

The Multi-Stage Investment Timing Game in Offshore Petroleum Production: Preliminary results from an econometric model

C.-Y. Cynthia Lin¹
Preliminary—Comments welcome

Abstract

This paper uses a structural econometric model to analyze the investment timing game in offshore petroleum production that ensues on wildcat tracts in U.S. federal lands off the Gulf of Mexico. When individual petroleum-producing firms make their exploration and development investment timing decisions, there are two types of externalities that they do not internalize: an information externality and an extraction externality. The model I develop enables me to estimate the structural parameters governing each firm's investment timing decisions and therefore to assess the net effect of these externalities. According to my results, the extraction externality appears to dominate the information externality. Moreover, decreasing the lease term may increase ex ante tract value and hence government profits. The econometric methodology presented in this paper can be employed to analyze any problem of dynamic multi-stage strategic decision making in the presence of externalities.

First draft: April 2, 2004

This draft: October 14, 2004

¹Department of Economics, Harvard University; cclin@fas.harvard.edu. I thank Gary Chamberlain, William Hogan, Dale Jorgenson, David Laibson, Markus Mobius, and Ariel Pakes for their support and guidance throughout this project. Victor Aguirregabiria, Michael Ostrovsky, Jack Porter, Marc Rysman, and Jim Stock provided useful suggestions on how to develop my econometric model. This paper also benefited from discussions with Joe Aldy, Glenn Ellison, Scott Farrow, Drew Fudenberg, Jerry Green, Kenneth Hendricks, Ali Hortacsu, Darby Jack, Adrian Lajous, Julie Mortimer, Bijan Mossavar-Rahmani, Rohan Pitchford, Robert Porter, Juan Rosellon, Al Roth, Lori Snyder, Robert Stavins, Sasha Volokh, and Martin Weitzman. I received helpful comments from participants at the First Bonzenfreies Colloquium on Market Dynamics and Quantitative Economics in Alessandria, Italy in September 2004; and from participants at student workshops in industrial organization, in microeconomic theory and in environmental economics at Harvard University. I thank Ken Hendricks and Rob Porter for sharing their lease sale data with me. Bill Jones from J.S. Herold generously provided proprietary data that will be useful for future versions of this paper. Jeff Obermiller (American Petroleum Institute), Jon Jeppesen (Sr. VP, Apache), Tom Chambers (VP, Apache), Aloulou Fawzi (Energy Information Administration), and Michael Lynch (President, Strategic Energy and Economic Research, Inc.) provided leads for possible sources of data. Tricia Fiore, April Larson and John Shaw helped me to understand the geology of oil production. Shelby Gerking, James Griffin, Omowumi Iledare, John List, Adele Morris, David Nissen, and James Smith provided useful leads for sources of data and other information. Debra Armond, John Bratland, Barney Congdon, George Dellagiarino, Barry Dickerson, Larry Slaski, John Rodi, Marshall Rose, Ted Tupper, Steve Waddell, and Robert Zainey, all from the Minerals Management Service, provided valuable information about the OCS leasing program. I am indebted to Bijan Mossavar-Rahmani (Chairman, Mondoil Corporation) and Bill 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 thank Mark Bauer (Reservoir Engineering Manager, Apache), Robert Dye (VP, Apache), Steve Farris (President, CEO & COO, Apache), Richard Gould (VP, Wells Fargo), Paul Griesedieck (Manager, Apache), Thomas Halsey (Corporate Strategic Research, ExxonMobil Research and Engineering), Becky Harden (Land Manager, Apache), David Higgins (Director, Apache), Jon Jeppesen (Sr. VP, Apache), Adrian Lajous (President, Oxford Institute for Energy Studies), Kenneth McMinn (Offshore District Production Manager, Apache), Gregg Olson (Director, Apache), and Bob Tippee (Editor, Oil and Gas Journal) for enlightening discussions about the petroleum industry. I thank Derrick Martin for flying me offshore by helicopter, Mike Thibodeaux for giving me a tour of the production platform, Joey Bridges for giving me a tour of the drilling rig, and especially Billy Ebarb (Production Superintendent, Apache) for accompanying me throughout the entire offshore trip. I received financial support from an EPA Science to Achieve Results (STAR) graduate fellowship, a National Science Foundation (NSF) graduate research fellowship, and a Repsol YPF - Harvard Kennedy School Pre-Doctoral Fellowship in energy policy. All errors in this draft (and there are many!) are my own.

1 Introduction

Petroleum production is a multi-stage process involving sequential investment decisions. When a firm acquires a previously unexplored tract of land, it must first decide whether and when to invest in the rigs needed to begin exploratory drilling. 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, as they do in U.S. federal offshore lands in the Gulf of Mexico (Hendricks & Porter, 1993), there are two types of externalities that add a strategic dimension to firms' investment timing decisions.² The first type of externality is an *information externality*: if firms learn information about their own tracts when other firms drill exploratory wells or install production platforms on neighboring tracts, then firms may play a noncooperative 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 (Hendricks & Porter, 1993).³ A second type of externality is an *extraction externality*: when firms have competing rights to a common-pool resource, strategic considerations may lead firms to extract at an inefficiently high rate (Libecap & Smith, 1999; Libecap & Wiggins, 1985). Because these two externalities have opposing theoretical effects on the rate of production, empirical methods are needed to determine the net effect.

In this paper I develop and estimate a structural econometric model of dynamic strategic decision making in the presence of externalities, and apply it to the investment timing game in petroleum production. Estimates of the structural parameters of my model enable me to assess the net effect of the information and extraction externalities and to determine which externality dominates.

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. As will become clear when I describe my model, these parameters will include not only those governing the relationship between various state variables and the

²In 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.

³If 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, 1993; 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).

profits of firms, but also parameters governing the distribution of tract-specific private information. Second, by modelling the dynamic programming problem of the firms, a structural model enables one to estimate each firm's value function, and as a consequence, the ex ante expected tract values. Since firms' bids in the auctions of U.S. federal offshore lands in the Gulf of Mexico – and hence government revenue – are a function of the ex ante expected value of the tracts, an estimate of how ex ante expected values relate to the state variables in the model gives a measure of how one might expect government revenue to change if these state variables were changed. The third advantage to using a structural model is that, in principle, it enables the evaluation of the effects of changes in policy or in the environment.⁴ Finally, the parameters provide a foundation for estimating the net efficiency effect resulting from the two countervailing externalities.

I use my structural econometric model to answer the following questions, among others. First, *do firms care about what their neighbors do?* In other words, *how important are the information and extraction externalities described above, and which externality dominates?* Second, *can the federal government increase ex ante tract value by changing the lease term or by changing the tract size?*⁵

According to my results, the extraction externality appears to dominate the information externality. Moreover, decreasing the lease term may increase ex ante tract value and hence government profits.

The balance of the paper proceeds as follows. Section 2 describes the paper's contributions to the existing literature. I present my model of the investment timing game in offshore petroleum production in Section 3 and my structural econometric estimation strategy in Section 4. In Section 5, I describe my data. I present the results from preliminary reduced-form analyses in Section 6 and the results from my structural model in Section 7.

2 Contributions to Existing Literature

My research will make several contributions to the existing literature. The first strand of literature upon which my work innovates is that on the information externality that arises in the federal offshore leasing program, particularly those studies by Kenneth Hendricks, Robert Porter and their coauthors (see e.g., Hendricks & Kovenock, 1989; Hendricks & Porter, 1993). My research extends their work in several ways. First, while previous studies used stylized theoretical models or reduced-form empirical tests to analyze the information externality, my work estimates the structural parameters of the discrete choice dynamic game. With these structural

⁴As will be described in this paper, strong assumptions need to be made in my model in order for these policy evaluations to be valid. In particular, I need to assume that the equilibrium does not change when the policy environment changes. The development of a structural model that enables policy evaluations to be made without these strong assumptions will be the subject of future work.

