Do Firms Interact Strategically?
An Empirical Analysis of Investment Timing Decisions in Offshore Petroleum Production

C.-Y. Cynthia Lin[1]

JOB MARKET PAPER

Preliminary; Comments welcome

Abstract

When individual petroleum-producing firms make their exploration and development investment timing decisions, positive information externalities and negative extraction externalities may lead them to interact strategically with their neighbors. This paper examines whether these inefficient strategic interactions take place in U.S. federal lands in the Gulf of Mexico. In particular, it analyzes whether a firm's production decisions and profits depend on the decisions of firms owning neighboring tracts of land. Both reduced-form and structural econometric approaches are employed. The reduced-form approach is a discrete response model of a firm's exploration timing decision, using variables based on the timing of a neighbor's lease term as instruments for the neighbor's decision. The structural approach is a structural econometric model of the firms' multi-stage investment timing game. Although the models only permit the identification of the net effect of the two countervailing externalities, and not each individually, theory suggests that the importance of the extraction externality relative to the information externality should be greater on small tracts than on large tracts, and the data are consistent with this theory. Also as expected, the externalities intensify as the tract size decreases. The results suggest that the federal government best eliminates any inefficiencies in petroleum production that may result from non-cooperative strategic interactions when the tract size is large.

JEL Classification: C21, C25, C51, D92, L71

First draft: April 2, 2004
This draft: December 12, 2005

[1]Department of Economics, Harvard University; cclin@fas.harvard.edu. I am indebted to William Hogan, Dale Jorgenson and Robert Stavins for their support and encouragement throughout this project. This paper also benefited from discussions with Lori Bennear, Gary Chamberlain, Ulrich Doraszelski, Michael Greenstone, Joseph Harrington, Oliver Hart, David Laibson, Adrian Lajous (President, Petrometrica), Markus Mobius, Marcelo Moreira, Julie Mortimer, Michael Ostrovsky, Ariel Pakes, Jack Porter, James Stock, Stephen Weinberg, and Martin Weitzman, among numerous others. I received helpful comments from participants at the First Bonzenfreies Colloquium on Market Dynamics and Quantitative Economics in Alessandria, Italy in September 2004; at the Harvard Environmental Economics Workshop in September 2005; at the Harvard University Environmental Economics and Policy Seminar in October 2005; at student workshops in econometrics, in industrial organization, in microeconomic theory, in energy economics, and in environmental economics at Harvard University; and at the student workshop in econometrics at MIT. I thank Kenneth Hendricks and Robert Porter for sharing their lease sale data with me. Patricia Fiore, April Larson and John Shaw helped me to understand the geology of oil production. Robert Yantosca provided assistance with the IDL software. I thank the Minerals Management Service, and especially John Rodi, Marshall Rose and Robert Zainey, for answering my many questions about the OCS leasing program. I am indebted to Bijan Mossavar-Rahmani (Chairman, Mondoil Corporation) and William Hogan for arranging for me to visit Apache Corporation's headquarters in Houston and a drilling rig and production platform offshore of Louisiana, and for their support of my research, and I thank the Repsol YPF - Harvard Kennedy School Energy Fellows Program for providing travel funds. I received financial support from an EPA Science to Achieve Results graduate fellowship, a National Science Foundation graduate research fellowship, and a Repsol YPF - Harvard Kennedy School Pre-Doctoral Fellowship in energy policy. All errors are my own.

1 Introduction

Petroleum production is a multi-stage process involving sequential investment decisions. The first stage is exploration: when a firm acquires a previously unexplored tract of land, it must first decide whether and when to invest in the drilling rigs needed to begin exploratory drilling. The second stage is development: after exploration has taken place, a firm must subsequently decide whether and when to invest in the production platforms needed to develop and extract the reserve. Because the profits from petroleum production depend on market conditions such as the oil price that vary stochastically over time, an individual firm producing in isolation that hopes to make dynamically optimal decisions would need to account for the option value to waiting before making either irreversible investment (Dixit & Pindyck, 1994).

The dynamic decision-making problem faced by a petroleum-producing firm is even more complicated when its profits are affected not only by exogenous market conditions, but also by the actions of other firms producing nearby. When firms own leases to neighboring tracts of land that may be located over a common pool of reserve, there are two types of externalities that add a strategic (or non-cooperative)[2] dimension to firms' investment timing decisions and may render these decisions socially inefficient.[3] The first type of externality is an *information externality*: if tracts are located over a common pool or share common geological features so that their ex post values are correlated, then firms learn information about their own tracts when other firms drill exploratory wells or install production platforms on neighboring tracts (Hendricks & Porter, 1996). The information externality is a positive one, since a firm benefits from its neighbors' information.[4] A second type of externality is an *extraction externality*: when firms have competing rights to a common-pool resource, strategic considerations may lead them to extract at an inefficiently high rate (Libecap & Smith, 1999a; Libecap & Wiggins, 1985). The extraction externality is a negative one, since it induces a firm to produce inefficiently. Owing to both information and extraction externalities, the dynamic decision-making problem faced by a petroleum-producing firm is not merely a single-agent problem, but rather can be viewed as a multi-agent, non-cooperative game in which firms behave strategically and base their exploration and development policies on those of their neighbors.

Since 1954, the U.S. government has leased tracts from its federal lands in the Gulf of Mexico to firms interested in offshore petroleum production by means of a succession of lease sales. A lease sale is initiated

[2] In this paper, I use the terms "strategic" and "non-cooperative" interchangeably.

[3] In my broad definition of an externality, I say that an externality is present whenever a non-coordinated decision by individual firms is not socially optimal.

[4] Although the information externality has a positive effect on a firm's profits, it is socially inefficient. For example, it may cause firms to play a non-cooperative timing game that leads them to inefficiently delay production, since the possibility of acquiring information from other firms may further enhance the option value to waiting. If firms are subject to a lease term by the end of which they must begin exploratory drilling, or else relinquish their lease, then the information externality would result in too little exploration at the beginning of the lease term and duplicative drilling in the final period of the lease (Hendricks & Porter, 1996; Porter, 1995). In contrast, the optimal coordinated plan would entail a sequential search in which one tract would be drilled in the first period and, if productive, a neighboring tract is drilled in the next (Porter, 1995).

when the government announces that an area is available for exploration, and nominations are invited from firms as to which tracts should be offered for sale. In a typical lease sale, over a hundred tracts are sold simultaneously in separate first-price, sealed-bid auctions. Many more tracts are nominated than are sold, and the nomination process probably conveys little or no information (Porter, 1995). A tract is typically a block of 5000 acres or 5760 acres (Marshall Rose, Minerals Management Service, personal communication, 9 November 2005). The size of a tract is often less than the acreage required to ensure exclusive ownership of any deposits that may be present (Hendricks & Kovenock, 1989), and tracts within the same area may be located over a common pool (Hendricks & Porter, 1993). To date, the largest petroleum field spanned 23 tracts, and 57-67 percent of the fields spanned more than one tract depending on water depth (Marshall Rose, Minerals Management Service, personal communication, 31 March 2005). Because neighboring tracts of land may share a common pool of petroleum reserve, information and extraction externalities that lead firms to interact strategically may be present. As a consequence, petroleum production on the federal leases may be inefficient.

In this paper, I analyze whether a firm's investment timing decisions and profits depend on the decisions of firms owning neighboring tracts of land. Do the positive information externalities and negative extraction externalities have any net strategic effect that may cause petroleum production to be inefficient?

To answer this question, I use two econometric approaches. In the first approach, I estimate a discrete response model of a firm's exploration investment decision using variables based on the timing of a neighbor's lease term as instruments for the neighbor's decision. In the second approach, I develop and estimate a structural econometric model of the firms' multi-stage investment timing game. These econometric approaches are implemented on data on petroleum production on U.S. federal lands in the Gulf of Mexico.

The research presented in this paper is important for several reasons. First, an empirical analysis of investment timing decisions enables one to examine whether the strategic interactions that are predicted in theory actually occur in practice. Second, the estimation of strategic interactions, especially those that arise in dynamic decision-making, is of methodological interest. In particular, my structural econometric model enables the investigation of results that do not appear in the reduced-form analysis. Third, my results have implications for leasing policy: if the strategic effects and externalities turn out to be large, then the program by which the U.S. government leases tracts to firms may be inefficient, and possible modifications should be considered.

The exploration timing game in offshore petroleum production in the Gulf of Mexico has been examined in a seminal series of papers by Kenneth Hendricks, Robert Porter and their co-authors (see e.g. Hendricks & Kovenock, 1989; Hendricks & Porter, 1993; Hendricks & Porter, 1996). These papers focus on the information externality associated with exploratory drilling. They analyze this externality and the learning and strategic delay that it causes by developing theoretical models of the exploration timing game. In addition, Hendricks and Porter (1993, 1996) calculate the empirical drilling hazard functions for cohorts in specific areas, and study the determinants of the exploration timing decision and of drilling outcomes. According to their results, equilibrium predictions of plausible non-cooperative models are reasonably accurate and more descriptive than those of cooperative models of drilling timing.

This paper extends upon Hendricks and Porter's empirical work in several ways. My reduced-form discrete response model of firms' exploration timing decision improves upon that of Hendricks and Porter by using instruments to address endogeneity problems and by defining neighbors based on geographic distances rather than by the area boundaries drawn by the federal government. My structural model of the firms' multi-stage investment timing game further improves upon the Hendricks and Porter analyses by explicitly modeling a firm's multi-stage dynamic decision-making process and by allowing for negative extraction externalities as well as positive information externalities that arise during both exploration and development.

The results from both the reduced-form and the structural econometric models do not indicate that externalities from exploration have any net strategic effect. Neither a firm's exploration decision nor its profits from development depend significantly on the exploration decisions of its neighbors. In contrast, externalities from development do have a net strategic effect. A firm's real profits from developing increase when its neighbor develops, perhaps because this is a signal to the firm that the neighbor's exploratory efforts were successful, and therefore that there may be deposits present.

There are several possible explanations why the results reject strategic, non-cooperative behavior during exploration. One is that the tract sizes are large enough that cross-tract externalities are insignificant or even nonexistent. A second is that cross-tract externalities exist, but firms owning neighboring tracts cooperate to jointly internalize the inefficient externalities they impose on each other, for example by forming joint ventures in exploration[5] or by consolidating their production rights through purchase or unitization.[6] A third is that cross-tract externalities are significant, but the positive information externality exactly cancels the negative extraction externality, resulting in zero net strategic effects.

To distinguish among these three explanations for the lack of strategic interactions during exploration, I

[5]Joint ventures in exploration occur less frequently than one might expect, however, because negotiations are contentious, because firms fear allegations of pre-sale anti-trust violations (Marshall Rose, Minerals Management Service, personal communication, 3 May 2005), and because prospective partners have an incentive to free ride on a firm's information gathering expenditures (Hendricks and Porter, 1992). In their theoretical model of the persuasion game, Hendricks and Kovenock (1989) find that, even with well-defined property rights, bargaining does not eliminate all the inefficiencies of decentralized drilling decisions. As a consequence, the information externality may not be fully internalized.

[6]Under a unitization agreement, a single firm is designated as the unit operator to develop the entire reservoir, while the other firms share in the profits according to negotiated formulas (Libecap & Smith, 1999b). There are many obstacles to consolidation, however, including contentious negotiations, the need to determine relative or absolute tract values, information costs, and oil migration problems (Libecap & Wiggins, 1984). In addition, another free rider problem that impedes coordination is that firms may fear that if they reveal to other firms their information or expertise, for example about how to interpret seismic data, then they may lose their advantage in future auctions (Hendricks & Porter, 1996). Thus, despite various means of coordination, firms may still behave strategically and non-cooperatively, and information and extraction externalities may not be fully internalized.

estimate the strategic interactions by tract size. If externalities are insignificant when tracts are large enough, then one would expect to see strategic, non-cooperative behavior only on small tracts. This is because the smaller the tract size, the more likely the tracts are located over a common pool, and therefore the more acute the information and extraction externalities faced by the firms. Evidence for significant strategic interactions on small tracts but not on large tracts would thus be consistent with the first explanation.

If externalities exist even for the largest tracts in the sample, but are eliminated through coordination, then, assuming firms would coordinate regardless of tract size, one would not expect to see any strategic, non-cooperative behavior even on small tracts. Insignificant strategic interactions regardless of tract size would thus be consistent with the second explanation.

If strategic interactions do not occur on net because the positive information externality exactly cancels the negative extraction externality, then one may not expect the exact cancellation to still take place when the tract size is small. This is because the geographical span of the information externality is larger than that of the extraction externality: while the former only requires that tracts may share common geological features, the latter requires that tracts may be located over a common pool. As a consequence, it is possible for the information externality to be present on all the tracts in the sample, but for the extraction externality to be present on only the smaller tracts. Theory therefore suggests that the importance of the extraction externality relative to the information externality should be greater on small tracts than on large tracts. Thus, if the externalities cancel when all the tract sizes are considered, one might expect that the negative extraction externality would dominate the positive information externality when the sample is limited to small tracts only. Strategic interactions that are more significantly negative on small tracts than on large tracts would therefore be consistent with the third explanation.

According to the structural estimation, the importance of strategic interactions depends on tract size. As expected, strategic interactions are more likely to take place on smaller tracts, where the externalities are more acute. When the tract size is large enough, the net strategic effects of the externalities from both exploration and development disappear. Also as predicted by theory, the relative importance of the extraction externality from exploration with respect to the information externality is greater on small tracts than on large tracts; on large tracts, the two externalities cancel each other out.

The results suggest that, by selling predominantly large tracts, the federal government has minimized inefficiencies in petroleum production that may have resulted from non-cooperative strategic interactions.

The balance of the paper proceeds as follows. Section 2 presents a reduced-form analysis of the exploration investment timing decision. Section 3 presents a structural analysis of the multi-stage investment timing game. Section 4 concludes.

2 The Exploration Investment Timing Decision

2.1 Discrete Response Model

The first econometric approach I employ to investigate whether firms interact strategically when they make their investment timing decisions focuses on the exploration stage of production. For this approach, I estimate a reduced-form discrete response model of a firm's exploratory drilling decision. The question I pursue is: *How does a firm's exploratory drilling decision depend on the exploratory drilling decisions of its neighbors?*

Let t denote the time since the lease sale. As will be formalized in my structural model of the firm's dynamic decision-making problem, a firm will invest in the drilling rigs needed to begin exploratory drilling at time t if its profits from exploration exceed the continuation value from waiting. This decision thus depends on the firm's time-t estimate of the quantity of reserves on its tract, which, in turn, is affected by whether or not the firm's neighbors drilled in the previous period and also by the outcome of the neighbors' drilling. A firm may be more likely to explore following a positive outcome from its neighbor, but less likely to explore following a neighbor's dry hole.

I define a tract's "neighbors" by the following criteria. For each tract, its set of neighbors at any point in time consists of tracts nearby on which exploratory drilling could potentially have been first initiated in the previous period. In particular, a tract j is considered a neighbor of tract i at time t if (i) it is located within a certain distance of tract i, (ii) its lease began before time t, (iii) it has not been explored before time $t-1$, and (iv) it is owned by a different firm from the firm owning tract i. In most cases I also impose the additional restriction that (v) all of tract i's neighbors at time t must be sold on a different date from tract i. This additional restriction enables cleaner identification because neighboring tracts each have a different value of the instrument; as a consequence, the instrument is less likely to be correlated with spatially correlated unobservables.

Measuring neighbors' effects is difficult owing to two sources of endogeneity. One source is the simultaneity of the strategic interaction: if tract i is affected by its neighbor j, then tract j is affected by its neighbor i. The other arises from spatially correlated unobservable variables (Manski, 1993; Manski, 1995; Robalino & Pfaff, 2004).[7] To address these endogeneity problems, I exploit a unique feature of the federal lease sales. When a tract is won, a firm must begin exploration before the end of the five-year lease term, or else relinquish the lease. As a consequence, the hazard rate of exploratory drilling has a U-shaped pattern,

[7] Simultaneity of the strategic interaction poses less of a problem because I am using the lag of the neighbors' decisions. However, if errors are serially correlated, spatially correlated unobservable variables remain a concern even with the lag specification.

with high rates both at the beginning of the lease and at the end of the lease term (Hendricks & Porter, 1996). I therefore instrument for the fraction of neighbors who explore with the fraction of neighbors in the first year of their respective leases and the fraction of neighbors in the last year of their respective lease terms. The timing of a neighbor's lease term, which is related to the exogenous timing of the neighbor's sale date, is exogenous to a firm's exploratory drilling decision, as it is unlikely to have an effect except through its effect on the neighbor's exploratory drilling. Moreover, it is unlikely to be correlated with spatially correlated unobservables. The timing of lease sales is exogenous, especially for wildcat tracts, because the federal government chooses only a small subset of nominated tracts to be sold in any given lease sale (Porter, 1995); firms therefore have little influence over the timing of the lease sales.

To assess how a firm's exploratory drilling decision depends on the exploratory drilling decisions of its neighbors, I estimate a discrete response model by regressing the probability of exploration on tract i at time t on the fraction of neighbors j who explored at time $t-1$ and on other covariates x_{it}. I control for the following covariates: estimated pre-sale value per acre, winning bid per acre, the number of years since the lease sale, a dummy for being in the last year of the lease term, the size of the tract, the year the lease was sold, firm fixed effects, area fixed effects, and year effects. I estimate both a linear probability model:

$$\Pr(I_{it}^e = 1) = \beta_0 \frac{\sum_{j=1}^{J_{it}} I_{j,t-1}^e}{J_{it}} + x_{it}'\beta_1 \tag{1}$$

and a probit model:

$$\Pr(I_{it}^e = 1) = \Phi\left(\beta_0 \frac{\sum_{j=1}^{J_{it}} I_{j,t-1}^e}{J_{it}} + x_{it}'\beta_1\right), \tag{2}$$

where I_{it}^e is an indicator for whether exploration began on tract i at time t, J_{it} is the number of neighbors that tract i has at time t, $\Pr(\cdot)$ denotes probability, $\Phi(\cdot)$ denotes the standard normal cumulative distribution function, β_0 is a scalar, and β_1 is a vector of the same length as x_{it}. The coefficient of interest is the coefficient β_0 on the fraction $\frac{\sum_{j=1}^{J_{it}} I_{j,t-1}^e}{J_{it}}$ of neighbors who explored at time $t-1$. The instrumental variables analogs of the two models are two-stage least squares and Amemiya generalized least squares, respectively. The Amemiya generalized least squares estimator is formed by first estimating reduced-form parameters and then solving for the structural parameters; this estimator is asymptotically more efficient than a two-stage estimator (Newey, 1987).

