes from the bringing the development of the first 22.43 blocks forward in time, increasing the discounted value of revenue. The remaining 55% comes from the development of additional blocks before 1990 that were not developed in the baseline.

Table 8: Decomposition of Effects

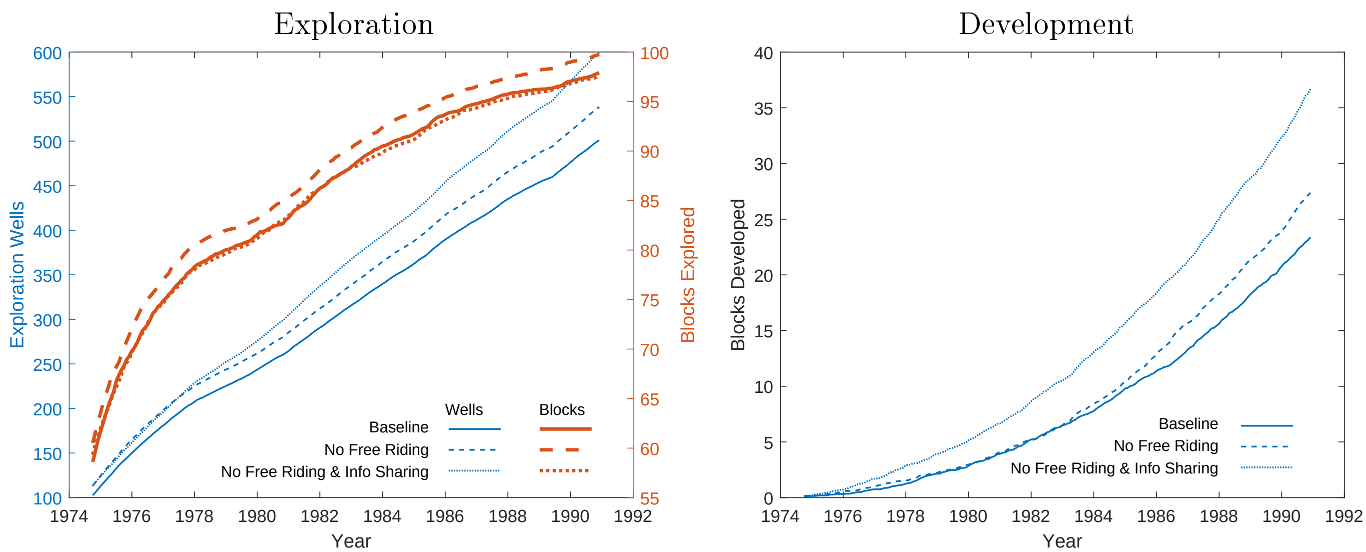
	Baseline	No Free Riding	Info. Sharing	Both
Q^E	0.0223	0	0.0223	0
α	0.3661	0.3661	1	1
Exp. Wells	503.65	541.15	567.30	604.83
Blocks Dev.	22.43	28.45	35.48	38.18
Exp. Wells/Dev	22.45	19.02	15.99	15.84
Exp. Wells on Dev. Blocks.	3.91	4.01	4.01	4.09
Exp. Wells on Undev. Blocks.	13.19	14.17	14.80	15.00
Revenue	24.09	30.30	37.74	40.15
Profit	13.85	18.12	23.59	25.06

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 using the posterior success probabilities computed using the true outcomes of all wells drilled before 1990. Revenue and profits are in 2015 dollars. Profits are computed using estimates of exploration well and development cost from OGA data on capital expenditure. PDV revenue and profit are 1964 values where the annual discount factor is 0.9.

The effect of removing free riding on the timing of exploration and development is illustrated by comparing the solid and dashed lines in Figure 8. The left panel records the average number of exploration wells and blocks explored each month from 1975 to 1990. The right panel records the average number of blocks developed for the same period. Removing the free riding incentive shifts the date that a block is first explored back in time by around one year. This increase in exploration speed translates to more rapid development. In the baseline simulation, 22.43 blocks are developed by the end of 1990. Under no free riding, this development level is attained 13 months earlier, at the end of 1989.

The second quantification 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 (and, implicitly, their policy functions) 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, 0.3661. This isolates the direct effect of increased flow of information from the equilibrium effects of setting $\alpha = 1$ on firms' drilling decisions.

Figure 8: Decomposition of Effects



Notes: The left panel plots the cumulative number of exploration wells drilled and blocks explored (blocks on which at least one exploration well has been drilled) for each month from 1975 to 1990 for three simulations. Thick red lines plot the number of blocks explored 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 no free riding counterfactual. The dotted lines are the average of 40 simulations under the no free riding and information sharing counterfactual. The right panel plots the number of blocks developed for the same three simulations.

The third column of Table 8 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 143 relative to the baseline and increases the number of blocks developed by 58% to 35.48. The efficiency of exploration improves substantially - the number of exploration wells drilled per block developed is reduced to 15.99 from 22.45 in the baseline. This increase in efficiency is also reflected in an increased concentration of exploration wells on productive blocks - the average number of exploration wells on developed blocks increases by 12% from 13.19 to to 14.80 while the average number of exploration wells on undeveloped blocks increases by only 3% from a much lower base of 3.91.

Perfect information flow increases discounted industry profit by 70% to \$23.59 billion from \$13.85 billion in the baseline simulation. This effect is about 2.28 times as large as the effect of removing free riding. This change in industry surplus can be decomposed into 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 relative importance of these two effects can be examined using the following back of the envelope calculation. In the information sharing counterfactual profit is 62.5% of revenue, while in the baseline the margin is 57.5%. Applying the information sharing margin to the baseline revenue results in a profit increase of $1.2 billion. This suggests that increased cost efficiency is responsible for about 19% of the increase in profit from information sharing, with the rest coming from faster development.

Finally, I run a counterfactual simulation that removes both free riding and wasteful exploration. That is, I set $\alpha = 1$ and $Q^E = 0$.[27] The results of this simulation are recorded in the fourth column of Table 8. Eliminating both sources of inefficiency increases exploration drilling by 20% and development before 1990 by 70%. The dotted lines in Figure 8 illustrate the path of exploration and development over time when both sources of inefficiency are removed. Relative to the baseline, development is brought forward in time by about three years. However, notice that the speed at which new blocks are explored is actually *reduced* relative to the no free riding counterfactual - the thick red dotted line in the left panel is below the thick red dashed line. Because of the increased information flow, fewer blocks are explored more intensively and wasteful exploration is reduced. The combination of bringing development forward in time and reducing inefficient exploration increases discounted profits by $11.21 billion, or 81% of the baseline.

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. Indeed, the Coase theorem suggests that firms should be able to achieve the first-best outcome by sharing information through bilateral contracts, eliminating both inefficient exploration and free riding by allowing firms to internalize the benefits of their discoveries to other firms. The empirical evidence indicates that this efficient exchange of information does not take place in reality. Furthermore, 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. There is a large literature which documents the role of such asymmetric information in preventing efficient trade (Myerson and Satterthwaite, 1983; Farrell, 1987; Bessen, 2004). Beyond the standard problem of trade under asymmetric information, there is an additional set of barriers to efficient trade when the object being traded *is information*. For example, it is difficult to signal the value of information to a buyer without revealing that information (Anton and Yao, 2002), and the potential for information to be costlessly resold prevents the original seller from capturing the entire social surplus that it generates (Ali, Chen-Zion, and Lillethun, 2017).

[27] Note that this is not equal to the first best outcome where firms jointly maximize industry profit. In this counterfactual, firms do not internalize the benefit of their drilling activity on other firms' profit.

8 Counterfactual Property Rights Policy

The results indicate that the presence of a free riding incentive and the limited spillover of information between firms both have significant effects on industry surplus. Removing both of these sources of inefficiency would result in a 81% increase in the present discounted value of 1964-1990 profits by bringing development forward in time and increasing the efficiency of exploration. These large inefficiencies suggest that the design of drilling rights and property rights over well data should take information externalities into account. In this section I ask how much industry surplus could be increased in equilibrium through alternative design of property rights that minimize the inefficiencies resulting from information spillovers.

I consider two main regulatory levers which the government can use to manipulate the flow of information between firms. First, the regulator can define property rights over data on well outcomes. In particular, 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. Second, fixing the confidentiality window, the government can change the spatial distribution of property rights. When each firm's drilling licenses neighbor fewer other-firm licenses the incentive for firms to delay exploration is reduced.

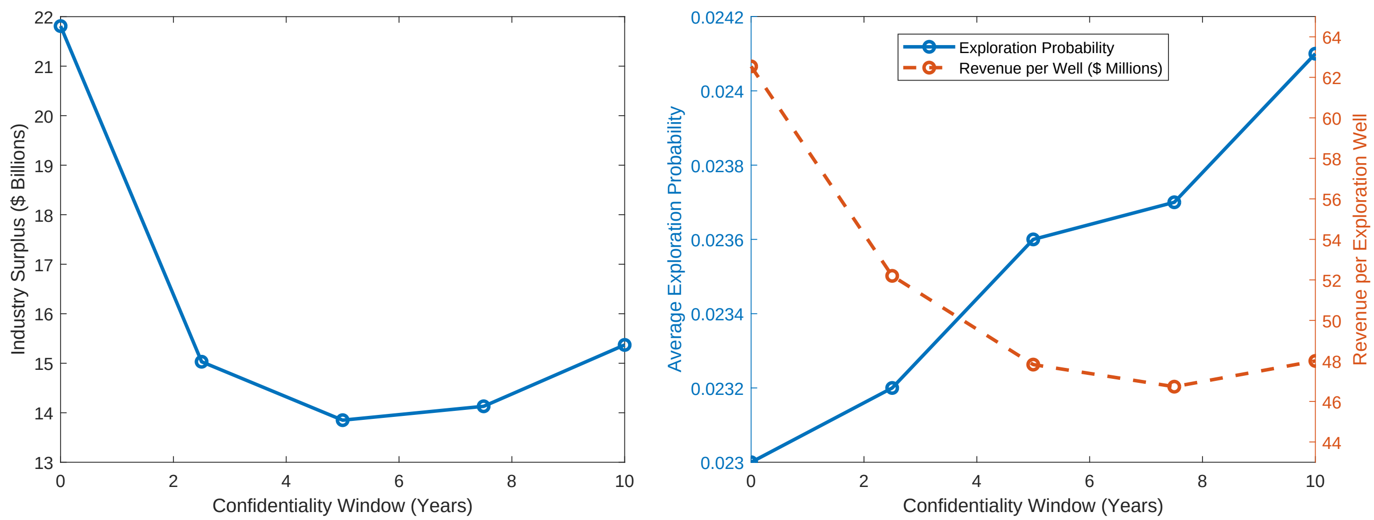
8.1 Confidentiality Window

UK regulations specify well outcomes are made public five years after the date a well is drilled. 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. For example, when licenses are issued on two neighboring blocks to two different firms, each firm's drilling strategy depends on their expectations about the flow of information from the other firm's wells. 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 probability of exploratory drilling should increase. On the other hand, lengthening the confidentiality window will reduce the efficiency of exploration by increasing wasteful drilling. When well data is held confidential for longer, firms are more likely to explore blocks that other firms already believe to be unproductive.