⁵In addition to changing the lease term and/or tract size, other possible modifications to the current federal offshore lease auction mechanism might include encouraging unitization of exploration programs, perhaps by limiting the amount of nonunitized acreage that a firm can possess; requiring firms to make their seismic reports publicly available (Hendricks & Kovenock, 1989); changing the quantity or location of the tracts offered in each lease sale; using multi-unit auctions; or making the contractual environment more conducive to coordination.

parameters, one can not only estimate the ex ante value of each of the tracts, but also analyze the effects of alternate policies.⁶

A second way in which my work contributes to the existing literature on the information externality in the federal leasing program 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).

The third innovation I make to the existing literature on the information externality in the federal leasing program is that while the existing literature focuses exclusively on spillovers that arise during exploratory drilling, my model allows for information spillovers that arise during both exploration and development, as well as for common pool extraction spillovers that arise during extraction. 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 actually induce firms to expedite rather than delay their production process.

In addition to the literature on the information externality, a second branch of literature upon which my work innovates is that on econometric models of discrete dynamic games. In particular, my work builds upon the methods developed by Pakes, Ostrovsky and Berry (2004) for estimating parameters of discrete dynamic games such as those involving firm entry and exit. This paper extends the work of Pakes et al. (2004) in several ways. First, unlike their paper, which uses simulation 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

⁶The validity of policy evaluation is subject to the caveats outlined in a previous footnote.

in my econometric model.

A third innovation I make to the literature on discrete dynamic games 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 exact 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 questions I hope to answer, I choose to estimate this relationship from the data rather than impose it *a priori*. However, the task of estimating these additional parameters poses an additional econometric challenge, not only because they require the development and use of additional moment conditions, but also because profits are only observed for tracts that are developed.⁷

In addition to the literatures on the information externality and on discrete dynamic games, the third strand of literature upon which my work innovates is that on oil production. My work makes several contributions to this literature. First, unlike previous work on the federal offshore lease auctions, my study will include not just revenue maximization, but also expeditious exploration, environmental protection, and strategic concerns in the government's objective function. Although the trade-offs between revenue and other objectives have been examined in timber and spectrum auctions, very little work has been done evaluating these trade-offs for oil auctions (Robert Porter, personal communication, March 12, 2002). Second, while the financial theory of option valuation is fairly well developed, structural models applying the theory to the oil production process still remain to be estimated. 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 neither compare the predicted investment timing policy with those that were actually followed nor account for the information externality. Hurn and Wright (1994) test reduced-form implications of the theory, but do not estimate a structural model. Third, previous work on socially optimal natural resource extraction has been primarily theoretical (Robert Porter, personal communication, March 12, 2002); in contrast, my work will also involve a structural estimation.

3 A Model of the Investment Timing Game

In my model, 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. For each market k , the state of the market t years after the

⁷I hope eventually to further extend the Pakes et al. framework by using both continuous and discrete variables. In contrast, the state variables in the framework developed by Pakes et al. must be discrete.

leases began is given by a vector Ω_{kt} of finite-valued state variables that are observed by all the firms in market k and as well as by the econometrician.

Each firm's time- t investment timing decision depends in part on the state of the market Ω_{kt} , which can be decomposed into endogenous state variables and exogenous state variables. There are two endogenous state variables upon which investment timing decisions depend: 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} upon which investment timing decisions depend include the exploration cost c^e and the development cost c^d and are assumed to evolve as a first-order Markov process: $X_{k,t+1} \stackrel{iid}{\sim} F_X(\cdot | X_{k,t})$.

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.⁸ 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 μ_{it} , which is only observed by the firm owning tract i , represents the outcome of the post-sale, pre-exploration seismic study conducted on tract i at time $t - 1$, and affects the exploration investment decision made on tract i at time t .⁹ I assume that $\mu_{it} \stackrel{iid}{\sim} \text{exponential}(\sigma_\mu)$.

The second source of private information is a post-exploration shock ε_{it} to an explored but undeveloped tract i at time t . This post-exploration shock ε_{it} , which is only observed by the firm owning tract i , represents the outcome of the exploratory drilling conducted on tract i at time $t - 1$, and affects the development investment decision made on tract i at time t . I assume that $\varepsilon_{it} \stackrel{iid}{\sim} \text{exponential}(\sigma_\varepsilon)$. In addition, I assume that the pre-exploration shocks μ_{it} and the post-exploration shocks ε_{it} are independent of each other.¹⁰

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 assumptions above on the observable state variables and on the unobservable shocks, firms can take expectations over their neighbors' private information. While each firm plays a pure strategy, from the point of view of their neighbors, they appear to play mixed strategies.¹¹

⁸I currently do not have firm fixed effects or correlations between tracts owned by the same firm, but should think about how I might incorporate them.

⁹Firms conduct and analyze seismic studies in order to help them decide whether or not to begin exploratory drilling (John Shaw, personal communication, April 18, 2003; Bob Dye, personal communication, January 21, 2004; Jon Jeppesen, personal communication, January 21, 2004; Mark Bauer, personal communication, January 21, 2004; Billy Ebarb, personal communication, January 22, 2004).

¹⁰The assumptions that both types of shocks are i.i.d. and independent of each other, while restrictive, is 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. In future work I hope to develop econometric techniques that allow these assumptions to be relaxed.

¹¹Thus, 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.

The sequential investment problem is solved backwards (Dixit & Pindyck, 1994). The second-stage problem is to determine the optimal timing for investing in the development of a tract that has already been explored. The first-stage problem is to determine the optimal policy for investment in exploration. I will now examine each stage in turn.

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

$$V^e(\Omega_{kt}, \varepsilon_{it}; \theta) = \max\{\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta), \beta V^{ce}(\Omega_{kt}; \theta)\}, \quad (1)$$

where $\theta \equiv (\sigma_\mu, \sigma_\varepsilon, \gamma)'$ denotes the vector of parameters to be estimated, $\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta)$ is the profit that a firm will get after developing tract i at time t , and where $V^{ce}(\Omega_{kt}; \theta)$ is the continuation value to waiting instead of developing at time t .

I will assume that the profit $\pi^d(\cdot)$ 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:

$$\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) = \Omega'_{kt} \gamma + \varepsilon_{it} \quad (2)$$

where the deterministic component of profit is linear the state variables, which include the oil price and the development cost c^d .

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

From the exponential distribution for ε_{it} and the functional form assumption 2 for profits, the continuation value $V^{ce}(\cdot)$ to waiting instead of developing reduces to:

$$V^{ce}(\Omega_{kt}; \theta) = E[\beta V^{ce}(\Omega_{k,t+1}; \theta) + g^d(\Omega_{k,t+1}) \cdot \sigma_\varepsilon | \Omega_{kt}], \quad (3)$$

and the development probability $g^d(\cdot)$ reduces to the following function of the continuation value, the state

variables and the parameters:

$$g^d(\Omega_{kt}) = \exp\left(-\frac{\beta V^{ce}(\Omega_{kt}; \theta) - \Omega'_{kt}\gamma}{\sigma_\varepsilon}\right). \quad (4)$$

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

$$V^n(\Omega_{kt}, \mu_{it}; \theta) = \max\{\pi^e(\Omega_{kt}, \mu_{it}; \theta), \beta V^{cn}(\Omega_{kt}; \theta)\}. \quad (5)$$

where $\pi^e(\Omega_{kt}, \mu_{it}; \theta)$ is the payoff to exploring tract i at time t , and where $V^{cn}(\Omega_{kt}; \theta)$ is the continuation value to waiting instead of exploring at time t .

