An Empirical Investigation of Investment under Uncertainty with Sunk Costs: Oil Production in Oklahoma

Boris Molls
Department of Economics
The Pennsylvania State University

November, 2000

Abstract

In the oil industry, significant numbers of well openings and closings are only observed during periods of very high and very low oil prices. This pattern of hysteresis can be explained by the combination of sunk costs and uncertainty of future market conditions. This paper develops an empirical model of firm investment that can measure the role of sunk costs and uncertainty of future market conditions on investment and production. In my empirical dynamic discrete choice model firms make the decision to produce oil, drill, abandon, or temporarily close a well. Each activity requires some non-recoverable cost. The model, which allows testing if sunk costs and price uncertainty are determinants of the investment choice, is used to explain the supply response of oil producers to fluctuations in the market price of oil. The model is estimated using a simulation estimator that allows to control for persistent time-invariant firm effects and serial correlation in profit shocks. Using a micro-level data set for oil wells in Oklahoma my estimates for a preliminary model in which firms make the decision to produce or not produce, show that sunk costs are indeed an important determinant of production status. Larger sunk costs decrease the probability of a well entering production, but increase the probability of an existing well remaining in production.

Correspondence: 603 Kern Graduate Bldg., Department of Economics, University Park, PA 1802. Email: bmolls@psu.edu. Office: (814)861-7744. Fax: (814)863-4775.

I am grateful to Susumu Imai, James Tybout, and Tor Winston for helpful discussions. Special thanks go to my dissertation advisor Mark Roberts for invaluable guidance, support, and advice. The help of Tim Dunne and Sarkeys Energy Center at the University of Oklahoma for providing the data are greatly appreciated. This research was supported by NSF grant #9631766.

1. Introduction

Within just the last two years the price of a barrel of crude oil has varied from a low of $11 to more than $35 and this has had a substantial impact on the profitability of individual U.S. oil wells. At the high prices, even the least productive wells are quite profitable, yet oil producers are reluctant to drill new wells.[1] During the periods of low prices, many wells cannot cover their variable costs of production, yet oil companies are slow to abandon these unprofitable wells. In general, large numbers of well openings are only observed when the price of oil is very high and large numbers of well closing are only observed when the price of oil is very low. Between the two prices that trigger entry and exit is a band of low well drilling or plugging activity. This pattern, in which domestic oil production does not respond quickly or smoothly to changes in the world price of oil, is important for understanding the domestic crude oil supply.

Recent economic theory has found that the combination of sunk costs and uncertainty in future market conditions is capable of generating the sluggish or lumpy patterns of investment observed in many industries. Dixit (1989) analyses a firm's entry and exit decision when future output prices are uncertain and market entry and exit require sunk costs. The optimal investment strategy is a pair of entry and exit trigger prices, where the exit threshold is lower than the entry threshold. There is a band of inaction, referred to as the hysteresis band, between the two thresholds in which price movements do not lead to any investment activity by the firm. The entry trigger price is increasing, and the exit trigger price is decreasing, in sunk costs and uncertainty. Intuitively, uncertainty gives firms an incentive to delay investment and wait for new information to arrive before sinking entry and exit costs. This model can be extended in various ways, for example to a model of entry and exit with temporary shutdown or mothballing, as in Dixit and Pindyck (1994).

In the oil industry, where there can be substantial uncertainty of future market prices and significant sunk costs of investment in new wells, there are incentives for firms to delay profitable investments, such as opening new wells, when oil prices rise.

The same forces give firms an incentive to continue unprofitable activities, such as keeping wells open in periods of low prices, in the expectation that future market conditions will improve.

In order to understand the supply response of the domestic U.S oil industry to fluctuations in market price, this paper develops an empirical, discrete choice model of a firm's decision to produce oil, drill, abandon, or mothball oil wells. Current and expected future profits of the firm depend on the current investment decision so that sunk costs and uncertainty affect the firm's investment choices. Using Bellman's equation technique the model illustrates under which conditions the empirically observable choices are made and derives an empirical estimation equation.

The model, which allows me to test if sunk costs and price uncertainty are determinants of the investment choice, is estimated using a micro-level data set for oil wells and oil leases in Oklahoma. This very unique data set allows me to see which firm's execute their investment opportunity and which ones do not. In addition, the data allows me to distinguish between different levels of sunk cost. I use the empirical model to explain the supply response of oil producers to fluctuations in the market price of oil.

The discrete choice model of investment is estimated with a dynamic multinomial probit model. Using a simulation estimator, developed by Keane (1994), I estimate the model allowing for correlation between the firm's investment and production choices, a firm level profit shock, and serial correlation in profit shocks. My estimates show that the probability of being a producer is largely affected by the production status in the previous period. Controlling for other possible sources of profit persistence, this phenomenon is attributed to the importance of sunk costs.

2. Entry, Exit, and Oil Production in Oklahoma and Texas

Out of a total of 573,504 producing wells in operation in the U.S. in 1997, 436,084 oil wells were classified as stripper oil wells; wells that produce less than

---

[1] The highest estimate, for the year 1998, of operating cost per barrel I was able to find is $15.35. See "*Assessing the cost of Environmental Compliances*", by the Interstate Oil and Gas Compact Commission, 1999.

10bbls/day.[2] Although these wells produced, on average, only 2.07 bbls/day in 1998 their total combined production accounted for more than 316 million bbls, or 26.6% of the oil produced in the continental U.S. in the same year.[3] Understanding the investment and production patterns of these small wells is important for several reasons. During periods of low prices the high variable cost stripper wells are in jeopardy of being plugged and abandoned. Because of their high variable costs of production and the substantial fixed re-opening costs, stripper wells are unlikely to be reopened once plugged and, thus, oil reserves are lost. In order to develop sound policies that prevent the closing of wells during periods of low oil prices and foster the opening of wells during times of high prices it is important to understand the determinants of well abandonment.

In 1997 in Oklahoma, the state that will be examined in this paper, 67,498 out of a total of 88,144 wells, were stripper wells and had an average production of 2.05 bbls/day.[4] Estimates of the average operating cost of oil production for marginal wells in Oklahoma range between $9.64/bbl and $15.34/bbl.[5] Especially in the years 1986 and 1998, when Oklahoma well-head crude oil prices were in the range of $10 to $12, marginal oil producers came under severe economic pressure to abandon wells and indeed large numbers of well closings were observed.

Plot 1 through *Plot 3* shows the patterns of entry and exit of wells for the states of Oklahoma and Texas, where much of the U.S. stripper well production is concentrated.[6] The plots graph the numbers of entries and exits in the states of Oklahoma and Texas, and the average annual real oil price for the years 1971 to 1997.[7] *Plot 1* are graphs of the

---

[2] *API Basic Petroleum Data Book* (1999).

[3] The percentage of 26.6% refers only to continental U.S. oil production. Excluded are Alaska, Florida, and federal offshore, which have no stripper well production. This information is from "*Marginal Oil and Gas: Fuel for Economic Growth*," Interstate Oil and Gas Compact Commission, 1999. Note, that in total the U.S produce 2,354 million bbls of oil in 1997, of which 15% were produced by stripper wells. See *Api Basic Petroleum Data Book* (1999).

[4] See *API Basic Petroleum Data Book* (1999).

[5] See *API Basic Petroleum Data Book* for the lower estimate, $9.64/bbl. $15.34/bbl from "*Assessing the Cost of Environmental Compliance*," by the Interstate Oil and Gas Compact Commission, December 1999.

[6] Total U.S. stripper well production in 1997 was 352.9 million bbs of oil. Texas and Oklahoma, the largest and second largest stripper oil producers, accounted for 33.4%and 15.7%, respectively, of that.

[7] The well time series are from the *API Basic Petroleum Data Book* (1999): Section III, Table 10 (United States Total Wells Reported as Completed by State); Section III, Table 13 (Producing Oil Wells in the United States by State); Section III, Table 15 (United States Stripper Oil Wells by State); and Section III, Table 16 (United States Stripper Oil Wells Abandonments by State).

number of well entries in Oklahoma and Texas and the real price of oil. *Plot 2* displays stripper well exits in these two states and the real price of oil. Finally, *Plot 3* displays the total number of wells and the total number of stripper wells in Oklahoma and Texas. The first graph within a row plots data for Oklahoma. Looking at *Plot 1*, the number of well completions and the price of oil are highly correlated and exhibit the same pattern over time. In both states the number of well completions is the highest when the oil price is the highest and the number of well completions is lowest when the oil price is low. *Plot 2* displays the number of abandoned stripper wells and the price of oil over time. Again, the price of oil is plotted as the solid line. The number of stripper wells abandoned and the oil price are inversely related and we observe the lowest number of exits in the year when the price of oil peaks. The first two panels indicate clearly that entry and exit respond to changes in the oil price, although the exit pattern is less closely related to the price than the entry pattern is. Nevertheless, the price of oil appears to be an important factor in the decision to drill or to abandon an oil well.

The final two graphs, Plot 3, show the total number of producing wells and the number of stripper wells over time in Oklahoma and Texas. The two time series evolve similarly. This indicates that the movements over time in the total number of wells are mainly driven by entries and exits of stripper wells. Moreover, we learn from the last two graphs that the number wells increases during the time of high oil prices, however, it does not fall back to its initial level when the price falls below the level that triggered the drilling of new wells. The numbers of producing wells and stripper wells in Oklahoma and Texas are always higher after the second oil price shock than before, although the real oil prices in 1993, 1994 and 1995 are lower than in the years before the second oil price shock during which the price of oil more than doubled.