The regulatory problem of setting the optimal confidentiality window is therefore a case of trading off these two effects. If the free riding effect dominates and there is "too much" information flow between firms, then it may be optimal to lengthen the confidentiality window. On the other hand if the wasteful exploration effect dominates, and there is "too little" information flow between firms, then it may be optimal to shorted the confidentiality window. Whether one effect or the other dominates at the current window length of five years is an empirical question.

To determine the effect of changing the confidentiality window on industry surplus, I run counterfactual simulations of the model under different window lengths. For each window length, I first compute the approximate equilibrium choice probabilities implied by the estimated model parameters using the fixed point algorithm described in Appendix E. I then simulate the model using these choice probabilities, imposing the relevant confidentiality window lengths. The left panel of Figure 9 records the average over 40 simulations of industry surplus under confidentiality windows of 0, 2.5, 5 (the baseline), 7.5, and 10 years.

Figure 9: Confidentiality Window



Notes: The left panel records the 1964 present discounted value of 1964-1990 profit in counterfactual simulations with different confidentiality window lengths. In the right panel, the blue line, corresponding to the left y-axis, records the average exploration probability over firms, blocks, and dates using equilibrium exploration choice probabilities computed under different window lengths. The exploration probabilities are computed at the baseline distribution of states. That is, the reported numbers are the average *counterfactual* drilling probabilities at the states realized in a simulation that uses the *baseline* drilling probabilities. The dashed red line, corresponding to the right y-axis, records the average present discounted value of revenue per exploration well in equilibrium under different window lengths. Revenue and profit are in 2015 dollars. All figures are average over 40 simulations.

The results suggest that moving the confidentiality window in *either* direction from the 5 year baseline will increase expected industry surplus. In particular, lengthening the confidentiality window to 7.5 raises surplus by 2% of the baseline value of $13.44 billion. Lengthening the confidentiality further to 10 years increases surplus to $15.37 billion, 11% higher than the baseline. At 10 years, the gain in industry surplus is 36% of the gain from eliminating free riding recorded in Table 8. The no free riding counterfactual provides a theoretical maximum on the increase in surplus that can be obtained by increasing the confidentiality window. Surplus under longer confidentiality windows is less than this maximum because the no free riding counterfactual holds information flow fixed at the baseline level, while longer confidentiality windows reduce the flow of information between firms and therefore reduce the efficiency of exploration.

Reducing the length of the confidentiality window leads to a steeper rise in surplus, increasing to \$15.03 billion at 2.5 years. Surplus increases to \$21.81 billion, or 57% higher than the baseline, when the window is reduced to 0 years and well data is released immediately. When well data is released immediately, the gain in surplus is 82% of the gain in the information sharing counterfactual. Surplus is lower than under the information sharing counterfactual because of the additional free rising incentive induced by reducing the exploration window. The information sharing counterfactual in Table 8 held firm choice probabilities fixed at the baseline, while the 0 confidentiality window simulation uses counterfactual equilibrium exploration choice probabilities.

The U-shaped relationship between the length of the confidentiality window and industry surplus suggests that at window lengths greater than 5 years, the effect of limiting information flow on the free riding incentive dominates the effect on the efficiency of drilling, and that at window lengths less that 5 years the efficiency effect dominates. The right panel of Figure 9 illustrates these two effects separately. The solid blue line records the average probability of exploration (Q^E) for each confidentiality window. To illustrate the free riding effect independently from the effect of improved information flow on the speed of learning I *fix* the distribution of states at the baseline - the figure indicates that for any given state the probability of exploration decreases with shorter confidentiality window lengths. The dashed red line records revenue per exploration well at the equilibrium distribution of states under each confidentiality window. This measure of drilling efficiency is higher and the marginal effect of window length on efficiency is greatest for shorter window lengths. Indeed, for window lengths greater than 5 years, the effect of extending the window approaches 0. At these longer window lengths the effect on free riding dominates - extending the window increases the rate of exploration without substantially decreasing the rate at which exploration is converted into development.

The result that the true confidentiality window is close to the *least* optimal length begs the question of why this length was chosen by the regulator. 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, who wanted information to be made public earlier, and the major oil companies, who were resistant to any regulation that diminished their property rights over well data. The results reported in Figure 9 suggest that the settlement the parties arrived at, limiting well data confidentiality to five years, actually reduced industry surplus. The regulator's imposition of a five-year window was not short enough for the efficiency effect to substantially kick in, but did increase firms' incentive to strategically delay exploration relative to the no-regulation default of total confidentiality.

Although the results indicate that it is optimal to set the confidentiality window to 0, this historical background suggests that the optimal *politically feasible* policy change might be to extend the confidentiality window. This finding is specific to the UK setting, and is a function of the political process that determined the initial regulations. In other regulatory environments where confidentiality periods are already short, for example the Bakken Shale fields of North Dakota where well data is confidential for 6 months, lengthening the confidentiality would likely have a negative effect on industry surplus.[28]

8.2 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 8, 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 their blocks. First, since there are fewer inter-firm boundaries in the spatial allocation of licenses there is less incentive for firms to delay exploration in order to learn from other firms' exploration. Second, the spatial correlation of well outcomes means that value of exploration to the firm is higher when a block is surrounded by more same-firm licenses. Finally, the efficiency of exploration should be improved under a clustered license assignment since each well provides more information to the firm about the probability of success on its blocks, and fewer wells are therefore required to obtain a given amount of information.[29]

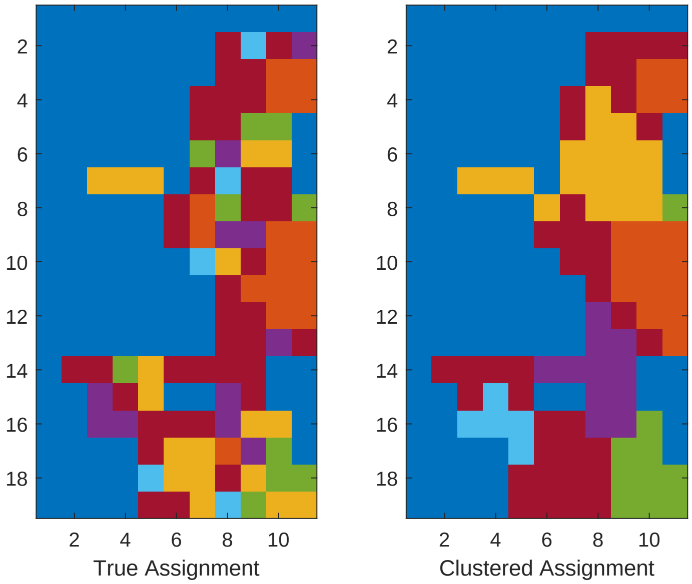
To quantify the effect of spatial reallocation of licenses, I construct an alternative license allocation for each month in the data using an algorithm that maximizes the spatial clustering of firms' licenses. Each year, the algorithm reallocates the licenses that are issued to year to firms using a deferred acceptance algorithm in which blocks propose to firms and are accepted or rejected. The algorithm increases clustering because blocks prefer to be allocated to firms with more existing licenses nearby, and firms would like to be assigned the blocks that are nearest to their existing blocks. 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 provided in Appendix F.

Figure 10 illustrates the true and counterfactual license assignments in January 1975. The left panel maps the licenses held by the largest 5 firms, with licenses held by other firms in red. The

[28] Of course, other oil and gas producing regions such as the Bakken Shale are subject to different drilling technology, geology, tract sizes etc. and the shape of the effects illustrated in Figure 9, which are a function of the underlying model parameters, are likely different.

[29] Note that clustering licenses has an additional effect on drilling capacity. For instance, if a set of four neighboring blocks are licensed to four different firms, the drilling capacity for that set of blocks is higher than if all four blocks are licensed to the same firm. Clustering licenses therefore reduces local drilling capacity, although total capacity across the entire region is held fixed. This effect is likely not of first order importance in practice since the average exploration probability per firm-block-month is around 2%, and the one block per month capacity constraint is far from binding.

Figure 10: 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.

right panel illustrates the counterfactual clustered license assignment in the same month. The difference between the allocations is visually clear - each of the largest 5 firms holds licenses on one or two contiguous regions in the counterfactual assignment, while in the true assignment these firms hold licenses on between 3 and 7 disconnected sets of blocks. The first two rows of Table 9 record how the clustering algorithm changes the average number of nearby own and other firm licenses (1st or second degree neighbors), where the average is taken across firms, blocks, and months.

The third through seventh rows of Table 9 record statistics on exploration wells, development of blocks, revenue and profit in equilibrium under the baseline and counterfactual license assignments.[30] Clustering firms' licenses increases the total number of exploration wells drilled between 1964 and 1990 by 8% and increases the number of blocks developed by 28%. The discounted value of industry profit increases by 42% from \$13.85 billion to \$19.62 billion. 13% of this increase in profit is from cost savings - the number of exploration wells drilled per developed blocks falls from

[30]Equilibrium choice probabilities change under the counterfactual license assignment because of the definition of equilibrium given by Assumption A.2 in Section 5. The equilibrium value of Q^E, firms' beliefs about the rate of exploration of other firms, is defined as the average exploration rate at the equilibrium distribution of states. Under a different allocation of licenses the equilibrium distribution of states changes. I estimate a new license allocation process, $P(j \in J_{ft+1}|J_t, \{J_{gt}\}_{\forall g \in F})$, using the counterfactual licenses, which I use when forward simulating in the equilibrium algorithm detailed in Appendix E.

Table 9: Clustered Licenses

Licenses	Baseline	Clustered
Nearby Own Licenses	0.371	0.583
Nearby Other Licenses	3.270	2.873
Exp. Wells	503.65	543.88
Blocks Dev.	22.43	28.78
Exp. Wells/Dev	22.45	18.90
Revenue	24.09	32.40
Profit	13.85	19.62

Notes: Results are averages over 40 simulations that cover 1964-1990. Oil price is set at its realized values. Well outcomes and development revenue are drawn using the posterior success probabilities computed using the true outcomes of all wells drilled before 1990. Revenue and profits are in 2015 dollars. Profits are computed using estimates of exploration well and development cost from OGA data on capital expenditure. PDV revenue and profit are 1964 values where the annual discount factor is 0.9. In the first column, the assignment of blocks to firms is set to the true assignment. I the second coulmn, the assignment of blocks to firms is set to the counterfactual clustered assignment.

22.45 to 18.90 - with the remaining 87% due to increased revenue. Industry surplus is greater than in the counterfactual that eliminates free riding reported in Table 8, and achieves 59% of the gain in surplus from the information sharing counterfactual.

Under this counterfactual assignment, firms have less incentive to free ride and are able to learn more quickly from the results of their own wells, since each well provides more information about other blocks owned by the same firm than under the baseline. By taking advantage of these effects, the results suggest that the government could substantially increase industry surplus through a simple rearrangement of the spatial allocation of blocks to firms. Indeed, 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. Within the limits of the model, which for example rules out any firm specific knowledge about particular blocks before exploration, these results provide a lower bound on the potential gain from spatial reassignment of licenses.

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. The allocation mechanism that has been in place since the first licenses were issued in 1964 has relied on firms submitting applications for specific blocks. 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 firm with a constant prior mean that is even slightly risk averse would prefer to be allocated licenses that are spread over a wide area. Under risk aversion, clustered license allocations are therefore likely to be industry-optimal but not optimal in expectation for the individual firms. Application data is confidential, so I cannot empirically verify whether firms' applications are spatially dispersed. However, in my conversations with the regulator I learned that the government has occasionally recommended firms take on licenses for blocks for which they did not apply in order to create contiguous blocks of licenses like those generated by the clustering algorithm. 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.