I assume that the payoff $\pi^e(\cdot)$ to exploring tract i at time t can also be separated into a deterministic component and a stochastic component as follows:

$$\pi^e(\Omega_{kt}, \mu_{it}; \theta) = E_\varepsilon [V^e(\Omega_{kt}, \varepsilon_{it}; \theta) | \Omega_{kt}] - c^e + \mu_{it} \quad (6)$$

where, owing to the sequential nature of the investments, the deterministic component of the payoff to exploring is equal to the expected value of an explored but undeveloped tract net the cost of exploration.

Let $g^e(\Omega_{kt})$ 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.

From the exponential distribution for μ_{it} and the functional form assumption 6 for profits, the continuation value $V^{cn}(\cdot)$ to waiting instead of exploring reduces to:

$$V^{cn}(\Omega_{kt}; \theta) = E [\beta V^{cn}(\Omega_{k,t+1}; \theta) + g^e(\Omega_{k,t+1}) \cdot \sigma_\mu | \Omega_{kt}], \quad (7)$$

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

$$g^e(\Omega_t^n) = \exp\left(-\frac{\beta V^{cn}(\Omega_{kt}; \theta) - (\beta V^{ce}(\Omega_{kt}; \theta) + g^d(\Omega_{kt}) \cdot \sigma_\varepsilon) + c^e}{\sigma_u}\right). \quad (8)$$

Further details about my model are given in Appendix A.

4 Structural Econometric Estimation Technique

My econometric estimation technique employs a two-step estimation procedure for each of the two investment stages in petroleum production. This model builds upon recently developed methods for estimating parameters of one-stage discrete dynamic games such as those involving firm entry and exit (Pakes et al., 2004) by allowing the decision making game to involve multiple stages rather than only one.

In the first step, estimates of the continuation values $V^{ce}(\Omega_{kt}; \theta)$ and $V^{cn}(\Omega_{kt}; \theta)$ are derived by solving sequentially for fixed points in equations (3) and (7), 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.

After obtaining estimates of the continuation values in the first step of my econometric estimation technique, I estimate the parameters θ in the second step using generalized method of moments. The moments I construct involve matching the probabilities of exploration and development predicted by my model, as given by substituting the estimates for the continuation values into equations (8) and (4), with the respective empirical probabilities in my data. Similarly, I construct moments that match the profits for tracts that develop predicted by my model with the actual profits observed in the data.

Further details about my econometric estimation technique are given in Appendix B.

5 Data

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. In a typical lease sale, there are over a hundred tracts which are sold simultaneously in separate first-price, sealed-bid auctions. The size of a tract is often less than the acreage required to ensure exclusive ownership of any deposits which may be present (Hendricks & Kovenock, 1989), and tracts within the same area may be located over a common pool (Hendricks & Porter, 1993). When a tract is won, a firm must begin exploration before the end of the lease term, or else relinquish its lease. A lease term is usually five years long. Wildcat tracts are located in regions where no exploratory drilling has occurred previously. While seismic studies are permitted prior to the sale of a wildcat tract, on-site drilling is not (Porter, 1995).

Although the government captures a reasonable share of the rents in wildcat auctions (Porter, 1995), the federal leasing program may induce inefficient investment timing policies owing to information externalities and extraction externalities. While information externalities may induce firms to inefficiently delay production,

extraction externalities may induce firms to produce too quickly.¹² Are the investment timing policies inefficient on net? Because these two externalities have opposing theoretical effects on the rate of production, empirical methods are needed to determine the net effect.

I use a data set on federal lease sales in the Gulf of Mexico between 1954 and 1990 compiled by Ken Hendricks and Rob Porter using U.S. Department of Interior data. For each tract receiving at least one bid, the data set includes information on the date of sale; location and acreage; the identity of bidders and their bids; whether the government accepted the high bid; the number and date of wells that were drilled; and monthly production of oil, condensate, natural gas, and other hydrocarbons (Porter, 1995). Their data set also include predicted ex post revenues and costs (in 1982 U.S. dollars) calculated from the drilling and production data and from the annual survey of drilling costs conducted by the American Petroleum Institute (Hendricks, Porter & Boudreau, 1987).

Fifty-nine lease sales took place offshore of Louisiana and Texas between 1954 and 1990; of these, 48 included wildcat tracts. There are 9537 tracts in the entire data set, 2510 of which are wildcat tracts offered for sale between 1954 and 1979, and another 5920 of which are wildcat or proven tracts offered for sale between 1980 to 1990, for a total of 8340 wildcat tracts. For the remainder of my data summary, I will focus on these wildcat tracts only. Of these wildcat tracts, 2255 and 5644 of the tracts were sold for sale dates from 1954 to 1979 and from 1980 to 1990, respectively, for a total of 7899 wildcat tracts sold; the high bids on the remaining tracts were rejected by the government. All the tracts were offshore; 3599 were in Louisiana state waters, 2042 were in Texas state waters, 2412 were in federal waters off the coast of Louisiana, and 361 were in federal waters off the coast of Texas. On average, a wildcat tract covers 5099 acres.

In Figure 1, I plot the location of each of the wildcat tracts. 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 is approximated by map inspection.

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.¹³ The cost is average cost per well over all wells(oil wells, gas wells, dry holes), in

¹²In addition to information externalities and extraction externalities, a third reason why the federal leasing program is socially inefficient is that it does not account for the environmental costs of the oil production process. Such costs might include the ecosystem damage caused by exploratory drilling or the pollutants emitted during reserve development. To the extent that they affect social welfare, these environmental costs affect the socially optimal policy for exploration, development and extraction.

A fourth reason why the oil production rates induced by the federal leasing program may be inefficient is that they do not account for the need for a domestic strategic oil reserve. Such a need may dictate that the socially optimal policy involve less delay than would occur otherwise, or it may tie the production rate to current political factors. In this paper I focus exclusively on the information and extraction externalities; environmental and strategic concerns will be the subject of future work.

¹³The 1982 issue *Joint Association Survey on drilling costs* of the was out of print.

nominal dollars, for Total Offshore.¹⁴ I convert the nominal costs to real costs in 1982-1984 U.S. dollars using the consumer price index (CPI). I discretize the real drilling cost into three bins: low (0 million - 2 million 1982-1984 U.S. dollars), medium (2 million - 3 million 1982-1984 U.S. dollars), high (over 3 million 1982-1984 U.S. dollars). Figure 2 plots the real drilling cost data, along with the bins.¹⁵

For the real oil price, I use the U.S. average crude oil domestic first purchase price from the EIA Annual Energy Review; the oil price is in 1996 U.S. dollars. I discretize the real oil price into three bins: low (0-13 1986 dollars), medium (13-25 1986 dollars), and high (over 25 1986 dollars). Figure 3 plots the real oil price, along with the bins.¹⁶

6 Preliminary reduced-form analysis

The offshore region off the coasts of Louisiana and Texas is divided into 51 separate geographical areas (Hendricks & Porter, 1996); the wildcat tracts in this data set span 37 areas. For my preliminary analysis, I will assume, as did Hendricks and Porter (1996), that all tracts in a given area are potential neighbors. I define an area-cohort as a set of leases sold at the same time and located in the same area. There are 530 area-cohorts containing wildcat tracts. On average, there are 15.9 wildcat tracts per area-cohort; the number of wildcat tracts per area-cohort ranges from 1 to 150. One firm can have the highest percentage share of between 1 and 110 wildcat tracts per area-cohort, with an average of 3.7 wildcat tracts.