This pattern of asymmetric adjustment with respect to price shocks is an example of hysteresis. The pattern is particularly strong when we take into account that oil well production rates generally decline over time and that the variable costs of production increase over time. The plots provide evidence that the number of wells in production is

---

The average annual real oil price is calculated as the annual average of the monthly spot oil price for West Texas Intermediate, Dollar per Barrel, deflated by the seasonally adjusted producer price index for intermediate energy goods as calculated by the U.S. Department of Labor, Bureau of Labor Statistics. The data source for both time series is the Federal Reserve Economic Database, St. Louis.

subject to hysteresis and suggest that sunk costs and uncertainty may be important determinants of the timing of investment.

In summary the time-series patterns for the number of wells, entries, and exits in Oklahoma and Texas show:

- There is a strong positive relation between the number of entries and the oil price.
- Stripper well exits and the oil price are inversely related.
- The total number of wells does not fall back to its initial level when the price plunges below the entry triggering level.
- The time series evolve similarly in Oklahoma and Texas.

Oil production in Oklahoma is an ideal test case for the theory of "Investment under Uncertainty." The many small producers in Oklahoma are price takers and strategic considerations among producers do not play a role in the production decision. As needed for testing the theory, there were large fluctuations in oil prices and the costs of drilling an oil well are sunk. Most wells in Oklahoma are only marginally profitable; for many wells the variable costs of production are between the lowest and highest oil price observed over the period under investigation. If the theory is correct we should only observe a significant increase in the number of well drillings when oil prices are substantially higher than the variable production costs. Similar, significant well closing should only be observed at oil prices substantially lower than the variable production costs.

In the next section I develop a theoretical model of investment for stripper well producers.

3. A Theoretical Model of Investment, Mothballing and Shutdown

3.1 An Overview of Investment Under Uncertainty

Most investment decisions have three underlying features. First, investment is not reversible and costs cannot be recovered. Second, the environment in which the investment is undertaken changes over time. Investment involves risk and new information arrives as time passes. Third, the decision does not only invoke the question whether investment should be undertaken or not, it also involves the aspect of when to invest. Generally an investment opportunity does not vanish if not taken at once but can be postponed. Sunk costs, uncertainty, and timely persistence together give value to an investment opportunity itself and one can view an investment opportunity as an option to invest. The holder of the investment opportunity gives up the value of the option and spends the sunk costs to receive an investment project with an uncertain future profit stream. Not exercising the investment option helps to avoid possible bad future states of the world.

Traditional investment models are not able to explain the sluggish and lumpy patterns of investment and abandoning described above. The recent theoretical work in the field of investment has focused on the importance of sunk costs and uncertainty for investment decisions.[8] Within the framework of "Investment under Uncertainty" investing is viewed as exercising a call option on an investment opportunity, where the owner of the investment opportunity has to decide when to exercise, e.g. invest. In these continuous time models, firms invest when the output price or the value of the investment strikes an upper entry threshold and scrap when it hits a lower exit threshold. Higher uncertainty or larger sunk costs, everything else being equal, raise the entry threshold and lower the exit threshold.

Analyzing real investment based on the options approach is relatively new in economics. In contrast to the real options view, earlier models of investment, for example neoclassical models, rely in some form on the net present value rule. A firm should invest if the net present value (NPV) of investment is positive. Using the NPV rule implies that firms have static expectations. If the demand (price) evolution follows a stochastic process, demand (price) changes come as a surprise to firms. Their behavior is irrational. The options approach on the other hand assumes that firms take stochastic demand (price) changes into account and make rational decisions based on the stochastic process governing the demand (price) evolution. Generally, the options approach adds rational behavior to the firm's decision.

---

[8] See Dixit & Pindyck (1994) for a very comprehensive discussion.

Another assumption of many older investment models is that investment is largely reversible and investment opportunities are once-and-for-all. Since this assumption implies that sunk costs can be recovered they do not matter for the investment decision. Because the investment opportunity vanishes there is no intrinsic option value to it. Further, again assuming away the importance of sunk costs, some models assume that investment is divisible (in combination with convex adjustment costs), which results in firms making marginal investment decisions. Contrary to this assumption compelling evidence has been presented that speaks in favor of lumpy investment[9]. The real options model is a model of lumpy investment. From an empirical point of view it is important to recognize that no matter how sophisticated the older models are, econometric models based on them have had limited success in explaining investment behavior and have done poorly in forecasting changes in investment spending.[10]

While traditional investment models have largely failed in forecasting investment, empirical verification of the theory of investment under uncertainty has been very limited. Paddock et al. (1988) evaluate and compare oil investments using the net present value of an oil lease and the lease's value using the options approach to investment. Comparing, but not statistically testing the theory, they find the latter estimate is a better predictor of the winning industry bid in an off-shore oil field auction. Using market level data, Campa (1993) finds that the number of foreign firms entering the U.S. market is positively correlated with the exchange rate and negatively correlated with the exchange rate volatility and measures of entry sunk costs. Hurn and Wright (1994) develop a hazard rate model to analyze the sequential investment for 108 North Sea oil rigs, and find that higher oil prices "speed up" investment. However, in their study oil price variance is an insignificant investment determinant. Roberts and Tybout (1997) develop an empirical discrete choice model and examine the entry and exit decision of Columbian manufacturers into the export market. They find that sunk costs are an important entry and exit determinant. Nevertheless, they do not examine how much investment is affected by different levels of uncertainty and sunk costs.

---

[9] See for example Doms and Dunne (1998), and Cooper et al. (1999).

[10] See Chirinko (1993)

3.2 Opening and Closing, Production or Non-Production of Oil Wells

Below I develop a dynamic, discrete choice model of well drilling and plugging, and oil production that is based on the new theories of investment under uncertainty.[11] Using data for oil producers in Oklahoma the model is used to explain the observed their investment and production behavior. Moreover, taking the theoretical model to the data I examine some empirical implication of the theory. Do sunk cost matter for investment and production decisions? Does higher uncertainty increase the importance of sunk costs? The unit of observation in my data set is an oil lease; a geographically confined area for which the owner of the lease has the rights to extract the oil that lies underneath the surface. Oil on a lease is produced from a well, the discrete unit of production capital. The opening of a well requires non-recoverable investment cost, such as the costs of drilling the borehole. The production capacity of a lease depends on the number of wells in lieu, which can be increased by drilling a new well or decreased by shutting down an existing one. Shutdown is accomplished by filling (plugging) the well with concrete. Since re-opening of a shutdown well is very costly, wells may be temporarily taken out of production, or mothballed, without plugging. However, there are variable costs of mothballing and there is a legal maximum time length, generally one year, for mothballing. After this time the well has to go back into production or it must be plugged.

I start by specifying the profit maximization problem for a lease. From that I derive an empirical model to explain the discrete production stages for a lease. The model can be used to estimate the importance of the oil price level, sunk costs, and uncertainty for the investment and production decisions on the lease.

The data set used in this research includes the amount of oil produced from a lease during a six-month interval and the number of wells drilled or plugged on the lease during that time period. The discrete choice model developed in this section allows the firm to undertake one of five actions in each period. The owner of an oil lease chooses among 1) producing with its existing capacity, 2) producing and increasing production capacity, 3) producing and decreasing production capacity, 4) temporarily shutting down

---

[11] See Dixit & Pindyck (1994), Chapter 7, for a continuous time version of a related investment problem with the decisions entry/exit and mothball/produce.

by mothballing all wells on the lease, or 5) permanently shutting down by plugging all wells on the lease.

Consider an active, producing lease i and define $\Pi_{it}(C_{it}, P_t, X_{it})$ as the period t production profits from the lease. Current profits at time t are a function of market conditions P_t, the characteristics of the lease X_{it}, which measure lease specific cost and demand factors, and the capital stock (number of wells) on the lease C_{it}. The firm can adjust its capital by changing the number of wells on a lease. Increasing the capital stock by one unit requires the expenditure of the sunk investment cost I_i. The net cost of scrapping a well (scrap value less scrapping costs) is SC_t.

Switching the entire lease from producing to non-producing entails sunk fixed costs and these are summarized in Table 1. The costs of switching from producing to mothballing are M_i. For a mothballed lease there are variable costs of mothballing, which are vm_i per period. A mothballed lease can be switched back to production or to shutdown permanently. The cost to switch from mothballed to shutdown is S_i. A producer who has mothballed a lease can re-enter production by paying the fixed re-entry costs RM_i. Finally, a producer can skip the mothballing stage and move directly between production and shutdown. A producer re-entering from shutdown has to sink (RM_i+RS_i), while a production move from production to shutdown incurs cost (M_i+S_i). Accordingly, moving from mothballing to shutdown cost S_i and moving from shutdown to mothballing cost RS_i.