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 exploration and development forward in time, increasing the number of exploration wells drilled between 1965 and 1990 by 7.4% and increasing industry surplus in the same time period by 31%. Holding the free riding incentive fixed and allowing perfect information flow between firms increases surplus by 70% 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 tradeoff 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, although extending the confidentiality window increases surplus at the baseline of 5 years.

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. In this setting, the ability of firms to profit from their discoveries is held fixed across alternative policies. Firms always have the right to extract the oil they find on their blocks, with only the ability to benefit from other firms' investments changing across alternative policies. The specific features of this setting mean that the *information externality* effects of variation in property rights are not conflated with changes in the ability of a firm to profit from its own discoveries.[31]

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.

The final set of results quantifies the effect of changing the spatial allocation of licenses to firms. By clustering licenses, the regulator is able to reduce the incentive to free ride and increase the speed of learning, since each firm learns more about its own blocks from a single well. The effects of clustering on industry surplus are large, increasing surplus by more than the no free riding counterfactual. This finding is related to the theoretical literature on learning in teams (Holmstrom, 1982; Campbell, Ederer, and Spinnewijn, 2013), and suggests in settings where research is decentralized but a social planner is able to assign projects to each researcher (here, oil firms), surplus can be enhanced by designing the assignment to minimize the extent to which each team member can free ride off the others' research and maximize the extent to which each team member's research is cumulative. This insight could, for example, have applications to the organization of publicly funded research efforts which involve many independent researchers and labs contributing to a common project.

Methodologically, this paper makes two contributions that are applicable to other settings. First,

[31] Similarly, in none of the counterfactual experiments I examine do firms internalize the benefit their exploration to other firms. In particular, simplifying assumption A1 prevents firms from internalizing the effect of their own exploration on other firms' future behavior. Relaxing this assumption would complicate the model but would allow me to compute, for example, first-best exploration behavior in a scenario with full information sharing in which firms collude to maximize industry surplus.

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 (Nikolova and Jaworska, 2003), 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.*

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.

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.

Campbell, A., F. Ederer, and J. Spinnewijn. 2013. "Delay and Deadlines: Freeriding and Information Revelation in Partnerships." *American Economic Journal: Microeconomics.* 6 (2), 163-204.

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.*

Cox, D. R. 1975. "Partial Likelihood." *Biometrika.* 62 (2), 269-276.

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.*

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

Gill, R. D. 1992. "Marginal Partial Likelihood." *Scandinavian Journal of Statistics.* 19, 133-137.

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.

Gourieroux, C., and A. Monfort. 1993. "Simulation-Based Inference." *Journal of Econometrics.* 59, 5-33.

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. and G. Lewis. 2018. "You Can Lead a Horse to Water: Consumer Search with Spatial Learning." *Working Paper.*

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

Holmstrom, B. 1982. "Moral Hazard in Teams." *The Bell Journal of Economics.* 13 (2), 324-340.

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.

Jones, C. I. 2002. "Sources of U.S. Economic Growth in a World of Ideas." *American Economic Review.* 92 (1), 220-239.

Kellogg, R. 2011. "Learning by Drilling: Interfirm Learning and Relationship Persistence in the Texas Oilpatch." *Quarterly Journal of Economics.* 126 (4), 1961-2004.

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.

Kullback, S., and R. A. Leibler. 1951. "On Informationa and Sufficiency." *The Annals of Mathematical Statistics.* 22 (1) 79-86.

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.

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.

Nikolova, N., and J. Jaworska. 2004. “Approaches to Measure Chemical Similarity - a Review.” *Molecular Informatics.* 22 (9-10), 1006-1026.

Osborne, M. A., R. Garnett, and S. J. Roberts. 2009. “Gaussian processes for global optimization”. *3rd international conference on learning and intelligent optimization (LION3),* 1-15.

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.

Romer, P. M. 1990. “Endogenous Technological Change.” *Journal of Political Economy.* 98 (5), 71-102.

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.

Tipping, M. E. 2001. “Sparse Bayesian Learning and the Relevance Vector Machine.” *Journal of Machine Learning Research.* 1, 211-244.

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

Wong, W. H. 1986. “Theory of Partial Likelihood.” *The Annals of Statistics*. 14 (1), 88-132.

Wooldridge, J. M. 2002. “Econometric Analysis of Cross Section and Panel Data.” Cambridge: The MIT Press.

Appendix

A Theoretical Framework

In this section, I present a simple model of exploration to illustrate the effects of information externalities on firms' drilling decisions and structure the subsequent empirical analysis. Consider a two period drilling game played by two firms, i and j, who control adjacent blocks. In the first period, firms simultaneously decide whether to drill an exploration well on their respective blocks. Exploration wells on block i provide a binary signal about the presence of oil, and are successful with probability $\rho_i \in (0,1)$, which is a primitive determined by technology and the geology of the region being explored. Each firm always observes whether their own well is successful, and observes whether or not a well drilled by the rival firm is successful with probability $\alpha \in [0,1]$. In the second period, firms decide whether or not to develop the block at cost κ. Development yields a payoff $\pi(\rho_i) > 0$ with $\pi'(\rho_i) > 0$, $\pi(0) < \kappa$, and $\pi(1) > \kappa$, which can be thought of as the expected present discounted profit from the flow of oil over the block's lifetime. 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 exploring, or abandon a block. See for example Lerche and MacKay (1995) and Bickel and Smith (2006) who present models of optimal sequential exploration decisions based on binary signals.

Firm i's decision in each period depends on their beliefs about $\rho_i \in [0,1]$, the probability of exploration well success on their block. Suppose that firms have a common prior belief that the vector $\rho = (\rho_i, \rho_j)$ is drawn from a distribution $F(\rho)$. Let σ_{ij} be the correlation between ρ_i and ρ_j implied by $F(\rho)$. Let $I_{it} = (own_{it}, other_{it})$ be firm i's information at the beginning of period t. $own_{it} \in \{-1, 0, 1\}$ records firm i's exploration well outcomes from period $t-1$. If $own_{it} = 1$, firm i drilled a successful exploration well, if $own_{it} = -1$, firm i drilled an unsuccessful well, and if $own_{it} = 0$, firm i did not drill an exploration well. $other_{it} \in \{-1, 0, 1\}$ is firm i's information about firm j's exploration well outcomes, defined analogously except that $other_{it} = 0$ if firm j drilled a well and firm i did not observe it. Let $G(\rho|I)$ be the Bayesian posterior distribution of ρ given observed outcomes I. Assume $I_{i1} = (0,0)$ and therefore $G(\rho|I_{i1}) = F(\rho)$ for both firms. Firms start period 1 with identical information and beliefs. Firms then decide whether to drill an exploration well, and the results of wells are observed, with the results of a rival firm's well being observed with probability α. At the beginning of period 2, firm i's beliefs are represented by the posterior distribution $G(\rho|I_{i2})$. At this stage, firms' posterior beliefs can differ because of differences in their information sets.

Let $\tilde{\rho}(I) = \int_0^1 \rho dG(\rho|I)$ be the expected success probability, and $\tilde{\pi}(I) = \int_0^1 \pi(\rho) dG(\rho|I)$ be the expected development profit for a given information set, I. Let $\rho_0 = \tilde{\rho}(0,0)$. In period 2, firm i

will drill a development well at cost κ if and only if the expected return to doing so is positive. That is, $\tilde{\pi}(I_{i,2}) - \kappa \geq 0$. Therefore, define a firm's value function at the beginning of period 2 as:

$$V(I) = max\{\tilde{\pi}(I) - \kappa, 0\}$$

Let $W_{n,m}$ be the period 1 expectation of $V(I)$ conditional on the firm observing the results of $n \in \{0,1\}$ of their own and $m \in \{0,1\}$of the other firm's exploration wells. That is,

$$\begin{aligned}
&W_{0,0} = V(0,0) \\
&W_{0,1} = \rho_0 V(0,1) + (1-\rho_0)V(0,-1) \\
&W_{1,0} = \rho_0 V(1,0) + (1-\rho_0)V(-1,0) \\
&W_{1,1} = \rho_0\tilde{\rho}(0,1)V(1,1) + \rho_0(1-\tilde{\rho}(0,1))\left(V(-1,1)+V(1,-1)\right) + (1-\rho_0)(1-\tilde{\rho}(0,-1))V(-1,-1)
\end{aligned}$$

In the first stage, firms choose whether or not to drill an exploration well at cost $c+\epsilon_i$. I assume ϵ_i private information to firm i, and is drawn from a type-I extreme value distribution with variance parameter σ_ϵ. It is then straightforward to show that the unique Bayes-Nash equilibrium of the exploration game is for each firm to drill an exploration well with probability p^* given by the solution to equation 22. In what follows I assume $W_{0,0} = 0$. This assumption means that if not exploration results are observed it is not optimal to develop the block. This assumption can be relaxed without changing the nature of the equilibrium.

$$p^* = \frac{exp\left(\frac{1}{\sigma_\epsilon}\left(p^*\alpha(W_{1,1}-W_{1,0}) + W_{1,0} - c\right)\right)}{exp\left(\frac{1}{\sigma_\epsilon}p^*\alpha W_{0,1}\right) + exp\left(\frac{1}{\sigma_\epsilon}\left(p^*\alpha(W_{1,1}-W_{1,0}) + W_{1,0} - c\right)\right)} \tag{22}$$

Note that the value of additional information is always positive, so $W_{1,1} > W_{1,0} > W_{0,1} > W_{0,0}$. I will focus on the case of diminishing marginal value of information where $W_{1,1} - W_{1,0} < W_{0,1}$. That is, I assume the marginal value to firm i of observing the outcome of firm j's well is higher when firm i does not drill a well itself.[32] Under this assumption, it is straightforward to demonstrate the following proposition.

Proposition 1. *If* $W_{1,1} - W_{1,0} < W_{0,1}$ *then* $\frac{\partial p^*}{\partial \alpha} < 0$. *If in addition,* $0 < \frac{\partial W_{1,1}}{\partial \sigma_{ij}} < \frac{\partial W_{0,1}}{\partial \sigma_{ij}}$, *then* $\frac{\partial p^*}{\partial \sigma_{ij}} < 0$

Proof. Let P_1 denote the right hand side of equation 22. Let $P_0 = 1 - P_1$.

[32]That the value of additional signals should be diminishing is intuitive - in the limit additional signals have no value as the posterior variance goes to zero. However, returns to information are not necessarily diminishing everywhere, and it is possible to construct settings in which the second signal to be more valuable than the first (see Radner and Stiglitz (1984) for a discussion of non-concavities in the returns to information).

Applying the implicit function theorem to equation 22 yields

$$\frac{\partial p^*}{\partial \alpha} = -\left(\frac{p^* P_1 P_0 \left(W_{11} - W_{10} - W_{01}\right)}{\alpha P_1 P_0 \left(W_{11} - W_{10} - W_{01}\right) - \sigma_\epsilon}\right),$$

which is < 0 if $W_{1,1} - W_{1,0} < W_{0,1}$.