My discrete response model of firms' exploration timing decision extends that of Hendricks and Porter

(1996) in two ways. First, Hendricks and Porter do not instrument for the variables they use to capture local drilling experience, and therefore do not address the endogeneity problems that arise when measuring neighbors' effects. In contrast, I instrument for the neighbors' decisions. Second, Hendricks and Porter define a "neighborhood" as one of the 51 "areas" into which the U.S. government has divided the federal lands in the Gulf of Mexico offshore of Louisiana and Texas. Using their definition, any two adjacent tracts that are located along either side of an area boundary belong to two different neighborhoods, even though they are right next to each other and may be located over a common pool of petroleum reserve. To address this problem, I define neighbors based on geographic distances, not on the arbitrary area boundaries drawn by the federal government.

2.2 Data for the Discrete Response Model

I use a data set on federal lease sales in the Gulf of Mexico between 1954 and 1990 compiled by Kenneth Hendricks and Robert Porter from U.S. Department of Interior data. There are three types of tracts that can be offered in an oil and gas lease sale: wildcat, drainage, and developmental. Wildcat tracts are located in regions where no exploratory drilling has occurred previously and therefore where the geology is not well known. Exploration on wildcat tracts entails searching for a new deposit. In contrast, both drainage and developmental tracts are adjacent to tracts on which deposits have already been discovered; developmental tracts, in addition, are tracts that have been previously offered in an earlier lease sale but either whose previous bids were rejected as inadequate or whose leases were relinquished because no exploratory drilling was done (Porter, 1995).

I focus my attention on wildcat tracts offshore of Louisiana and Texas that were auctioned between 1954 and 1979, inclusive. I do so for several reasons. First, my restrictions are similar to those made by Hendricks and Porter (1996), thus enabling me to best compare my results with theirs. Second, since wildcat tracts are tracts on which no exploratory drilling has occurred previously, information externalities are likely to be most acute. Third, because the data set only contains production data up until 1990, the restriction to tracts sold before 1980 eliminates any censoring of either drilling or production.[8] Additional restrictions I impose for a tract to be included in my data set are that it must be a tract for which location data is available, for which the first exploration occurred neither before the sale date nor after the lease term,[9] and for which production did not occur before exploration.

[8] Another reason to focus on the earlier lease sales is that post-auction lease transfers occurred less frequently in the past (Porter, 1995; John Rodi, Minerals Management Service, personal communication, 8 May 2003; Robert Porter, personal communication, 21 May 2003).

[9] It is possible for a lease to receive a suspension of production (SOP) or suspension of operations (SOO) which will extend the life of the lease beyond its primary term (Jane Johnson, Minerals Management Service, personal communication, October 29, 2003). Exploratory drilling first occurred after the lease term on 77 (or 3.1 %) of the 2481 wildcat tracts sold before 1980.

In total, there are 2404 tracts in my data set satisfying the above criteria, from 26 different lease sales. Table 1a presents summary statistics for these tracts; Figure 1 displays the tract size distribution. The maximum tract size, as stipulated by a provision in section 8(b) of the Outer Continental Shelf Lands Act (OCSLA), 43 U.S.C. 1337(b)(1), is 5760 acres, or 3 miles by 3 miles (Marshall Rose, Minerals Management Service, personal communication, 17 April 2003). Most tracts are either 2500 acres, 5000 acres or 5760 acres in size. The average tract size is 4790 acres (s.d. = 1100) and the median tract size is 5000 acres. The average real pre-sale value (in 1982 $) of these tracts, as estimated by the U.S. Department of Interior Minerals Management Service, is $360 per acre, while the average winning bid is $2520 per acre. Exploratory drilling eventually occurred on 1721 of the tracts. The U.S. government divides the federal lands in the Gulf of Mexico offshore of Louisiana and Texas into 51 areas; the data includes tracts from 26 of these areas.

For the panel data set, each time observation is a year.[10] Tracts enter the panel when they are sold. Tracts that were eventually explored exit the panel after the first drilling occurs. Tracts that were never explored exit the panel after five years, which is the length of the lease term a firm is given to begin exploration, or else relinquish its lease. Tracts have on average 2.03 time observations in the panel. The panel spans the years 1954 to 1983.

Figure 2 plots the aggregate hazard rate of exploration. The aggregate hazard rate H_t at time t is computed as the number of tracts that explored at time t divided by the risk set R_t at time t, where the risk set is simply the set of tracts that have not explored before time t. Following Hendricks and Porter (1996), the standard deviation of the hazard is $\sqrt{H_t\left(1-H_t\right)/R_t}$; the error bars in the figure indicate plus or minus one standard deviation. The aggregate hazard rate exhibits a U-shaped pattern: the hazard rate of exploration is monotonically decreasing with time except in the year right before the lease term expires, when there is a spike in exploration. It is this U-shaped feature that I exploit to construct my instruments for the neighbors' exploration decision. Since the hazard rate of exploration is high in both the first and last years of the lease term, my instruments for the fraction of neighbors who explored at any time t are the fraction of neighbors in the first year of their lease at time t and the fraction of neighbors in the last year of their lease term at time t.

Although the set of possible tracts i is limited to wildcat tracts sold before 1980 for which exploration did not occur after the lease term, the set of possible neighbors for these tracts i is larger. In particular, any tract sold between 1954 and 1983, inclusive, for which location data is available, for which the first exploration did not occur before the sale date and for which production did not occur before exploration is eligible as a potential neighbor for a tract i. Figure 3 plots the location of each of the tracts used in my

[10] As a robustness check, I also repeat my analyses on a panel in which each time period is a quarter, or 91 days.

reduced-form analyses.[11] Wildcat tracts i included in my sample are denoted with a filled circle. Other tracts that may serve as potential neighbors to these tracts are denoted with an open diamond; these tracts include drainage and developmental tracts sold between 1954 and 1983, inclusive, as well as wildcat tracts that were sold between 1980 and 1983, inclusive, and wildcat tracts that began exploration after the lease term. For a sense of the geographic span of a neighborhood, the grey asterisk in Figure 3 denotes a wildcat tract and the circle around it encompasses all tracts located within 5 miles from it.

To be considered an actual neighbor for a given tract i at time t, a potential neighbor must also satisfy the conditions listed in the previous section: namely, that (i) it is located within a certain distance of tract i, (ii) its lease began before time t, (iii) it has not been explored before time $t-1$, and (iv) it is owned by a different firm from the firm owning tract i,[12] and, in most cases, (v) all of tract i's neighbors at time t must be sold on a different date from tract i. In the base case, the distance used to define neighbors was 5 miles; for robustness, the analyses were also run using 4 miles, 6 miles and 10 miles.

Table 1b presents the summary statistics for variables that varied over both tract and time. Each column represents a different case. The base case is column (1). In the base case, the time period is one year in length, neighbors are located within 5 miles, and the additional restriction (v) that all of tract i's neighbors at time t must be sold on a different date from tract i is imposed. There are 1139 observations in the base case. For the other cases, I vary the time period, the distance of the furthest neighbor, and whether or not all neighbors have to be sold on a different sale date. In the base case, a tract has on average 1.74 (s.d. = 1.00) neighbors at any time t.[13] Of its neighbors, on average 43% of them began exploration in the previous period. Moreover, during the previous period an average of 36% and 15% of a tract's neighbors were in the first and last years of their lease term, respectively.

2.3 Results from the Discrete Response Analysis

I now present the results from the discrete response analysis, in which the probability of exploration on tract i at time t is regressed on the fraction of neighbors j who explored at time $t-1$ and other covariates x_{it}. As a benchmark, Table 2 presents the results from running the discrete response models without the use of instruments, when neighbors must be located within 5 miles of a tract. In specifications (1) and (2),

[11] I use the latitude-longitude coordinates provided by Hendricks and Porter. The longitude and latitude are compiled from the well-bore tape as the average of the longitude (or latitude) over all wells recorded in the block for the entire coverage period of the tape. The tape includes spud dates from January 13, 1947 to July 6, 1991. This is intended to give a "representative location" of the tract. The location for blocks not on the well-bore tape was approximated by map inspection.

[12] For cases in which multiple firms share the lease for a tract, I define the owner of the tract as the firm with the highest share in the bid, as this firm is likely to have the primary decision-making authority.

[13] The base case cutoff of 5 miles was chosen so that a neighborhood would be at least as large as the size of most petroleum fields. To date, 79% and 70% of the fields spanned 3 or fewer leases for blocks with maximum water depth of 0-199 meters and 200-399 meters, respectively (Marshall Rose, Minerals Management Service, personal communication, 31 March 2005).

the linear probability model and the probit model, respectively, are run with the base case sample in which the time period is one year in length, neighbors are located within 5 miles, and the additional restriction (v) that all of tract i's neighbors at time t must be sold on a different date from tract i is imposed. For robustness, specification (3) runs the linear probability model on the sample in which restriction (v) is relaxed; specification (4) runs the linear probability model on quarterly data. The coefficient of interest is that on the fraction of neighbors who drilled at time $t-1$. For all four specifications, neighbors do not have a significant effect. However, the other covariates have coefficients of the expected sign: the probability of exploring increases with the estimated pre-sale value, the winning bid and the size (acreage) of the tract. Also as expected, the probability of exploring often decreases with the number of years since the lease sale, with the coefficient being either significantly negative or insignificant at a 5% level, and increases at the last year of the lease term.

To test for the endogeneity of the neighbors' drilling, a Durbin-Wu-Hausman test is used for the linear probability model and a Smith-Blundell test is used for the probit model. Both are tests of whether the residual from a regression of the variable in question on all the exogenous variables has a significant coefficient when added to the original model; the table reports the p-value from the tests under the null hypothesis that the variable is exogenous. According to these tests, neighbors' decisions are not significantly endogenous at a 5% level in any of the four specifications. Although the tests fail to reject an exogeneity assumption, it still seems plausible, at least in theory, that neighbors' decisions are endogenous owing to simultaneity and/or spatially correlated unobservables.

To guard against any potential endogeneity of the neighbors' decisions, I instrument for the fraction of neighbors who drill with the fraction of neighbors in the first year of the lease and the fraction of neighbors at the last year of the lease term. Table 3 presents the results from first-stage regression of the endogenous variable on the instruments and the covariates for the base case (specification (1)), as well as for three variants. The first-stage F-statistic from a joint test of the two instruments is over 10 in all specifications, so weak instruments should not be a concern (Stock & Watson, 2003). The instruments are thus correlated with the neighbors' decisions.

Table 4a presents the results from running the discrete response analysis with the use of instruments, when neighbors must be located within 5 miles of a tract. Irrespective of the probability model (linear or probit), the time period (year or quarter) and whether or not the restriction (v) that all of tract i's neighbors at time t must be sold on a different date from tract i is imposed, the effect of neighbors' decisions is statistically insignificant and small: according to the results from the base case specification (1), a change in the percent of neighbors who explored in the previous period from 0 percent to 100 percent would only increase a firm's probability of exploration by a statistically insignificant 0.14. Negative effects greater than

0.15 and positive effects greater than 0.43 on a firm's probability of exploration can be rejected at a 5% level.

As with the uninstrumented regressions, the signs of the coefficients on the other covariates are as expected. According to the results from the base case specification (1), an increase in the real pre-sale value of a tract of $10,000 per acre would increase the probability of exploration by 0.66. Thus, pre-sale values, which vary greatly from $0 to $18,230 per acre, have a very large and statistically significant effect on exploration decisions. Similarly, an increase in the winning bid of $10,000 per acre would increase the probability of exploration by 0.34. Thus, winning bids, which vary from $450 to $60,800 per acre, have a large and statistically significant effect as well. The coefficients on the pre-sale value and on the winning bid indicate that, all else equal, tracts are more likely to be explored if their ex ante estimated values are high.

The coefficient on acreage indicates that an increase in tract size of 100 acres increases the probability of exploration by 0.40. Thus, larger tracts are more likely to be explored. One likely explanation is that because larger tracts cover more surface area, the probability that oil and gas reserves are present is higher.

The coefficient on the dummy for being in the last year of the lease term indicates that, all else equal, a tract's probability of being explored is higher by 0.13 when it is in the last year of its lease term.

In addition to being robust to the probability model, the time period and whether or not the restriction (v) that all of tract i's neighbors at time t must be sold on a different date from tract i is imposed, the discrete response results are also robust to the distance used to delineate neighbors; this can be readily seen in Table 4b, which compares the results from using 5 miles as the cutoff in the base case to those from using 4 miles, 6 miles and 10 miles in the linear probability model.

To examine whether strategic interactions have a larger effect on smaller tracts, where externalities arising from common-pool considerations are likely to be more prevalent, the sample is also stratified by tract size (in acres). The results are presented in Table 5. Specification (1) displays the results for larger tracts (i.e., tracts greater than or equal to 5000 acres in size); specification (2) displays the results for smaller tracts (i.e., tracts less than 5000 acres in size). As the results indicate, neighbors do not have a significant effect in either sample. Specification (3) includes a term that interacts the fraction of neighbors who drilled at time $t-1$ with the tract size, and is run on all tract sizes.[14] The coefficient on the fraction of neighbors who drilled at time $t-1$ and the coefficient on the interaction term are both statistically insignificant. The result that neighbors do not have a significant effect on a firm's probability of exploratory drilling is therefore robust to tract size.

Thus, regardless of whether instruments are used or not, the results of my discrete response analysis

[14] This interaction term is instrumented with interactions between the original instruments and acreage.

do not indicate that a firm's exploratory drilling decision depends on those of its neighbors.[15] Test results rejecting the endogeneity of neighbors' decisions to a firm's own decision provide further evidence that neighbors do not base their decisions on each other. My results are consistent with the weak results of Hendricks and Porter (1996): in their regressions of the probability of initial exploration, the coefficients on the variables they use to capture the neighborhood exploratory drilling experience are, for the most part, not significant. Thus, even though one may expect information externalities to be particularly acute on wildcat tracts, information and extraction externalities do not appear to induce firms to interact strategically on net.

My discrete response model of a firm's exploratory drilling decision is a reduced-form specification of the firm's dynamic decision-making problem. The advantages of this reduced-form approach lie in its simplicity and its ability to enable identification of the parameter of interest via instrumental variables. One drawback is that the connection between the reduced-form parameters and the actual structural parameters governing the firm's dynamic decision-making problem is unclear. A second drawback is that it ignores the firm's development investment decision.[16] Like exploration, development – or the installation of production platforms – is a costly and irreversible investment, and a firm's development decision might depend on the actions of its neighbors. Moreover, a neighbor's development decision is likely to impose substantial information and extraction externalities on a firm. A third drawback is that while the reduced-form model generates results on how a firm's drilling decision is affected by neighbors' decisions, it does not indicate how the firm's profits are affected by neighbors' decisions. Although the externalities may not have any net effect on a firm's exploratory drilling decision, they might have effects on its profits, and therefore on the efficiency of petroleum production. To address these concerns, I use a structural econometric approach as a second means of analyzing whether firms interact strategically. It is to this approach that I now turn.

3 The Multi-Stage Investment Timing Game

3.1 Contributions to Existing Literature

The second econometric approach I employ to investigate whether firms interact strategically when they make their investment timing decisions is a structural econometric model of the firms' dynamic strategic

[15] Results from a Cox proportional hazards model without instruments also do not indicate that firms interact strategically with their neighbors. The development of an instrumental variables analog of the Cox model using nonlinear instrumental variables techniques will be the subject of future work.

[16] The main reason the development decision is excluded from the reduced-form analysis is that, just as for exploration, a neighbor's development decision is potentially endogenous. Development is not subject to a lease term, however, and thus, unlike for exploration, the timing of the lease term could not be exploited for instruments. The design of suitable instruments for the development decisions of one's neighbors will be the subject of future work.

multi-stage decision-making game. The question I pursue is: *How does a firm's profits depend on the exploration and development decisions of its neighbor?*

My structural econometric analysis improves upon the existing literature on the information externality that arises in offshore petroleum production in several ways. First, unlike the theoretical models and reduced-form empirical analyses conducted by Hendricks, Porter and their co-authors, a structural approach yields estimates of the structural parameters of the discrete choice dynamic game. With these structural parameters, one can identify the effects of a neighbor's exploration and development decisions on the profits a firm would get from developing its tract.

A second way in which my work contributes to the existing literature on the information externality in offshore petroleum production is that it combines the externality problem with real options theory. Oil production is a multi-stage process involving sequential investment decisions. Since the decision to explore a reserve entails an irreversible investment, the value of an unexplored reserve is the value of the option to invest in exploration. Similarly, the value of an explored but undeveloped reserve is the value of the option to invest in development. There is thus an option value to waiting before making either investment because the value of a developed reserve can change, either because exogenous conditions such as the oil price might change, or because there is a chance that neighboring firms might explore or develop first. Moreover, because these two types of investment are made sequentially, they act as compound options: completing one stage gives the firm an option to complete the next (Dixit & Pindyck, 1994).

While literature on the financial theory of option valuation is abundant, structural models applying the theory to the oil production process that account for strategic considerations have yet to be developed. Hurn and Wright (1994) test reduced-form implications of the theory via a hazard model, but neither estimate a structural model nor account for possible strategic interactions. Paddock, Siegel and Smith (1998) compare the option valuation estimate of the market value of selected offshore petroleum tracts with estimates from other valuation methods and with the winning bids, but do not account for either information externalities or extraction externalities. Similarly, Pesaran (1990) estimates an intertemporal econometric model for the joint determination of extraction and exploration decisions of a "representative" profit-maximizing oil producer, but does not examine the case of multiple producers that may interact strategically.

The third innovation I make to the existing literature on the information externality in offshore petroleum production is that while the existing literature focuses exclusively on externalities that arise during exploratory drilling, my model allows for extraction externalities as well as information externalities that arise during both exploration and development. If firms do indeed learn about the value of their own tracts from the actions of their neighbors, then one would expect firms to update their own beliefs not only if their neighbors begin exploratory drilling, but also if, after having already begun exploring, the neighbors then decide to install a production platform. That a neighbor has decided to begin extracting after it explored should be at least as informative as the initiation of exploration in the first place. Furthermore, extraction externalities are another form of spillover that is not accounted for by previous studies of the investment timing game, and, unlike the information externality, is one that may have a negative effect on a firm's profits.