The actual period t profits realized by the lease in each of the five possible states will depend on profits from production and the production status in $t\text{-}1$:

$$
\begin{aligned}
&\text{produce with existing capacity}: \Pi_{it}(C_{it}, P_t, X_{it}) - m_{it-1}RM_i - s_{it-1}(RM_i + RS_i)\\
&\text{produce and increase capacity}: \Pi_{it}(C_{it}, P_t, X_{it}) - I_i - m_{it-1}RM_i - s_{it-1}(RM_i + RS_i)\\
&\text{produce and decrease capacity}: \Pi_{it}(C_{it}, P_t, X_{it}) + SC_i - m_{it-1}RM_i - s_{it-1}(RM_i + RS_i)\\
&\text{mothball the lease}: -vm_i - (1 - m_{it-1})M_i - s_{it-1}(RS_i - M_i)\\
&\text{shutdown the lease}: -(1 - s_{it-1})(M_i + S_i) + m_{it-1}M_i.
\end{aligned}
$$

The indicator variable m_{it-1} takes the value 1 if the lease is mothballed in period $t\text{-}1$, zero otherwise, and the indicator variable s_{it-1} takes the value 1 if the lease is shutdown in period *t-1*, zero otherwise. The first three equations are straightforward. The lease generates profits from production, Π_{it}, net of the sunk capacity adjustment costs, I_i and SC_t. In addition, if the lease was non-producing in the previous period, the appropriate re-entry costs are RM_i if returning from mothballing and (RM_i+RS_i) if returning from shutdown. The last two equations describe current profits from non-production. They entail variable costs, vm_{it}, only if mothballed. However, there are sunk costs for changing the production status, M_i for changing from production to mothballed and M_i+S_i for changing from to production to shutdown. The terms $s_{it-1}(RS_i-M_i)$ and $m_{it-1}M_i$ in these two equations adjust the switching cost such that switching from shutdown to mothballed cost RS_i and switching from mothballed to shutdown costs S_i.[12]

Let x_{it} and y_{it} be indicator variables that take the value 1 for "invest in a new well" and "scrap an existing well," respectively. The current profit equations can be nested as:

$$\begin{aligned} R_{it}(\bullet) = {} & (1-m_{it})(1-s_{it})* \\ & \{\Pi_{it}(C_{it}, P_t, X_{it}) - x_{it} I_i + y_{it} SC_i - m_{it-1} RM_i - s_{it-1}(RM_i + RS_i\} \\ & - m_{it}\{vm_i + (1-m_{it-1})M_i - s_{it-1}(RS_i - M_i)\} \\ & - s_{it}\{(1-s_{it-1})(S_i + M_i) - m_{it-1}M_i\}. \end{aligned} \tag{1}$$

Capital adjustments are assumed to be instantaneous. A firm can add or subtract one well each period.[13] Consequently, the number of wells on a lease evolves over time according to

$$C_{it} = C_{it-1} + x_{it} - y_{it}. \tag{2}$$

Define $Y_{it} \equiv (x_{it}, y_{it}, m_{it}, s_{it})'$. At most one of the dummy variables can take the value 1 for firm *i* in period *t*. The vector Y_{it} takes the values

---

[12] The re-entry costs from mothballing are RM_i and re-entry costs from shutdown are RM_i+RS_i. Note, that the mothballing equation is set up such that it will never be profitable to switch from shutdown to mothballing.

[13] Considering the data, this is a reasonable assumption. On an annual level, 75% of the leases that increase capacity add just one well. Of the leases that close one or more wells, 75% of these closings were of a single well.

$$Y_{it} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, or \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

In period *t* the lease owner chooses the infinite sequence $Y_{it}^{(+)} = \{(x_{i,t+j}, y_{i,t+j}, m_{i,t+j}, s_{i,t+j})' \mid j \geq 0\}$ that maximizes the expected present value of payoffs. The maximized expected present value is represented by

$$V_{it}(\Omega_{it}) = \max_{Y_{it}^{(+)}} E_t \left(\sum_{j=t}^{\infty} \beta^{j-t} R_{ij}(C_{ij}, P_j, X_{ij}, Y_{ij}) \middle| \Omega_{it} \right), \tag{3}$$

where β it the one period discount rate and expectations are conditioned on the lease specific information set Ω_{it}. Using Bellman's equation the value function in period *t* is

$$V_{it}(\Omega_{it}) = \max_{Y_{it}} \left(R_{it}(C_{ij}, P_j, X_{ij},, Y_{it}) + \beta E_t \left[V_{i,t+1}(\Omega_{i,t+1}) \mid Y_{it} \quad \right] \right). \tag{4}$$

The values of the Bellman equation corresponding to the five possible choices for Y_{it} are

produce with existing capacity:

$$\begin{aligned} Z_{it}^{*1} &= \Pi_{it}(C_{it-1}, P_t, X_{it}) - m_{it-1}RM_i - s_{it-1}(RM_i + RS_i) \\ &\quad + \beta E_t\left[V_{i,t+1}(\Omega_{i,t+1}) | C_{it}, Y_{it} = \boldsymbol{0}\right] \end{aligned}$$

produce and increase capacity:

$$\begin{aligned} Z_{it}^{*2} &= \Pi_{it}(C_{it-1} + 1, P_t, X_{it}) - I_i - m_{it-1}RM_i - s_{it-1}(RM_i + RS_i) \\ &\quad + \beta E_t\left[V_{i,t+1}(\Omega_{i,t+1}) | C_{it}, x_{it} = 1\right] \end{aligned}$$

produce and decrease capacity:

$$\begin{aligned} Z_{it}^{*3} &= \Pi_{it}(C_{it-1} - 1, P_t, X_{it}) + S_i - m_{it-1}RM_i - s_{it-1}(RM_i + RS_i) \\ &\quad + \beta E_t\left[V_{i,t+1}(\Omega_{i,t+1}) | C_{it}, y_{it} = 1\right] \end{aligned}$$

mothball the lease:

$$\begin{aligned} Z_{it}^{*4} &= -vm_i - (1 - m_{it-1})M_i - s_{it-1}(RS_i - M_i) \\ &\qquad\qquad + \beta E_t\left[V_{i,t+1}(\Omega_{i,t+1}) \mid m_{it} = 1\right] \end{aligned}$$

shutdown the lease:

$$\begin{aligned} Z_{it}^{*5} &= -(1 - s_{it})(S_i + M_i) + (1 - s_{it-1})m_{it-1}M_i \\ &\qquad\qquad + \beta E_t\left[V_{i,t+1}(\Omega_{i,t+1}) \mid s_{it} = 1\right]. \end{aligned} \tag{5}$$

A lease owner will chose the action with the largest current expected value. Thus choice j is observed in period t if $Z_{it}^{*j} > Z_{it}^{*k}\ \forall\, k \neq j$. Note, that subtracting one choice from the other choices would not change anything in the relative profit ordering. Thus, since only the actual choice made in period t is observed, this set of equations is empirically not identified. To generate a profit relation that can be estimated I subtract the first category, "produce with existing capacity," from the other choices.

To facilitate the algebraic expressions, define the difference in current profits between "produce and increase capacity" or "produce and decrease capacity" and "produce with existing capacity" as

$$\begin{aligned} &\Delta\Pi_{it}(x_{it} = 1) \equiv \Pi_{it}(C_{it-1} + 1,, P_t, X_{it}) - \Pi_{it}(C_{it-1}, P_t, X_{it}) \text{ and} \\ &\Delta\Pi_{it}(y_{it} = 1) \equiv \Pi_{it}(C_{it-1} - 1, P_t, X_{it}) - \Pi_{it}(C_{it-1}, P_t, X_{it}), \end{aligned} \tag{6}$$

and define the current expected future lease value given choice j=2,3,4, or 5 minus the value if the choice is "produce with existing capacity" as

$$
\begin{aligned}
\Delta V_{i,t+1}^{e}(x_{it}=1) &\equiv \beta E_t\left[V_{i,t+1}(\Omega_{i,t+1})\,|\,Y_{it}=(1,0,0,0)'\right]-\beta E_t\left[V_{i,t+1}(\Omega_{i,t+1})\,|\,Y_{it}=(0,0,0,0)'\right],\\
\Delta V_{i,t+1}^{e}(y_{it}=1) &\equiv \beta E_t\left[V_{i,t+1}(\Omega_{i,t+1})\,|\,Y_{it}=(0,1,0,0)'\right]-\beta E_t\left[V_{i,t+1}(\Omega_{i,t+1})\,|\,Y_{it}=(0,0,0,0)'\right],\\
\Delta V_{i,t+1}^{e}(m_{it}=1) &\equiv \beta E_t\left[V_{i,t+1}(\Omega_{i,t+1})\,|\,Y_{it}=(0,0,1,0)'\right]-\beta E_t\left[V_{i,t+1}(\Omega_{i,t+1})\,|\,Y_{it}=(0,0,0,0)'\right],\\
\Delta V_{i,t+1}^{e}(s_{it}=1) &\equiv \beta E_t\left[V_{i,t+1}(\Omega_{i,t+1})\,|\,Y_{it}=(0,0,0,1)'\right]-\beta E_t\left[V_{i,t+1}(\Omega_{i,t+1})\,|\,Y_{it}=(0,0,0,0)'\right].
\end{aligned}
\tag{7}
$$

Rearranging terms and recognizing that m_{it-1}=1 implies that (1-s_{it-1})=1 the values, normalized to choice 1, are

$$
\begin{aligned}
\text{produce with existing capacity}: Z_{it}^{1} &= 0\\
\text{produce and increase capacity}: Z_{it}^{2} &= -I_i + \Delta\Pi_{it}(x_{it}=1) + \beta\Delta V_{i,t+1}^{e}(x_{it}=1)\\
\text{produce and decrease capacity}: Z_{it}^{3} &= SC_i + \Delta\Pi_{it}(y_{it}=1) + \beta\Delta V_{i,t+1}^{e}(y_{it}=1)\\
\text{mothball the lease}: Z_{it}^{4} &= -vm_i - M_i - \Pi_{it} + \beta\Delta V_{i,t+1}^{e}(m_{it}=1)\\
&\quad + m_{it-1}(M_i + RM_i) + s_{it-1}(M_i + RM_i)\\
\text{shutdown the lease}: Z_{it}^{5} &= -(S_i + M_i) - \Pi_{it} + \beta\Delta V_{i,t+1}^{e}(s_{it}=1)\\
&\quad + m_{it-1}(M_i + RM_i)\\
&\quad + s_{it-1}(S_i + M_i + RM_i + RS_i).
\end{aligned}
\tag{8}
$$

This system of equations takes the form of a multinomial discrete choice model and is the basis for the empirical estimates in this paper.