Applying the same approach to obtain the derivative with respect to σ_{ij}, noting that $\frac{\partial W_{0,1}}{\partial \sigma_{ij}} \neq 0$, $\frac{\partial W_{1,1}}{\partial \sigma_{ij}} \neq 0$, and $\frac{\partial W_{1,0}}{\partial \sigma_{ij}} = 0$, yields

$$\frac{\partial p^*}{\partial \sigma_{ij}} = -\left(\frac{p^* P_1 P_0 \left(\frac{\partial W_{1,1}}{\partial \sigma_{ij}} - \frac{\partial W_{0,1}}{\partial \sigma_{ij}}\right)}{\alpha P_1 P_0 \left(W_{11} - W_{10} - W_{01}\right) - \sigma_\epsilon}\right),$$

which is < 0 if $W_{1,1} - W_{1,0} < W_{0,1}$ and $0 < \frac{\partial W_{1,1}}{\partial \sigma_{ij}} < \frac{\partial W_{0,1}}{\partial \sigma_{ij}}$. $\square$

The first part of this theorem says that as the probability of information spillover between firms increases, the equilibrium exploration probability falls. If firms are more likely to observe the results of their rival's exploration wells, then firms have more of an incentive to free ride since the relative expected value of drilling their own well falls. The second part of this theorem says that the equilibrium probability of exploration is negatively related to the correlation between ρ_i and ρ_j, as long as $0 < \frac{\partial W_{1,1}}{\partial \sigma_{ij}} < \frac{\partial W_{0,1}}{\partial \sigma_{ij}}$. This property applies, for example, if ρ_i and ρ_j are distributed according to a transformation of a multivariate Niormal distribution, as in the Gaussian process model developed in Section 3 of the paper. Intuitively, increased correlation between firms' signals has a larger effect of a firm's continuation value when they only observe the other firm's signal and not their own. There is more incentive for firms to free ride when the signals generated by exploration wells on different blocks are more correlated. In particular, if $\rho_i = \rho_j$ (perfect correlation) then information generated by firm j's exploration well is of equal value to firm i as information generated by its own exploration well. In this case, $W_{1,0} = W_{0,1}$. If there is no correlation, then signals generated by firm j are not informative about ρ_i, and $W_{1,1} = W_{1,0}$ and $W_{0,1} = 0$. In this case, the equilibrium exploration rate, p^*, is identical to the equilibrium exploration rate that obtains when $\alpha = 0$.

This result illustrates that the extent to which firms have an incentive to free ride in exploration depends on the *information flow* between firms - parameterized by α - and the *covariance* of signals generated by exploration wells on different blocks - parameterized by σ_{ij}. Information flow is largely a function of technology and regulation - for example, the information confidentiality period imposed by the UK regulator. Correlation of exploration well outcomes at different locations is a function of underlying geology and the size and arrangement of license blocks. The remainder of this paper uses the UK data to estimate empirical analogues of these objects in the context of

North Sea oil exploration and quantifies the effect of information externalities on industry surplus using an econometric model that builds on the simple theoretical model presented here.

A final theoretical result illustrates the trade off faced by the social planner in manipulaitng information flow between firms.

Proposition 2. *Let $\tilde{p}$ be the probability of exploration that maximizes the joint expected surplus of the two firms. Let $\overline{p}$ be the equilibrium probability when $\alpha = 0$, and $\underline{p}$ be the equilibrium probability when $\alpha = 1$. If $W_{10} > c$ and $W_{1,0} + W_{0,1} - c > 2W_{1,1} - 2c$, then for some value of σ_ϵ, $\underline{p} < \tilde{p} < \overline{p}$.*

Proof. First, note that if $W_{1,0} > c$, then $\overline{p} \to 1$ and $\underline{p} \to 1$ as $\sigma_\epsilon \to \infty$ and $\overline{p} \to 0.5$ and $\underline{p} \to 0.5$ as $\sigma_\epsilon \to 0$. Note also that $\overline{p} > \underline{p}$ for any value of $\sigma_\epsilon \in (0, \infty)$ by Proposition 1. Since equation 22 is continuous in σ_ϵ, for any $\tilde{p} \in (0.5, 1)$ there exists a value $\tilde{\sigma}_\epsilon \in (0, \infty)$ such that $\overline{p} > \tilde{p} > \underline{p}$.

Now, write the objective function of the planner who can set the probability of exploration and observes all well outcomes as:

$$\tilde{p} = \arg\max_{p \in [0,1]} p^2(2W_{1,1} - 2c) + 2p(1-p)(W_{1,0} + W_{0,1} - c).$$

The planner's optimum is given by:

$$\tilde{p} = \frac{1}{2}\left(\frac{W_{1,0} - W_{0,1} - c}{W_{1,0} + W_{0,1} - W_{1,1}}\right).$$

If $W_{1,0} > c$, then $W_{1,1} > c$ and therefore $\tilde{p} > 0.5$. furthermore, if $W_{1,0} + W_{0,1} - c > 2W_{1,1} - 2c$ then $\tilde{p} < 1$. $\square$

The condition $W_{10} > c$ says that the social planner would prefer to drill a well on one of the blocks than none of the blocks. The condition $W_{1,0} + W_{0,1} - c > 2W_{1,1} - 2c$ holds when the value of information is sufficiently concave such that the social planner would like to drill only one well on one of the blocks. This result shows that the decentralized equilibrium can generate either too many or too few wells in expectation, and information flow between firms can be "too high" or "too low". Values of α that are too close to one induce too much free riding, such that the expected number of exploration wells is too low. On the other hand, low values of α make is more likely that more than one exploration well is drilled. This result illustrates the countervailing effects of information flow between firms on social surplus in equilibrium. Too little information flow results in socially inefficient exploration, since the social value of additional exploration wells beyond the first is lower than c. On the other hand, too much information flow between firms increases the free riding incentive and results in too little exploration in equilibrium.

This result suggests that exploration behavior is a decentralized equilibrium may be suboptimal, and that government policy that manipulates the arrangement of licenses (and thus the correlation of signals between firms) or the information flow between firms might bring equilibrium exploration rates closer to the social optimum.

B 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.[33]

The latent variable, $\lambda(X)$ is assumed to be distributed according at 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 $\kappa(X_j, X_k) \rightarrow \kappa(X_j, X_j)$ as $|X_j - X_k| \rightarrow 0$.

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:

$$\begin{aligned} p_1(\lambda|y) &= \frac{p_0(\lambda)p(y|\lambda)}{p(y)} \qquad (23) \\ p(y|\lambda) &= \prod_{w \in W} \left(1(s(w)=1)\rho(\lambda(X_w)) + 1(s(w)=0)\left(1 - \rho(\lambda(X_w))\right)\right) \\ p(y) &= \prod_{w \in W} \left(1(s(w)=1)\int \rho(\lambda(X_w))p_0(\lambda)d\lambda + 1(s(w)=0)\left(1 - \int \rho(\lambda(X_w))p_0(\lambda)d\lambda\right)\right) \end{aligned}$$

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:

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

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

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

The posterior marginal distribution of $\lambda(X)$ given by equation 24 is non-gaussian and has no

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

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 Willaims and Barber (1998). This method is widely used for Bayesian classification problems in computer science (Tipping, 2001) 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.1 Gaussian Process Likelihood

Let s be a vector of well outcomes and X be a vector of well locations, both random variables. Vectors are arranged in chronological order so that the first element of each vector corresponds to the first well drilled, the second to the second well drilled etc. Write the wth element of each vector as s_w and X_w. Let $\rho(\cdot) : \boldsymbol{X} \rightarrow [0, 1]$ be the random function which defines the probability of success at each location in the space $\boldsymbol{X}$, drawn form a logistic Gaussian process with density $g(\rho, \theta)$ where θ is a parameter vector. s_w is a Bernoulli random draw with probability $P(s_w = 1) = \rho(X_w)$. Adopt the following assumption about the process that generates X:

- Assumption A.3: X_w is drawn from a distribution $F(X_w|\theta, \{(X_y, s_y)\}_{y<w})$. That is, the distribution of X_w depends only on the parameters θ, and the locations and outcomes of past wells, and *not* on the random function ρ directly.

The joint distribution of (ρ, s, X) is then given by:

$$F(\rho, s, X) = \left[g(\rho, \theta) \prod_w \rho(X_w)^{1(s_w=1)} (1 - \rho(X_w))^{1(s_w=0)} \right] \left[\prod_w f(X_w|\theta; \{(X_y, s_y)\}_{y<w}) \right].$$

In the language of Cox (1975) the joint distribution is the product of two *partial likelihood* functions. One that is the product of the probabilities of outcomes s_w conditional on locations X_w, (the left brackets) and one that is the product of the probabilities of locations X_w conditional on past locations and outcomes $\{(X_y, s_y)\}_{y<w}$ (the right brackets). Wong (1986) shows that consistent estimates of the parameters θ can be obtained by maximizing partial likelihood functions with this *nested conditioning* structure. That is, once can omit one or the other of the two partial likelihood functions and obtain consistent estimates of the parameters θ. Chapter 13.8 of Wooldridge (2002) discusses this partial likelihood approach in detail for a panel data setting (of which this is a special case).

To obtain the likelihood function given in equation 3, the random function ρ is integrated out of the partial likelihood given by the left brackets. Gill (1992) shows that such a *marginalized partial likelihood* function has the same properties as the partial likelihood provided that the omitted term that appears in the full but not the partial likelihood does not depend on the variable that is integrated out. This is exactly assumption A.3.

B.2 KL Divergence

I compute the *expected* KL divergence for each (j,t) according to the following equation:

$$\begin{aligned} KL_{jt} = & E_t(\rho_j) \int g_t(\rho|\{j,1\}) \log\left(\frac{g_t(\rho|\{j,1\})}{g_t(\rho)}\right) d\rho \\ & + (1 - E_t(\rho_j)) \int g_t(\rho|\{j,0\}) \log\left(\frac{g_t(\rho|\{j,0\})}{g_t(\rho)}\right) d\rho \end{aligned} \tag{26}$$

Where $g_t(\rho)$ is the density of the firm's posterior beliefs over the vector ρ after observing all wells up to date t, $g_t(\rho|\{j,1\})$ is the updated posterior after observing an additional successful well on block j, and $g_t(\rho|\{j,0\})$ is the updated posterior after observing an additional unsuccessful well on block j. The first term in the expression is the expected probability of success on block j multiplied by the information gain from a successful well on that block. The second term is the expected probability of failure on block j multiplied by the information gain from a failed well.

C Estimation Details

C.1 First Step: Estimating Conditional Choice Probabilities

In the first step, I estimate CCPs $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$ - the probabilities that a firm takes an action j in the exploration and development stages of the game conditional on its state $\mathcal{S}$. With a sufficiently large data set, these probabilities could be estimated as empirical means for each state. However, since the number of possible states is large relative to the data, I impose some additional structure. Consider first the exploration decision. Notice that equation 12 can be rewritten as

$$P(a_f^E = j|\mathcal{S}) = \frac{exp\left(\tilde{v}_f^E(j,\mathcal{S})\right)}{1 + \sum_{k\in J_{ft}} exp\left(\tilde{v}_f^E(k,\mathcal{S})\right)\big)} \tag{27}$$

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

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

- Summary statistics of the firm's beliefs: $E(\rho_j|G_{ft})$, $E(\rho_j|G_{ft})^2$, $Var(\rho_j|G_{ft})$, $Var(\rho_j|G_{ft})^2$, and $E(\rho_j|G_{ft})Var(\rho_j|G_{ft})$.
- The number of licenses held near block j by firm f and by other firms: $|\{k : k \in J_{ft}$ and $d(j,k) \leq 1\}|$, $|\{k : k \in \cup\{J_{gt}\}_{g\neq f}$ and $d(j,k) \leq 1\}|$, and $|\{k : k \in \cup\{J_{gt}\}_{g\in F}$ and $d(j,k) \leq 2\}|$, 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 wells within one year of being made public: $|\{w : o_f(w) = 0$ and $t(w) + \tau - 12 \leq t \leq t(w) + \tau\}|$.
- A quadratic in the price level: P_t and P_t^2.
- Block j and firm f fixed effects.