In addition to the literature on the information externality, a second branch of related literature is that on econometric models of discrete dynamic games (see e.g. Aguirregabiria & Mira, 2004; Bajari, Benkard & Levin, 2004 and references therein). In particular, my work applies a method developed by Pakes, Ostrovsky and Berry (2005) for estimating parameters of discrete dynamic games such as those involving firm entry and exit. This paper builds upon the work of Pakes et al. (2005) in several ways. First, unlike their paper, which uses simulated data, my paper estimates a discrete dynamic game using actual data. Second, while the entry and exit decisions they examine are two independent investments, the exploration and development decisions I examine are sequential investments: the decision to invest in development can only be made after exploration has already taken place. Thus, unlike the one-stage entry and exit games, the investment timing game is a two-stage game. The sequential nature of the investments is an added complexity that I address in my econometric model. Third, whereas the estimators Pakes et al. propose are for infinite-horizon dynamic games, the exploration stage of petroleum production is a finite-horizon dynamic optimization problem: firms must begin exploration before the end of the five-year lease term, or else relinquish their lease. As a consequence, an appropriate modification to the estimation algorithm is required. A fourth innovation I make is that, unlike Pakes et al., I do not assume that the profit function is a known function of the underlying state variables, but instead estimate its parameters from data. While Pakes et al. are able to appeal to economic theory to posit an exact form for profits as a function of state variables such as the number of firms in the industry, I cannot. No economic theory predicts the relationship between profits and such state variables as whether or not a firm's neighbors explore or develop. Indeed, since the relationship between profits and the actions of one's neighbor is among the very parameters I hope to identify, I choose to estimate this relationship from the data rather than impose it *a priori*. The task of estimating these additional parameters requires the use of additional moment conditions.

There are several advantages to using a structural model. First, a structural model enables the estimation of all the structural parameters of the underlying dynamic game. These parameters include not only those governing the relationship between various state variables and the profits of firms, but also parameters governing the distribution of tract-specific private information. Second, the structural model addresses the endogeneity problems without the need for instruments. Because the structural model is based on the equilibrium of the underlying dynamic game, it addresses the simultaneity problem directly by explicitly modeling the firms' strategies. Moreover, the problem of spatially correlated unobservables can be addressed by interpreting the profits in the model as expected profits conditional on observables. A third advantage to a structural model is that it enables one to estimate how a firm's profits are affected by the decisions of its neighbors; the sign of the effect indicates the net sign of the information and extraction externalities. Fourth, the structural model enables one to explicitly model each of the stages of the multi-stage dynamic decision-making problem faced by petroleum-producing firms. It is to this model that I now turn.

3.2 A Model of the Investment Timing Game

In my model of the investment timing game, each "market" k consists of an isolated neighborhood of adjacent tracts i that were each leased to a petroleum-producing firm on the same date. Time t denotes the number of years after the lease sale date. Firms must begin exploration before time T, the length of the lease term, or else relinquish their lease. Let the "lease term time" τ_{kt} of market k at time t be given by:

$$\tau_{kt} = \begin{cases} t & \text{if} \quad t = 0, 1, ..., T-1 \\ T & \text{if} \quad t \geq T \end{cases}$$

For each market k, the state of the market t years after the leases began is given by a vector Ω_{kt} of discrete and finite-valued state variables that are observed by all the firms in market k and as well as by the econometrician. Let θ denote the vector of parameters to be estimated.

At the beginning of each period t, the owner of each tract i must make one of two investment decisions. If tract i has not been explored before time t, its owner must decide whether to invest in exploration at time t. If tract i has been explored but has not been developed before time t, its owner must decide whether to invest in development at time t. For each period t, all firms make their time-t investment decisions simultaneously.

Each firm's time-t investment timing decision depends in part on the state of the market $\Omega_{kt} \equiv (N_{kt}, X_{kt}, \tau_{kt})$, which can be decomposed into endogenous state variables N_{kt}, exogenous profit-shifting state variables X_{kt}, and the lease term time τ_{kt}. Investment decisions depend on N_{kt} and X_{kt} because these state variables are assumed to affect profits. Because of the finite-horizon nature of the firm's exploration investment problem, the finite-valued and exogenous lease term time τ_{kt} affects investment decisions as well, as will be explained below. In the present model, there are two endogenous state variables N_{kt}: the total number of tracts in market k that have been explored before time t, and the total number of tracts in market k that have been developed before time t. These endogenous state variables capture the strategic component of the firms' investment timing decisions. The exogenous state variables X_{kt} include the drilling cost and the oil price and are assumed to evolve as a finite state first-order Markov process: $X_{k,t+1} \overset{iid}{\sim} F_X(\cdot|X_{kt})$. In other words, the next period's value $X_{k,t+1}$ of the exogenous state variables are assumed to be independently and identically distributed (i.i.d.) with a probability distribution that depends only on the time-t realization X_{kt} of the exogenous state variables, and not additionally on what happened before time t (Dixit & Pindyck, 1994).[17]

In addition to the publicly observable state variables Ω_{kt}, each firm's time-t investment timing decision also depends on two types of shocks that are private information to the firm and unobserved by either other firms or by the econometrician. The first source of private information is a pre-exploration shock μ_{it} to an unexplored tract i at time t. This pre-exploration shock, which is only observed by the firm owning tract i, represents any and all private information that affects the exploration investment decision made on tract i at time t. Such private information may include, for example, idiosyncratic shocks to exploration costs and the outcome of the post-sale, pre-exploration seismic study conducted on tract i at time $t-1$.[18] Following Pakes, Ostrovsky and Berry (2005), assume that the pre-exploration shock μ_{it} is an independently and identically distributed random variable with an exponential distribution and mean σ_μ. That is, $\mu_{it} \overset{iid}{\sim} exponential(\sigma_\mu)$.

The second source of private information is a pre-development shock ε_{it} to an explored but undeveloped tract i at time t. This pre-development shock, which is only observed by the firm owning tract i, represents any and all private information that affects the development investment decision made on tract i at time t. Such private information may include, for example, the outcome of the exploratory drilling conducted on tract i at time $t-1$. Following Pakes, Ostrovsky and Berry (2005), assume that the pre-development shock ε_{it} is an independently and identically distributed random variable with an exponential distribution and mean σ_ε. That is, $\varepsilon_{it} \overset{iid}{\sim} exponential(\sigma_\varepsilon)$. In addition, assume that the pre-exploration shocks μ_{it} and the pre-development shocks ε_{it} are independent of each other.[19] All distributions are common knowledge.

In the absence of strategic considerations, the firm owning tract i would base its investment timing decisions on only the exogenous state variables X_{kt}, the lease term time τ_{kt}, and the private shocks μ_{it} and

[17] The lease term time τ_{kt} evolves as a finite state first-order Markov process as well. I include this exogenous finite-valued variable τ_{kt} as a separate argument distinct from X_{kt} both because it does not affect profits and also to elucidate my later exposition of the finite-horizon nature of the exploration stage.

[18] Firms conduct and analyze seismic studies in order to help them decide whether or not to begin exploratory drilling (John Shaw, personal communication, 18 April 2003; Bob Dye, Apache, personal communication, 21 January 2004; Jon Jeppesen, Apache, personal communication, 21 January 2004; Mark Bauer, Apache, personal communication, 21 January 2004; Billy Ebarb, Apache, personal communication, 22 January 2004).

[19] The assumptions that both types of shocks are i.i.d. and independent of each other, while restrictive, are needed in order for the estimation technique used in this paper to work. If either type of shock were serially correlated (or if, at the extreme, there were tract fixed effects), then firms would base their decisions not only on the current values of the state variables and of their shocks, but also on past values of the state variables and shocks as well. The state space would then be too large. If the distribution of the pre-development shock ε_{it} depended on the realization of the pre-exploration shock μ_{it} (e.g., the μ_{it} at the time of exploration), then μ_{it} would be a state variable in the development stage of production. As a consequence, the econometrician would need to observe μ_{it}, which she does not. The i.i.d. assumption is reasonable if the shocks are interpreted to encompass all idiosyncratic factors affecting investment decisions, including managerial shocks and technological shocks. Moreover, since one of my state variables is the average winning bid, which is a measure of tract value, it is reasonable to assume that, conditional on tract value, shocks are i.i.d.

ε_{it}. To derive its dynamically optimal investment policy, it would solve a single-agent dynamic programming problem.

If information and extraction externalities were present, however, then strategic considerations would become important. As a consequence, the exploration and development investment decisions of the firm owning tract i in market k would depend on the exploration and development investment decisions of the firms owning the other tracts in market k. In other words, the firm owning tract i would base its investment timing decisions not only on the exogenous state variables X_{kt}, the lease term time τ_{kt}, and the private shocks μ_{it} and ε_{it}, but also on the endogenous state variables N_{kt} as well, namely the total number of tracts in its market k that have been explored before time t and the total number of tracts in market k that have been developed before time t. Each firm would then no longer solve merely a single-agent dynamic programming problem, but rather a multi-agent dynamic game.

The equilibrium concept used in the model is that of a Markov perfect equilibrium. Each firm is assumed to play a Markov "state-space" strategy: the past influences current play only through its effect on the state variables. A firm's dynamically optimal investment policy is then the Markov strategy that it plays in the Markov perfect equilibrium, which is a profile of Markov strategies that yields a Nash equilibrium in every proper subgame (Fudenberg & Tirole, 1998).

While each firm's time-t investment decision depends on both the publicly available endogenous and exogenous state variables Ω_{kt} as well as the firm's own private information μ_{it} or ε_{it}, its perception of its neighbor's time-t investment decisions depend only on the publicly observable state variables Ω_{kt}. This is because, owing to the above assumptions on the observable state variables and on the unobservable shocks, firms can take expectations over their neighbors' private information.[20] In equilibrium, firms' perceptions of their neighbors' investment probabilities should be consistent with those that are actually realized (Starr & Ho, 1969).

The model has at least one Markov perfect equilibrium, and each equilibrium generates a finite state Markov chain in Ω_{kt} tuples (Pakes, Ostrovsky & Berry, 2005).[21] Although model assumptions do not guarantee a unique equilibrium, they do insure that there is only one set of equilibrium policies that is consistent with the data generating process. It is thus possible to use the data itself to pick out the equilibrium that is played. For large enough samples, the data will pick out the correct equilibrium and the estimators for the parameters in the model will be consistent (Pakes, Ostrovsky & Berry, 2005).

The firm's dynamic decision-making problem is as follows. The first-stage problem is to determine the

[20] While each firm plays a pure strategy, from the point of view of their neighbors, they appear to play mixed strategies. Thus, as with Harsanyi's (1973) purification theorem, a mixed distribution over actions is the result of unobserved payoff perturbations that sometimes lead firms to have a strict preference for one action, and sometimes a strict preference for another.

[21] A Markov chain is a Markov process on a finite state space (Stokey, Lucas & Prescott, 1989).

optimal policy for investment in exploration. Because firms must begin exploration before the end of their lease term, or else relinquish their lease, this is a finite-horizon problem. As a consequence, firms' decisions will depend not only on the profit-shifting state variables N_{kt} and X_{kt}, but also on time t. However, since firms can only make exploration decisions at the beginning of periods $t = 0, ..., T-1$, the time dependence only applies until time $t = T-1$, after which exploration can no longer begin and the endogenous variable that counts the total number of tracts in the market that have been explored stays constant. It is for this reason that the exogenous and finite-valued state variable "lease term time" τ_{kt} captures the entire time dependence of the problem.

The second stage of the firm's dynamic decision-making problem is to determine the optimal timing for investing in the development of a tract that has already been explored. This second-stage problem has both a finite-horizon component and an infinite-horizon component. A firm's development strategy depends in part on its perceptions of the future exploration policies of the firms in the market. Since exploration policies depend on time until time $t = T-1$, this means that perceptions, and therefore development strategies, will depend on time for $t < T$. As a consequence, the dynamic programming problem for time $t < T$ is a finite-horizon problem. However, because the lease term only applies to the exploration stage of production, and because the endogenous variable that counts the total number of tracts in the market that have been explored – a variable that depends on the time-dependent exploration policies of the firms in the market – stays constant after the lease term expires, the dynamic programming problem for the development stage from time T onwards is an infinite-horizon problem that does not depend on time. Thus, once again, the lease term time τ_{kt} sufficiently captures the entire time dependence of the problem.

The firm's sequential investment problem is a two-stage optimization problem, and can be solved backwards using dynamic programming (Dixit & Pindyck, 1994). In the second, or development, stage of oil production, a firm with an explored but undeveloped tract i must decide if and when to invest in a production platform. Assume that the profit $\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta)$ that a firm will get after developing tract i at time t can be separated into a deterministic component and a stochastic component as follows:

$$\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) = \pi_0^d(\Omega_{kt}; \theta) + \varepsilon_{it} \ , \tag{3}$$

where the deterministic component of profit is linear in the publicly observable state variables:

$$\pi_0^d(\Omega_{kt}; \theta) \equiv N'_{kt}\gamma_N + X'_{kt}\gamma_X \ , \tag{4}$$

and where the stochastic component is the privately observed pre-development shock ε_{it}.[22] The development

[22] If there were additional market state variables that affected profits but were unobserved by the econometrician, then

profit is therefore independent of time (and lease term time) except through the state variables (N_{kt}, X_{kt}) and the shock ε_{it}.

Let $\gamma \equiv (\gamma_N, \gamma_X)$ denote the vector of all the coefficients in the development profit function. The coefficients γ_N in the profit function on the endogenous state variables N_{kt} – the total number of tracts in the market that have been explored and the total number of tracts in the market that have been developed – indicate whether and how one firm's profits depend on the production decisions of its neighbors. If a neighbor explores, then the state variable counting the total number of tracts in the market that have been explored increases by one and the value of the development profits increase by the value of its coefficient. Similarly, if a neighbor develops, then the state variable counting the total number of tracts in the market that have been developed increases by one and the value of the development profits increase by the value of its coefficient. The coefficients γ_N on the endogenous variables thus measure the net effects of the information and extraction externalities, and therefore indicate whether firms interact strategically on net. Positive values of the coefficients γ_N would indicate that the information and extraction externalities were positive on net, and therefore that the information externality was dominant. Negative values would indicate that the externalities were negative on net, and therefore that the extraction externality was dominant.

The value V^e of an explored but undeveloped tract i in market k at time t is given by:

$$V^e(\Omega_{kt}, \varepsilon_{it}; \theta) = \max\{\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta), \beta V^{ce}(\Omega_{kt}; \theta)\}, \tag{5}$$

where $\beta \in (0,1)$ is the discount factor and $V^{ce}(\Omega_{kt}; \theta)$ is the continuation value to waiting instead of developing at time t. For the structural estimation, I set the discount factor β to 0.9. The continuation value to waiting is the expectation over the state variables and shocks of next period's value function, conditional on not developing this period:

$$V^{ce}(\Omega_{kt}; \theta) = E\left[V^e(\Omega_{k,t+1}, \varepsilon_{i,t+1}; \theta) | \Omega_{kt}, I^d_{it} = 0\right], \tag{6}$$

where I^d_{it} is an indicator for whether development began on tract i at time t.

Let $g^d(\Omega_{kt}; \theta)$ denote the probability of developing an explored but undeveloped tract i at time t conditional on the publicly available information Ω_{kt} on time t, but not on the private information ε_{it}. The development probability $g^d(\Omega_{kt}; \theta)$ function represents a firm's perceptions of the probability that a neighbor owning an explored but undeveloped tract will decide to develop its tract in period t, given that the state of their market at time t is Ω_{kt}. Moreover, a firm's expectation of its own probability of development in the

$\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta)$ can be interpreted as the expected profits conditional on the available information Ω_{kt} (Pakes, Ostrovsky & Berry, 2005). Under this interpretation, spatially correlated unobservables do not pose a concern.

next period is simply the expected value of the next period's development probability, conditional on this period's state variables: $E[g^d(\Omega_{k,t+1};\theta)|\Omega_{kt}]$.

Using the exponential distribution for ε_{it} and equation (3) for development profits as shown in Appendix A, the continuation value $V^{ce}(\cdot)$ can be reduced to:

$$V^{ce}(\Omega_{kt};\theta) = E[\beta V^{ce}(\Omega_{k,t+1};\theta) + \sigma_\varepsilon g^d(\Omega_{k,t+1};\theta)|\Omega_{kt}, I^d_{it} = 0], \tag{7}$$

and the development probability $g^d(\cdot)$ can be reduced to the following function of the continuation value, the state variables and the parameters:

$$g^d(\Omega_{kt};\theta) \ = \exp\left(-\frac{\beta V^{ce}(\Omega_{kt};\theta) - \pi^d_0(\Omega_{kt};\theta)}{\sigma_\varepsilon}\right). \tag{8}$$

In the first, or exploration, stage of oil production, a firm with an unexplored tract i must decide if and when to invest in exploratory drilling. Owing to the sequential nature of the investments, the publicly observable deterministic component of the payoff $\pi^e_0(\cdot)$ to exploring in the first stage is equal to the expected value of having an explored but undeveloped tract in the second stage, net the cost of exploration $c^e(\cdot)$:

$$\pi^e_0(\Omega_{kt};\theta) \equiv E_\varepsilon\left[V^e(\Omega_{kt},\varepsilon_{it};\theta)|\Omega_{kt}\right] - c^e(\Omega_{kt};\theta), \tag{9}$$

where the exploration cost is assumed to be linear in the exogenous cost-shifting state variables:[23]

$$c^e(\Omega_{kt};\theta) = -X'_{kt}\alpha\ . \tag{10}$$

Assume that the actual payoff $\pi^e(\cdot)$ to exploring tract i at time t also includes a privately observed stochastic component as well:

$$\pi^e(\Omega_{kt},\mu_{it};\theta) = \pi^e_0(\Omega_{kt};\theta) + \mu_{it}\ , \tag{11}$$

where the stochastic component is the pre-exploration shock μ_{it}.

The value V^n of an unexplored tract i in market k and time t is given by:

$$V^n(\Omega_{kt},\mu_{it};\theta) = \max\{\pi^e(\Omega_{kt},\mu_{it};\theta), \beta V^{cn}(\Omega_{kt};\theta)\}, \tag{12}$$

where $V^{cn}(\Omega_{kt};\theta)$ is the continuation value to waiting instead of exploring at time t. The continuation value

[23] I define costs with a negative sign so that the coefficients can be interpreted as coefficients in the exploration profit function. Variables that increase cost will decrease profit, and vice versa.

to waiting is the expectation over the state variables and shocks of next period's value function, conditional on not exploring this period:

$$V^{cn}(\Omega_{kt};\theta) = E\left[V^n(\Omega_{k,t+1},\mu_{i,t+1};\theta)|\Omega_{kt}, I^e_{it}=0\right], \tag{13}$$

where I^e_{it} is an indicator for whether exploration began on tract i at time t. The lease term imposes the following boundary condition:

$$V^n\left((N,X,T),\mu;\theta\right) = 0 \quad \forall N,X,\mu\ . \tag{14}$$

Let $g^e(\Omega_{kt};\theta)$ denote the probability of exploring an unexplored tract i at time t conditional on the publicly available information Ω_{kt} on time t, but not on the private information μ_{it}. As with the development probability, the current value of the exploration probability represents a firm's perceptions of the probability that a neighbor owning an unexplored tract will decide to explore its tract in period t, given that the state of their market at time t is Ω_{kt}; its expected value at time $t+1$ represents a firm's expectation of its own probability of exploration in the next period.