According to crude reduced-form regression results (not shown), the number of tracts in an area-cohort owned by a firm has a significant effect on gross profits, net profits and unit net profits, but has no significant effect on unit gross profits, average gross profits, average net profits, area-cohort-level unit gross profits, or area-cohort-level unit net profits. That is, the number of tracts in an area-cohort owned by a firm affects the profitability of each individual tract it owns in the area cohort, but does not affect the aggregate profitability of all the tracts it owns in the area-cohort. Thus, the data provides some weak evidence for a possible cross-tract externality within area cohorts. A structural model of the dynamic investment timing game would better assess the extent of the externalities.

¹⁴I also have data for Offshore Louisiana (LA) and offshore Texas (TX).

¹⁵There are 3 bins: 0=low, 1=medium, 2=high.

¹⁶There are 3 bins: 0=low, 1=medium, 2=high.

7 Structural Econometric Estimation

For my structural estimation, I focus on exclusive 2-tract markets whose sale dates occurred before 1980. More specifically, 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 did not occur before the sale date, and for which production did not occur before exploration. 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,¹⁷ both the sale dates and the lease terms are the same for both tracts, and no other wildcat tracts from the same sale date are within 6 miles of either tract. There are 175 such markets in my data set, which range in duration from 5 years to 22 years, with an average length of 7.01 years (s.d.=3.61). Figure 4 maps out the tracts used.

I restrict the size of the market N to two tracts for two reasons. First, since the number of combinations of state variables is quadratic in N ,¹⁸ limiting the market size to two minimizes the state space. 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, my econometric model becomes a lot more complicated.

Each discrete period t will be a year. The time of exploration will be the year of the tract's first spud date. The time of development will be the year of the tract's first production date. I assume that development and production are one decision. It is possible that development begins in the same year as exploration does. I choose my state variables based on considerations of state space, policy analysis and data availability. For my preliminary empirical analysis, the exogenous state variables chosen were the number of years left until the least term expires, a dummy for whether the tracts belong to the same firm, the discretized real drilling cost, and the discretized real oil price.¹⁹

Table 1 presents summary statistics of the state variables.

¹⁷I choose 6 miles because the maximum tract size is 5760 acres, or 3 miles by 3 miles (Marshall Rose, personal communication, 17 April 2003; Larry Slaski, personal communication, 25 April 2003). I convert latitude and longitude to miles using the following following factors from the Louisiana 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.

¹⁸This is because the supports for the two endogenous variables – the total number of tracts in a market that have been explored (tot_e_{kt}) and the total number of tracts in a market that have been developed (tot_d_{kt}) – both increase in the number of firms in the market, and the number of possible combinations of state variables is the product of the supports of each of the state variables.

¹⁹Because I only have drilling cost data, I use as my exploration cost:

$$c_t^e = 1000000 * (c_bin_t + 1) + 500000,$$

where c_bin_t is the discretized real drilling cost, and I use as my development cost the binned drilling cost:

$$c_t^d = c_t^e.$$

Table 1. Summary Statistics

Variable	obs	mean	s.d.	min	max
# tracts in market that have been explored	1402	1.01	0.85	0	2
# tracts in market that have been developed	1402	0.21	0.51	0	2
# years left until lease expires	1402	1.87	1.83	0	5
dummy for the same firm owning both tracts	1402	0.49	0.50	0	1
discretized real drilling cost	1402	0.34	0.67	0	2
discretized real oil price	1402	0.32	0.56	0	2

Table 2 presents, for the tracts that are eventually developed, summary statistics of the values of the state variables at the time of development, as well as the profits and gross revenues from development. Among the 350 tracts I considered in my estimation, 130 were eventually developed.

Table 2. Summary statistics for tracts that were developed

Variable	obs	mean	s.d.	min	max
# tracts in market that have been explored	130	1.65	0.57	0	2
# tracts in market that have been developed	130	0.18	0.39	0	1
# years left until lease expires	130	0.98	1.39	0	5
dummy for the same firm owning both tracts	130	0.46	0.50	0	1
discretized real drilling cost	130	0.61	0.86	0	2
discretized real oil price	130	0.48	0.66	0	2
revenue (million 1982 \$)	130	52.56	70.34	0.02	405.00
gross profits (million 1982 \$)	130	-8.04	9.41	-38.10	18.80

The preliminary results from my structural estimation are shown in Table 3.²⁰ Standard errors were estimated by bootstrapping 100 simulated panels of 175 markets each, as the original data set contained 175 markets. In the bootstrap simulations, whenever a market is drawn, all its time observations are used.

²⁰These estimates were derived from using the "G" estimators of both continuation values; I am in the process of obtaining results for the other estimators as well.

Table 3. Results from econometric model

Parameter	Estimate	Standard Error
σ_μ	19.13	1.79E-14
σ_ε	1.42	3.79E-15
coefficients in profit function on:		
# tracts in market that have been explored	-10.08	1.79E-14
# tracts in market that have been developed	-0.77	4.46E-15
# years left until lease expires	-3.93	5.36E-15
dummy for the same firm owning both tracts	-2.36	2.23E-15
discretized real drilling cost	-3.48	1.79E-15
discretized real oil price	-2.91	6.69E-15
constant	-5.43	8.93E-16

The first feature of my results to note is that all the standard errors are quite small, and thus all my parameters are precisely estimated. The negative coefficient on the total number of tracts in the market that have been explored is evidence *against* a positive exploration information externality, as it suggests that it is better for a firm to explore before its neighbor does, rather than wait to learn from the outcome of its neighbor's exploration. The negative coefficient on the total number of tracts in the market that have been developed is evidence *against* a positive development information externality, but evidence *for* a negative extraction externality, as it suggests that the benefits from extracting before a neighbor in order to enhance a firm's own production outweigh the benefits from delaying production in order to learn from a neighbor's development decision. The negative coefficient on the number of years left in the lease term suggests that if one shortens the lease term, which limits the number of years a firm can delay exploring, one would increase the profits from developing, increase the ex ante tract value, increase the bids in the auctions, and, as a consequence, increase the government revenue from the auction. The negative coefficient on whether or not the two tracts in the market are owned by the same firm suggests that it would be undesirable for both tracts to be owned by the same firm, and perhaps therefore undesirable to increase the tract size. A higher drilling cost decreases profits, as is expected. Perhaps unexpectedly, a higher oil price decreases profits as well.

For preliminary robustness checks that use additional moments (with an optimal weight matrix), that calculate the standard errors analytically, and that vary the first stage estimators of the continuations values, see Table 4.

My results therefore suggest that the extraction externality appears to dominate the information externality. Moreover, decreasing the lease term may increase ex ante tract value and hence government profits. The econometric methodology presented in this paper can be employed to analyze any problem of dynamic strategic decision-making in the presence of externalities.

References

- [1] Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics*, 34, 305-334.
- [2] Dixit, A., & Pindyck, R. (1994). *Investment under uncertainty*. Princeton, NJ: Princeton University Press.
- [3] Harsanyi, J. (1973). Games with randomly disturbed payoffs. *International Journal of Game Theory*, 2, 1-23.
- [4] Hendricks, K., & Kovenock, D. (1989). Asymmetric information, information externalities, and efficiency: The case of oil exploration. *RAND Journal of Economics*, 20 (2), 164-182.
- [5] 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.
- [6] Hendricks, K., & Porter, R. (1996). The timing and incidence of exploratory drilling on offshore wildcat tracts. *The American Economic Review*, 86 (3), 388-407.
- [7] 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.
- [8] Hurn, A., & Wright, R. (1994). Geology or economics? Testing models of irreversible investment using North Sea oil data. *Economic Journal*, 104 (423), 363-371.
- [9] Libecap, G., & Smith, J. (1999). Regulatory remedies to the common pool: The limits to oil field unitization. Mimeo, University of Arizona and Southern Methodist University.
- [10] 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.
- [11] 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.