Estimating $\hat{P}(a_f^E = j|\mathcal{S})$ is then a case of estimating the parameters of this approximation to $\tilde{v_f}^E(j,\mathcal{S})$.

The approximation to $\tilde{v_f}^E(j,\mathcal{S})$ depends on the distribution of licenses and wells "near" block j. Intuitively, the *difference* between the value of drilling on block j and taking no action should not depend on the distribution of licenses and wells at distant locations. Fixed effects are included to account for block level heterogeneity in drilling costs or beliefs not accounted for by well results and firm level heterogeneity is drilling costs. If block level fixed effects are not included, block level heterogeneity can lead to biased estimates of the logit coefficients on firms' beliefs. In particular, blocks that have idiosyncratically low drilling costs or on which there is additional public information indicating potential productivity are likely to be explored more intensively. Firm beliefs about these blocks are likely to have lower variance on average because of this high exploration rate. Across-block variation in average drilling rates and beliefs would therefore lead to the spurious conclusion that greater uncertainty in beliefs reduces the probability of exploration. Since there is no explicit block or firm level heterogeneity in the model, I estimate the parameters of the polynomial approximation to $\tilde{v_f}^E(j,\mathcal{S})$ once including fixed effects, then I find the intercept that matches the average exploration probability without fixed effects, holding other parameters at their estimated level. I use this intercept in generating predicted choice probabilities.

If the state variable were observable in the data, then $\hat{P}(a_f^E = j|\mathcal{S})$ could be estimated using the likelihood function implied by equation 27. 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} and W^U_{ft}. The data is therefore consistent with a *set* of possible states $\tilde{\mathcal{S}}_f$ for each firm.[34]

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 $\mathcal{S}^1_{ft}$ be the state if $o_f(w) = 1$ and $\mathcal{S}^0_{ft}$ be the state if $o_f(w) = 0$. From the econometrician's perspective, $P(\mathcal{S}^1_{ft}|\{\mathcal{S}^1_{ft}, \mathcal{S}^0_{ft}\}, \alpha) = \alpha$. I provide a formal definition of the distribution $P(\mathcal{S}_f|\tilde{\mathcal{S}}_f, \alpha)$ in subsection C.3 below

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

$$\mathcal{L}^E_f = \sum_{\mathcal{S}_f \in \tilde{\mathcal{S}}_f} \left[\left(\prod_{t=1}^{T} \prod_{j \in J_{ft} \cup \{0\}} 1(a^E_{ft} = j) \frac{exp\left(\tilde{v}^E_f(j, \mathcal{S}_{ft})\right)}{1 + \sum_{k \in J_{ft}} exp\left(\tilde{v}^E_f(k, \mathcal{S}_{ft})\right))} \right) P(\mathcal{S}_f|\tilde{\mathcal{S}}_f, \alpha) \right]. \qquad (28)$$

I maximize this likelihood to jointly estimate the coefficients of the approximation to $\tilde{v}^E_f(j, \mathcal{S}_{ft})$ 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 in equation 28.

I derive a similar expression for the likelihood of a sequence of development choices. Fixing α at the previously estimated value, I maximize the development likelihood to estimate the coefficients of the approximation to $\tilde{v}^D_f(j, \mathcal{S}_{ft})$. Because development of a block is a rare event (it occurs only 20 times in the estimation sample), I include fewer statistics in the approximation to the state variable to avoid overfitting. In particular, I omit fixed effects, quadratic terms in firms' beliefs about ρ_j and the oil price, and statistics on the number of nearby licenses and nearby unobserved wells. Adding higher order terms in beliefs about ρ_j leads to imprecise coefficient estimates, suggesting that extrapolation of the predicted choice probabilities to unobserved states would be unreliable. The estimated coefficients imply conditional choice probability estimates, $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$.

I use $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$ to estimate the firms beliefs about the average exploration

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

rate Q^E and Q^D defined in equation 16 with the mean CCPs across realized states in the data,

$$\hat{Q}_E = \frac{1}{TF}\sum_{t=1}^{T}\sum_{f=1}^{F}\frac{1}{|J_{ft}|}\sum_{j\in J_{ft}}\hat{P}(a^E = j|\mathcal{S}_{ft}) \tag{29}$$
$$\hat{Q}_D = \frac{1}{TF}\sum_{t=1}^{T}\sum_{f=1}^{F}\frac{1}{|J_{ft}|}\sum_{j\in J_{ft}}\hat{P}(a^D = j|\mathcal{S}_{ft}).$$

Logit coefficients and marginal effects for he estimated CCPs are recorded in Table A1.

Table A1: Conditional Choice Probabilities: Logit Coefficients

	Exploration			Development		
	Coefficient	SE	Marginal Effect	Coefficient	SE	Marginal Effect
Beliefs about ρ_j						
Mean	14.526	2.292	0.1764	3.022	1.680	0.0049
Variance	4.916	1.149	0.0216	-5.582	2.201	-0.0089
Mean Squared	-9.041	1.967				
Variance Squared	-1.733	0.396				
Mean $*$ Variance	-1.461	1.741				
Oil Price (\$100s)	3.272	1.337	0.0134	-0.233	0.704	-0.0004
Oil Price Squared (\$100s)	-0.020	0.009				
Licenses						
Own Firm Neighboring	0.129	0.029	0.0028			
Other Firm Neighboring	0.015	0.030	0.0003			
Total Nearby	0.105	0.016	0.0023			
Unobserved Wells	-0.153	0.034	-0.0033			
Mean Exploration Probability ($\hat{Q}^E$)	0.0223					
Mean Development Probability ($\hat{Q}^D$)				0.0017		
N Firms	44			44		
N Firm-Months	5977			5977		

Notes: Table records logit coefficients on state var summary statistics the enter the approximation to the state for the firm's exploration and development decisions. Standard errors are comuted using the outer product of the gradients of the log likelihood. Marginal effects are the predicted change in exploration and development probability from a marginal change in each of the listed statistics. Effects are calculated using the first derivatives of the logit choice probability expression. All statistics are for the case of a firm with drilling rights on a single block, j, for which the statistics that enter the approximation to the state variable are set to the mean observed values from the data.

C.2 Second Step: Estimating Dynamic Parameters

In the second step, I use the estimated conditional choice probabilities $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$ 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

$$V_f^E(\mathcal{S}, \theta_2) = E\left[\sum_{t=0}^{\infty} \beta^t \sum_{j=J_{ft}} \left(1(a_{ft}^D = j)\,(\pi_j - (\kappa_0 - \nu_{ftj})) - 1(a_{ft}^E = j)\,(c(j, \mathcal{S}_{ft}) - \epsilon_{ftj})\right)\right]. \quad (30)$$

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 $\mathcal{S}$, and $c(\mathcal{S}_{ft}, j)$ is given by equation 17. To estimate this expectation, I forward simulate the model from initial state $\mathcal{S}$ using the CCP estimates $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$ to draw firm f's actions and estimates of firm f's beliefs about other firms actions $\hat{Q}^E$ and $\hat{Q}^D$ to draw other firms' actions.[35] Simulation proceeds as follows:

1. Draw an exploration action using probabilities $\hat{P}(a_{ft}^E = j|\mathcal{S}_t)$. 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* $\mathcal{S}_t$.
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 $\hat{\alpha}$.
4. Update state to $\mathcal{S}'_{ft}$.
5. Draw a development action using $\hat{P}(a_{ft}^D = j|\mathcal{S}'_{ft})$. 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 at state* $\mathcal{S}'_{ft}$.
6. Draw other firms' development actions using $\hat{Q}_D$.
7. Update state to $\mathcal{S}_{ft+1}$. Go to step 1.[36]

[35] 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).

[36] Notice that since cost parameters θ_2 enter equation 30 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 $V_{fr}^{E}(\mathcal{S},\theta_2)$ be the present discounted sum of firm f's payoffs and costs from run r. Given R simulations from state $\mathcal{S}$, estimates of the value functions given by equation 30 are:

$$\hat{V}_f^E(\mathcal{S},\theta_2) = \frac{1}{R}\sum_{r=1}^{R}\left[V_{fr}^{E}(\mathcal{S},\theta_2)\right]. \tag{31}$$

A similar procedure is used to compute estimates of development stage value functions $\hat{V}_f^D(\mathcal{S},\theta_2)$ where the simulation algorithm is started at step 5. In practice I set $R = 500$ and run each simulation for 480 periods (40 years). Plugging estimated value functions into equation 10 yields estimates of choice-specific value functions, $\hat{v}_f^E(a^E,\mathcal{S},\theta_2)$ and $\hat{v}_f^D(a^D,\mathcal{S},\theta_2)$, which can be combined with equation 12 to generate model-implied choice probabilities

$$\tilde{P}(a_f^E = j|\mathcal{S},\theta_2) = \frac{exp\left(\frac{1}{\sigma_\epsilon}\hat{v}_f^E(j,\mathcal{S},\theta_2)\right)}{\sum_{k\in J_{ft}\cup\{0\}} exp\left(\frac{1}{\sigma_\epsilon}\hat{v}_f^E(j,\mathcal{S},\theta_2))\right)}. \tag{32}$$

With a similar expression for $\tilde{P}(a_f^D = j|\mathcal{S},\theta_2)$. Dropping the E and D for simplicity, I write the relationship between the model-implied probabilities and the empirical first-step probabilities, $\hat{P}_j(\mathcal{S})$, as:

$$\hat{P}(a = j|\mathcal{S}) = \tilde{P}(a = j|\mathcal{S},\theta_2) + \xi_{j\mathcal{S}} \tag{33}$$

Where, at the true parameters, ξ_{jS} contains the error due to sample size and approximation of the state variables in $\hat{P}(a = j|\mathcal{S})$ and the simulation error in $\tilde{P}(a = j|\mathcal{S},\theta_2)$. I estimate the parameters θ_2 by non-linear least squares, stacking exploration choice and development choice probabilities for each state $\mathcal{S}$. Note that I can compute both $\hat{P}(a = j|\mathcal{S}))$ and $\tilde{P}(a = j|\mathcal{S},\theta_2)$ for any state $\mathcal{S}$, including those not directly observed in the data. In practice I select a random 25% subset of the states observed in the data to include in the regression.

Since the simlated value functions enter non-linearly in the model implied probabilities, $\tilde{P}(a = j|\mathcal{S},\theta_2)$, non-linear least squares estimation based on equation 33 is asymptotically biased if the number of simulation draws, R, is fixed (Laffont, Ossard, and Vuong, 1995). To ensure consistency, it is necessary either to add a bias correction term, or to assume that R goes to infinity faster than the square root of the number of observations (Gourieroux and Monfort, 1993) - here the number of states included in the regression. Due to computational difficulty in obtaining the bias correction term, I rely on the assumption of an asymptotically increasing number of simulation draws.