As Roberts and Tybout (1997) show for a binary investment choice model in which a firm decides to export or not to export, I can test for sunk costs by testing for state dependence. Note, that in the presence of sunk costs, namely M_i>0, S_i>0, RM_i>0, and RS_i>0, the lagged production state, indicated by m_{it-1} and s_{it-1}, matters for the current choice of production and investment. By estimating the above set of equations and testing for state dependence, in particular testing whether coefficients on the lagged state dummy variables are significant and positive, I can test whether sunk costs matter for production decisions in the oil industry.

What if there are no sunk costs involved in moving from one state to another and increasing or decreasing production capacity, that is M_i=S_i=RM_i=RS_i=I_i=SC_i=0?[14] In this case changes in expected future values vanish ($\Delta V_{i,t+1}^{e}=0$) and the investment decision collapses to a purely static one. Because changes in the parameters governing the oil price evolution, and thus changes in expected future oil prices, have implications for

---

[14] Note, with this sunk cost specification the mothballing state will not be used at all.

investment decisions only through the $\Delta V_{i,t+1}^{e}$ terms, changes in these parameters matter only if sunk costs are non zero. In particular, changes in oil price volatility will affect the option value of investment and the investment decision only if sunk costs matter. Thus, including oil price volatility in the estimation and testing for its significance is an indirect way to test whether sunk costs matter.

A general theoretical finding is that the investment option value increases with uncertainty. For example, the choices j=2 and j=3 can be interpreted as exercising the investment option on adjusting the capital stock, which implies for my model that the term $E_t\left[V_{i,t+1}(\Omega_{i,t+1})|Y_{it}=0\right]$, the value if no change in lease status takes place, is increasing in uncertainty. Hence, the probability of observing choices j=2 or 3 in the data should be decreasing in uncertainty. Interpretation of the investment opportunity is different for choice j=4 and 5. Here the value of the opportunity depends on the lagged production status of the lease. A non-producing lease has the opportunity to return to production and vice versa. Consequently, the probability of being mothballed or shutdown will be increasing in uncertainty if the lease wasn't producing in the previous period and deceasing in uncertainty if the lease was producing in the last period.

4. Data

Before discussing the empirical model and presenting results, I discuss the data and provide some aggregate statistics on the patterns of investment and production. Regarding the variables used in the analytical model, for each lease I observe the choice made in period t, Y_{it}, and the lagged choices, including s_{it-1} and m_{it-1}, which indicate whether the lease was shutdown or mothballed in the previous period. In addition I observe a set of lease characteristics X_{it}, the number of wells on a lease C_{it}, and the market level variables P_t, oil price and oil price volatility.

4.1 The Data Set

With the cooperation of the Sarkeys Energy Center at the University of Oklahoma I have gained access to two data sets containing the investment and production records of virtually all oil and gas wells in Oklahoma. A well-level data set provides the complete investment history of individual wells, with first observations in the early 1900's. As oil producers generally do not own the land on which their wells are located, they lease the right to produce oil from it. Monthly lease production rates for the years 1979 to 1999 are recorded in a lease data set.

The well data set provides information on the dates that define each well's investment history, so that I am able to construct annual time-series that indicate when the well was (re-)drilled, which years it was in operation, and, if applicable, in which year the well was officially plugged. This data set provides information on a well's characteristics. The lease data set provides information on the geographic size of the lease, and the monthly oil and gas production rates, which I aggregate to the annual level.

To date I have been working with a small non-representative subset of leases and wells. The subset focuses on a region that has experienced higher than average drilling activity in the past 20 years. The investment and production data for this region has been organized into an unbalanced panel data set at the annual level.[15] I use this panel data set for estimation. It describes the well opening and closing, and production histories of 1144 leases for a total of 19126 observations over the period 1980 to 1997. It indicates the number of open wells, the lease's age, the average well depth, the amount of oil that was produced, the current investment and production choice, and the past production status for each lease in each year.[16] Finally, I add market level variables: the logarithm of the average, annual, real price of oil and its annual standard deviation. I use monthly spot price data for West Texas intermediate [$/bbl] and discount it with the seasonally adjusted producer price index for intermediate energy goods. After taking the natural

---

[15] Only wells that can be matched to a lease and vice versa are used for constructing the data set. If there is no well data for a lease I do not know the lease's investment history. For wells without matching lease data I do not know if a lease stopped producing before 1979 or if lease data is missing. Not using these observations implies that the data contains only leases that produce at some point during the years 1979 to 1997. Thus, there are no wells in the data set that were plugged prior to 1979 and do not come back to production. This implies that the in-sample probability of re-entry by shutdown leases is higher than it is for the population of all leases. However, if my estimates find that the lagged shutdown status is significant, then it would be even more significant if the same model were estimated with the true population of leases and their respective production rates.

[16] The first year with production data, 1979, is used to create the lagged production status dummy variable.

logarithm the annual means are compute and the standard deviation from the mean in each year[17]

4.2 Patterns of Investment and Production in the Data

The analytical model of investment and production predicts that the presence of sunk costs and uncertainty creates state dependence in the lease production history. For example, a lease that currently produces has already sunk some costs for being in the producing state and, thus, is more likely to produce next period. The transition rates displayed in Table 1 provide first evidence that this conjecture is correct. Each cell describes the conditional percentage of transitions observed in the data, e.g. the percentage of leases that move from the status on the left hand side to the status in a given column. For instance, 87.5% of all leases that produce and decrease their capacity in year *t* continue to produce in the following year (94.5% if there is no change in the number of wells, and 98% if there is an increase in the number of wells on the lease). The shutdown category displays persistence, too. 91% of the leases that are shutdown in year *t* remain shutdown in year *t+1*. Mothballing, the intermediate state, does not create as much state persistence. Nevertheless, the probability of remaining in the mothballed category for another period is still 44%. Leases in the category "produce and decrease capacity" have a higher transition rate to the categories "mothballed" or "shutdown" than lease that "produce and don't change capacity" and lease that "produce and increase capacity." Transition patterns of this type can also be explained with profits shocks that are correlated over time and this will be recognized in the econometric model developed in the next question.[18]

The analytical model is set up so that switches from any producing state to any non-producing state, and vice versa, involve sunk costs. Thus the transition rates between the states "production" and "no production" for a given year should display state

---

[17] I use monthly spot price data for West Texas intermediate [$/bbl of oil] as provided by the Federal Reserve Bank of St. Louis and discount it with the seasonally adjusted producer price index for intermediate energy goods (base year=1982), provided by the same data source.

[18] To make sure that serially correlated profit shocks do not show up as sunk costs, e.g. significant coefficients on lagged "mothballed" and "shutdown" dummy variables, I employ an estimation strategy that allows for auto-correlation in the error terms.

persistence, too.[19] These transition rates are summarized in Table 2. The third and forth columns in Table 2 are the probabilities of staying in the "production" state and switching from "production" to "non production," respectively. Column five displays the probability of switching from "no production" to "production," and column six contains the probability of being in the state "no production" for another period. Providing further evidence for the importance of sunk costs, the probabilities of staying in the lagged production status are very high. In any year at least 90% of the leases in operation continue producing in the following year. Considering non-producing leases, apart from 1988, at least 80% of the leases that do not produce remain out of production in the following year.

Table 3 displays the annual transition rates between the categories "production," "mothballed," and "shutdown." The sunk costs involved with switching between the categories "production," "mothballed," and "shutdown" depend on the lagged state. For example switching from "production" to "mothballed" is less costly than switching from "production" to "shutdown." Consistent with this, producing leases that stop producing are more likely to mothball than to shutdown. Moreover, a large percentage of mothballed leases switch to either "production" or "shutdown." Again, as expected, the lagged states "production" and "shutdown" generate a lot of persistence. In most years, about 90% of the leases that were in "production" or "shutdown" do not change their status. It will be important in the empirical model to control for all sources of profit persistence that can also lead to persistence in a lease's production status over time.

5. The Empirical Model

This section discusses the estimation strategy and derives the estimation equation for the analytical model developed in section 3.2. Without making specific functional form assumptions, the model outlined in equation (8) identifies the conditions that determine which investment and production category the lease is in.

---

[19] The new category "production" is defined as the union of "produce and don't change capacity," "produce and increase capacity," and "produce and decrease capacity." The category "no production" is defined as the union of "mothballed" and "shutdown."