- [12] Pakes, A., Ostrovsky, M., & Berry, S. (2004). Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). Mimeo. Harvard University.
- [13] Porter, R. (1995). The role of information in U.S. offshore oil and gas lease auctions. *Econometrica*, 63 (1), 1-27.

8 Appendix A: Solving the model

In this Appendix, I provide the details for solving my model of the investment timing game.

8.1 Stage 2: Development

Equation (3) for the continuation value V^{ce} to waiting instead of developing is derived as follows. Substituting in the functional form 2 for profits, the expected truncated profits from development conditional on development can be written as:

$$\begin{aligned}
& E [\pi^d(\Omega_{k,t}, \varepsilon_{i,t}; \theta) | \pi^d(\Omega_{k,t}, \varepsilon_{i,t}; \theta) > \beta V^{ce}(\Omega_{k,t}; \theta)] \\
&= E [\pi_0^d(\Omega_{k,t}; \theta) + \varepsilon_{i,t} | \pi_0^d(\Omega_{k,t}; \theta) + \varepsilon_{i,t} > \beta V^{ce}(\Omega_{k,t}; \theta)] \\
&= \pi_0^d(\Omega_{k,t}; \theta) + E [\varepsilon_{i,t} | \varepsilon_{i,t} > \beta V^{ce}(\Omega_{k,t}; \theta) - \pi_0^d(\Omega_{k,t}; \theta)] \\
&= \beta V^{ce}(\Omega_{k,t}; \theta) + \sigma_\varepsilon,
\end{aligned} \tag{9}$$

where the final step comes from the exponential distribution for ε_{it} . Using equation 9 for the expected truncated profits conditional on development, the continuation value can thus be written as:

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

which yields equation (3) as desired.

Equation (4) for the development probability $g^d(\Omega_{kt})$ is derived as follows. The probability $g^d(\Omega_{kt})$ of developing tract i at time t given that development of tract i has not occurred previously is given by:

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

Substituting in the functional form assumption 2 for profits, we get:

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

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

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

which yields equation (4) as desired.

8.2 Stage 1: Exploration

Equation (7) for the continuation value V^{cn} to waiting instead of exploring is derived as follows. Substituting in

$$E_\varepsilon [V^e(\Omega_{kt}, \varepsilon_{it}; \theta) | \Omega_{kt}] = \beta V^{ce}(\Omega_{kt}; \theta) + g^d(\Omega_{kt}) \cdot \sigma_\varepsilon \quad (10)$$

into the equation for the continuation value $V^{cn}(\cdot)$ to waiting instead of exploring, we get:

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

which yields equation (7) as desired.

Equation (8) for the exploration probability $g^e(\Omega_{kt})$ is derived as follows. The exploration policy function $g^e(\Omega_t^n)$, which is the probability of exploring tract i at time t given that exploration of tract i has not occurred previously, is given by:

$$g^e(\Omega_{kt}) = \Pr(\pi^e(\Omega_{kt}, \mu_{it}; \theta) > \beta V^{cn}(\Omega_{kt}; \theta)). \quad (11)$$

Substituting in the functional form assumption 6 for $\pi_0^e(\Omega_{kt}; \theta)$, we get:

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

which yields equation (8) as desired.

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

$$E_\mu[V^n(\Omega_{k0}, \mu; \theta)|\Omega_{k0}] = \beta V^{cn}(\Omega_{k0}; \theta) + g^e(\Omega_{k0}) \cdot \sigma_\mu. \quad (12)$$

9 Appendix B: Details About Econometric Estimation

Suppose that there are a maximum of N firms in each market. Then the total number of tracts that have been explored $tot_e_{kt} \in \{0, 1, \dots, N\}$ and the total number of tracts that have been developed $tot_d_{kt} \in \{0, 1, \dots, N\}$. Suppose that each of the J discrete exogenous state variable X_j takes on num_j discrete values. Then, there are $n_tuples = (N + 1) \cdot (N + 1) \cdot X_1 \cdot X_2 \cdot \dots \cdot X_J$ possible combinations of the state variables $\Omega_{kt} = (tot_e_{kt}, tot_d_{kt}, X_{kt})$.

My econometric estimation technique employs a two-step estimation procedure for each of the two investment stages in oil production. In the first step, continuation values are estimated, in some cases as functions of the parameters. In the second step, the parameters are estimated via generalized method of moments (GMM).

9.1 Step 1: Estimating the continuation values

9.1.1 The continuation value V^{ce} to waiting instead of developing

- Let $\overrightarrow{V^{ce}}$ be an $n_tuples \times 1$ vector, each component of which is $V^{ce}(\cdot)$ evaluated at a different tuple of state variables.
- Let $\overrightarrow{(g^d \times E[\pi^d | \pi^d > \beta V^{ce}])}$ be an $n_tuples \times 1$ vector, where the component corresponding to the tuple Ω_{kt} of state variables is:

$$g^d(\Omega_{kt}) \cdot E[\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) | \pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) > \beta V^{ce}(\Omega_{kt}; \theta)].$$

- Let $\overrightarrow{((1 - g^d) \times V^{ce})}$ be an $n_tuples \times 1$ vector, where the component corresponding to the tuple Ω_{kt} of state variables is:

$$(1 - g^d(\Omega_{kt})) \cdot \beta V^{ce}(\Omega_{kt}; \theta).$$

- Let $\overrightarrow{(g^d \times \sigma_\varepsilon)}$ be an $n_tuples \times 1$ vector, where the component corresponding to the tuple Ω_{kt} of state variables is:

$$g^d(\Omega_{kt}) \cdot \sigma_\varepsilon.$$

- Let $\overrightarrow{(\exp(-\frac{\beta V^{ce}(\Omega_{k,t+1}; \theta) - \pi_0^d(\Omega_{k,t+1}; \theta)}{\sigma_\varepsilon}) \cdot \sigma_\varepsilon)}$ be an $n_tuples \times 1$ vector, where the component corresponding to the tuple Ω_{kt} of state variables is:

$$\exp\left(-\frac{\beta V^{ce}(\Omega_{k,t+1}; \theta) - \pi_0^d(\Omega_{k,t+1}; \theta)}{\sigma_\varepsilon}\right) \cdot \sigma_\varepsilon.$$

- Let M^e be an $n_tuples \times n_tuples$ transition matrix where 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.

There are three possible estimators of the continuation value for the development stage, each derived from a different equation for the continuation value V^{ce} .

No functional form assumptions in V^{ce} equation (N) This estimator is derived from the following general form of the V^{ce} equation, before any functional form assumptions are made.

$$\begin{aligned}
& V^{ce}(\Omega_{kt}; \theta) \\
= & E \left[\begin{array}{l} g^d(\Omega_{k,t+1}) \cdot E \left[\pi^d(\Omega_{k,t+1}, \varepsilon_{i,t+1}; \theta) | \pi^d(\Omega_{k,t+1}, \varepsilon_{i,t+1}; \theta) > \beta V^{ce}(\Omega_{k,t+1}; \theta) \right] \\ + (1 - g^d(\Omega_{k,t+1})) \cdot \beta V^{ce}(\Omega_{k,t+1}; \theta) \end{array} \middle| \Omega_{k,t} \right]. \quad (13)
\end{aligned}$$

Using the notation from above, the continuation value can be rewritten in vector form as:

$$\overrightarrow{V_N^{ce}} = M^e \left(\overrightarrow{(g^d \times E[\pi^d | \pi^d > \beta V^{ce}])} + \beta \overrightarrow{((1 - g^d) \times V_N^{ce})} \right) \quad (14)$$

I use empirical historical frequencies to estimate M^e and g^d . Moreover, because I only observe profits when firms develop (i.e., when $\pi^d > \beta V^{ce}$), the empirical average of observed profits conditional on state variables can be used to estimate $E[\pi^d | \pi^d > \beta V^{ce}]$. Plugging these estimates into equation (14), I can then solve for a fixed point to get our nonparametric (N) estimate $\widehat{V_N^{ce}}$ of the continuation value.