C.3 Technical Details on Distribution of States

Define a period t observation as

$$X_t = \{\{(j(w), s(w), f(w)) : t(w) < t\}, \{J_{ft}\}_{f \in F \cup \{0\}}, P_t\}, \tag{34}$$

where 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} and W_{ft}^U. 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:

$$P(\mathcal{S}_f|X, \alpha) = \sum_{o \in s^{-1}(\mathcal{S}_f|X)} \left(\alpha^{\sum_w o(w)} (1-\alpha)^{\sum_w (1-o(w))}\right). \tag{35}$$

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

$$\mathcal{L}_f^E = \sum_{\mathcal{S}_f \in \tilde{\mathcal{S}}_f(X)} \left[\left(\prod_{t=1}^T \prod_{j \in J_{ft} \cup \{0\}} 1(a_{ft}^E = j) \frac{exp\left(\tilde{v}_f^E(j, \mathcal{S}_{ft})\right)}{1 + \sum_{k \in J_{ft}} exp\left(\tilde{v}_f^E(k, \mathcal{S}_{ft})\right))} \right) P(\mathcal{S}_f|X, \alpha) \right]. \tag{36}$$

Note that the summation in equation 36 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 methods.

C.4 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. π_j is drawn from a distribution $\Gamma(\pi; \rho_j, P)$. I assume that development payoff is given by $\pi_j = R_j \mu(P)$ where R_j is the quantity of oil reserves on block j (in barrels), and

$\mu(P)$ is a multiplier that depends on the price per barrel. I assume that reserves are drawn from a log normal distribution: $R_j \sim logN(\alpha_R + \mu_R \rho_j, \sigma_R)$. Note that the mean parameter depends on the true exploration success probability of the block, ρ_j.

Note that 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 for two reasons. First, 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. Second, older fields may have undergone several rounds of redevelopments (so-called "enhanced oil recovery". See Jahn, Cook, and Graham, 1998).

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 initially perceived as recoverable by the firm, 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:

$$r_j(t) = r_j(t^*(j))exp(-b_j(t - t^*(j))) + \epsilon_{jt} \tag{37}$$

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

$$R_j = \sum_{t=t_0(j)}^{t^*(j)} r_j(t) + \sum_{t=0}^{\infty} r_j(t^*(j))exp(-\hat{b}_j t) \tag{38}$$

Where the first term is the realized pre-peak production, and the second term is the extrapolated post-peak production.

Figure 2 illustrates the relationship between exploration success rate and log estimated reserves. Notice that the expected size of the reserves is monotonically increasing in the success rate of exploration wells on the same block, and the relationship is approximately log-linear. I assume log-linearity and estimate the parameters of the distribution of R_j by OLS using the following regression specification:

$$log(R_j) = \alpha_R + \mu_R \rho_j + \epsilon_j \tag{39}$$

Where $\epsilon_j \sim N(0, \sigma_R)$ and I measure ρ_j using the realized pre-development exploration well success rate on block j. The estimated parameters are reported in Table A2.

Finally, note that $\pi_j = R_j \mu(P)$ where $\mu(P) = P(1-0.125)\frac{1-\beta^{40}}{40(1-\beta)}$, This multiplier converts the total reserves in barrels to the present discounted value of revenue at the current price level, less the 12.5% royalty paid to the government, where oil is assumed to flow at a constant rate for 40 years, at which point the reserves, R_j are exhausted.

Table A2: Distribution of Development Payoffs

Parameter	Estimate	SE
α_R	1.594	0.420
μ_R	5.990	0.964
σ_R^2	1.949	0.115
N	80	

Notes: Reported coefficients are from OLS estimation of regression specification given by equation 39. Sample includes one observation for each of the 80 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.

C.5 Estimation of License Issuing Process

Firm f has beliefs about the evolution of the distribution of drilling licenses described by a two step process that takes place at the beginning of each period. First, the set of all blocks that will be licensed to any firm that period is drawn. Next the identities of the firms who receive licenses on each block are drawn. The process is described by the following equations:

$$\begin{aligned} P(j \in \cup\{J_{gt}\}_{g\in F} | \mathcal{S}_{ft-1}) &= \Phi(\beta_0 + \beta_1 Lic_{jt-1} + \beta_2 LicNeighbors_{jt-1}) \\ P(j \in J_{ft} | j \in \cup\{J_{gt}\}_{g\in F}, \mathcal{S}_{ft-1}) &= \Phi(\beta_3 + \beta_4 Lic_{fjt-1} + \beta_5 Lic_{jt-1} + \beta_6 LicNeighbors_{fjt-1}) \end{aligned} \tag{40}$$

Where Lic_{jt-1} is an indicator for whether block j was licensed to any firm at date $t-1$, Lic_{fjt-1} is an indicator for whether block j was licensed to firm f at date $t-1$, $LicNeighbors_{jt-1}$ is the number of blocks neighboring block j that were licensed to any firm at date $t-1$, $LicNeighbors_{fjt-1}$ is the number of blocks neighboring block j that were licensed to firm f at date $t-1$, and $\Phi(\cdot)$ is the standard Normal distribution function.

The first equation describes the probability that block j is licensed to some firm in date t. The second equation describes the probability that block j is licensed to firm f, conditional on it being licensed to some firm at date t. Notice that this specification does not rule out multiple firms receiving licenses on the same block. However, I allow the probability block j is licensed to firm f in period t to be a function of whether it was licensed to another firm in the previous period, $t-1$.

Table A3: License Issuing Process

	Probability of Assignment to Any Firm	Conditional Probability of Assignment to f
Dependent Variable	$1(j \in \cup\{J_{gt}\}_{g \in F})$	$1(j \in \cup J_{ft})$
Conditional on	$\forall j \in J$	$\forall j \in \cup\{J_{ft}\}_{f \in F}$
Constant	-3.004***	-2.001***
	(.039)	(.036)
Licensed in $t-1$	5.334***	-1.780***
	(.056)	(.050)
Licensed to f in $t-1$	.	6.611***
	.	(.056)
Neighbors Licensed in $t-1$	.366***	.
	(.055)	.
Neighbors Licensed to f in $t-1$	.	.099
	.	(.066)
N	81270	860112

Notes: Reported coefficients are from probit regressions of equations 40. The first column reports coefficients from the first equation. An observation is a block-month. The left hand side variable is an indicator for whether block j is licensed to any firm $f \in F$ in month t. The sample includes all block-month combinations for 1965-1990 on the set of blocks used in the structural estimation, including those never licensed. The second column records coefficients from the second equation. An observation is a firm-blopck-month. The left hand side is an indicator fro whether block j is licensed to firm f in month t. The sample includes all possible firm-block-month combinations for those block-months where j is licensed to some firm $f \in F$. This is, if block j was licensed to firm f in month t, the regresison would include a (g, j, t) observation for every firm $g \in F$.

I estimate the parameters of equations 40 by running two probit regressions. The first equation is estimated using a panel at the block-month level. The sample includes of all blocks for every month from 1965 to 2000. The left hand side variable is an indicator for whether block j was licensed to any firm in month t. The second equation is estimated using a panel at the firm-block-month level. The sample includes an observation for every possible (f, j, t) combination for months t in which block j was licensed to *some* firm.

The estimated parameters for both equations are recorded in Table A3.

D 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. I use the notation developed in Section 6 of the main paper and in Appendix D. In addition, let $\boldsymbol{X}$ be the space of possible data points, where $X \in \boldsymbol{X}$ is an observation as defined by equation 34.

Proposition 3. *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 35 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. □

E Simulation Details

In this section, I describe the simulation algorithm used to compute approximate 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, and first step conditional choice probability (CCP) estimates, $\hat{P}(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$. The output of the simulation are equilibrium CCPs, $P^*(a^E = j|\mathcal{S})$. Note that I hold development choice probabilities fixed.

The algorithm works by taking a set of CCPs as input and forward simulating those probabilities from each state $\mathcal{S}_f$. The simulation generates model-implied choice probabilities. If the probability of exploration is, on average, higher (lower) according to the model implied probabilities than the input CCPs then the CCPs are adjusted by increasing (decreasing) the intercept term in the linear approximation to the relative continuation values, $\tilde{v}_f^E(j, \mathcal{S})$, that enter the logit expression of CCPs given by equation 27. The procedure is repeated using the adjusted CCPs and and adjusted value of Q^E until the difference in implied probability of exploration between the model-implied probabilities and the input CCPs converges to 0. In particular,

Note that this procedure adjusts the average exploration probability, allowing the rate of exploration to vary under different counterfactual scenarios for example because of increased or decreased incentive to free ride, but holds *fixed* the response of relative continuation values, $\tilde{v}_f^E(j, \mathcal{S})$, to variation in the state variable. I make this simplification to improve the stability of the procedure while using a computationally feasible number of simulation runs.

The algorithm proceeds as follows:

1. Fix a set of states, $\boldsymbol{S}$ and use first step CCPs $\hat{P}^1(a^E = j|\mathcal{S})$ and $\hat{P}(a^D = j|\mathcal{S})$ and first step estimates of $\hat{Q}^{E1}$ and $\hat{Q}^{D1}$ to perform the forward simulation described in Appendix Section

D.2 for each $S \in \boldsymbol{S}$. This procedure generates model implied exploration probabilities, $\tilde{P}^1(a_f^E = j|\mathcal{S}, \theta)$.

2. Compute the the average deviation between the first step and model implied CCPs, $\Delta^1 = \sum_{S \in \boldsymbol{S}} \left(\tilde{P}^1(a_f^E = j|\mathcal{S}, \theta) - \hat{P}^1(a^E = j|\mathcal{S}) \right)$. Adjust the first step CCPs according to:

$$\hat{P}^2(a_f^E = j|\mathcal{S}) = \frac{exp\left(\hat{v}_f^{E1}(j, \mathcal{S}) + \delta\right)}{1 + \sum_{k \in J_{ft}} exp\left(\hat{v}_f^{E1}(k, \mathcal{S}) + \delta\right)}$$

Where Δ is the adjustment to the estimated first step continuation values. $\delta > 0$ if $\Delta^1 > 0$ and $\delta < 0$ if $\Delta^1 < 0$. Let $\hat{v}_f^{E2}(j, \mathcal{S}) = \hat{v}_f^{E1}(j, \mathcal{S}) + \delta$.

3. Simulate the model for all months from 1965 to 1990 using the distribution of licenses $\{J_{ft}\}_{f \in F}$ and the new CCPs $\hat{P}^2(a_f^E = j|\mathcal{S})$. Generate a new average exploration and developemnt probabilities, $\hat{Q}^{E2}$ and $\hat{Q}^{D2}$.

4. Go back to step 1 and repeat with new exploration CCPs $\hat{P}^2(a_f^E = j|\mathcal{S})$ and new average probabilities $\hat{Q}^{E2}$ and $\hat{Q}^{D2}$. Repeat the algorithm k times until

$$\sum_{S \in \boldsymbol{S}} \left(\tilde{P}^k(a_f^E = j|\mathcal{S}, \theta) - \hat{P}^k(a^E = j|\mathcal{S}) \right) \approx 0.$$

F License Clustering Algorithm

In this section I describe the algorithm used to generate the clustered license assignment. Let $\{J_{fy}\}_{f \in F}$ be the license assignment at the end of year y. Let $\tilde{J}_y$ be the set of licenses that were *issued* in year y. An element of $\tilde{J}_y$ is a triple (X_j, t_1, t_2) where X_j identifies the block coordinates, t_1 is the start date and t_2 is the end date of the license as observed in the data. Let $\tilde{J_{fy}} \subset \tilde{J}_y$ be the set of subset of year y licenses that were assigned to firm f in the data. Finally, let $\{\tilde{J'_{fy}}\}_{f \in F}$ be the counterfactual license assignment for year y.