Using the exponential distribution for μ_{it} and equation (11) for exploration profits as shown in Appendix A, the continuation value $V^{cn}(\cdot)$ to waiting instead of exploring can be reduced to:

$$V^{cn}(\Omega_{kt};\theta) = E\left[\beta V^{cn}(\Omega_{k,t+1};\theta) + \sigma_\mu g^e(\Omega_{k,t+1};\theta)|\Omega_{kt}, I^e_{it}=0\right], \tag{15}$$

and the exploration policy function $g^e(\cdot)$ can be reduced to the following function of the continuation values, state variables and parameters:

$$g^e(\Omega_{kt};\theta) \;= \exp\left(-\frac{\beta V^{cn}(\Omega_{kt};\theta) - \left(\beta V^{ce}(\Omega_{kt};\theta) + \sigma_\varepsilon g^d(\Omega_{kt};\theta)\right) + c^e(\Omega_{kt};\theta)}{\sigma_u}\right). \tag{16}$$

Owing to the sequential nature of the investment decisions, the continuation value $V^{ce}(\cdot)$ and the investment probability $g^d(\cdot)$ from the development stage appear in the expression for the investment probability $g^e(\cdot)$ in the exploration stage.

The ex ante expected value of an unexplored tract at time $t=0$, where expectations are taken over the pre-exploration shock μ, is given by:

$$E_\mu\left[V^n(\Omega_{k0},\mu_{i0};\theta)|\Omega_{k0}\right] = \beta V^{cn}(\Omega_{k0};\theta) + \sigma_\mu g^e(\Omega_{k0};\theta). \tag{17}$$

3.3 The Structural Econometric Model

The econometric estimation technique I use employs a two-step semi-parametric estimation procedure. It is an extension of the estimator proposed by Pakes, Ostrovsky and Berry (2005) to finite-horizon, multi-stage games. In the first step, the continuation values are estimated non-parametrically and these estimates are used to compute the predicted probabilities of exploration and development. In the second step, the parameters $\theta \equiv (\sigma_\mu, \sigma_\varepsilon, \gamma, \alpha)'$ are estimated by matching the predicted probabilities with the actual probabilities in the data. I will now describe each step in turn.

3.3.1 Step 1: Estimating continuation values and predicted probabilities

The first step entails computing the non-parametric[24] estimators $V^{ce}\widehat{(\Omega_{kt};\theta)}$ and $V^{cn}\widehat{(\Omega_{kt};\theta)}$ for the continuation values $V^{ce}(\Omega_{kt};\theta)$ and $V^{cn}(\Omega_{kt};\theta)$, respectively. To do so, historical empirical frequencies are used to estimate the elements of the Markov transition matrix governing the evolution of the finite-valued state variables from one period to the next. Estimators for the continuation values are subsequently derived from equations (7) and (15) using dynamic programming. These estimators are then substituted into equations (8) and (16) to obtain predicted probabilities for development and exploration, respectively.[25]

Formally, the non-parametric estimator $V^{ce}\widehat{(\Omega_{kt};\theta)}$ for $V^{ce}(\Omega_{kt};\theta)$ is derived from equation (7) and is computed as follows. For each period t, let each component of the vector $\overrightarrow{V_t^{ce}}$ be $V^{ce}(\Omega_{kt};\theta)$ evaluated at a different tuple of state variables. Similarly, for each period t, let each component of the vector $\overrightarrow{g_t^d}$ be $g^d(\Omega_{kt};\theta)$ evaluated at a different tuple of state variables. Finally, for each lease term time period τ, let M_τ^e be a transition matrix from the point of view of an owner of an explored but undeveloped tract who decides not to develop at time t. The element in the i^{th} row and j^{th} column is the probability that the state tuple next period will be the j^{th} tuple, given that the state tuple this period is the i^{th} tuple, given that the tract has already been explored but not yet developed at time t, and conditional on not developing at time t.

The estimator $V^{ce}\widehat{(\Omega_{kt};\theta)}$ is obtained from rewriting equation (7) in vector form:

$$\overrightarrow{V_t^{ce}} = M_\tau^e \left(\beta \overrightarrow{V_{t+1}^{ce}} + \sigma_\varepsilon \overrightarrow{g_{t+1}^d} \right). \tag{18}$$

The estimator is obtained after further substituting in the empirical average $\widehat{M_\tau^e}$ for M_τ^e. For $t \geq T$, since $\overrightarrow{V_t^{ce}} = \overrightarrow{V_{t+1}^{ce}}\ \forall t \geq T$, we can solve for a fixed point $\widehat{V^{ce}}\left((N_{kt}, X_{kt}, T;\theta\right)$, which, from Blackwell's Theorem, is

[24] The continuation values are non-parametric functions of the state variables Ω_{kt} conditional on the parameters θ.

[25] Rather than use historical empirical frequencies to estimate the Markov transition matrix, it is possible to compute an estimator for the matrix using the estimators for the exploration and development probabilities. However, because the latter, more complicated approach imposes a computational burden and because Pakes, Ostrovsky and Berry (2005) find that it did not improve the performance of their estimator, I choose the former, simpler approach.

unique. To obtain the estimator of the value function for $t < T$, we then iterate backwards in time from $t = T$ using $\overrightarrow{V_T^{ce}} = \widehat{V^{ce}}\left((N_{kt}, X_{kt}, T; \theta\right)$ as a boundary condition. The predicted probability of development is then given by:[26]

$$g^d(\widehat{\Omega_{kt}}; \theta) \ = \exp\left(-\frac{\beta V^{ce}(\widehat{\Omega_{kt}}; \theta) - \pi_0^d(\Omega_{kt}; \theta)}{\sigma_\varepsilon}\right). \tag{19}$$

The non-parametric estimator $V^{cn}(\widehat{\Omega_{kt}}; \theta)$ for the continuation value to waiting instead of exploring is derived from equation (15) and is computed in a similar fashion as the estimator of $V^{ce}(\cdot)$. For each period t, let each component of the vector $\overrightarrow{V_t^{cn}}$ be $V^{cn}(\cdot)$ evaluated at a different tuple of state variables. For each period t, let each component of the vector $\overrightarrow{g_t^e}$ be $g^e(\cdot)$ evaluated at a different tuple of state variables. Finally, for each lease term time period τ, let M_τ^n be a transition matrix from the point of view of an owner of an unexplored tract who decides not to explore at time t. The element in the i^{th} row and j^{th} column is the probability that the state tuple next period will be the j^{th} tuple, given that the state tuple this period is the i^{th} tuple, given that the tract has yet to be explored, and conditional on not exploring at time t.

The estimator $V^{cn}(\widehat{\Omega_{kt}}; \theta)$ is obtained from rewriting equation (15) in vector form:

$$\overrightarrow{V_t^{cn}} = M_\tau^n \left(\beta \overrightarrow{V_{t+1}^{cn}} + \sigma_\mu \overrightarrow{g_{t+1}^e}\right). \tag{20}$$

Substituting in the empirical average $\widehat{M_\tau^n}$ for M_τ^n, we can solve backwards in time from the boundary condition $\overrightarrow{V_{T-1}^{cn}} = 0$ implied by equation (14) to obtain $\widehat{V^{cn}}(\Omega_{kt}; \theta)$ for all $t \leq T-1$. The predicted probability of exploration is then given by:[27]

$$g^e(\widehat{\Omega_{kt}}; \theta) \ = \exp\left(-\frac{\beta V^{cn}(\widehat{\Omega_{kt}}; \theta) - \left(\beta V^{ce}(\widehat{\Omega_{kt}}; \theta) + g^d(\widehat{\Omega_{kt}}; \theta) \cdot \sigma_\varepsilon\right) + c^e(\Omega_{kt}; \theta)}{\sigma_u}\right). \tag{21}$$

Owing to the sequential nature of the investment decisions, the estimators $V^{ce}(\widehat{\Omega_{kt}}; \theta)$ and $g^d(\widehat{\Omega_{kt}}; \theta)$ of the continuation value and the investment probability, respectively, from the development stage are needed to form the estimator of the investment probability $g^e(\widehat{\Omega_{kt}}; \theta)$ in the exploration stage.

[26] In practice, I use: $g^d(\widehat{\Omega_{kt}}; \theta) \ = \min\left\{\exp\left(-\frac{\beta V^{ce}(\widehat{\Omega_{kt}}; \theta) - \pi_0^d(\Omega_{kt}; \theta)}{\sigma_\varepsilon}\right), 1\right\}$.

[27] In practice, I use: $g^e(\widehat{\Omega_{kt}}; \theta) \ = \min\left\{\exp\left(-\frac{\beta V^{cn}(\widehat{\Omega_{kt}}; \theta) - \left(\beta V^{ce}(\widehat{\Omega_{kt}}; \theta) + g^d(\widehat{\Omega_{kt}}; \theta) \cdot \sigma_\varepsilon\right) + c^e(\Omega_{kt}; \theta)}{\sigma_u}\right), 1\right\}$.

3.3.2 Step 2: Generalized Method of Moments

After obtaining estimates of the continuation values and predicted probabilities as functions of the state variables Ω_{kt} and the parameters θ in the first step, I estimate the parameters θ in the second step using generalized method of moments (GMM). The moments I construct involve matching the probabilities of exploration and development predicted by the model, as given by equations (21) and (19), with the respective empirical probabilities $\overline{g^e(\Omega_{kt})}$ and $\overline{g^d(\Omega_{kt})}$ in the data. I also form moments that match, for those tracts that developed, the expected development profits conditional on development predicted by the model with the actual average realized profits $\overline{\pi^d(\Omega_{kt})}$ in the data. Additional moments are constructed by interacting the above moments with the state variables. The moment function $\Psi(\Omega_{kt}, \theta)$ is therefore:

$$
\begin{array}{l}
\left(g^e(\widehat{\Omega_{kt}};\theta) \ - \overline{g^e(\Omega_{kt})}\right) \cdot n^{not\_yet\_e}(\Omega_{kt}) \\
\left(g^d(\widehat{\Omega_{kt}};\theta) - \overline{g^d(\Omega_{kt})}\right) \cdot n^{e\_not\_yet\_d}(\Omega_{kt}) \\
\left(\left(E\left[\pi^d(\Omega_{kt}, \varepsilon_{it};\theta) | \pi^d(\Omega_{kt}, \varepsilon_{it};\theta) > \beta V^{ce}(\Omega_{kt};\theta)\right] - \overline{\pi^d(\Omega_{kt})}\right) \cdot n^d(\Omega_{kt})\right) \\
\Omega_{kt}' \left(\left(g^e(\widehat{\Omega_{kt}};\theta) \ - \overline{g^e(\Omega_{kt})}\right) \cdot n^{not\_yet\_e}(\Omega_{kt})\right) \\
\Omega_{kt}' \left(\left(g^d(\widehat{\Omega_{kt}};\theta) - \overline{g^d(\Omega_{kt})}\right) \cdot n^{e\_not\_yet\_d}(\Omega_{kt})\right)
\end{array}
$$

where, for each state of the market Ω_{kt}, $n^{not\_yet\_e}(\Omega_{kt})$ is the number tracts that have yet to be explored, $n^{e\_not\_yet\_d}(\Omega_{kt})$ is the number of tracts that have been explored but not developed, and $n^d(\Omega_{kt})$ is the number of tracts that have been developed. The GMM estimator $\widehat{\theta}$ is then the solution to the problem:

$$
\min_{\theta} \left[\frac{1}{n}\sum_{k,t}\Psi(\Omega_{kt}, \theta)\right]' W_n^{-1} \left[\frac{1}{n}\sum_{k,t}\Psi(\Omega_{kt}, \theta)\right], \tag{22}
$$

where n is the number of observations (i.e., the number of market-time pairs) and W_n is a weight matrix. In the present estimation, the number of moments that involved interactions with state variables is chosen so that the number of moments is equal to the number of parameters. Because the system is therefore exactly identified, the weight matrix used is the identity matrix.[28]

Standard errors are formed by a parametric bootstrap (Pakes, Ostrovsky & Berry, 2005), as follows. Estimates of the continuation values and the parameters are used to solve for the exploration and development probabilities arising in the Markov perfect equilibrium. These probability functions are then used in

[28] If the system is overidentified, a two-step GMM estimator can be used (Hansen, 1982; Graham, 2005). In the first step, a preliminary estimate of θ is obtained using the identity matrix as the weight matrix. In the second step, this preliminary estimate of θ is used to construct the optimal weighting matrix as specified by Chamberlain (1987), which is then used to obtain the final estimate of θ.

conjunction with the empirical transition matrix for the exogenous variables and with random draws, with replacement, from the empirical distribution of initial conditions to simulate 100 independent panels of size equal to the actual sample size. Finally, the structural econometric model is run on each of the simulated panels. The standard error is then formed by taking the standard deviation of the estimates from each of the pseudo random samples.

To assess the finite sample distribution of the estimators, I ran two Monte Carlo experiments. The results from both experiments indicate that the estimators recover the actual parameter values fairly well. Details and results of these experiments are provided in Appendix B.

3.4 Data for the Structural Model

For my structural estimation, I focus on two-tract markets.[29] As before, in order for a tract to be included in my data set, it must be a wildcat tract for which location data is available, for which the first exploration occurred neither before the sale date nor after the lease term, for which production did not occur before exploration, and which was sold prior to 1980. In order for two tracts to qualify as a market, the two tracts must be within 6 miles north and south of each other or 6 miles east and west of each other,[30] both the sale dates and the lease terms must be the same for both tracts, the tracts must be owned by different firms, and no other wildcat tracts from the same sale date can be within 6 miles of either tract. There are 87 such markets in my data set. Figure 4 maps out the tracts used.

Each discrete period t represents a year. The time of exploration is the year of the tract's first spud date. The time of development is the year of the tract's first production date.[31] It is possible that development begins in the same year as exploration does.

The panel spans the years 1954 to 1990. A market enters the panel when its tracts are sold. If both tracts are eventually developed, the market exits the panel when the second tract to develop first begins development. If neither tract has explored by the end of the five-year lease term, the market exits the panel when the leases expire.

[29] I restrict the size of the market to two tracts for two reasons. First, limiting the market size to two minimizes the state space. This is because the number of possible combinations of state variables, which is the product of the cardinality of the supports of each of the state variables, is quadratic in the market size. Second, when there are only two tracts in the market, each tract is equidistant to all its neighbors. Since the federal tracts in the Gulf of Mexico form a grid, the next sensible size of a market is four. With four tracts, however, diagonal tracts are not as close together as tracts that share a side are; as a consequence, the distance between each pair of neighbors in the market is not the same. It is plausible that firms may weight the behavior of their neighbors by their distance; when neighbors are no longer symmetric, the econometric model becomes more complicated.

[30] I choose 6 miles because the maximum tract size is 5760 acres, or 3 miles by 3 miles (Marshall Rose, Minerals Management Service, personal communication, 17 April 2003; Larry Slaski, Minerals Management Service, personal communication, 25 April 2003). I convert latitude and longitude to miles using the following following factors from the Lousiana Sea Grant web site (http://lamer.lsu.edu/classroom/deadzone/changedistance.htm): 1 minute longitude in Louisiana offshore = 60.5 miles; 1 minute latitude = 69.1 miles.

[31] More specifically, for both the time of exploration and the time of development, I take the floor of the number of years since the lease began.

There are two endogenous state variables N_{kt}: the total number of tracts in the market that have been explored and the total number of tracts in the market that have been developed. Because there are two tracts in each market, there are three possible values for the the total number of tracts that have been explored: 0, 1 and 2. As for the total number of tracts that have been developed, because a market ends once both tracts have been developed, there are only two possible values: 0 and 1.

The coefficients γ_N in the development profit function on the endogenous state variables indicate whether and how one firm's profits depend on the production decisions of its neighbors. Positive values of the coefficients γ_N would indicate that the information and extraction externalities were positive on net, and therefore that the information externality was dominant. Negative values would indicate that the externalities were negative on net, and therefore that the extraction externality was dominant. Moreover, one would expect each coefficient γ_N to be less than the cost of development. Otherwise, if the coefficients were greater than or equal to the development cost, this would mean that having a neighbor explore or develop would offset the cost of development, making development essentially costless. The effects of strategic interaction are unlikely to be that large.

I use three exogenous state variables X_{kt}; these variables were chosen based on considerations of state space and data availability. The first exogenous state variable is the discretized average winning bid per acre over the two tracts in market k at time t. Because the winning bid is a measure of the value of the tract, the average winning bid over the tracts in the market captures any fixed market-specific variables such as geological structures that may affect profits. To construct this variable I average the winning bids per acre over the two tracts in the market, and then discretize the average into three bins: 0 = low (0 to 1 thousand 1982 \$/acre), 1 = medium (1 thousand to 5 thousand 1982 \$/acre), and 2 = high (over 5 thousand 1982 \$/acre). One expects that profits would increase in the value of the tract, and therefore that the coefficient on the winning bid in the development profit function is positive.

The second exogenous state variable is the discretized real drilling cost at time t. I use data on annual drilling costs from the American Petroleum Institute's *Joint Association Survey of the U.S. Oil & Gas Producing Industry* for the 1969-1975 data and its *Joint Association Survey on Drilling Costs* for the 1976-1990 data. The cost is average cost per well over all offshore wells (oil wells, gas wells, dry holes), in nominal dollars. I convert the nominal costs to real costs in 1982-1984 dollars using the consumer price index (CPI). I discretize the real drilling cost into two bins: 0 = low (0 to 2.5 million 1982-1984 \$/well), 1 = high (over 2.5 million 1982-1984 \$/well). The drilling cost is a measure of both exploration costs and development costs, and therefore enters into both the extraction profit function and the development profit function. In particular, from (10), the exploration cost is assumed to be the following function of the discretized drilling cost $drill\_cost_t$:

$$c^{e}(\Omega_{kt};\theta) = -\alpha \cdot (drill\_cost_{t} + 1), \tag{23}$$

where α is now a scalar, so that exploration profits are:

$$\pi^{e}(\Omega_{kt},\mu_{it};\theta) = E_{\varepsilon}\left[V^{e}(\Omega_{kt},\varepsilon_{it};\theta)|\Omega_{kt}\right] + \alpha \cdot (drill\_cost_{t} + 1) + \mu_{it}. \tag{24}$$

The discretized drilling cost is incremented by one so that costs are non-zero even when they fall in the low bin. Figure 5 plots the real drilling cost data, along with the bins.[32] The expected sign of the coefficient α on costs in the exploration profit function is negative; higher costs should lower exploration profits. Similarly, the expected sign of the coefficient on costs in the development profit function is negative as well.