Parametric estimate of expected truncated profits leading to inversion (I) The second estimator uses the equation for the continuation value derived after solving for the expected truncated development profits $E[\pi^d(\Omega_{k,t+1}, \varepsilon_{i,t+1}; \theta) | \pi^d(\Omega_{k,t+1}, \varepsilon_{i,t+1}; \theta) > \beta V^{ce}(\Omega_{k,t+1}; \theta)]$ using functional form assumptions on profits and distributional assumptions on ε_i :

$$V^{ce}(\Omega_{kt}; \theta) = E[\beta V^{ce}(\Omega_{k,t+1}; \theta) + g^d(\Omega_{k,t+1}) \cdot \sigma_\varepsilon | \Omega_{kt}]. \quad (15)$$

This equation can be rewritten in vector form as:

$$\overrightarrow{V_I^{ce}} = M^e \left(\beta \overrightarrow{V_I^{ce}} + \overrightarrow{(g^d \cdot \sigma_\varepsilon)} \right). \quad (16)$$

Moreover, this equation can be inverted to:

$$\overrightarrow{V_I^{ce}} = (I - \beta M^e)^{-1} M^e \overrightarrow{(g^d \cdot \sigma_\varepsilon)}. \quad (17)$$

Once again, empirical averages $\widehat{M^e}$ and $\widehat{g^d}$ can be used to estimate M^e and g^d , respectively, so that my inversion (I) estimate for $\overrightarrow{V_I^{ce}}$ becomes:

$$\widehat{V_I^{ce}}(\Omega_{kt}; \theta) = (I - \beta \widehat{M^e})^{-1} \widehat{M^e} \overrightarrow{(\widehat{g^d} \cdot \sigma_\varepsilon)}. \quad (18)$$

Fixed point for policy function using all functional form assumptions (G) The third estimator uses the equation for the continuation value derived after solving for both the expected truncated development profits $E[\pi^d(\Omega_{k,t+1}, \varepsilon_{i,t+1}; \theta) | \pi^d(\Omega_{k,t+1}, \varepsilon_{i,t+1}; \theta) > \beta V^{ce}(\Omega_{k,t+1}; \theta)]$ and the policy function $g^d(\cdot)$:²¹

$$V^{ce}(\Omega_{kt}; \theta) = E \left[\beta V^{ce}(\Omega_{k,t+1}; \theta) + \exp \left(-\frac{\beta V^{ce}(\Omega_{k,t+1}; \theta) - \pi_0^d(\Omega_{k,t+1}; \theta)}{\sigma_\varepsilon} \right) \cdot \sigma_\varepsilon | \Omega_{kt} \right]. \quad (19)$$

This equation can be rewritten in vector form as:

$$\overrightarrow{V_G^{ce}} = M^e \left(\beta \overrightarrow{V_G^{ce}} + \overrightarrow{\left(\exp \left(-\frac{\beta V_G^{ce} - \pi_0^d(\theta_\pi)}{\sigma_\varepsilon} \right) \cdot \sigma_\varepsilon \right)} \right). \quad (20)$$

Substituting in the empirical average \widehat{M}^e for M^e , we can solve for a fixed point $\widehat{V_G^{ce}}(\Omega_{kt}; \theta)$, where θ_π is the component of the entire parameter vector θ that consists of parameters of the deterministic component $\pi_0^d(\cdot)$ of profits.

9.1.2 The continuation value V^{cn} to waiting instead of exploring

- Let $\overrightarrow{V^{cn}}$ be an $n_tuples \times 1$ vector, each component of which is $V^{cn}(\cdot)$ evaluated at a different tuple of state variables.
- Let $\overrightarrow{(g^e \times E[\pi^e | \pi^e > \beta V^{cn}])}$ be an $n_tuples \times 1$ vector, where the component corresponding to the tuple Ω_{kt} of state variables is:

$$g^e(\Omega_{kt}) \cdot E[\pi^e(\Omega_{kt}, \mu_{it}; \theta) | \pi^e(\Omega_{kt}, \mu_{it}; \theta) > \beta V^{cn}(\Omega_{kt}; \theta)].$$

- Let $\overrightarrow{((1 - g^e) \times V^{cn})}$ be an $n_tuples \times 1$ vector, where the component corresponding to the tuple Ω_{kt} of state variables is:

$$(1 - g^e(\Omega_{k,t+1})) \cdot \beta V^{cn}(\Omega_{k,t+1}; \theta).$$

- Let $\overrightarrow{(g^e \times \sigma_\mu)}$ be an $n_tuples \times 1$ vector, where the component corresponding to the tuple Ω_{kt} of state variables is:

$$g^e(\Omega_{kt}) \cdot \sigma_\mu.$$

- Let M^n be an $n_tuples^n \times n_tuples^n$ transition matrix where 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.

²¹A fourth estimator would be to solve for the policy function but not for expected truncated profits.

There are two possible estimators of the continuation value for the development stage, each derived from a different equation for the continuation value V^{cn} .²²

Parametric estimate of expected truncated profits leading to inversion (I) This estimator uses

the equation for the continuation value derived after solving for the expected truncated exploration profits $E[\pi^e(\Omega_{kt}, \mu_{it}; \theta) | \pi^e(\Omega_{kt}, \mu_{it}; \theta) > \beta V^{cn}(\Omega_{kt}; \theta)]$ using functional form assumptions on profits and distributional assumptions on μ_{it} :

$$V^{cn}(\Omega_{kt}; \theta) = E[\beta V^{cn}(\Omega_{k,t+1}; \theta) + g^e(\Omega_{k,t+1}) \cdot \sigma_\mu | \Omega_{kt}] \quad (21)$$

This equation can be rewritten in vector form as:

$$\overrightarrow{V_I^{cn}} = M^n \left(\beta \overrightarrow{V_I^{cn}} + \overrightarrow{(g^e \cdot \sigma_\mu)} \right). \quad (22)$$

Moreover, this equation can be inverted to:

$$\overrightarrow{V_I^{cn}} = (I - \beta M^n)^{-1} M^n \overrightarrow{(g^e \cdot \sigma_\mu)}. \quad (23)$$

Once again, empirical averages \widehat{M}^n and \widehat{g}^e can be used to estimate M^n and g^e , respectively, so that our inversion (I) estimate for $\overrightarrow{V_I^{cn}}$ becomes:

$$\widehat{V_I^{cn}}(\Omega_{kt}; \theta) = (I - \beta \widehat{M}^n)^{-1} \widehat{M}^n \overrightarrow{(\widehat{g}^e \cdot \sigma_\mu)}. \quad (24)$$

Fixed point for policy function (G) This estimator uses the equation for the continuation value derived after solving for both the expected truncated exploration profits $E[\pi^e(\Omega_{kt}, \mu_{it}; \theta) | \pi^e(\Omega_{kt}, \mu_{it}; \theta) > \beta V^{cn}(\Omega_{kt}; \theta)]$ and the policy function $g^e(\cdot)$.²³ The intuition is similar to that for the development stage, so I won't repeat it here.

$$V^{cn}(\Omega_{kt}; \theta) = E \left[\beta V^{cn}(\Omega_{k,t+1}; \theta) + \exp \left(- \frac{\beta V^{cn}(\Omega_{kt}; \theta) - (\beta V^{ce}(\Omega_{kt}; \theta) + g^d(\Omega_{kt}) \cdot \sigma_\varepsilon) + c^e}{\sigma_u} \right) \cdot \sigma_\mu | \Omega_{kt} \right]. \quad (25)$$

²²The nonparametric (N) estimator does not work for estimating V^{cn} because one cannot estimate $E[\pi^e(\Omega_{kt}, \mu_{it}; \theta) | \pi^e(\Omega_{kt}, \mu_{it}; \theta) > \beta V^{cn}(\Omega_{kt}; \theta)]$.