Licenses and firms have preferences over each other given by a distance metric, Ω_{fjy}. The distance metric is chosen such that new licenses want to be assigned to firms which hold a larger number of nearby licenses, and firms want to be assigned the licenses that are close to many of their existing licenses. In particular,

$$\Omega_{fjy} = \sum_{k \in J'_{fy-1}} exp\left(-|X_k - X_j|\right). \tag{41}$$

Notice that Ω_{fjy} is increasing in the number of licenses held by f at a given distance from block j, and decreasing in the distance of any one license from block j, holding the locations of the other licenses fixed.

The algorithm proceeds as follows.

1. Start with the initial assignment $\{\tilde{J}'_{f0}\}_{f\in F} = \{\tilde{J}_{f0}\}_{f\in F}$.
2. Let F_0 be the set of firms for which $\tilde{J}'_{f0} \neq \{\}$. Let $F_{-0} = F \backslash F_0$.
3. Run a deferred acceptance matching algorithm between the set of firms F_0 and the set of licenses $\tilde{J}_1$. Each firm f ranks blocks according to a distance metric Ω_{fj1}. Each license j ranks firms according to Ω_{fj1}. Each license j can only be matched to one firm. Each firm has a quota given by $Q_{f1} = |\tilde{J}_{f1}|$.
   (a) Each license j proposes to its highest ranked firm.
   (b) Firm f accepts the highest ranked Q_{f1} licenses from those that propose to it. If fewer than Q_{f1} licenses propose to it it accepts all of them. Licenses that are not accepted are rejected.
   (c) Rejected licenses propose to their second highest ranked firm.
   (d) Firm f accepts the highest ranked Q_{f1} licenses from those that propose to it and those that it has already accepted. Licenses that are not accepted are rejected (including those previously accepted).
   (e) Repeat until all licenses are either accepted by some firm or have been rejected by all firms.
   (f) For each firm $f \in F_0$, the set of licenses that were accepted is then $\tilde{J}'_{f1}$.
4. Denote the licenses rejected at year 1 by $\tilde{J}_1^R$.
5. Take the firm $f \in F_{-0}$ with the largest quota, Q_{f1}. Assign firm f a random license $j \in \tilde{J}_1^R$. Compute Ω_{fj1} for the remaining licenses given this assignment.
6. Assign firm f its $Q_{f1} - 1$ top ranked licenses. The set of licenses assigned in then $\tilde{J}'_{f1}$. Repeat steps 5 and 6 for all other firms $f \in F_{-0}$ in order of quota size.
7. Repeat for each year.

The algorithm generates a license assignment that holds fixed the number of blocks assigned to each firm each year and the length that each license was active. As recorded in Table 9, the average number of nearby own-firm blocks is higher and the average number of nearby other-firm blocks is lower under the clustered license assignment. I do not claim that this assignment is in any way "optimal", but this algorithm provides a method for systematically assigning blocks to firms in a way that increases the average number of same-firm neighbors.

G Additional Tables and Figures

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

	Exploration Well			
$BlocksOwn_{fjt}$	2.467***	2.479***	2.505***	2.401***
	(.875)	(.858)	(.851)	(.868)
$BlocksOpEquity_{fjt}$	-.514	.	.	-1.026
	(1.277)	.	.	(1.304)
$BlocksEquityOp_{fjt}$	.	1.351	.	1.220
	.	(.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 6 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
	(.002)	(.003)	(.003)
Year	-.005***	.005**	.
	(.001)	(.002)	.
N	2105	2105	2105
Block FE	No	Yes	Yes

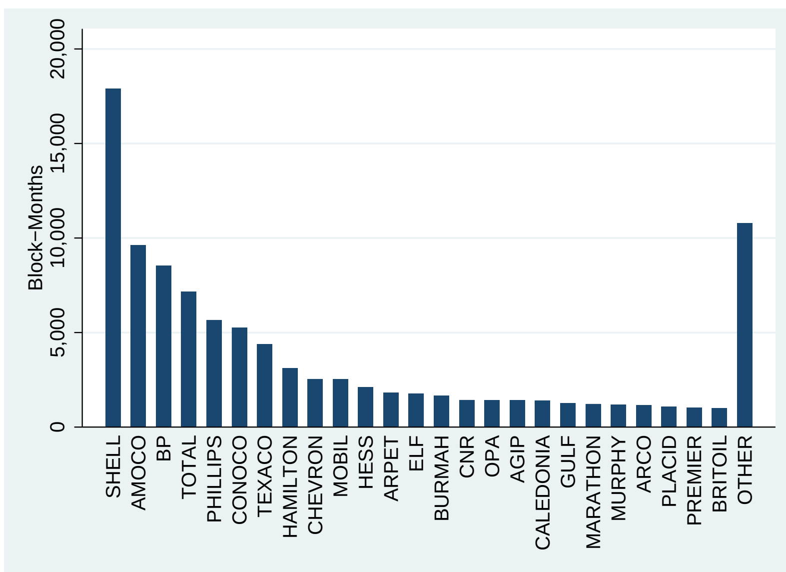
Notes: Sample includes all exploration wells drilled before 1991 on the region north of 55°N and east of 2°W. Left hand side variable is an indicator for whether the well was successful. Well sequence number records the order in which wells were drilled on a block. The first well on block j has well 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

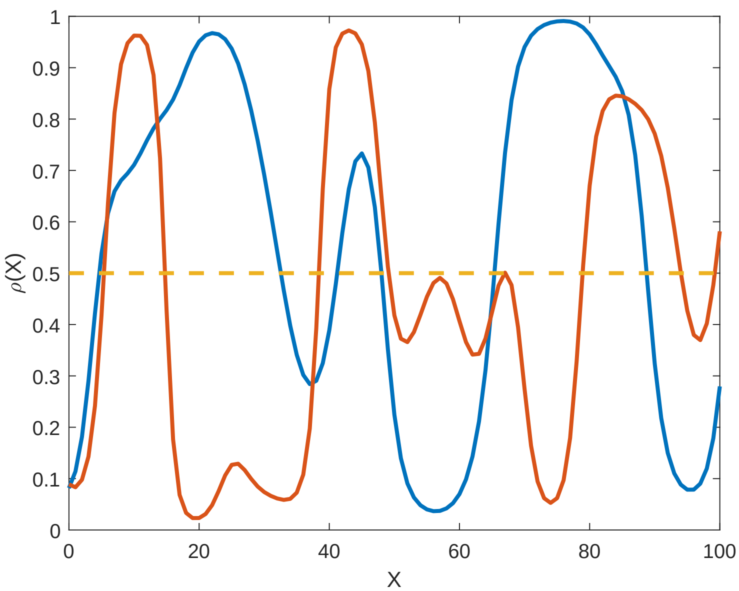
Notes: Table reports the ratio of the estimated marginal effect of past wells on nearby blocks (1-3 blocks away) to past wells on the same block on $Explore_{fjt}$ from the specification given by equation 6 where $g_{do}(\cdot)$ is quadratic in each of the arguments. Marginal effect is computed for the 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.

Figure A1: Top 25 Firms



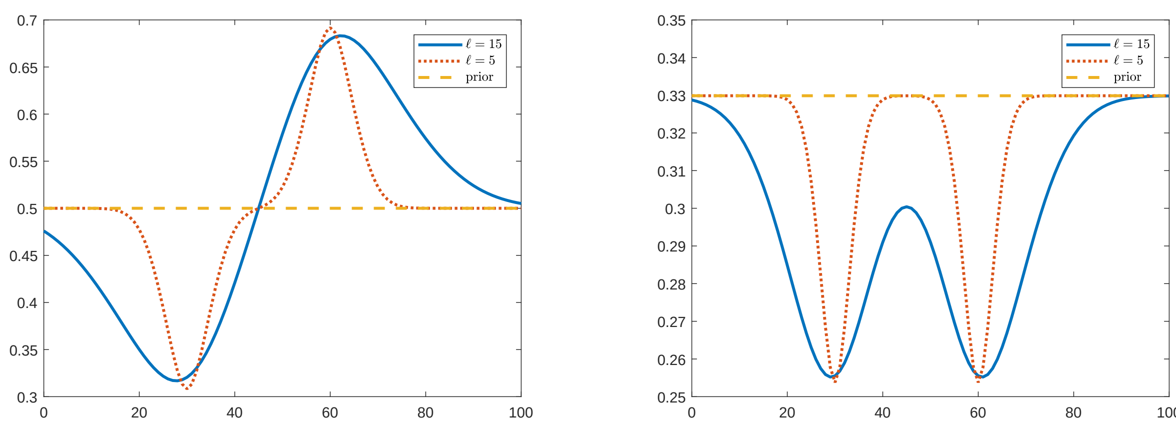
Notes: Figure plots the number of block-month pairs for 1964-1990 licensed to each of the top 25 firms, and the set of all other firms.

Figure A2: Gaussian Process Draws



Notes: Figure plots two draws (solid lines) from a logistic Gaussian process with parameters $\mu = 0$, $\omega = 5$, and $\rho = 5$ on a one-dimensional space. The dashed line corresponds to the prior mean.