The third exogenous state variable is the discretized real oil price. I use the U.S. average crude oil domestic first purchase price from the EIA Annual Energy Review and deflate the time series to 1982-1984 dollars per barrel using the CPI. I discretize the real oil price into three bins: 0 = low (0 to 13 1982-1984 $/barrel), 1 = medium (13 to 25 1982-1984 $/barrel), and 2 = high (over 25 1982-1984 $/barrel). Figure 6 plots the real oil price, along with the bins. The expected sign on oil price in the development profit function is positive: higher oil prices should increase revenues.

Table 6 presents summary statistics for the panel data used for the structural estimation. There are 1646 observations spanning 174 tracts and 87 markets. The markets range in duration from 2 years to 36 years, with an average length of 17.92 years (s.d. = 11.17). Of the 174 tracts, 122 were eventually explored and 66 were eventually developed. The average number of years to exploration, conditional on exploring, is 1.21 (s.d. = 1.42). The average number of years to development, conditional on developing, is 5.79 (s.d. = 3.51). For the 66 tracts that developed, the predicted ex post revenues, as calculated by Hendricks, Porter and Boudreau (1987), range from $22,000 to $298 million, with an average of $49.34 million (s.d. = 65.53 million). The real gross profits from development, which are the predicted ex post revenues times the government royalty rate minus costs, but not net of the bid, also as calculated by Hendricks, Porter and Boudreau (1987), range from -$38.10 million to $18.80 million, with an average of -$10.83 million (s.d. = 9.57 million). Table 6 also provides the summary statistics for the state variables in the panel. The number of possible combinations of state variables is the product of the cardinality of the supports of each of the state variables, or $3 \times 2 \times 3 \times 2 \times 3 = 108$.

The distribution of tract size is presented in Figure 7. Most tracts are either 2500 acres, 5000 acres

[32] Before 1969, the real drilling cost is assumed to fall into the low bin. The 1982 real drilling cost, which was unavailable because the 1982 issue of the *Joint Association Survey on Drilling Costs* was out of print, is assumed to fall in the high bin.

or 5760 acres in size. The mean tract size is 4460 acres (s.d. = 1300), and the median tract size is 5000 acres.

3.5 Results from the Structural Model

There are three different types of parameters to be estimated: the parameters $(\sigma_\mu, \sigma_\varepsilon)$ governing the distribution of the private information, the coefficient α on drilling costs in the exploration profit function, and the coefficients γ on the state variables in the development profit function.

The results from running the structural model on all tracts in the panel regardless of tract size are shown in Table 7. The coefficients $\gamma_N \equiv (\gamma_{tote}, \gamma_{totd})$ on the two endogenous state variables can be interpreted as follows. Since the total number of tracts in the market that have been explored increases by one if a neighbor explores, the coefficient γ_{tote} on this variable measures how the profits from development change when a neighbor explores. Similarly, since the total number of tracts in the market that have been developed can only take values of 0 and 1, with 0 indicating that the neighbor has not developed and 1 indicating that the neighbor has developed, the coefficient γ_{totd} on this variable measures how the profits from development change when a neighbor develops first. An important result is that the coefficient on the total number of tracts in the market that have been explored is statistically insignificant, which means that firms do not interact strategically on net during exploration. This result is consistent with the results from the reduced-form models, which also do not indicate that strategic, non-cooperative behavior occurs during the first stage of petroleum production. In contrast, the coefficient on the total number of tracts in the market that have been developed is statistically significant and positive, which means that a firm's profits increase when its neighbor develops. This seems reasonable, because when a neighbor develops following exploration, this is a signal to the firm that the neighbor's exploratory efforts were successful, and therefore that there may be deposits present.

To assess the economic significance of the strategic interactions, I compare the coefficients γ_N on the endogenous state variables with the development cost $|\gamma_{drill} \cdot drill\_cost_t|$, where γ_{drill} denotes the coefficient on the discretized real drilling cost $drill\_cost_t$ in the development profit function. As noted above, one would expect each coefficient γ_N to be less than the cost of development. Since the maximum development cost is given by $|\gamma_{drill} \cdot \max(drill\_cost_t)| = |\gamma_{drill}|$, one would therefore expect $\gamma_{tote} < |\gamma_{drill}|$ and $\gamma_{totd} < |\gamma_{drill}|$. The results are consistent with the expectations. Further comparison of the coefficients γ_N with the mean development cost $\left|\gamma_{drill} \cdot \overline{drill\_cost_t}\right|$ can give a measure of the economic importance of a neighbor's decisions to a firm's profits. In particular, the relative importance of a neighbor's exploration decision as a fraction of a firm's costs is given by $\frac{\gamma_{tote}}{\left|\gamma_{drill} \cdot \overline{drill\_cost_t}\right|} = \frac{\gamma_{tote}}{0.49|\gamma_{drill}|} = 0.06$. Similarly, the relative importance of a neighbor's development decision as a fraction of a firm's costs is given by $\frac{\gamma_{totd}}{\left|\gamma_{drill} \cdot \overline{drill\_cost_t}\right|} = \frac{\gamma_{totd}}{0.49|\gamma_{drill}|} = 0.19$. The small values of both of these numbers indicate that the effects of neighbors' decisions are second-order compared to costs.

In absolute terms, strategic interactions in exploration are not only statistically insignificant, but economically insignificant as well: changes in profits (in 1982 \$) from a neighbor's exploration less than \$80,000 and greater than \$640,000 can be rejected at a 5% level. These values are small relative to predicted ex post revenues, which average \$49.34 million. Strategic interactions in development are statistically significant but only moderately economically significant: changes in profits from a neighbor's development less than \$700,000 and greater than \$980,000 can be rejected at a 5% level.

As for the values of the parameters governing the distribution of private information, both the parameter σ_μ from the distribution of the pre-exploration shock μ_{it} and the parameter σ_ε from the distribution of the pre-development shock ε_{it} are statistically significant. This suggests that private information has a statistically significant impact on both the exploration decision and the development decision.

In terms of economic significance, one way to interpret the mean σ_μ of the pre-exploration shock μ_{it} is to compare it with exploration costs. Since both the pre-exploration shock μ_{it} and the exploration cost function $c^e(\Omega_{kt};\theta)$ enter linearly into the exploration profit function (11), the importance of private information in the exploration decision can be measured by comparing the mean σ_μ of the shock with the mean $\overline{c^e(\Omega_{kt};\theta)}$ of the costs. Expressed as a fraction of the mean exploration costs, where the mean exploration costs are computed by substituting the mean value of the discretized real drilling cost into the equation (23) for the exploration costs, the relative importance of private information is therefore given by: $\frac{\sigma_\mu}{\overline{c^e(\Omega_{kt};\theta)}} = \frac{\sigma_\mu}{-\alpha\cdot\left(\overline{drill\_cost_t}+1\right)} = \frac{\sigma_\mu}{-1.49\alpha}$. A large value of $\frac{\sigma_\mu}{-1.49\alpha}$ would indicate that private information plays a large role relative to costs in the first-stage exploration decision; a small value would indicate that costs are more relatively more important. In this case, the value is 0.33. Private information is about a third as important as costs. Thus, the role of private information in the exploration decision is both economically and statistically significant.

The mean σ_ε of the pre-development shock ε_{it} can be similarly compared with development costs to assess the importance of private information in the second-stage development decision. Since both the pre-development shock ε_{it} and the development cost $|\gamma_{drill} \cdot drill\_cost_t|$ enter linearly into the development profit function (3), the importance of private information in the development decision can be measured by comparing the mean σ_ε of the shock with the mean $\left|\gamma_{drill} \cdot \overline{drill\_cost_t}\right|$ of the costs. Expressed as a fraction of the mean development costs, the relative importance of private information is therefore given by: $\frac{\sigma_\varepsilon}{\left|\gamma_{drill} \cdot \overline{drill\_cost_t}\right|} = \frac{\sigma_\varepsilon}{0.49|\gamma_{drill}|}$. A large value of $\frac{\sigma_\varepsilon}{0.49|\gamma_{drill}|}$ would indicate that private information plays a large role relative to costs in the second-stage development decision; a small value would indicate that costs are more relatively more important. In this case, the value is 0.91. Private information is almost as important as costs in the development decision.

The coefficients on the other covariates are all significant and have the expected sign. As expected, the coefficient on the discretized drilling cost is negative in both the exploration profit equation and the development profit equation. Also as expected, the profits from development increase in both the average winning bid and in the real oil price.

Table 7 also reports estimates of the ex ante expected value of an unexplored tract (million 1982 \$). This is the expected value of an unexplored tract at time $t = 0$ when neither of the tracts in the market have been explored and when the discretized average winning bid per acre, the discretized real drilling cost and the discretized real oil price take on the given values at time $t = 0$. Of the values that were precisely estimated, the real ex ante expected value of an unexplored tract (in million 1982 \$) when both the average winning bid and the drilling cost are in their low bins while the oil price is in the medium bin is 5.39 (s.e. = 0.09); the value (in million 1982 \$) when the average winning bid is in its high bin, the drilling cost is in the low bin, and the oil price is in the medium bin is 111.73 (s.e. = 0.26).

There are several possible explanations why the results do not provide evidence for strategic, non-cooperative behavior during exploration. One is that the tract sizes are large enough that cross-tract externalities are insignificant or even nonexistent. A second is that cross-tract externalities exist, but firms owning neighboring tracts cooperate to jointly internalize the externalities they impose on each other, for example through joint ventures or unitization. A third is that cross-tract externalities are significant, but the positive information externality exactly cancels the negative extraction externality, resulting in zero net strategic effects.

To distinguish among these three explanations for the lack of strategic interactions during exploration, I estimate the strategic interactions by tract size. If externalities are insignificant when tracts are large enough, then one would expect to see strategic, non-cooperative behavior only on small tracts. This is because the smaller the tract size, the more likely the tracts are located over a common pool, and therefore the more acute the information and extraction externalities faced by the firms. Evidence for significant strategic interactions on small tracts but not on large tracts would thus be consistent with the first explanation.

If externalities exist even for the largest tracts in the sample, but are eliminated through coordination, then, assuming firms would coordinate regardless of tract size, one would not expect to see any strategic, non-cooperative behavior even on small tracts. Insignificant strategic interactions regardless of tract size would thus be consistent with the second explanation.

If strategic interactions do not occur on net because the positive information externality exactly cancels the negative extraction externality, then one may not expect the exact cancellation to still take place when the tract size is small. This is because the geographical span of the information externality is larger than that of the extraction externality: while the former only requires that tracts may share common geological features, the latter requires that tracts may be located over a common pool. As a consequence, it is possible for the information externality to be present on all the tracts in the sample, but for the extraction externality to be present on only the smaller tracts. Theory therefore suggests that the importance of the extraction externality relative to the information externality should be greater on small tracts than on large tracts. Thus, if the externalities cancel when all the tract sizes are considered, one might expect that the negative extraction externality would dominate the positive information externality when the sample is limited to small tracts only. Strategic interactions that are more significantly negative on small tracts than on large tracts would therefore be consistent with the third explanation.

Running the structural model on subsamples of the data set that differed in the acreage of the tracts in the panel would therefore enable one to distinguish among these three explanations for the lack of strategic interactions in the pooled sample. Assuming that the tract sizes differ for exogenous reasons, the results by acreage will also give a sense of whether or not the government can change the extent the which firms behave strategically by changing the size of the tracts.

Table 8a presents the results from running the structural model on three subsamples consisting of larger tracts, defined as tracts that are greater than or equal to 5000 acres, 4000 acres, and 3000 acres in size, respectively. As the coefficients on the endogenous variables indicate, strategic interactions are neither economically nor statistically significant for any of these subsamples. For tracts greater than or equal to 5000 acres in size, the 95% confidence interval for the effect of a neighbor's exploration on profits (in 1982 $) ranges from -$0.18 million to $0.10 million, and the 95% confidence interval for the effect of a neighbor's development on profits ranges from -$0.07 million to $0.05 million. These values are small relative to predicted ex post revenues, which average $49.34 million in the pooled sample. Investment decisions and profits are instead driven primarily by private information, the average winning bid, the real drilling cost, and the real oil price.

Table 8b presents the results from running the structural model on three subsamples consisting of smaller tracts, defined as tracts that are less than 5000 acres, 4000 acres, and 3000 acres in size, respectively. Strategic interactions are statistically and economically significant in all three subsamples. The coefficients in the three specifications on the number of tracts in the market that have been explored indicate that real development profits decrease by a statistically significant $10.73 million to a statistically significant $15.08 million when a neighbor explores; a neighbor's exploration is roughly as important to profits as maximum development costs. The negative extraction externality thus appears to dominate: when a neighbor explores, a firm's profits decrease because the neighbor has begun production and is likely to eventually compete with the firm for the same common pool. The coefficients in the three specifications on the number of tracts in the market that have been developed indicate that real development profits increase by a statistically significant \$1.09 million to a statistically significant \$3.29 million when a neighbor develops. In this case, the positive information externality dominates: a firm benefits when its neighbor develops after it explores because this is a signal to the firm that the neighbor's exploratory efforts were successful, and therefore that there may be deposits present. The magnitude of the positive net strategic effect that results from a neighbor's development is one order of magnitude smaller than the negative net strategic effect that results from a neighbor's exploration. As expected, the magnitude of the strategic interactions in both exploration and development increase monotonically as the tract sizes get smaller.

For the smaller tract sizes, investment decisions and profits are also driven by private information, the average winning bid, the real drilling cost, and the real oil price, as before, with two differences. First, pre-development private information plays a larger role in decision-making and profits for the smaller tracts than it does for the larger tracts. Second, the coefficient on the average winning bid is now negative, instead of positive, and is of smaller magnitude. On the smaller tracts, investment decisions and profits are governed more by post-auction private and public information, including that from the actions of one's neighbors, than by the initial estimate of tract value.

The results therefore indicate that when a neighbor explores, there are no strategic effects on larger tracts even though information and extraction externalities exist because these externalities cancel each other out. The cancelling no longer occurs on smaller tracts, however, where the negative extraction externality dominates. When a neighbor develops, there are no strategic effects on larger tracts because the tract size is large enough to enable a firm to internalize the externalities on its own. On small tracts, however, externalities are acute and the positive information externality dominates. Small tracts make up a small percentage of the tracts sold: of the tracts used in the structural estimation, only 37%, 27%, and 24% of tracts were less than 5000 acres, 4000 acres and 3000 acres in size, respectively. Thus, for the majority of tracts, externalities do not cause inefficient strategic interactions on net. The results suggest that, by making most of the tracts at least 5000 acres in size, the federal government has minimized the net effects of any externalities that may be present, and has thus eliminated any potential inefficiencies in petroleum production.

To assess whether the federal government can increase the ex ante tract value and therefore government revenue by changing the length of the lease term, Table 9 presents the results of simulations of different lease terms ranging from 2 years to 15 years, and compares these simulation results with the actual base case results (highlighted by the dotted lines) from the actual lease term of 5 years. For each lease term, I run 100 simulations of 87 markets keeping the parameters fixed at their estimated base case values.[33] To generate the simulated data, the Markov perfect equilibrium is first computed using dynamic programming. The exploration and development policy functions arising from the equilibrium are then used in conjunction with the empirical transition matrix for the exogenous variables and with the actual empirical initial conditions to simulate sample paths for each of 87 markets to form a simulated panel data set. The structural econometric model is then run on each of the 100 simulated panels to obtain the distribution of the estimators.

As seen from Table 9, the real ex ante tract value of an unexplored tract is not significantly greater under any of the alternate lease terms than it is under the actual lease term of 5 years for any of the tuples of average winning bid, drilling cost and oil price. It thus appears that the federal government can not do significantly better by modifying the lease term.

4 Conclusion

When individual petroleum-producing firms make their exploration and development investment timing decisions, information externalities and extraction externalities may lead them to interact strategically with their neighbors. A positive information externality arises if tracts are located over a common pool or share common geological features so that their ex post values are correlated, since firms learn information about their own tracts when other firms drill exploratory wells or install production platforms on neighboring tracts. A negative extraction externality arises when tracts are located over a common pool, since firms are competing for the same stock of petroleum. Owing to both information and extraction externalities, the dynamic decision-making problem faced by a petroleum-producing firm is not merely a single-agent problem, but rather can be viewed as a multi-agent, non-cooperative game in which firms behave strategically and base their exploration and development policies on those of their neighbors.

This paper examines whether strategic considerations arising from information and extraction externalities are present. In particular, it analyzes whether a firm's investment timing decisions and profits depend on the decisions of firms owning neighboring tracts of land. Both reduced-form and structural econometric approaches are employed. The reduced-form approach is a discrete response model of a firm's exploration timing decision, with instruments for the decisions of its neighbors. The structural approach is a structural econometric model of the firms' multi-stage investment timing game.

The research presented in this paper is important for several reasons. First, an empirical analysis of investment timing decisions enables one to examine whether the strategic interactions that are predicted in theory actually occur in practice. Second, the estimation of strategic interactions, especially those that

[33] I chose to simulate 87 markets because there are 87 markets in my actual data set.

arise in dynamic decision-making, is of methodological interest. In particular, my structural econometric model enables the investigation of results that do not appear in the reduced-form analysis. This structural econometric methodology can be used to analyze externalities in a variety of contexts, including spillovers that arise during research and development. Third, my results have implications for leasing policy: if the strategic effects and externalities turn out to be large, then the program by which the U.S. government leases tracts to firms may be inefficient, and possible modifications should be considered.

Do the positive information externalities and negative extraction externalities have any net strategic effect that may cause petroleum production to be inefficient? The answer depends on tract size. When the tract sizes are large, firms do not impose externalities on each other on net when choosing to explore or develop, and, as a consequence, strategic considerations are second-order. This is the case with most of the tracts in the federal leasing program. However, in the few cases where the tract size is small, externalities do matter, and they cause firms to interact strategically with their neighbors. As expected, these externalities intensify as the tract size decreases. Also as predicted by theory, the relative importance of the extraction externality from exploration with respect to the information externality is greater on small tracts than on large tracts. The results suggest that, in making most tracts at least 5000 acres in size, the federal government has minimized inefficiencies in petroleum production that may have resulted from non-cooperative strategic interactions.