²³Another estimator would be to solve for the policy function but not for expected truncated profits.

This equation can be rewritten in vector form as:

$$\overrightarrow{V_G^{cn}} = M^n \left(\beta \overrightarrow{V_G^{cn}} + \left(\exp \left(- \frac{\beta V_G^{cn} - (\beta \widehat{V}^{ce} + g^d(\widehat{\Omega}_{kt}) \cdot \sigma_\varepsilon) + c^e}{\sigma_\mu} \right) \cdot \sigma_\mu \right) \right). \quad (26)$$

where we can use any of the above estimates for $V^{ce}(\widehat{\Omega}_{kt}; \theta)$ and $g^d(\widehat{\Omega}_{kt})$. Substituting in the empirical average \widehat{M}^n for M^n , we can solve for a fixed point $\widehat{V}_G^{cn}(\Omega_{kt}; \theta)$.

9.2 Step 2: Generalized Method of Moments

After obtaining estimates of the continuation values in the first step of my econometric estimation technique, I estimate the parameters θ in the second step using generalized method of moments. To do so, I first substitute my estimates $\widehat{V}^{cn}(\Omega_{kt}; \theta)$ and $\widehat{V}^{ce}(\Omega_{kt}; \theta)$ for the continuation values into equations (8) and (4) for the exploration and development probabilities to get the predicted probabilities:

$$g^d(\widehat{\Omega}_{kt}; \theta) = \exp \left(- \frac{\beta V^{ce}(\widehat{\Omega}_{kt}; \theta) - \pi_0^d(\Omega_{kt}; \theta)}{\sigma_\varepsilon} \right) \quad (27)$$

and

$$g^e(\widehat{\Omega}_{kt}; \theta) = \exp \left(- \frac{\beta V^{cn}(\widehat{\Omega}_{kt}; \theta) - (\beta V^{ce}(\widehat{\Omega}_{kt}; \theta) + g^d(\widehat{\Omega}_{kt}) \cdot \sigma_\varepsilon) + c^e}{\sigma_u} \right). \quad (28)$$

The moments I construct involve matching the probabilities of exploration and development predicted by my model, as given equations (28) and (27), with the respective empirical probabilities in my data. Similarly, I construct moments that match the profits for tracts that develop predicted by my model with the actual profits observed in the data.

For the exactly identified case, my moment function $\Psi(z_{ikt}, \theta)$ is:

$$\begin{aligned}
& (g^e(\Omega_{kt}) - I_{it}^e) \cdot I_{it}^{not_yet_e} \\
& (g^d(\Omega_{kt}) - I_{it}^d) \cdot I_{it}^{e_not_yet_d} \\
& E [\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) | \pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) > \beta V^{ce}(\Omega_{kt}; \theta)] - \pi_i^d \cdot I_{it}^d \\
& tot_ekt \cdot (g^e(\Omega_{kt}) - I_{it}^e \cdot I_{it}^{not_yet_e}) \\
& tot_dkt \cdot (g^d(\Omega_{kt}) - I_{it}^d \cdot I_{it}^{e_not_yet_d}) \\
& X'_{kt} (E [\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) | \pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) > \beta V^{ce}(\Omega_{kt}; \theta)] - \pi_i^d \cdot I_{it}^d)
\end{aligned}$$

where I_{it}^e is an indicator for whether exploration began on tract i at time t ; $I_{it}^{not_yet_e}$ is an indicator variable for tract i not yet being explored before period t ; I_{it}^d is an indicator for whether development began on tract i at time t ; $I_{it}^{e_not_yet_d}$ is an indicator variable for tract i being explored but not yet developed before period t ; and where π_i^d is the profits earned on tract i after development.

For the overidentified case, my moment function is:

$$\begin{aligned}
& (g^e(\Omega_{kt}) - I_{it}^e) \cdot I_{it}^{not_yet_e} \\
& (g^d(\Omega_{kt}) - I_{it}^d) \cdot I_{it}^{e_not_yet_d} \\
& E [\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) | \pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) > \beta V^{ce}(\Omega_{kt}; \theta)] - \pi_i^d \cdot I_{it}^d \\
& \Omega'_{kt} (g^e(\Omega_{kt}) - I_{it}^e \cdot I_{it}^{not_yet_e}) \\
& \Omega'_{kt} (g^d(\Omega_{kt}) - I_{it}^d \cdot I_{it}^{e_not_yet_d}) \\
& \Omega'_{kt} (E [\pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) | \pi^d(\Omega_{kt}, \varepsilon_{it}; \theta) > \beta V^{ce}(\Omega_{kt}; \theta)] - \pi_i^d \cdot I_{it}^d)
\end{aligned}$$

and the optimal weighting matrix as specified by Chamberlain (1987) is used.

FIGURE 1

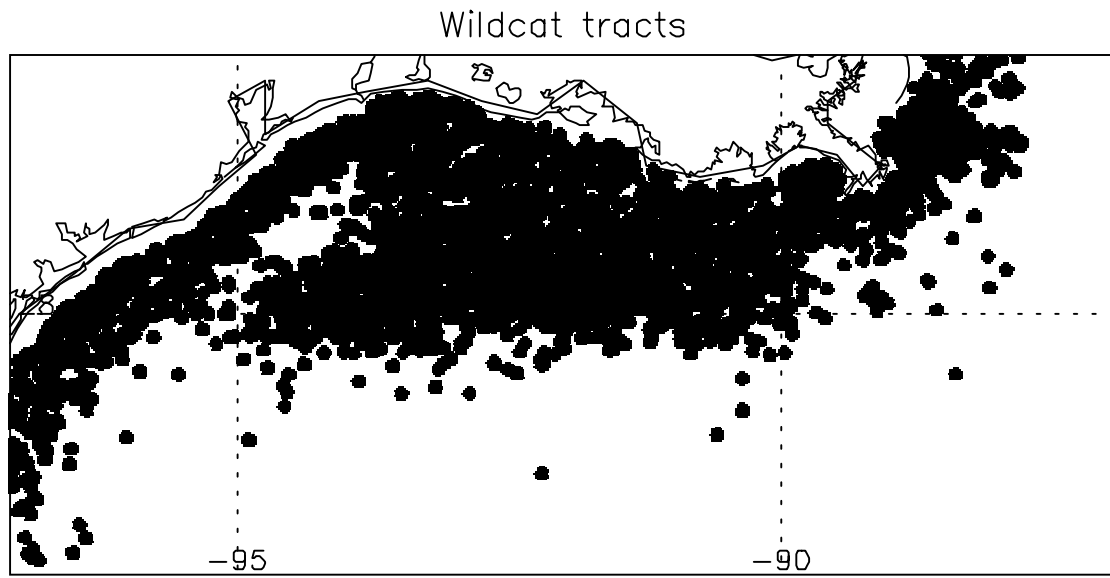
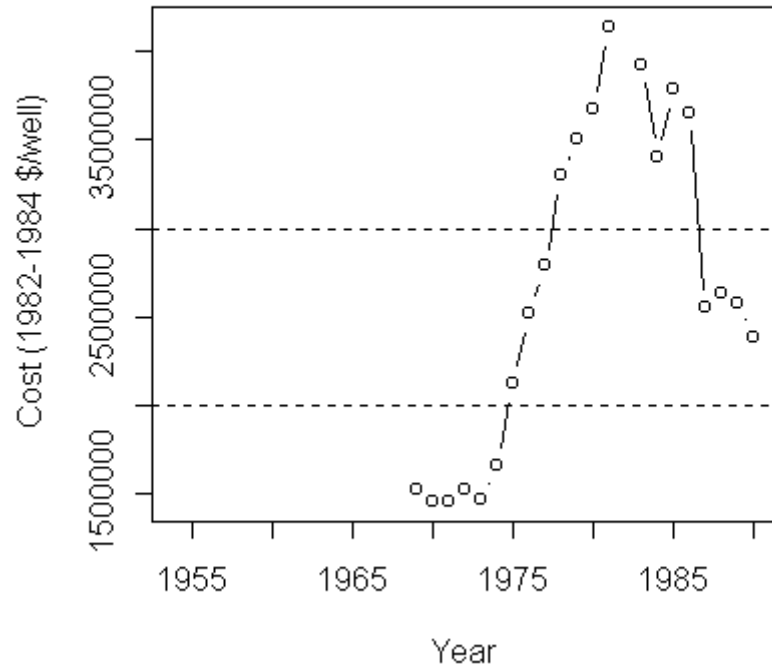