In the choice model, equation (8), s_{it-1} and m_{it-1} enter the choices directly. C_{it}, X_{it}, and P_t enter the model through the current production profit function Π_{it}, and the expected, next period value function $V^e_{i,t+1}$. Instead of assuming a functional form for current production profits and the value function I assume that they are sufficiently approximated using a reduced from expression. Additionally, there are unobservable sunk and variable cost terms. In an empirical discrete choice model the sunk costs terms in equation (8) that are not attached to m_{it-1} and s_{it-1} can be interpreted as intercepts. Subsequently, for the estimation of the model I specify that the variations in the latent profit differences across firms and time are due to changes in the macro level variables $\boldsymbol{P_t}$, the lease characteristics $\boldsymbol{X_{it}}$, and some random noise ε_{jit}. Specifically I assume that the unobservable profits and value functions of Z^2_{it} to Z^5_{it} can be expressed as:

$$\begin{aligned}
\text{in } Z^2_{it} &: -I_i + \Delta\Pi_{it}(x_{it}=1) + \beta\Delta V^e_{i,t+1}(x_{it}=1) \\
&= \alpha_2 + \boldsymbol{\beta}_2\boldsymbol{P}_t + \boldsymbol{\delta}_2\boldsymbol{X_{it}} + \varepsilon_{2it} \\
\text{in } Z^3_{it} &: SC_i + \Delta\Pi_{it}(y_{it}=1) + \beta\Delta V^e_{i,t+1}(y_{it}=1) \\
&= \alpha_3 + \boldsymbol{\beta}_3\boldsymbol{P}_t + \boldsymbol{\delta}_3\boldsymbol{X_{it}} + \varepsilon_{3it} \\
\text{in } Z^4_{it} &: -vm_i - M_i - \Pi_{it} + \beta\Delta V^e_{i,t+1}(m_{it}=1) \\
&= \alpha_4 + \boldsymbol{\beta}_4\boldsymbol{P}_t + \boldsymbol{\delta}_4\boldsymbol{X_{it}} + \varepsilon_{4it} \\
\text{in } Z^5_{it} &: -(S_i + M_i) - \Pi_{it} + \beta\Delta V^e_{i,t+1}(s_{it}=1) \\
&= \alpha_5 + \boldsymbol{\beta}_5\boldsymbol{P}_t + \boldsymbol{\delta}_5\boldsymbol{X_{it}} + \varepsilon_{5it}\ .
\end{aligned} \tag{9}$$

The remaining terms in equation (8) are the lagged discrete production indicators s_{it-1} and m_{it-1} multiplied by sunk costs. Although sunk costs are not observable, I can test for their presence by estimating coefficients on s_{it-1} and m_{it-1}. The specification of sunk costs can be improved upon by recognizing that their importance will vary across leases with differences in the characteristics of the lease. Using the average depth of the wells on a lease as a proxy for sunk costs I combine the lagged dummy variables with this proxy in order to allow sunk costs to be different across leases. Since the hysteresis band increases in sunk costs the coefficient on this variable is expected to be positive.

As pointed out, the theory of investment under uncertainty implies that the options value of investment is increasing in uncertainty. For a producing lease, for example, the options are shutdown and mothballing. For a non-producing lease re-entry is an option. Consequently, uncertainty has a positive effect on the categories “mothball the lease” and “shutdown the lease” if the lease was mothballed or shutdown, respectively, in the previous period, and a negative effect if the lease was producing in the previous period. In order to pick up the right effects of uncertainty on the production decision I combine the lagged dummy variables with the current oil price volatility. Then, this variable and the oil price volatility by itself are included in the regression.

Finally, to estimate the coefficients which are used to test for state dependence and how it is affected by uncertainty and varying degrees of sunk costs, I organize an intercept term and the average well depth on a lease in the vector H_{it} and multiply it with the dummy variables on the lagged production status. The latent variables Z_{it}^{j} that represent the value of choice j are represented:

$$\begin{gathered}
\text{produce and increase capacity:} \\
Z_{it}^{2} = \alpha_2 + \boldsymbol{\beta}_2 \boldsymbol{P}_t + \boldsymbol{\delta}_2 \boldsymbol{X}_{\boldsymbol{it}} + \varepsilon_{2it} \\
\text{produce and decrease capacity:} \\
Z_{it}^{3} = \alpha_3 + \boldsymbol{\beta}_3 \boldsymbol{P}_{\boldsymbol{t}} + \boldsymbol{\delta}_3 \boldsymbol{X}_{\boldsymbol{it}} + \varepsilon_{3it} \\
\text{mothball the lease:} \\
Z_{it}^{4} = \alpha_4 + \boldsymbol{\beta}_4 \boldsymbol{P}_t + \boldsymbol{\delta}_4 \boldsymbol{X}_{\boldsymbol{it}} + \boldsymbol{\gamma}_4 (m_{it-1} \boldsymbol{H}_{\boldsymbol{it}}) + \boldsymbol{\lambda}_4 (s_{it-1} \boldsymbol{H}_{\boldsymbol{it}}) + \varepsilon_{4it} \\
\text{shutdown the lease:} \\
Z_{it}^{5} = \alpha_5 + \boldsymbol{\beta}_5 \boldsymbol{P}_{\boldsymbol{t}} + \boldsymbol{\delta}_5 \boldsymbol{X}_{\boldsymbol{it}} + \boldsymbol{\gamma}_5 (m_{it-1} \boldsymbol{H}_{\boldsymbol{it}}) + \boldsymbol{\lambda}_5 (s_{it-1} \boldsymbol{H}_{\boldsymbol{it}}) + \varepsilon_{5it} .
\end{gathered} \tag{10}$$

In the normalized form, choice j is observed if $Z_{it}^{j} > Z_{it}^{k} \ \forall\ j \neq k$, j,k=2,3,4,5.

5.1 A Special Case: The Decision to Produce

The model developed in the last section allows the firm to choose among five different activities and generates an empirical multinomial choice model. Before estimating the multinomial model, I will specify two simpler models that just distinguish two or three states. How does the model change if we only observe the decision to produce or not to produce? Recall that the nested, current profits in the model with five choices are represented by

$$\begin{aligned} R_{it}(\bullet) = {} & (1-m_{it})(1-s_{it})* \\ & \left\{\mathit{\Pi}_{it}(C_{it}, P_t, X_{it}) - x_{it}I_i + y_{it}SC_i - m_{it-1}RM_i - s_{it-1}(RM_i + RS_i)\right\} \\ & - m_{it}\left\{vm_i + (1-m_{it-1})M_i - s_{it-1}(RS_i - M_i)\right\} \\ & - s_{it}\left\{(1-s_{it-1})(S_i + M_i) - m_{it-1}M_i\right\}. \end{aligned} \tag{11}$$

This equation can be manipulated by assuming that some cost terms are zero. First, focusing in this section on the decision “produce” vs. “don’t produce” I assume that there are no costs of adjusting capital and no difference between mothballing and shut-down, specifically I_{it}=0, SC_{it}=0, vm_{it}=0, S_{it}=0, and RS_{it}=0.

Rearranging terms the static current profits are

$$\begin{aligned} R_{it}(\bullet) = {} & (1-m_{it})(1-s_{it})*\left\{\mathit{\Pi}_{it}(C_{it}, P_t, X_{it}) - (m_{it-1} + s_{it-1})RM_i\right\} \\ & - M_i\left\{(m_{it} + s_{it})(1 - m_{it-1} - s_{it-1})\right\}. \end{aligned} \tag{12}$$

In this case M_i and RM_i represent the sunk cost of moving out production and into production, respectively. Define the production indicator p_{it}=1 if m_{it}=0 and s_{it}=0 and zero otherwise. Equation (12) can be written as:

$$R_{it}(\bullet) = p_{it}*\left\{\mathit{\Pi}_{it}(C_{it}, P_t, X_{it}) - (1-p_{it-1})RM_i\right\} - M_i(1-p_{it})p_{it-1}. \tag{13}$$

Using the same logic as before the lease is producing oil if the current and expected future profits from production are larger than the current and expected future profits from not producing. The oil lease’s values for the choices “produce” and “don’t produce” are

$$\begin{aligned} \text{produce}: Z_{it}^1 & \equiv \mathit{\Pi}_{it}(\bullet) - (1-p_{it-1})RM_{it} + \beta E_t\left[V_{i,t+1}(\mathit{\Omega}_{i,t+1}) | p_{it} = 1\right] \\ \text{don't produce}: Z_{it}^2 & \equiv -p_{it-1}M_{it} + \beta E_t\left[V_{i,t+1}(\mathit{\Omega}_{i,t+1}) | p_{it} = 0\right] \end{aligned} \tag{14}$$

Define the latent variable representing the expected current value increment from producing today as:

$$\mathit{\Pi}_{it}^* \equiv \mathit{\Pi}_{it}(\bullet) + \beta E_t\left[V_{i,t+1}(\mathit{\Omega}_{i,t+1}) | p_{it} = 1\right] - \beta E_t\left[V_{i,t+1}(\mathit{\Omega}_{i,t+1}) | p_{it} = 0\right]. \tag{15}$$

The observable production decision in period *t* can then be summarized as

$$p_{it} = \begin{cases} 1 \text{ if } \quad \mathit{\Pi}_{it}^* - RM_{it} + p_{it-1}(RM_{it} + M_{it}) \geq 0 \\ 0 \text{ otherwise} \end{cases} \tag{16}$$

Interpreting latent profits and sunk costs as in the previous section this can be estimated with the same set of variables as the model with five investment choices

$$p_{it} = \begin{cases} 1 \text{ if } \alpha + \boldsymbol{\beta P}_t + \boldsymbol{\delta X_{it}} + \gamma(p_{it-1}\boldsymbol{H_{it}}) + \varepsilon_{it} \\ 0 \text{ otherwises.} \end{cases} \tag{17}$$

As in the model with five choices, the coefficients of interest are $\boldsymbol{\gamma}$, which are estimates of the magnitude of the sum of the entry and exit costs. Again, I expect these coefficients to be positive.

Unlike in the model with five choices, this model does not differentiate between the two possible types of non-production, and their respective costs, or the possible capital adjustments the producing firm can make. Although some facets of the investment and production decision are lost, estimating the simpler binary problem of produce/don't produce has several advantages. In many data sets it might not be possible to distinguish between mothballing and shutdown and data on changes in capital stock may not be available. Also, apart from data availability, in many cases estimating a binary choice model is much easier econometrically then estimating a multinomial choice model.