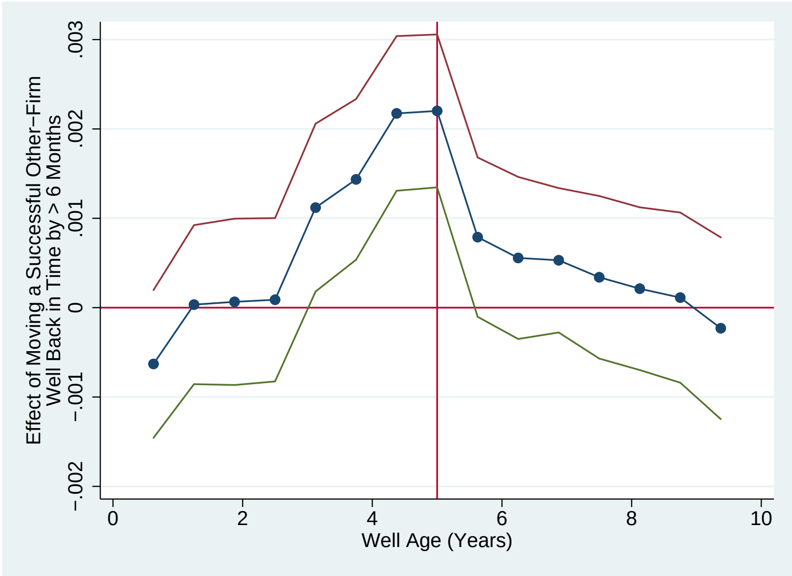
Figure A3: Gaussian Process Learning



Expected Probability | Standard Deviation of Probability

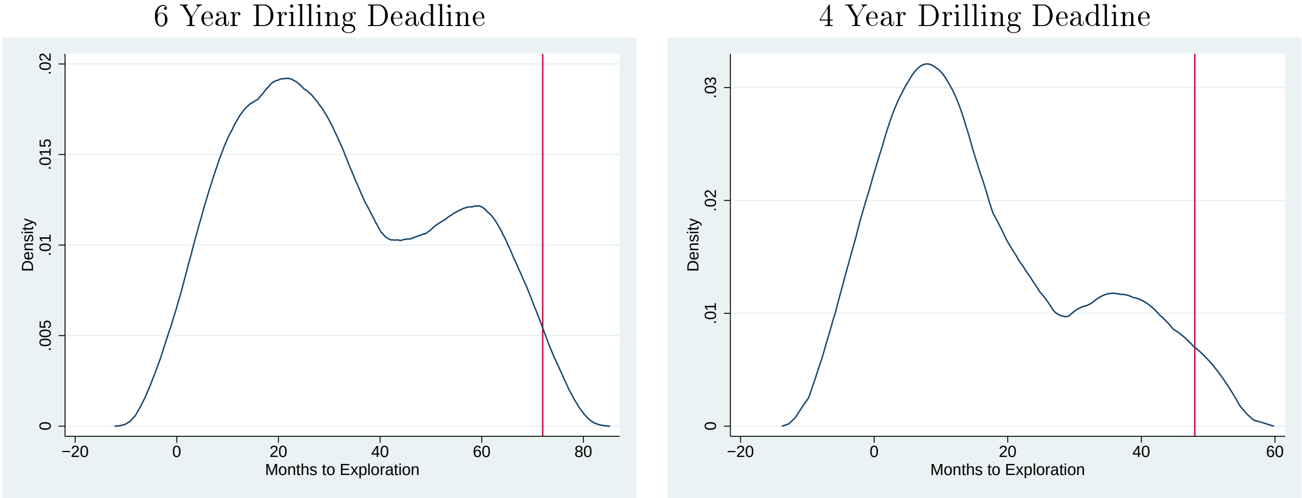
Notes: The x-axis of both panels represents the one dimensional space $[0, 1]$ on which the Gaussian process is defined. The dashed yellow line in the left panel plots the expected value of $\rho(X)$ for $X \in [0, 1]$ under prior beliefs represented by a logistic Gaussian process defined according to 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 represntes 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 A4: Effect of Well Age on Exploration



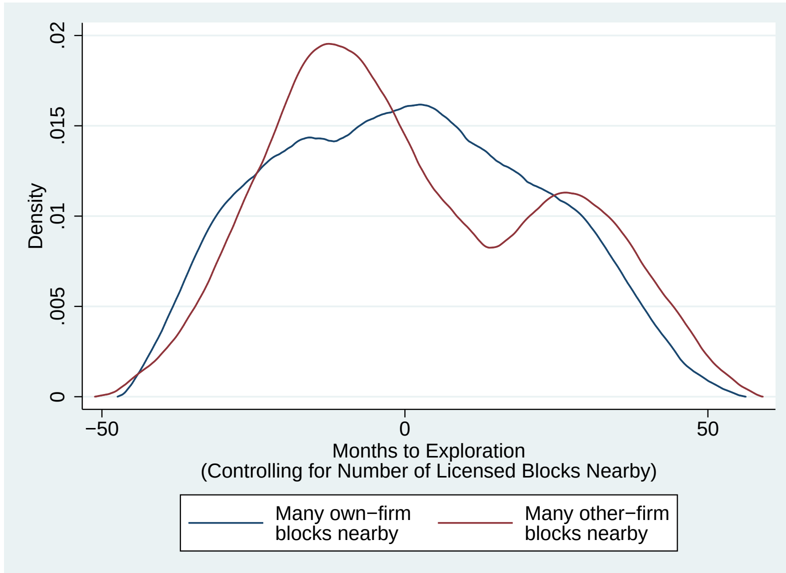
Notes: Figure plots OLS estimates of coefficients from a specification 6 with additional controls for the number of past successful other-firm wells 1-3 blocks away and more than T months old (Suc_T) and the number of such wells more than $T-6$ months old (Suc_{T-6}). Each point is the coefficient on Suc_T for a different regression, where the definition of T is given by the x-axis. For example, the first point plots the effect of increasing the number of successful other firm wells more than 1 year old, holding fixed the total number of past successful wells and the number of past successful wells more than 6 months old. It can therefore be interpreted as the effect of moving a well drilled 6-12 months ago back in time so it is more than 12 months old. Solid lines indicate a 95% confidence interval computed using robust standard errors. Vertical line indicates 5 year expiry date for well confidentiality.

Figure A5: Distribution of Months to First Exploration



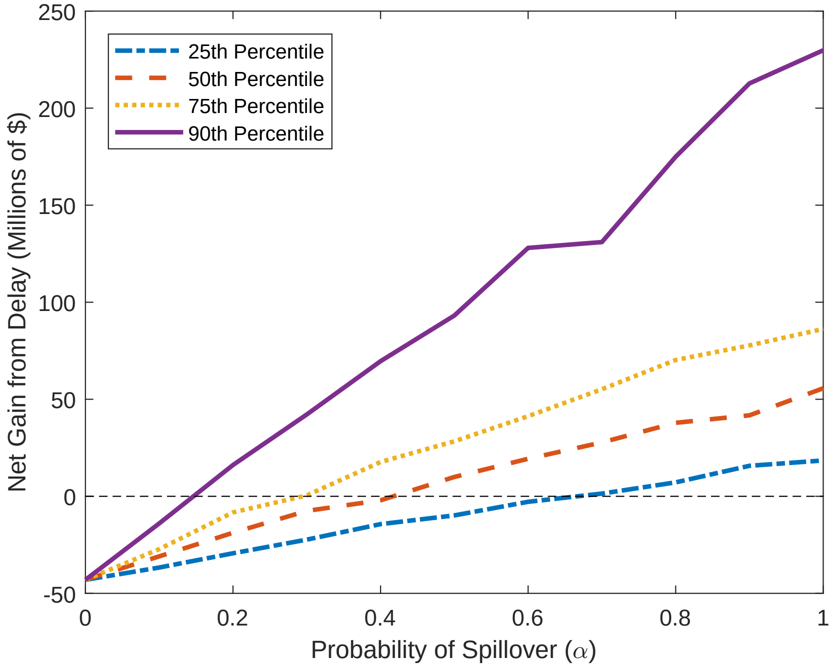
Notes: Each panel plots the distribution of time to first exploration across blocks. The left panel records this distribution for blocks with a 72 month initial drilling deadline, and the right panel records this distribution for blocks with a 48 month initial drilling deadline, with the deadlines indicated by vertical lines. The sample includes all blocks on the the region north of $55°N$ and east of $2°W$ first explored before 1990.

Figure A6: Distribution of Months to First Exploration by Distribution of Nearby Licenses



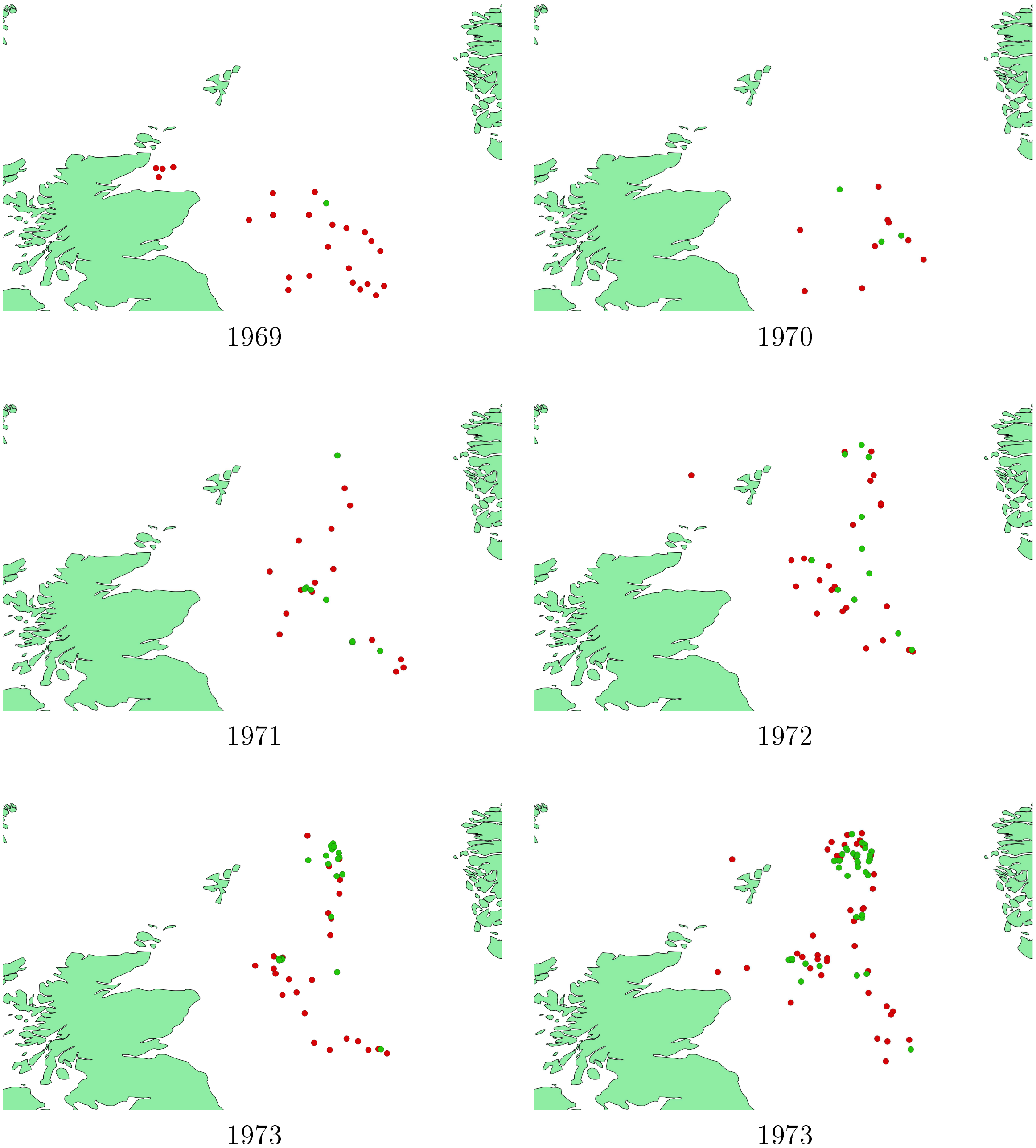
Notes: Figure plots the distribution of time to first exploration across blocks with 72 month drilling deadlines. I sort firm-blocks into quartiles according to the share of nearby licenses operated by the same firm at the date the drilling license was issued. I plot the distribution of time to first exploration for the top quartile - those block-licenses where more than 91% of nearby blocks are operated by the other firms - and the bottom quartile - those block-licenses where less than 70% of nearby blocks are operated by other firms. The sample includes all blocks with 72 motnh drilling deadlines on the the region north of $55°N$ and east of $2°W$ first explored before 1990. Time to drill is residualized against a cubic polynomial in the total number of nearby blocks licensed.

Figure A7: Incentive to Delay Exploration by One Year



Notes: Figure records the net gain from delaying exploration by 12 months for different license arrangements and levels of α. Computation of net gain is from 2000 simulations, as described in the text.

Figure A8: 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.