References

[1] Aguirregabiria, V., & Mira, P. (2004). Sequential estimation of dynamic discrete games. Working paper. Boston University and CEMFI.

[2] Bajari, P., Benkard, C.L., & Levin, J. (2004). Estimating dynamic models of imperfect competition. Working paper. Duke University and Stanford University.

[3] Casella, G., & Berger, R. (1990). Statistical inference. Belmont, CA: Duxbury Press.

[4] Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. Journal of Econometrics, 34, 305-334.

[5] Dixit, A., & Pindyck, R. (1994). Investment under uncertainty. Princeton, NJ: Princeton University Press.

[6] Fudenberg, D., & Tirole, J. (1998). Game theory. Cambridge: MIT Press.

[7] Graham, B. (2005). Finite sample properties of QML, GMM and GEL estimation and inference procedures: Social interaction models. Working paper. Harvard University.

[8] Hansen, L. (1982). Large sample properties of generalized method of moments estimators. Econometrica, 50 (4), 1029-1054.

[9] Harsanyi, J. (1973). Games with randomly disturbed payoffs. International Journal of Game Theory, 2, 1-23.

[10] Hendricks, K., & Kovenock, D. (1989). Asymmetric information, information externalities, and efficiency: The case of oil exploration. RAND Journal of Economics, 20 (2), 164-182.

[11] Hendricks, K., & Porter, R. (1992). Joint bidding in federal OCS auctions. American Economic Review Papers and Proceedings, 82 (2), 506-511.

[12] Hendricks, K., & Porter, R. (1993). Determinants of the timing and incidence of exploratory drilling on offshore wildcat tracts. NBER Working Paper Series (Working Paper No. 4605). Cambridge, MA.

[13] Hendricks, K., & Porter, R. (1996). The timing and incidence of exploratory drilling on offshore wildcat tracts. The American Economic Review, 86 (3), 388-407.

[14] Hendricks, K., Porter, R., & Boudreau, B. (1987). Information, returns, and bidding behavior in OCS auctions: 1954-1969. Journal of Industrial Economics, 35 (4), 517-542.

[15] Hurn, A., & Wright, R. (1994). Geology or economics? Testing models of irreversible investment using North Sea oil data. Economic Journal, 104 (423), 363-371.

[16] Libecap, G., & Smith, J. (1999a). Regulatory remedies to the common pool: The limits to oil field unitization. Working paper. University of Arizona and Southern Methodist University.

[17] Libecap, G., & Smith, J. (1999b). The self-enforcing provisions of oil and gas unit operating agreements: Theory and evidence. The Journal of Law, Economics, and Organizations, 15 (2), 526-548.

[18] Libecap, G., & Wiggins, S. (1984). Contractual responses to the common pool: Prorationing of crude oil production. The American Economic Review, 74 (1), 87-98.

[19] Libecap, G., & Wiggins, S. (1985). The influence of private contractual failure on regulation: The case of oil field unitization. The Journal of Political Economy, 93 (4), 690-714.

[20] Manski, C. (1993). Identification of endogenous social effects: The reflection problem. Review of Economic Studies, 60 (3), 531-542.

[21] Manski, C. (1995). Identification problems in the social sciences. Cambridge, MA: Harvard University Press.

[22] Newey, W. (1987). Efficient estimation of limited dependent variable models with endogenous explanatory variables. Journal of Econometrics, 36, 231-250.

[23] Paddock, J., Siegel, D., & Smith, J. (1988). Option valuation of claims on real assets: The case of offshore petroleum leases. Quarterly Journal of Economics, 103 (3), 479-508.

[24] Pakes, A., Ostrovsky, M., & Berry, S. (2005). Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). Working paper. Harvard University.

[25] Pesaran, M.H. (1990). An econometric analysis of exploration and extraction of oil in the U.K. Continental Shelf. The Economic Journal, 100 (401), 367-390.

[26] Porter, R. (1995). The role of information in U.S. offshore oil and gas lease auctions. Econometrica, 63 (1), 1-27.

[27] Robalino, J., & Pfaff, A. (2004). Estimating spatial interactions in forest clearing. Working paper. Columbia University.

[28] Starr, A.W., & Ho, Y.C. (1969). Nonzero-sum differential games. Journal of Optimization Theory and Applications, 3, 184-206.

[29] Stock, J., & Watson, M. (2003). Introduction to econometrics. Boston: Addison-Wesley.

[30] Stokey, N., Lucas, R., & Prescott, E. (1989). Recursive methods in economic dynamics. Cambridge, MA: Harvard University Press.

5 Appendix A: Derivations

In this Appendix, I provide the details for deriving the equations for the continuation values and investment probabilities used in the structural model of the investment timing game.

5.1 Stage 2: Development

Equation (7) for the continuation value V^{ce} to waiting instead of developing is derived as follows. Substituting in equation (3) for development profits, the expected truncated profits from development conditional on development can be written as:

$$\begin{aligned} & E\left[\pi^d(\Omega_{kt},\varepsilon_{it};\theta)|\pi^d(\Omega_{kt},\varepsilon_{it};\theta) > \beta V^{ce}(\Omega_{kt};\theta)\right] \\ = \; & E\left[\pi_0^d(\Omega_{kt};\theta)+\varepsilon_{it}|\pi_0^d(\Omega_{kt};\theta)+\varepsilon_{it} > \beta V^{ce}(\Omega_{kt};\theta)\right] \\ = \; & \pi_0^d(\Omega_{kt};\theta) + E\left[\varepsilon_{it}|\varepsilon_{it} > \beta V^{ce}(\Omega_{kt};\theta) - \pi_0^d(\Omega_{kt};\theta)\right] \\ = \; & \beta V^{ce}(\Omega_{kt};\theta) + \sigma_\varepsilon, \end{aligned} \tag{25}$$

where the final step comes from the exponential distribution for ε_{it}.[34] Using equation (25) for the expected truncated profits conditional on development, the continuation value can thus be written as:

$$\begin{aligned} & V^{ce}(\Omega_{kt};\theta) \\ = \; & E\left[\begin{array}{c} g^d(\Omega_{k,t+1};\theta)\cdot E\left[\pi^d(\Omega_{k,t+1},\varepsilon_{i,t+1};\theta)|\pi^d(\Omega_{k,t+1},\varepsilon_{i,t+1};\theta) > \beta V^{ce}(\Omega_{k,t+1};\theta)\right] \\ +\left(1-g^d(\Omega_{k,t+1};\theta)\right)\cdot\beta V^{ce}(\Omega_{k,t+1};\theta) \end{array} |\Omega_{k,t}, I_{it}^d = 0\right] \\ = \; & E\left[g^d(\Omega_{k,t+1};\theta)\cdot(\beta V^{ce}(\Omega_{k,t+1};\theta)+\sigma_\varepsilon) + \left(1-g^d(\Omega_{k,t+1};\theta)\right)\cdot\beta V^{ce}(\Omega_{k,t+1};\theta)|\Omega_{kt}, I_{it}^d = 0\right] \\ = \; & E[\beta V^{ce}(\Omega_{k,t+1};\theta) + \sigma_\varepsilon g^d(\Omega_{k,t+1};\theta)|\Omega_{kt}, I_{it}^d = 0], \end{aligned}$$

which yields equation (7) as desired.

Equation (8) for the development probability $g^d(\Omega_{kt};\theta)$ is derived as follows. The probability $g^d(\Omega_{kt};\theta)$ of developing tract i at time t given that development of tract i has not occurred previously can be expressed as:

$$g^d(\Omega_{kt};\theta) \quad = \Pr\left(\pi^d(\Omega_{kt},\varepsilon_{it};\theta) > \beta V^{ce}(\Omega_{kt};\theta)\right),$$

where $\Pr(\cdot)$ denotes probability.

Substituting in equation (3) for development profits, we get:

$$\begin{aligned} g^d(\Omega_{kt};\theta) \; &= \; \Pr\left(\pi_0^d(\Omega_{kt};\theta)+\varepsilon_{it} > \beta V^{ce}(\Omega_{kt};\theta)\right) \\ &= \; \Pr\left(\varepsilon_{it} > \beta V^{ce}(\Omega_{kt};\theta) - \pi_0^d(\Omega_{kt};\theta)\right). \end{aligned}$$

[34] If ε ~ *exponential*(σ_ε), then its pdf is $f(\varepsilon) = \frac{1}{\sigma_\varepsilon}\exp\left(-\frac{\varepsilon}{\sigma_e}\right)$ (Casella & Berger, 1990).

Substituting in the exponential distributional assumption for ε_{it}, we get:

$$g^d(\Omega_{kt};\theta) \; = \exp\left(-\frac{\beta V^{ce}(\Omega_{kt};\theta) - \pi_0^d(\Omega_{kt};\theta)}{\sigma_\varepsilon}\right),$$

which yields equation (8) as desired.

5.2 Stage 1: Exploration

Equation (15) for the continuation value $V^{cn}(\cdot)$ to waiting instead of exploring is derived as follows.

$$\begin{aligned}
& V^{cn}(\Omega_{kt};\theta) \\
= \; & E\left[\begin{array}{c} g^e(\Omega_{k,t+1};\theta)\cdot E\left[\pi^e(\Omega_{k,t+1},\mu_{i,t+1};\theta)|\pi^e(\Omega_{k,t+1},\mu_{i,t+1};\theta) > \beta V^{cn}(\Omega_{k,t+1};\theta)\right] \\ +(1-g^e(\Omega_{k,t+1};\theta))\cdot\beta V^{cn}(\Omega_{k,t+1};\theta)\end{array} |\Omega_{kt}, I_{it}^e = 0\right] \\
= \; & E\left[\begin{array}{c} g^e(\Omega_{k,t+1};\theta)\cdot(\pi_0^e(\Omega_{kt};\theta) + E\left[\mu_{it}|\mu_{it} > \beta V^{cn}(\Omega_{k,t+1};\theta) - \pi_0^e(\Omega_{kt};\theta)\right]) \\ +(1-g^e(\Omega_{k,t+1};\theta))\cdot\beta V^{cn}(\Omega_{k,t+1};\theta)\end{array} |\Omega_{kt}, I_{it}^e = 0\right] \\
= \; & E\left[g^e(\Omega_{k,t+1};\theta)\cdot(\beta V^{cn}(\Omega_{k,t+1};\theta) + \sigma_\mu) + (1-g^e(\Omega_{k,t+1};\theta))\cdot\beta V^{cn}(\Omega_{k,t+1};\theta)|\Omega_{kt}, I_{it}^e = 0\right] \\
= \; & E\left[\beta V^{cn}(\Omega_{k,t+1};\theta) + \sigma_\mu g^e(\Omega_{k,t+1};\theta)|\Omega_{kt}, I_{it}^e = 0\right].
\end{aligned}$$

The derivation of equation (16) for the exploration probability $g^e(\Omega_{kt};\theta)$ proceeds as follows. The exploration policy function $g^e(\Omega_t^n;\theta)$, which is the probability of exploring tract i at time t given that exploration of tract i has not occurred previously, is given by:

$$g^e(\Omega_{kt};\theta) \; = \Pr\left(\pi^e(\Omega_{kt},\mu_{it};\theta) > \beta V^{cn}(\Omega_{kt};\theta)\right). \tag{26}$$

Substituting in equations (11) and (9) for exploration profits, we get:

$$
\begin{aligned}
& g^e(\Omega_t^n;\theta) \\
= \; & \Pr\left(\pi_0^e(\Omega_{kt};\theta) + \mu_{it} > \beta V^{cn}(\Omega_{kt};\theta)\right) \\
= \; & \Pr\left(\mu_{it} > \beta V^{cn}(\Omega_{kt};\theta) - \pi_0^e(\Omega_{kt};\theta)\right) \\
= \; & \exp\left(-\frac{\beta V^{cn}(\Omega_{kt};\theta) - \pi_0^e(\Omega_{kt};\theta)}{\sigma_u}\right) \\
= \; & \exp\left(-\frac{\beta V^{cn}(\Omega_{kt};\theta) - E_\varepsilon\left[V^e(\Omega_{kt},\varepsilon_{it};\theta)|\Omega_{kt}\right] + c^e(\Omega_{kt};\theta)}{\sigma_u}\right) \\
= \; & \exp\left(-\frac{\beta V^{cn}(\Omega_{kt};\theta) - \left(\beta V^{ce}(\Omega_{kt};\theta) + \sigma_\varepsilon g^d(\Omega_{kt};\theta)\right) + c^e(\Omega_{kt};\theta)}{\sigma_u}\right),
\end{aligned}
$$

which yields equation (16).

6 Appendix B: Monte Carlo Experiments

To assess the finite sample distribution of the estimators, Table 10 presents the results from two Monte Carlo experiments. In each experiment, I run 100 simulations of 87 markets each for a given set of fixed parameters.[35] To generate the simulated data, the Markov perfect equilibrium is first computed for the given set of parameters using dynamic programming. The exploration and development policy functions arising from the equilibrium are then used in conjunction with the empirical transition matrix for the exogenous variables and with random draws, with replacement, from the empirical distribution of initial conditions to simulate sample paths for each of 87 markets to form a simulated panel data set. Finally, the structural econometric model is run on each of the 100 simulated panels to obtain the finite sample distribution of the estimators.

The table reports the results for two different experiments. In experiment (1), the parameters are chosen so that neighbors have a large, negative effect and so that private information plays a small role in the exploration decision and a moderate role in the development decision.[36] In particular, the relative importance of a neighbor's exploration decision as a fraction of a firm's average development cost is given by $\frac{\gamma_{tote}}{\left|\gamma_{drill}\cdot \overline{drill\_cost_t}\right|} = \frac{\gamma_{tote}}{0.49|\gamma_{drill}|} = -1.36$. Similarly, the relative importance of a neighbor's development decision as a fraction of a firm's average development cost is given by $\frac{\gamma_{totd}}{\left|\gamma_{drill}\cdot \overline{drill\_cost_t}\right|} = \frac{\gamma_{totd}}{0.49|\gamma_{drill}|} = -1.59$. Thus, when a neighbor explores or develops, this decreases profits by more than mean development costs. In terms of private information, the mean pre-exploration shock is only a small fraction of the mean exploration

[35] I chose to simulate 87 markets because there are 87 markets in my actual data set.
[36] See text for an explanation of how to interpret the relative values of the parameters.

costs: $\frac{\sigma_\mu}{c^e(\Omega_{kt};\theta)} = \frac{\sigma_\mu}{-\alpha\cdot\left(\overline{drill\_cost_t}+1\right)} = \frac{\sigma_\mu}{-1.49\alpha} = 0.08$. The mean pre-development shock is only a moderate fraction of mean development costs : $\frac{\sigma_\varepsilon}{\left|\gamma_{drill}\cdot\overline{drill\_cost_t}\right|} = \frac{\sigma_\varepsilon}{0.49|\gamma_{drill}|} = 0.91$. According to results of the first experiment, the estimators appear to recover the actual parameter values fairly well.

In experiment (2), the parameters are chosen so that neighbors have a small, positive effect and so that private information plays a large role in both the exploration and development decisions. In particular, the relative importance of a neighbor's exploration decision is a small fraction of a firm's development costs: $\frac{\gamma_{tote}}{\left|\gamma_{drill}\cdot\overline{drill\_cost_t}\right|} = \frac{\gamma_{tote}}{0.49|\gamma_{drill}|} = 0.51$. The relative importance of a neighbor's development decision is also a small fraction of a firm's costs: $\frac{\gamma_{totd}}{\left|\gamma_{drill}\cdot\overline{drill\_cost_t}\right|} = \frac{\gamma_{totd}}{0.49|\gamma_{drill}|} = 0.26$. In terms of private information, the mean pre-exploration shock is over two and a half times the mean exploration costs: $\frac{\sigma_\mu}{c^e(\Omega_{kt};\theta)} = \frac{\sigma_\mu}{-\alpha\cdot\left(\overline{drill\_cost_t}+1\right)} = \frac{\sigma_\mu}{-1.49\alpha} = 2.68$. The mean pre-development shock is nearly four times mean development costs : $\frac{\sigma_\varepsilon}{\left|\gamma_{drill}\cdot\overline{drill\_cost_t}\right|} = \frac{\sigma_\varepsilon}{0.49|\gamma_{drill}|} = 3.82$. According to results of the second experiment, the estimators appear to recover the actual parameter values fairly well. Because the variances σ_μ^2 and σ_ε^2 of the distribution of the stochastic shocks are larger than in the first experiment, the standard deviations are larger as well, as expected.

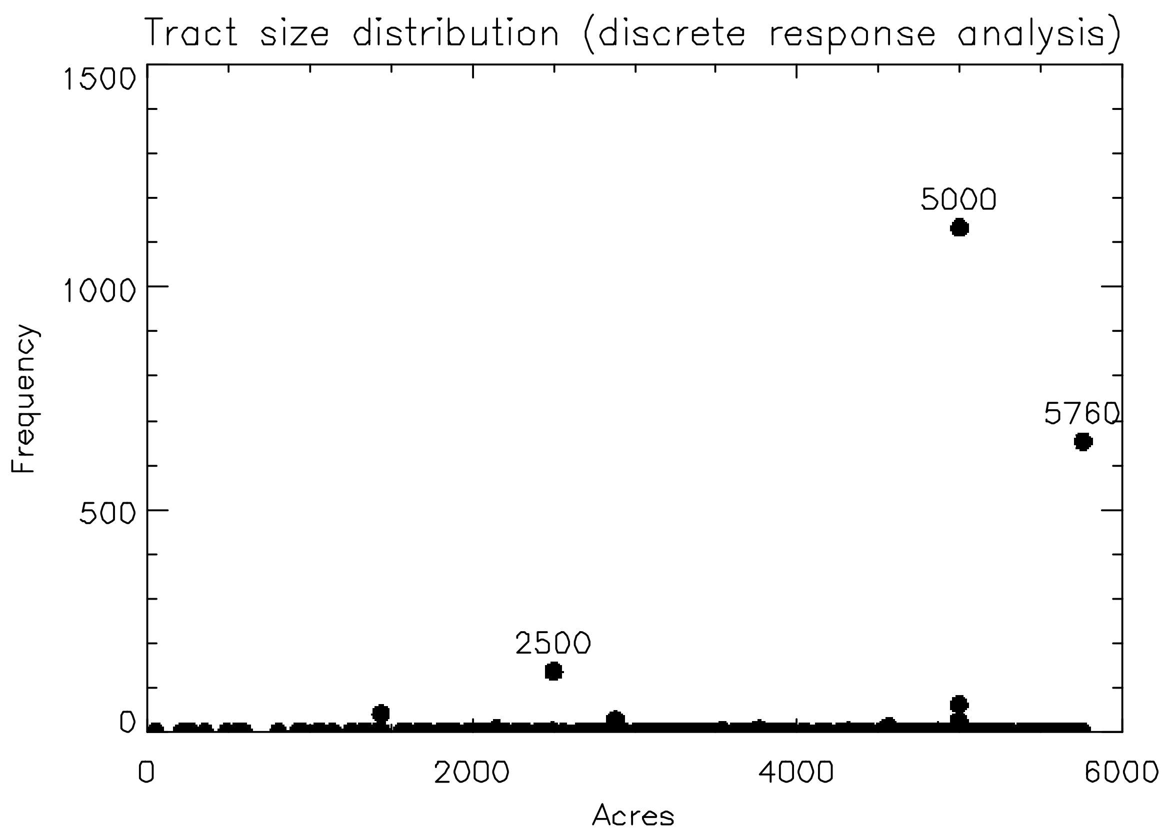
Thus, results from both Monte Carlo experiments indicate that the estimators recover the actual parameter values fairly well.