5.2 A Special Case: Produce, Mothball or Shutdown

A second simplified model ignores the firm's decision to change the capacity of its lease by investing in new wells or scraping existing wells, but allows for production, mothballing and shutdown decisions. If capital adjustments are unobservable and we can only observe whether a lease produces, is mothballed, or is shutdown the static expression for current profits changes slightly. First, assume that there are no capital adjustments or $I_{it}=SC_{it}=0$. The nested current profits equation is

$$\begin{aligned} R_{it}(\bullet) = {} & (1-m_{it})(1-s_{it}) * \{\Pi_{it}(C_{it}, P_t, X_{it}, Y_{it}) - m_{it-1}RM_i - s_{it-1}(RM_i + RS_i\} \\ & - m_{it}\{vm_i + (1-m_{it-1})M_i - s_{it-1}(RS_i - M_i)\} \\ & - s_{it}\{(1-s_{it-1})(S_i + M_i) - m_{it-1}M_i\}. \end{aligned} \tag{18}$$

Normalizing by the choice "produce" ($m_{it}=0$ and $s_{it}=0$), the differences in expected future values of the three choices are

$$\begin{aligned}
\text{produce}: Z_{it}^{1} &= 0 \\
\text{mothball the lease}: Z_{it}^{4} &= -vm_i - M_i - \Pi_{it} + \beta\Delta V_{i,t+1}^{e}(m_{it}=1) \\
&\quad + m_{it-1}(M_i + RM_i) + s_{it-1}(M_i + RM_i) \\
\text{shutdown the lease}: Z_{it}^{5} &= -(S_i + M_i) - \Pi_{it} + \beta\Delta V_{i,t+1}^{e}(s_{it}=1) \\
&\quad + m_{it-1}(M_i + RM_i) \\
&\quad + s_{it-1}(S_i + M_i + RM_i + RS_i)\,.
\end{aligned} \tag{19}$$

Choice j is observed if $Z_{it}^{j} \geq Z_{it}^{k}, \forall j \neq k,\ j,k = 1,4,5.$ This model can be estimated and interpreted in the same fashion as the models in the previous two sections. In addition to the model with only two choices ("produce"/"don't produce"), in this set up I can differentiate between the sunk costs involved with mothballing and shutdown. For example, the sunk costs involved with shutdown are larger than the ones for mothballing, so that state dependence is expected to be more pronounced among the leases that were shutdown. The latent variables representing the value of each option, relative to the producing option, are:

$$\begin{aligned}
&\text{mothball the lease}: \\
&Z_{it}^{4} = \alpha_1 + \boldsymbol{\beta}_1 \boldsymbol{P_t} + \boldsymbol{\delta}_1 \boldsymbol{X_{it}} + \boldsymbol{\gamma}_1 (m_{it-1}\boldsymbol{H_{it}}) + \boldsymbol{\lambda}_1 s_{it-1}\boldsymbol{H_{it}} + \varepsilon_{1it} \\
&\text{shutdown the lease}: \\
&Z_{it}^{5} = \alpha_2 + \boldsymbol{\beta}_2 \boldsymbol{P_t} + \boldsymbol{\delta}_2 \boldsymbol{X_{it}} + \boldsymbol{\gamma}_2 (m_{it-1}\boldsymbol{H_{it}}) + \boldsymbol{\lambda}_2 s_{it-1}\boldsymbol{H_{it}} + \varepsilon_{2it}\ .
\end{aligned} \tag{20}$$

For the same reasons as in the previous two models, I expect the estimates of $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ to be positive. This model can be estimated without data on capital investments.

6. Econometric Issues and Estimation Results

The dynamic structure of the empirical model requires the estimator to account for the fact that choices are linked through time. This section discusses an estimator that allows me to estimate probit models using micro panel data.

6.1 Econometric Issues

To estimate the importance of sunk costs correctly it is crucial that I control for other potential sources of persistence in the production status, such as the current and future expected profit level. Including the vector of observable lease characteristics X_{it} in the estimation controls for much of this persistence. Nevertheless, most likely there are some important characteristics that are unobserved, which can induce serial correlation in the error terms, ε_{jit}. Furthermore, it is very likely that errors are correlated among choices. For instance, a profit shock that induces capacity expansion should have a negative effect on the values of the choices "mothballing" and "shutdown." Therefore, the model is specified with an error structure that allows a serially-correlated choice component and a random effect lease component. With this structure the vector of error terms for the normalized choices is

$$\boldsymbol{\varepsilon_{it}} = \sqrt{1-\lambda} * \boldsymbol{\omega_{it}} + \sqrt{\lambda} * \boldsymbol{\Psi} * \boldsymbol{\upsilon_i}\text{, where } \boldsymbol{\omega}_{it} = \boldsymbol{I} * \boldsymbol{\rho} * \boldsymbol{\omega}_{it-1} + \boldsymbol{\Theta} * \boldsymbol{\eta}_{it}. \tag{21}$$

Let *J* be the number of choices in the initial, non-normalized problem.[20] For example, *J*=5 in the model with five choices. Then, since the estimates apply to the normalized model, $\boldsymbol{\varepsilon_{it}}, \boldsymbol{\omega_{it}}, \boldsymbol{\omega_{it-1}}, \boldsymbol{\upsilon_i}, \boldsymbol{\eta}_{it}$, and $\boldsymbol{\rho}$ are of dimension (*J*-1)×1. $\boldsymbol{I}$ is a (*J*-1)× (*J*-1) identity matrix. The lower-triangular matrices $\boldsymbol{\Psi}$ and $\boldsymbol{\Theta}$ are of dimension (*J*-1)× (*J*-1). The random vectors $\boldsymbol{\upsilon_i}, \boldsymbol{\eta}_{it}$ are *IIDN*(***0,I***) and the vector $\boldsymbol{\rho}$ contains the AR(1) serial correlation parameters. The covariances among the choices are determined by the matrices $\boldsymbol{\Psi}$ and $\boldsymbol{\Theta}$.[21] Finally, the parameter λ weights the contribution of the time-varying error and the lease error to the regression error $\boldsymbol{\varepsilon_{it}}$.

Assuming that $\boldsymbol{\upsilon_i}$ and $\boldsymbol{\eta}_{it}$ are normally distributed random variables, this error structure generate a multinomial probit model. Due to the serially correlated errors and the fact that in each period there are more than two choices, estimation of this probit model is non-trivial.[22] Keane (1994) has developed a feasible simulation estimator for estimating probit models on panel data. The estimator is based on McFadden's (1989) method of simulated moments (MSM) and recognizes that the probability of observing a particular choice sequence is the product of the conditional probabilities in each period. The MSM estimation strategy is based on these transition probabilities.[23]

---

[20] For instance, in the analytical model above with capital adjustment, mothballing, and shutdown: J=5.

[21] For reasons of identification, the variance of the first choice, and thus the parameter in the most upper, left corner of both matrices, is normalized to one. Note, if J=1 these two matrices collapse to a scalar, which is equal to one.

[22] Estimating this problem by maximum-likelihood involves T*(J-1) dimensional integration. In my data T=18.

[23] The simulations in this paper are based on the Geweke-Hajivassiliou-Keane (GHK) simulation algorithm.

Serial correlation in profit shocks is not the only econometric problem that can lead to biased estimates of coefficients on the lagged state indicators. I observe a lease's production and investment history in years 1 through *T*. Since the lagged structure reaches back one period, the models of equations (10), (17), and (20) can only be use to model the lease's decision for the years 2 through *T*. However, the production and investment choice in the pre-sample period, here t=1, cannot be treated as exogenously, because the profit shock in the first year of observations, ε_{i1}, is correlated with the previous period's choice component, $\varpi_{it=0}$, in the error and the random effects lease component ν_i. Heckman (1981) suggests a way of dealing with the "initials-condition" problem. He proposes to use an approximate representation for the choice in the pre-sample period and to allow for correlation between the disturbances from this pre-sample expression and the disturbances from the sample expression. This procedure can be incorporated in the simulation probit estimator I use.

6.2 Econometric Result for "Produce" or "Don't produce"

I begin the empirical application by estimating the simple model in which the firm makes the decision to produce or not produce on a lease. Table 4 reports results for the model of equation (17).[24] Columns one and two are coefficients and t-statistics, respectively, using standard probit. Standard probit does not allow for heterogeneity among lease profit levels and serial correlation in profit shocks. Columns three and four report random effect probit results. This model allows for a lease component in the error structure. In contrast to the standard probit model, this model allows leases to differ in unobserved profitability. The most general model, reported in the last two columns, is the probit model with random lease effects and serially-correlated profit shocks. It is estimated with MSM using the GHK-probability simulator.

The coefficient estimates are very similar across the different estimation methods. The first two variables measure market conditions. The oil price variable indicates that leases are more likely to produce during periods of higher oil prices. An increase in the within year volatility of the oil price has a positive, but not statistically significant effect

---

[24] This estimates are not yet corrected for the initial conditions. However, this can be done within the MSM framework.

on the probability of production. The characteristics of the individual lease are also statistically important. A higher past production rate increases the probability of production. The probability of production is decreasing in age of the lease and average depth of the wells.

The coefficients on the lagged production indicator variable, $prod_{t-1}$, are positive and significant across the different estimation methods. The probit estimates for the lagged production indicator are similar, ranging from 2.117 to 2.387. The coefficients on well depth and the lagged production indicator are positive regardless of which estimation method is used, too. Thus, a lease's current production status is heavily affected by its status in the previous period. These results imply that sunk costs are a very important determinant of current the production status. The estimates imply that the magnitude of sunk costs is increasing in the depth of the well.