FIGURE 2

U.S. offshore costs per well



Source: from API (1982 data out of print)

FIGURE 3

U.S. average crude oil price

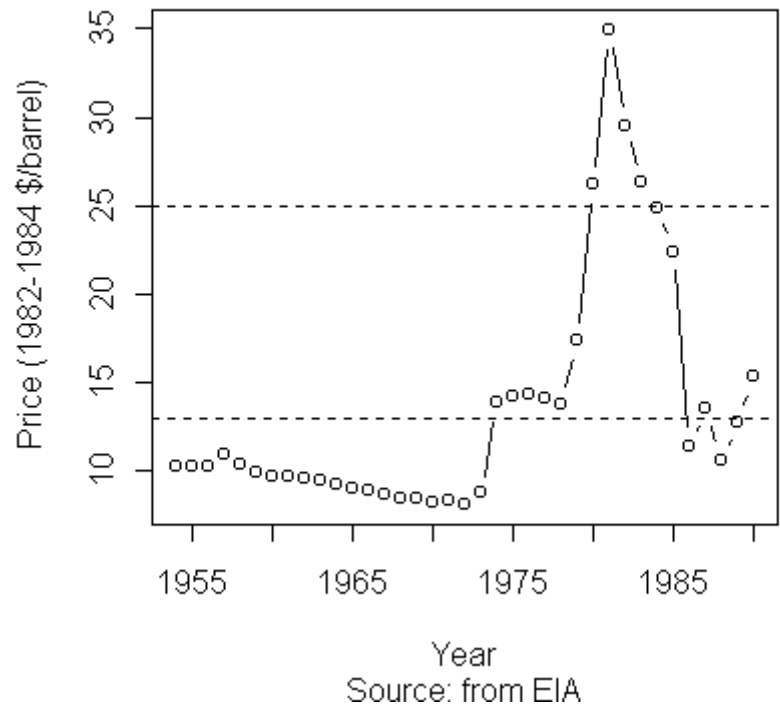


FIGURE 4

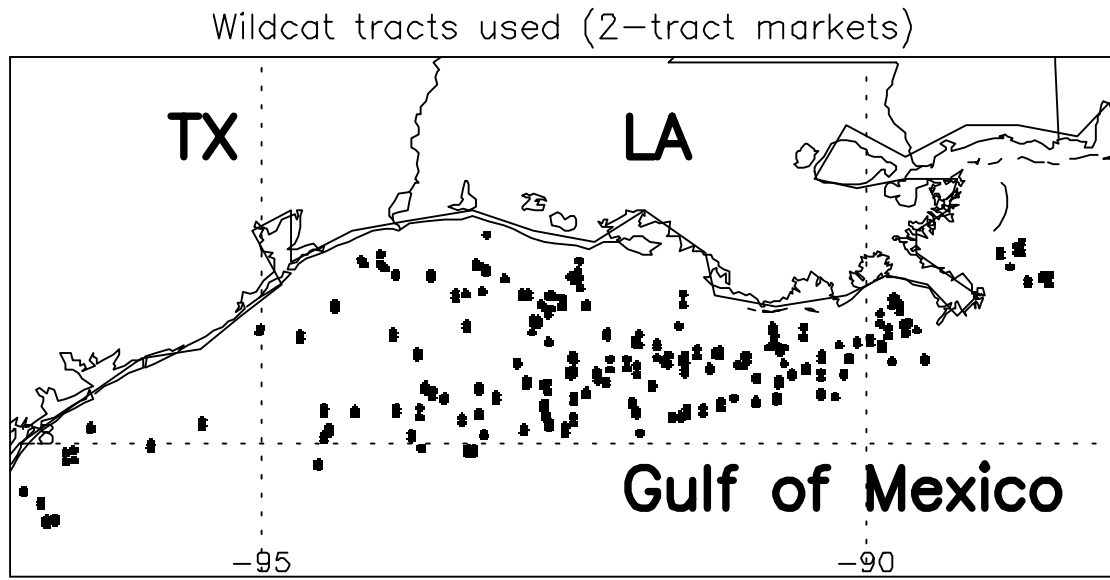


Table 4. Preliminary robustness checks from the econometric model

	(1)	(2)	(3)	(4)	(5)
σ_μ	19.13 (0.00)	38.92 (15.64)	3.92 (0.07)	48.97 (1.80)	36.54 (1.07)
σ_ε	1.42 (0.00)	7.38 (0.00)†	16.28 (0.03)	4.27 (0.00)†	10.47 (0.04)
<i>coefficients in the profit function on:</i>					
# tracts in market that have been explored	-10.08 (0.00)	-16.53 (2107)	-11.82 (0.81)	-65.20 (0.00)†	-7.62 (0.00)†
# tracts in market that have been developed	-0.77 (0.00)	-125.80 (3.3E9)	-0.60 (0.01)	-3.88 (0.00)†	0.38 (0.00)†
# years left until lease expires	-3.93 (0.00)	-0.61 (673.0)	-33.55 (16.19)	-192.49 (5E26)	-33.34 (53.01)
dummy for the same firm owning both tracts	-2.36 (0.00)	-16.86 (372.0)	-4.70 (0.82)	-28.98 (0.00)†	-2.78 (3.14)
discretized real drilling cost	-3.48 (0.00)	-33.87 (5.5E5)	-2.63 (1.01)	-20.74 (0.00)†	-1.28 (3.02)
discretized real oil price	-2.91 (0.00)	-26.22 (6.0E3)	-2.96 (1.02)	-21.50 (0.00)†	-1.56 (1.73)
constant	-5.43 (0.00)	-29.25 (3054)	-11.62 (1.31)	-62.48 (0.00)†	-6.30 (2.72)
estimator for:					
V^{cn}	G	I	G	G	G
V^{ce}	G	G	N	I	G
# moments	9	21	21	21	21
standard errors	bootstrap	analytical	analytical	analytical	analytical

Notes: Standard errors are in parentheses. There are 1402 observations spanning 175 markets. † indicates that the estimated variance was negative.