TABLE 1a. Discrete response analysis: Summary statistics by tract

Variable	# obs	mean	s.d.	min	max
pre-sale value (thousand 1982 $/acre)	2404	0.36	1.12	0.00	18.23
winning bid (thousand 1982 $/acre)	2404	2.52	4.92	0.45	60.80
acreage (1000 acres)	2404	4.79	1.10	0.05	5.76
number of time observations	2404	2.03	1.73	0	4
number of years to exploration, conditional on exploring	1721	1.24	1.42	0	4

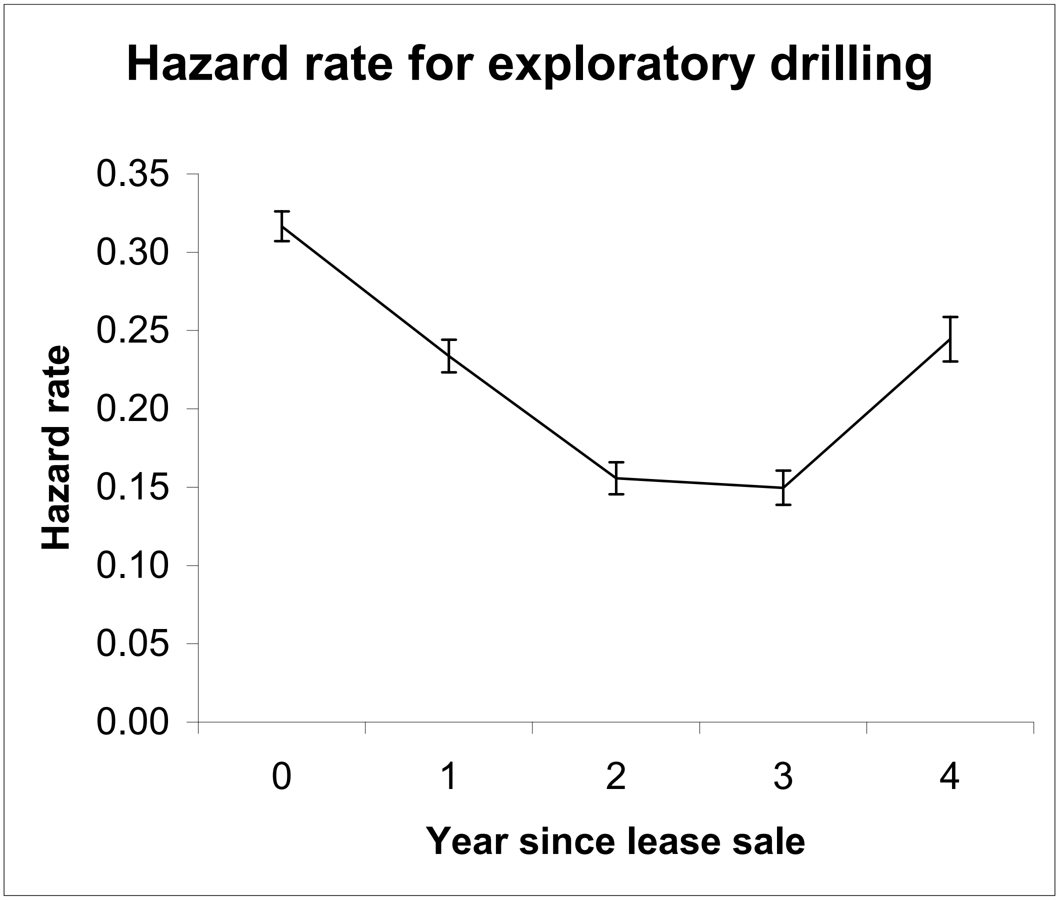
Note: The sample consists of wildcat tracts sold before 1980 whose date of first drilling did not occur after the lease term.

FIGURE 1.



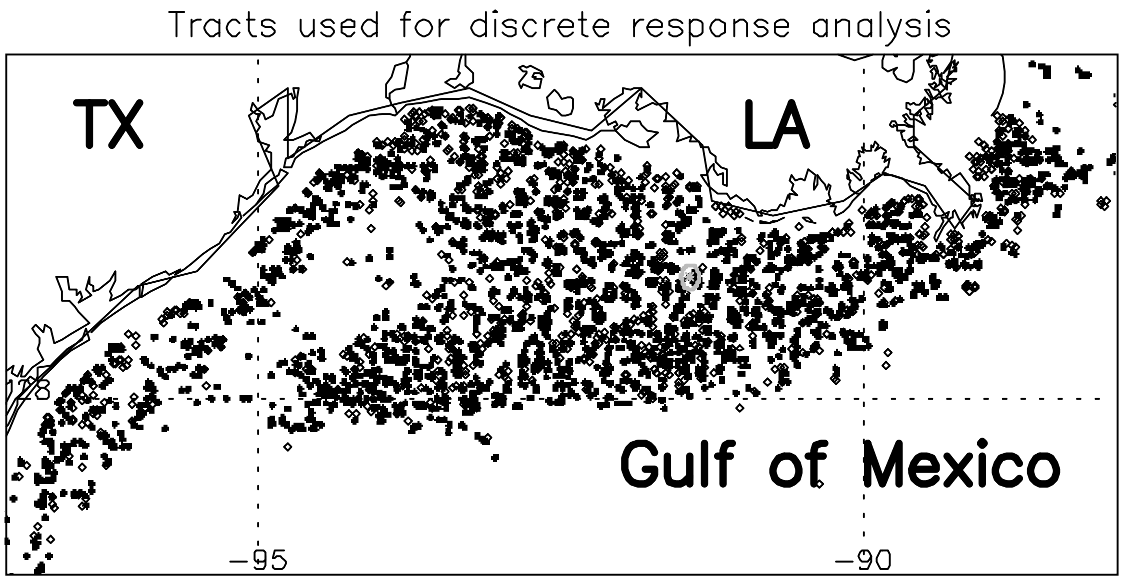
Note: The sample consists of wildcat tracts sold before 1980 whose date of first drilling did not occur after the lease term.

FIGURE 2.



Notes: The sample consists of wildcat tracts sold before 1980 whose date of first drilling did not occur after the lease term. Error bars indicate plus or minus one standard deviation.

FIGURE 3.



Notes: Filled circles denote the tracts used in the sample: these are wildcat tracts sold before 1980 whose date of first drilling did not occur after the lease term. Open diamonds denote other tracts that may be included as potential neighbors of the tracts in the sample. The grey asterisk highlights one of the wildcat tracts in the sample and the circle around it encompasses all tracts located within 5 miles of it.

TABLE 1b. Discrete response analysis: Summary statistics by tract-year

	Summary statistics					
Length of each period *t*	year	year	year	year	year	quarter
Definition of neighbors:						
distance of furthest neighbor	5	5	4	6	10	5
neighbors have different sale date	Y	N	Y	Y	Y	Y
	(1)	(2)	(3)	(4)	(5)	(6)
dummy for exploring at time *t*	0.21	0.23	0.20	0.22	0.26	0.01
	(0.41)	(0.42)	(0.40)	(0.42)	(0.44)	(0.12)
dummy for being in the last year of lease term at time *t*	0.16	0.13	0.17	0.16	0.13	0.12
	(0.37)	(0.34)	(0.37)	(0.36)	(0.34)	(0.33)
# neighbors at time *t*	1.74	2.41	1.50	2.03	3.58	2.67
	(1.00)	(1.61)	(0.76)	(1.33)	(2.90)	(1.97)
maximum # neighbors at time *t*	7	9	6	9	23	15
fraction of neighbors who explored at *t*-1	0.43	0.36	0.42	0.43	0.46	0.06
	(0.44)	(0.39)	(0.45)	(0.42)	(0.38)	(0.19)
fraction of neighbors in the first year of lease at *t*-1	0.36	0.42	0.37	0.35	0.34	0.04
	(0.45)	(0.47)	(0.46)	(0.44)	(0.42)	(0.17)
fraction of neighbors in the last year of lease term at *t*-1	0.15	0.13	0.13	0.14	0.12	0.04
	(0.33)	(0.32)	(0.32)	(0.32)	(0.28)	(0.16)
# observations	1139	3932	951	1345	1854	6013

Notes: The table reports the mean for the given variables. Standard errors are in parentheses. The maximum number of neighbors at time *t* is simply the maximum. The sample of tracts consists of wildcat tracts sold before 1980 whose date of first drilling did not occur after the lease term. A tract-year is included for a given definition of neighbors if the tract has at least one of the designated type of neighbor at time *t*.

TABLE 2. Discrete response analysis: Results without instruments

Dependent variable is probability of exploring at time *t*				
Probability model	linear	probit	linear	linear
Length of each period *t*	year	year	year	quarter
Definition of neighbors:				
distance of furthest neighbor (miles)	5	5	5	5
neighbors have different sale date	Y	Y	N	Y
	(1)	(2)	(3)	(4)
fraction of neighbors who explored at *t*-1	0.03	0.02	0.01	-0.01
	(0.03)	(0.02)	(0.02)	(0.01)
pre-sale value (million 1982 $/acre)	67.13 ***	34.64 **	59.19 ***	7.92 **
	(16.00)	(12.12)	(12.33)	(2.63)
winning bid (million 1982 $/acre)	34.64 ***	27.97 ***	42.52 ***	4.63 ***
	(3.57)	(3.95)	(2.70)	(0.52)
# years since lease sale	-0.03 *	0.00	-0.02 *	0.00
	(0.01)	(0.01)	(0.01)	(0.00)
dummy for being in last year of lease term	0.14 **	0.15 **	0.16 ***	0.00
	(0.05)	(0.06)	(0.03)	(0.01)
acreage (1000 acres)	0.04 ***	0.03 **	0.05 ***	0.008 ***
	(0.01)	(0.01)	(0.01)	(0.002)
year lease was sold	0.01	0.03 ***	0.00	-0.00
	(0.01)	(0.00)	(0.00)	(0.00)
firm fixed effects	Y	Y	Y	Y
area fixed effects	Y	Y	Y	Y
year effects	Y	Y	Y	Y
p-value (Pr > F for linear; Pr > χ^2 for probit)	0.000 ***	0.000 ***	0.000 ***	0.000 ***
adjusted R^2 (linear model) or pseudo R^2 (probit model)	0.22	0.31	0.16	0.04
# observations	1139	1139	3932	6013
Test of endogeneity of fraction of neighbors who explored at time t-1				
p-value	0.452	0.292	0.729	0.567

Notes: Standard errors are in parentheses. The probit results are the change in probability for a unit change in each dependent variable; continuous variables are evaluated at their mean. To test for the endogeneity of fraction of neighbors who explored at time *t*-1, a Durbin-Wu-Hausman test is used for the linear probability model and a Smith-Blundell test is used for the probit model. Both are tests of whether the residual from a regression of the variable in question on all the exogenous variables has a significant coefficient when added to the original model; the table reports the p-value from the tests under the null hypothesis that the variable is exogenous. Significance codes: * 5% level, ** 1% level, and *** 0.1% level.

TABLE 3. Discrete response analysis: First-stage regressions

Dependent variable is fraction of neighbors who explored at time *t*-1				
Length of each period *t*	year	year	quarter	year
Definition of neighbors:				
distance of furthest neighbor (miles)	5	5	5	10
neighbors have different sale date	Y	N	Y	Y
	(1)	(2)	(3)	(4)
fraction of neighbors in the first year of lease at *t*-1	-0.20 ***	-0.11 ***	-0.11 ***	-0.17 ***
	(0.03)	(0.02)	(0.01)	(0.02)
fraction of neighbors in the last year of lease term at *t*-1	-0.09 *	-0.03	-0.03 *	-0.14 ***
	(0.04)	(0.03)	(0.01)	(0.03)
pre-sale value (million 1982 $/acre)	2.15	13.64	-7.95 *	0.56
	(17.51)	(11.56)	(3.94)	(8.95)
winning bid (million 1982 $/acre)	4.46	3.31	2.84 ***	0.09
	(3.90)	(2.53)	(0.78)	(2.01)
# years since lease sale	-0.034 *	-0.05 ***	-0.00	-0.00
	(0.015)	(0.01)	(0.00)	(0.01)
dummy for being in last year of lease term	0.08	0.12 ***	-0.02	-0.01
	(0.05)	(0.03)	(0.01)	(0.04)
acreage (1000 acres)	0.00	0.02 ***	-0.00	-0.00
	(0.01)	(0.01)	(0.00)	(0.01)
year lease was sold	0.013 *	0.01 **	0.00	0.01 **
	(0.006)	(0.00)	(0.00)	(0.00)
firm fixed effects	Y	Y	Y	Y
area fixed effects	Y	Y	Y	Y
year effects	Y	Y	Y	Y
p-value (Pr > F)	0.000 ***	0.000 ***	0.000 ***	0.000 ***
adjusted R^2	0.20	0.15	0.14	0.34
# observations	1139	3932	6013	1854
Results from joint test of instruments				
F-statistic	18.05	16.37	13.03	28.70
p-value (Pr > F)	0.000 ***	0.000 ***	0.000 ***	0.000 ***

Notes: Standard errors are in parentheses. Significance codes: * 5% level, ** 1% level, and *** 0.1% level.

TABLE 4a. Discrete response analysis: Results with instruments

Dependent variable is probability of exploring at time *t*				
Probability model	linear	probit	linear	linear
Length of each period *t*	year	year	year	quarter
Definition of neighbors:				
distance of furthest neighbor (miles)	5	5	5	5
neighbors have different sale date	Y	Y	N	Y
	(1)	(2)	(3)	(4)
fraction of neighbors who explored at *t*-1	0.14	0.19	-0.06	0.04
	(0.15)	(0.16)	(0.19)	(0.09)
pre-sale value (million 1982 $/acre)	66.46 ***	52.90 **	60.11 ***	8.25 **
	(16.15)	(19.08)	(12.64)	(2.69)
winning bid (million 1982 $/acre)	34.15 ***	43.04 ***	42.72 ***	4.49 ***
	(3.65)	(4.92)	(2.77)	(0.58)
# years since lease sale	-0.03	0.01	-0.02 *	0.00
	(0.01)	(0.02)	(0.01)	(0.00)
dummy for being in last year of lease term	0.13 *	0.19 *	0.16 ***	0.01
	(0.05)	(0.09)	(0.04)	(0.01)
acreage (1000 acres)	0.04 ***	0.05 **	0.05 ***	0.008 ***
	(0.01)	(0.02)	(0.01)	(0.002)
year lease was sold	0.01	0.04 ***	0.01	-0.00
	(0.01)	(0.00)	(0.01)	(0.00)
firm fixed effects	Y	Y	Y	Y
area fixed effects	Y	Y	Y	Y
year effects	Y	Y	Y	Y
p-value (Pr > F for linear; Pr > χ^2 for probit)	0.000 ***	0.000 ***	0.000 ***	0.000 ***
adjusted R^2 (linear model) or pseudo R^2 (probit model)	0.21	0.26	0.15	0.03
# observations	1139	1037	3932	6013

Notes: Standard errors are in parentheses. The fraction of neighbors who explored at time *t*-1 is instrumented with the fraction of neighbors in the first year of their lease at *t*-1 and the fraction of neighbors in the last year of their lease term at *t*-1. The probit results are the change in probability for a unit change in each dependent variable; continuous variables are evaluated at their mean. Significance codes: * 5% level, ** 1% level, and *** 0.1% level.

TABLE 4b. Discrete response analysis: Results with instruments, varying distance of neighbors

Dependent variable is probability of exploring at time *t*				
Probability model	linear	linear	linear	linear
Length of each period *t*	year	year	year	year
Definition of neighbors:				
distance of furthest neighbor (miles)	5	4	6	10
neighbors have different sale date	Y	Y	Y	Y
	(1)	(5)	(6)	(7)
fraction of neighbors who explored at *t*-1	0.14	0.01	0.01	0.05
	(0.15)	(0.16)	(0.16)	(0.16)
pre-sale value (million 1982 $/acre)	66.46 ***	56.49 **	56.49 **	50.84 ***
	(16.15)	(18.12)	(18.12)	(11.04)
winning bid (million 1982 $/acre)	34.15 ***	36.56 ***	35.56 ***	28.94 ***
	(3.65)	(4.27)	(4.27)	(2.48)
# years since lease sale	-0.03	-0.03	-0.03	-0.04 **
	(0.01)	(0.02)	(0.02)	(0.01)
dummy for being in last year of lease term	0.13 *	0.12 *	0.12 *	0.14 **
	(0.05)	(0.05)	(0.05)	(0.05)
acreage (1000 acres)	0.04 ***	0.05 **	0.05 ***	0.04 ***
	(0.01)	(0.01)	(0.01)	(0.01)
year lease was sold	0.01	0.01	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.00)
firm fixed effects	Y	Y	Y	Y
area fixed effects	Y	Y	Y	Y
year effects	Y	Y	Y	Y
p-value (Pr > F)	0.000 ***	0.000 ***	0.000 ***	0.000 ***
adjusted R^2	0.21	0.21	0.21	0.23
# observations	1139	951	951	1854

Notes: Standard errors are in parentheses. The fraction of neighbors who explored at time *t*-1 is instrumented with the fraction of neighbors in the first year of their lease at *t*-1 and the fraction of neighbors in the last year of their lease term at *t*-1. Significance codes: * 5% level, ** 1% level, and *** 0.1% level.