Regarding the regression errors, the MSM estimate imply that about 4.5% of the variance is from lease-level errors. For the random effects probit this is 12.5%. Also, the MSM results show that contemporaneous errors are negatively serially correlated. While the additional parameters added by the more general error structure are statistically significant, they have little effect on the other coefficients. Note that these results are preliminary. They still need to be corrected for the initial condition.

6.3 Preliminary Econometric Results for Three and Five Choice Models

Estimation results for the three and five choice models with probit are not available yet. Because multinomial logit models are easy to estimate I use this method to estimate the three and five choice models. Tables 5 and 6 report preliminary estimation results, which I discuss briefly. Table 5 contains logit estimates for equation (20), with the choices "produce," "mothball," or "shutdown." In this estimation "produce" is the base category. Columns one and two are the coefficient estimates and t-statistics, respectively, for "mothballed." The last two columns report the estimates for "shutdown." In brief, a higher price of oil or larger past production decrease the probability of being either "mothballed" or " shutdown." Most importantly the coefficients on the mothballing and lagged shutdown indicators are significant and positive. Moreover, the model predicts that the estimated coefficient on the lagged mothballing indicator is smaller than the coefficient on the lagged shutdown indicator. My estimates indicate that this conjecture holds. In this preliminary estimates sunk costs play an important role in the investment decision, too.

Table 6 reports the logit result for equation (10), the least restrictive model. The choices are normalized to the category “produce and don’t change capacity.” The results for the choices “produce and increase capacity,” “produce and decrease capacity,” “mothball,” and “shutdown” are reported in consecutive order from left to right. Again, the preliminary results are as expected. The probability of investment is increasing in the price of oil and the probabilities of scrapping, mothballing, and shutdown are decreasing in the price of oil. Looking at the categories “mothballed” and “shutdown,” the coefficient estimates on the lagged indicator variables are positive and significant. Moreover, within the category “shutdown” the estimate on the lagged mothballing indicator is smaller than the one on the lagged shutdown indicator. Assuming that the sunk costs of shutdown are equal to the sunk costs of mothballing plus additional sunk costs, the model implies this result. In summary, the logit estimates of the two multinomial models indicate that sunk costs are a very important determinant of investment, mothballing, and shutdown.

7. Conclusion

This paper analyzes the investment and production behavior of oil producers in Oklahoma during the 1980’s to 1990’s. In particular, I estimate the effect of oil prices and sunk costs on investment and production. This paper finds that the current status of production is largely affected by the previous status of production. After controlling for other potential sources of profit persistence, I attribute this effect to the importance of sunk costs and conclude that sunk costs explain entry and exit hysteresis in the oil industry.

A higher price of oil increases the probability of oil production. However, the price of oil is not as important in explaining production as the prior production status. Because of the importance of sunk costs, producing oil leases have a low probability of exiting and non-producing leases have a low probability of entering production. This effect is increasing in a well’s depth. An increase in well depth decreases the probability of entering production, but increase the probability of a producer remaining in production. I conclude that larger sunk costs lead to higher state persistence.

8. References

American Petroleum Institute, *Basic Petroleum Data Book,* Vol. XIX, Number 1, Washington D.C.: American Petroleum Institute, January 1999.

Bresnahan, Timothy F. and Reiss, Peter C., "Measuring the Importance of Sunk Costs," *Annales d'Economie et de Statistique,* April-June 1994, No. 34, pp.181-217.

Campa, Jose M., "Entry by Foreign Firms in the United States under Exchange Rate Uncertainty," *Review of Economics and Statistics*, November 1994, 75(4), pp. 614-622.

Chirinko, Robert, "Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications," *Journal of Economic Literature*, December 1993, Vol. XXXI, pp. 1875-1911.

Cooper, Russel; Haltiwanger, John and Power, Laura, "Machine Replacement and the Business Cycle: Lumps and Bumps," *The American Economic Review*, September 1999, pp. 921-946.

Dixit, Avinash, "Entry and Exit Decisions under Uncertainty", *Journal of Political Economy,* June 1989, *97*, pp.620-638.

Dixit, Avinash and Pindyck, Robert, *Investment under Uncertainty,* Princeton: Princeton U. Press, 1994.

Doms, Mark and Dunne, Timothy, "Capital Adjustment Patterns in Manufacturing Plants," *Review of Economic Dynamics*, 1998, pp. 409-429.

Green, William H., *Econometric Analysis*, New Jersey: Prentic Hall, 2000.

Heckman, James J., “The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process,’ in Charles Manski and Daniel McFadden, eds., *The Structural Analysis of Discrete Data.* Cambridge, MA: MIT Press, 1981, pp. 179-195.

Hurn, A. S. and Robert E. Wright, “Geology or Economics? Testing Models of Irreversible Investment Using North Sea Oil Data,” *The Economic Journal*, March 1994, 104, pp. 363-371.

Keane, Michael P., "A Computationally Practical Simulation Estimator for Panel Data," *Econometrica*, 1994, 62(1), pp. 95-116.

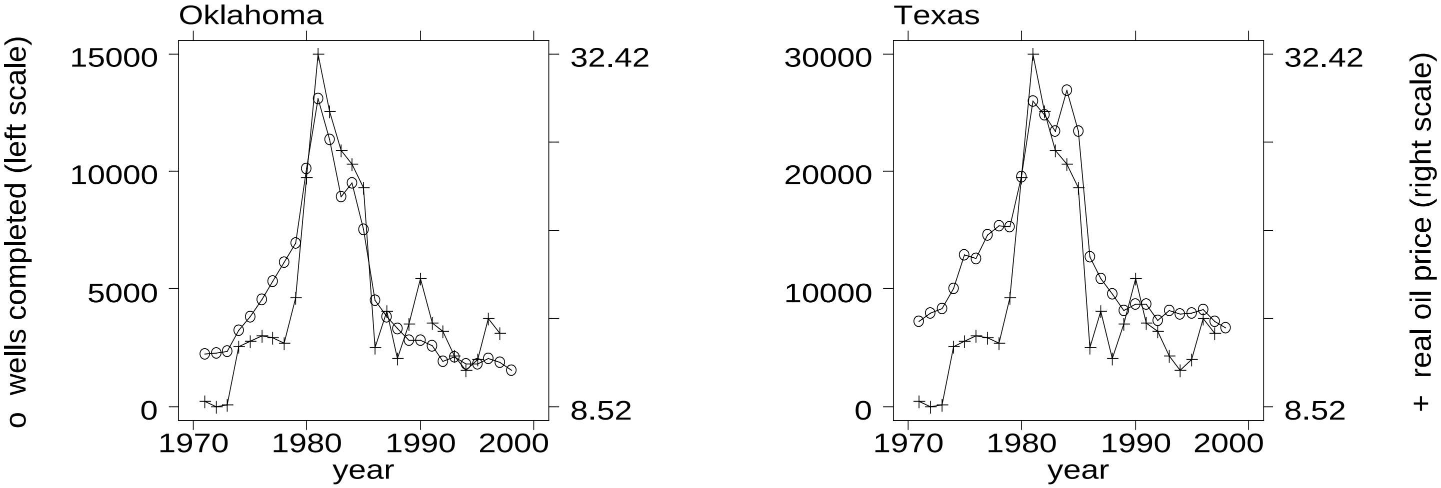
Geweke, John F.; Keane, Michael P. and Runkle, David E., "Statistical Inference in the Multinomial Multiperiod Probit Model," *Journal of Econometrics*, September 1997, 80(1), pp. 125-165.

McFadden, Daniel, “A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration,” *Econometrica*, 1989, 57(3), pp. 995-1026.

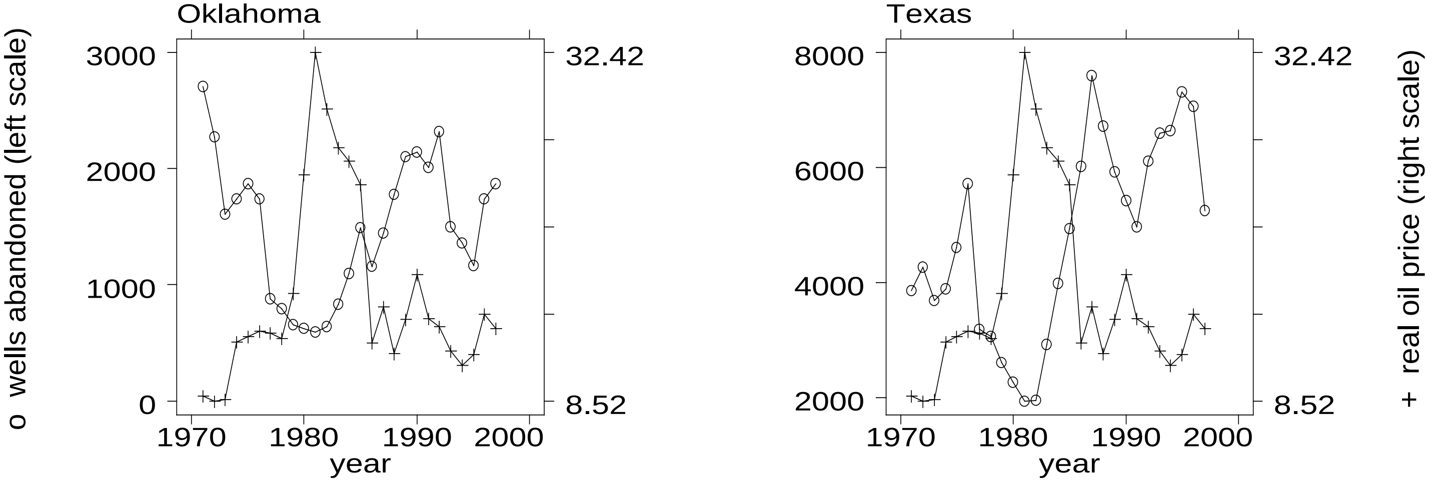
Paddock, James L.; Daniel R., Siegel and Smith, James L., “Option Valuation of Claims on Real Assets: The Case of Offshore Petroleum Leases,” *Quarterly Journal of Economics*, Aug. 1988, 103(3), pp. 479-508.