TABLE 5. Discrete response analysis: Results with instruments, by acreage

Dependent variable is probability of exploring at time *t*			
Probability model	linear	linear	linear
Length of each period *t*	year	year	year
Definition of neighbors:			
distance of furthest neighbor (miles)	5	5	5
neighbors have different sale date	Y	Y	Y
Acreage	≥ 5000	< 5000	all
	(1)	(2)	(3)
fraction of neighbors who explored at *t*-1	0.03	0.19	-0.01
	(0.15)	(0.28)	(0.52)
fraction of neighbors who explored at *t*-1 * acreage (1000 acres)			0.04
			(0.11)
pre-sale value (million 1982 $/acre)	44.60 *	84.85 *	65.37 ***
	(18.69)	(39.53)	(16.55)
winning bid (million 1982 $/acre)	52.76 ***	21.34 ***	34.00 ***
	(5.89)	(6.02)	(3.68)
# years since lease sale	-0.02	-0.01	-0.03
	(0.02)	(0.03)	(0.02)
dummy for being in last year of lease term	0.08	0.31 *	0.12 *
	(0.06)	(0.12)	(0.05)
acreage (1000 acres)			0.03
			(0.06)
year lease was sold	0.01	-0.00	0.01
	(0.01)	(0.01)	(0.01)
firm fixed effects	Y	Y	Y
area fixed effects	Y	Y	Y
year effects	Y	Y	Y
p-value (Pr > F)	0.000 ***	0.000 ***	0.000 ***
adjusted R^2	0.25	0.18	0.20
# observations	889	250	1139

Notes: Standard errors are in parentheses. The fraction of neighbors who explored at time *t*-1 is instrumented with the fraction of neighbors in the first year of their lease at *t*-1 and the fraction of neighbors in the last year of their lease term at *t*-1. For specification (3), the interaction between the fraction of neighbors who explored at time *t*-1 and acreage is instrumented with interactions between the above instruments and acreage. Significance codes: * 5% level, ** 1% level, and *** 0.1% level.

FIGURE 4.

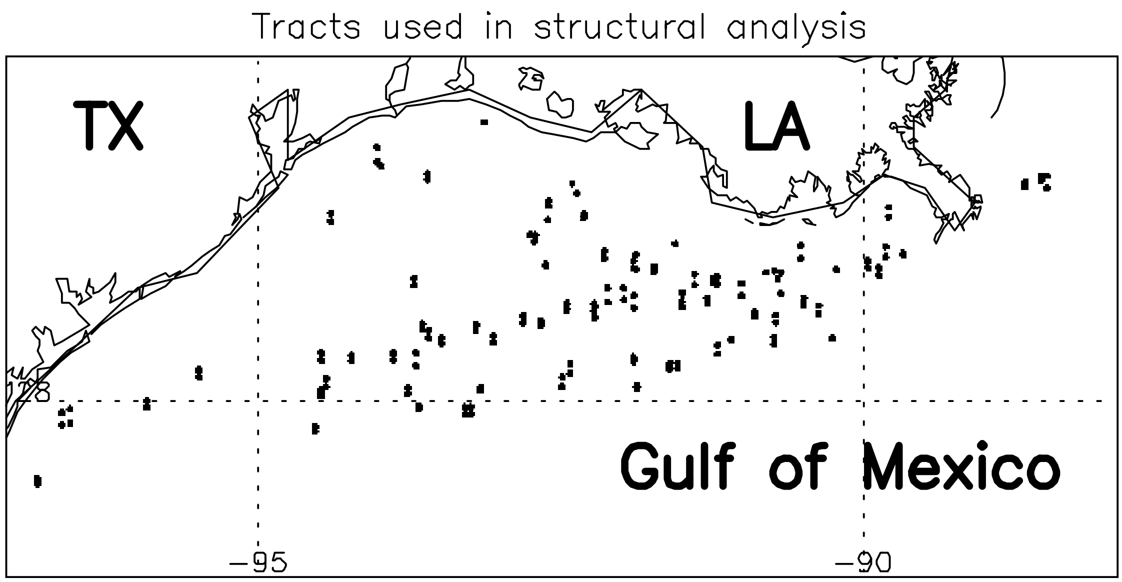

Tracts used in structural analysis
TX
LA
Gulf of Mexico
28
−95
−90

FIGURE 5.

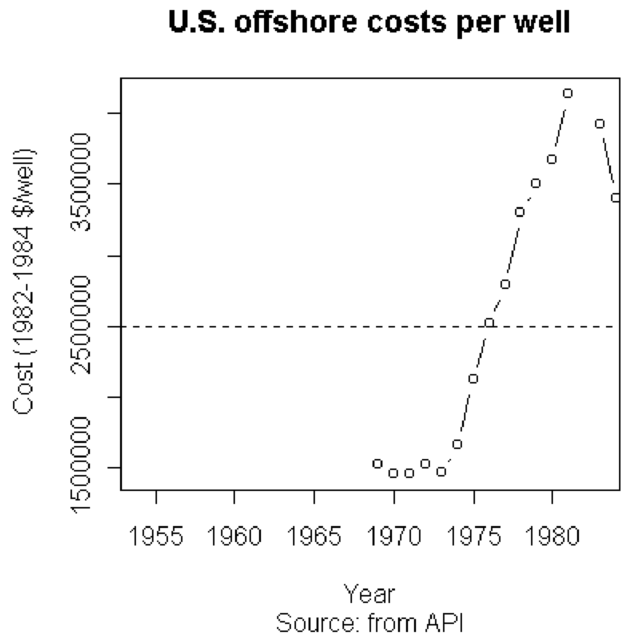


FIGURE 6.

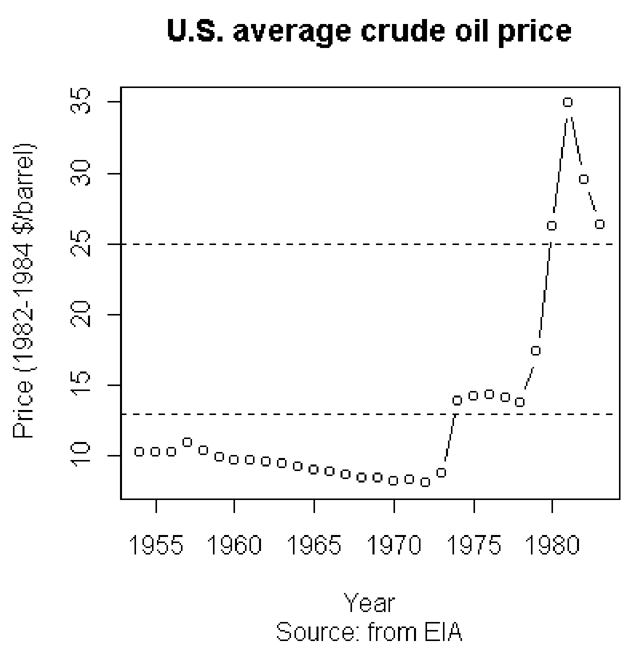


TABLE 6. Structural analysis: Summary statistics

	# obs	mean	s.d.	min	max
by tract					
acreage (1000 acres)	174	4.46	1.30	0.94	5.76
number of years to exploration, conditional on exploring	122	1.21	1.42	0	4
number of years to development, conditional on developing	66	5.79	3.51	0	15
revenue (million 1982 $), conditional on developing	66	49.34	65.53	0.02	298.0
gross profits (million 1982 $), conditional on developing	66	-10.83	9.57	-38.10	18.80
by market					
number of time observations	87	17.92	11.17	2	36
by market-year					
# tracts in market that have been explored	1646	1.31	0.68	0	2
# tracts in market that have been developed	1646	0.31	0.46	0	1
discretized average winning bid per acre	1646	0.70	0.69	0	2
discretized real drilling cost	1646	0.49	0.50	0	1
discretized real oil price	1646	0.65	0.72	0	2

FIGURE 7.

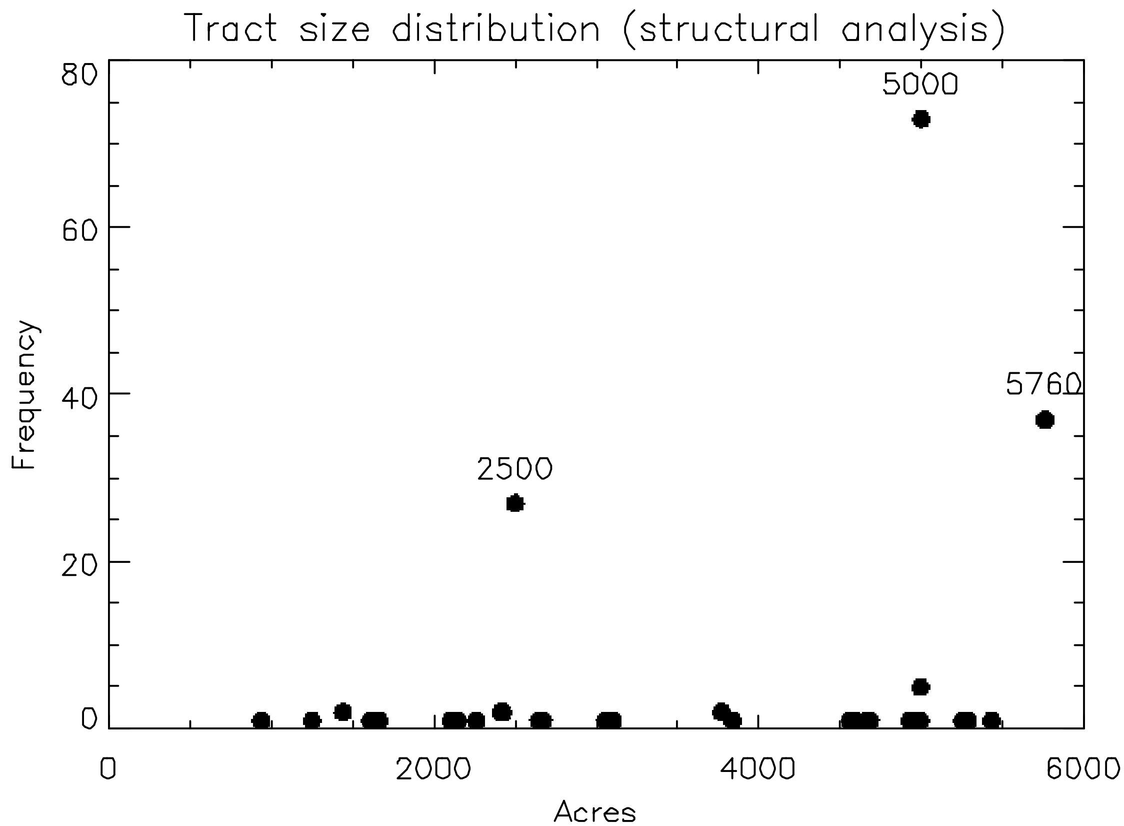

Tract size distribution (structural analysis)
80
60
40
20
0
Frequency
5000
5760
2500
0
2000
4000
6000
Acres

TABLE 7. Structural analysis: Pooled results

Parameter			Estimate	Standard Error
σ_{μ}			4.96	0.00
σ_{ε}			4.08	0.27
coefficient α in the exploration profit function on:				
discretized real drilling cost + 1			-10.00	0.00
coefficients γ in the development profit function on:				
# tracts in market that have been explored			0.28	0.18
# tracts in market that have been developed			0.84	0.07
discretized average winning bid per acre			5.56	0.09
discretized real drilling cost			-9.11	0.03
discretized real oil price			6.92	0.10
constant			5.00	0.09
Ex ante expected value of an unexplored tract (million 1982 $)				
winning bid bin	drilling cost bin	oil price bin		
0	0	0	8.07	1.23 E4
0	0	1	5.39	0.09
1	0	0	1.14 E5	1.84 E24
1	0	1	1.32 E5	4.29 E29
2	0	1	111.73	0.26

Notes: There are 1646 observations spanning 87 markets. The ex ante expected value of an unexplored tract (million 1982 $) is the expected value of an unexplored tract at time $t = 0$ when neither of the tracts in the market have been explored and when the discretized average winning bid per acre, the discretized real drilling cost and the discretized real oil price take on the given values at time $t = 0$. Standard errors are formed by bootstrapping 100 simulated panels of 87 markets each.

TABLE 8a. Structural analysis: Results for large tracts

			Acreage ≥ 5000	≥ 4000	≥ 3000
σ_μ			5.00	4.99	5.00
			(0.01)	(0.00)	(0.01)
σ_ε			4.96	4.96	4.96
			(0.05)	(0.04)	(0.04)
coefficient α in the exploration profit function on:					
discretized real drilling cost + 1			-10.00	-10.00	-10.00
			(0.00)	(0.00)	(0.00)
coefficients γ in the development profit function on:					
# tracts in market that have been explored			-0.04	-0.04	-0.04
			(0.07)	(0.07)	(0.05)
# tracts in market that have been developed			-0.01	-0.01	-0.01
			(0.03)	(0.02)	(0.03)
discretized average winning bid per acre			5.03	5.03	5.03
			(0.03)	(0.03)	(0.03)
discretized real drilling cost			-9.98	-9.98	-9.98
			(0.02)	(0.02)	(0.01)
discretized real oil price			5.08	5.08	5.08
			(0.02)	(0.03)	(0.02)
constant			4.96	4.96	4.96
			(0.04)	(0.04)	(0.03)
Ex ante expected value of an unexplored tract (million 1982 $)					
winning bid bin	drilling cost bin	oil price bin			
0	0	0	8.27	8.37	7.68
			(1.57)	(2.19)	(0.98)
0	0	1	5.34	5.33	5.36
			(0.02)	(0.02)	(0.02)
1	0	0	17.18	10.84	88.84
			(0.00)	(0.00)	(0.00)
1	0	1	15.00	15.01	203.41
			(0.27)	(0.24)	(0.18)
2	0	1	16.24	20.62	22.78
			(0.94)	(0.85)	(0.39)
# markets			55	61	64
# observations			1012	1119	1165

Notes: The acreage is the acreage of each tract in the market. The ex ante expected value of an unexplored tract (million 1982 $) is the expected value of an unexplored tract at time $t = 0$ when neither of the tracts in the market have been explored and when the discretized average winning bid per acre, the discretized real drilling cost and the discretized real oil price take on the given values at time $t = 0$.

TABLE 8b. Structural analysis: Results for small tracts

			Acreage < 5000	< 4000	< 3000
σ_μ			4.90	4.85	4.80
			(0.00)	(0.00)	(0.00)
σ_ε			16.37	12.96	17.80
			(0.03)	(0.01)	(0.02)
coefficient α in the exploration profit function on:					
discretized real drilling cost + 1			-9.96	-9.91	-9.86
			(0.00)	(0.00)	(0.00)
coefficients γ in the development profit function on:					
# tracts in market that have been explored			-10.73	-11.14	-15.08
			(0.01)	(0.01)	(0.01)
# tracts in market that have been developed			1.09	1.51	3.29
			(0.00)	(0.00)	(0.00)
discretized average winning bid per acre			-1.14	-0.03	-3.82
			(0.00)	(0.00)	(0.00)
discretized real drilling cost			-10.30	-9.97	-9.94
			(0.00)	(0.00)	(0.00)
discretized real oil price			4.82	4.80	4.85
			(0.00)	(0.00)	(0.00)
constant			-1.79	-1.17	-4.38
			(0.01)	(0.00)	(0.00)
Ex ante expected value of an unexplored tract (million 1982 $)					
winning bid bin	drilling cost bin	oil price bin			
0	0	0	39.27	13.74	18.00
			(1.39)	(0.67)	(1.44)
0	0	1	N/A	N/A	N/A
1	0	0	12.95	10.58	10.67
			(1.36)	(0.76)	(1.22)
1	0	1	8.41	6.36	8.04
			(4.89)	(0.00)	(4.79)
2	0	1	N/A	N/A	N/A
# markets			29	21	19
# observations			590	451	414

Notes: The acreage is the acreage of each tract in the market. The ex ante expected value of an unexplored tract (million 1982 $) is the expected value of an unexplored tract at time $t = 0$ when neither of the tracts in the market have been explored and when the discretized average winning bid per acre, the discretized real drilling cost and the discretized real oil price take on the given values at time $t = 0$.

TABLE 9. Structural analysis: Simulation results from varying the lease term

Ex ante expected value of an unexplored tract (million 1982 $)			Lease term (years) 2	4	5	6	8	10	15
bid bin	cost bin	oil price bin							
0	0	0	7.13	9.20	8.00	9.00	5.20	3.04	243.23
			(2.66)	(0.55)	(2 E4)	(0.13)	(2.20)	(0.40)	(1487.6)
0	0	1	5.81	6.99	5.39	6.73	3.50	1.05 E4	634.62
			(2.43)	(0.55)	(1.50)	(0.19)	(1.88)	(10.48 E5)	(4631.3)
1	0	0	5.81	6.99	1.91 E6	6.73	3.50	2.93	2.88
			(2.43)	(0.55)	(1 E23)	(0.19)	(1.88)	(0.46)	(0.13)
1	0	1	5.81	6.99	1.46 E6	6.73	3.50	1.72	1.62
			(2.43)	(0.55)	(9 E24)	(0.19)	(1.88)	(0.37)	(0.02)
2	0	1	5.81	6.99	228.49	6.73	3.50	1.72	1.62
			(2.43)	(0.55)	(4.03)	(0.19)	(1.88)	(0.37)	(0.02)

Notes: Actual results are reported for a lease term of 5 years; for each of the other values of the lease term, results are from 100 simulations of 87 markets each. The ex ante expected value of an unexplored tract (million 1982 $) is the expected value of an unexplored tract at time $t = 0$ when neither of the tracts in the market have been explored and when the discretized average winning bid per acre, the discretized real drilling cost and the discretized real oil price take on the given values at time $t = 0$.

TABLE 10. Structural analysis: Monte Carlo results

	(1) True value	(1) Mean	(1) Standard Deviation	(2) True value	(2) Mean	(2) Standard Deviation
σ_μ	1	2.17	0.14	40	40.00	0.00
σ_ε	4	4.25	0.03	15	14.01	1.41
coefficient α in the exploration profit function on:						
discretized real drilling cost + 1	-8	-7.70	0.04	-10	-10.00	0.01
coefficients γ in the development profit function on:						
# tracts in market that have been explored	-6	-6.00	0.00	2	1.81	0.22
# tracts in market that have been developed	-7	-7.00	0.00	1	0.93	0.08
discretized average winning bid per acre	8	8.00	0.00	5	4.97	0.02
discretized real drilling cost	-9	-9.00	0.00	-8	-8.00	0.01
discretized real oil price	10	10.00	0.00	5	4.98	0.01
constant	11	11.00	0.00	-15	-15.06	0.03

Notes: Results are from 100 simulations of 87 markets each.