Roberts, Mark J. and Tybout, James R., “The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Cost,” *The American Economic Review*, Sep. 1997, Vol. 87(4), pp. 545-564.

Number of Well Entries, Plot 1



Number of Stripper Well Exits, Plot 2



Total Number of Producing and Stripper Wells, Plot 3

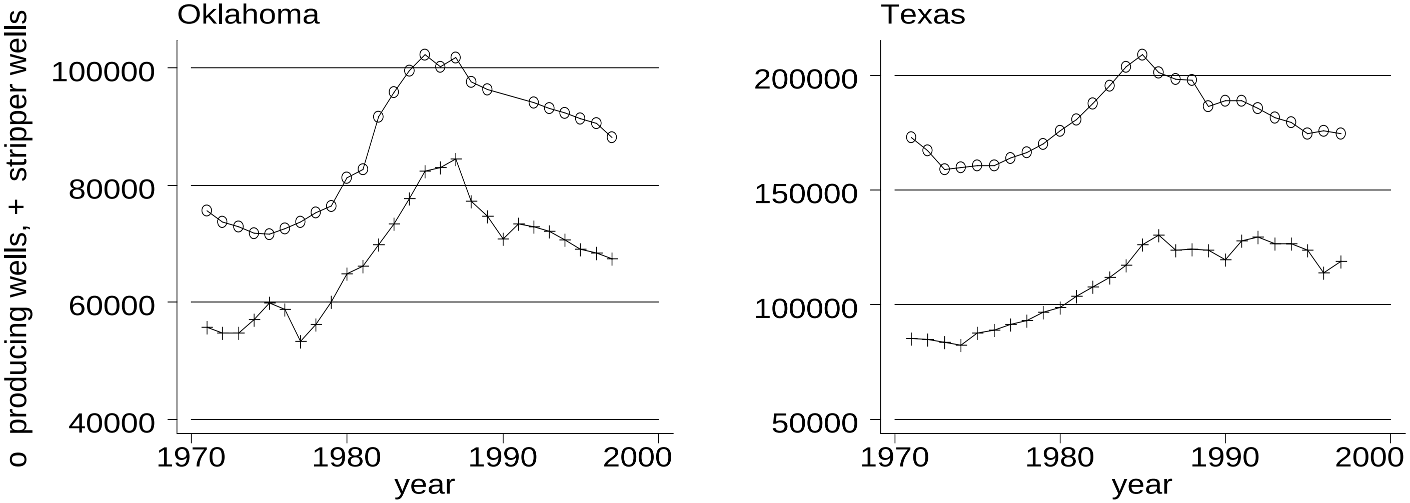


Table 1 – Sunk Costs Incurred When Changing Production Status

		Lease status in period *t*		
		Production	Mothballed	Shutdown
Lease status in period *t-1*	Production	0	M_i	M_i+S_i
	Mothballed	RM_i	0	S_i
	Shutdown	RM_i+RS_i	RS_i	0

Table 2 – Average Transition Rates

Year *t* status / Year *t-1* status	Produce and don't change capacity	Produce and increase capacity	Produce and decrease capacity	Mothball	Shutdown
Produce and don't change capacity	0.924	0.012	0.01	0.053	0.001
Produce and increase capacity	0.806	0.162	0.012	0.019	0
Produce and decrease capacity	0.822	0.016	0.038	0.06	0.065
Mothball	0.205	0.006	0.002	0.441	0.347
Shutdown	0.084	0.004	0.001	0	0.911
Unconditional probability of year *t* status	0.77	0.015	0.009	0.067	0.137

Table 3 - Transition Rates for Production/No production,” 1980-1998

Year *t*	$/bbl of oil	Year *t-1* status:	Production		No production		Number of producing leases, *t-1*	Number of non-producing leases, *t-1*
		Year *t* status:	Production	No production	Production	No production		
1980	37.32		0.980	0.020	0.019	0.981	808	106
1981	36.65		0.991	0.009	0.074	0.926	819	121
1982	33.57		0.983	0.017	0.200	0.800	848	120
1983	30.38		0.962	0.038	0.212	0.788	876	113
1984	29.24		0.970	0.030	0.285	0.715	889	123
1985	27.94		0.976	0.024	0.157	0.843	913	115
1986	14.83		0.961	0.039	0.134	0.866	932	119
1987	19.13		0.962	0.038	0.158	0.842	920	139
1988	15.91		0.957	0.043	0.336	0.664	921	152
1989	19.57		0.943	0.057	0.186	0.814	942	140
1990	23.71		0.982	0.018	0.143	0.857	930	168
1991	21.43		0.932	0.068	0.112	0.888	955	161
1992	20.52		0.952	0.048	0.139	0.861	912	209
1993	18.38		0.946	0.054	0.156	0.844	902	225
1994	17.11		0.921	0.079	0.146	0.854	890	240
1995	18.41		0.954	0.046	0.069	0.931	857	274
1996	22.06		0.900	0.100	0.099	0.901	841	294
1997	20.53		0.950	0.050	0.063	0.937	796	349

Table 4 - Transition Rates Production/Mothballed/Shutdown

Year *t*	$/bbl of oil	Year *t-1* status:	Production (**P**)			Mothballed (**M**)			Shutdown (**S**)			Number of leases in *t-1* that are:		
		Year *t* status:	**P**	**M**	**S**	**P**	**M**	**S**	**P**	**M**	**S**	**P**	**M**	**S**
1980	37.32		0.980	0.019	0.001				0.019	0.000	0.981	808	0	106
1981	36.65		0.991	0.009	0.000	0.375	0.625	0.000	0.029	0.010	0.962	819	16	105
1982	33.57		0.983	0.017	0.000	0.526	0.211	0.263	0.139	0.000	0.861	848	19	101
1983	30.38		0.962	0.038	0.000	0.381	0.476	0.143	0.174	0.000	0.826	876	21	92
1984	29.24		0.970	0.029	0.001	0.136	0.727	0.136	0.367	0.000	0.633	889	44	79
1985	27.94		0.976	0.023	0.001	0.224	0.310	0.466	0.088	0.000	0.912	913	58	57
1986	14.83		0.961	0.039	0.000	0.179	0.410	0.410	0.112	0.000	0.887	932	39	80
1987	19.13		0.962	0.038	0.000	0.308	0.442	0.250	0.069	0.000	0.931	920	52	87
1988	15.91		0.957	0.042	0.001	0.362	0.362	0.276	0.319	0.000	0.681	921	58	94
1989	19.57		0.943	0.056	0.001	0.305	0.475	0.220	0.099	0.000	0.901	942	59	81
1990	23.71		0.982	0.018	0.000	0.185	0.481	0.333	0.103	0.000	0.897	930	81	87
1991	21.43		0.932	0.066	0.002	0.286	0.179	0.536	0.019	0.000	0.981	955	56	105
1992	20.52		0.952	0.045	0.003	0.216	0.662	0.122	0.096	0.000	0.904	912	74	135
1993	18.38		0.946	0.050	0.004	0.209	0.275	0.516	0.119	0.000	0.881	902	91	134
1994	17.11		0.921	0.078	0.001	0.186	0.514	0.300	0.129	0.000	0.871	890	70	170
1995	18.41		0.954	0.043	0.002	0.144	0.529	0.327	0.024	0.000	0.976	857	104	170
1996	22.06		0.900	0.087	0.013	0.185	0.293	0.522	0.059	0.000	0.941	841	92	202
1997	20.53		0.950	0.050	0.000	0.110	0.680	0.210	0.044	0.000	0.956	796	100	249

Table 5 - Two Choices Model: "Produce" or "Don't Produce"

	Probit		Random effects probit		Probit with MSM/GHK	
	Coef.	t	Coef.	t	Coef.	t
ln(price)	0.4497	8.6140	0.4640	8.4770	0.4344	24.3734
std of ln(price)	0.6507	1.2110	1.0802	1.9320	0.5981	1.1261
production	0.0133	6.9940	0.0144	6.9400	0.0134	5.4747
age	-0.0210	-3.4380	-0.0316	-4.0880	-0.0221	-3.6526
age^2	0.0006	3.3770	0.0008	3.7120	0.0006	3.3305
depth	-0.0493	-4.5390	-0.0658	-5.1150	-0.0435	-5.7640
p_{t-1}-dummy	2.3055	28.2490	2.1166	23.1810	2.3870	32.5974
p_{t-1}-dummy*(std of ln(price))	0.4164	0.6460	-0.0763	-0.1130	0.5232	0.9812
p_{t-1}-dummy*depth	0.0941	7.3920	0.1145	8.0790	0.0835	8.1982
const.	-2.1344	-11.7450	-1.8658	-9.2670	-2.1351	-30.2319
lambda	-	-	0.1253	15.420	0.0486	7.3036
rho	-	-	-	-	-0.2766	-11.4682

“don’t produce” is the base category,

lambda: fraction of variance due to lease specific error component, rho: AR(1) parameter, p_{r-1}: production indicator variable

Table 6- Logit Estimates of a Three Choice Model

Choice	Mothball		Shutdown	
	Coef.	t	Coef.	t
const.	0.751254	1.72	-4.29751	-5.125
ln(price)	-1.06719	-8.143	-0.57118	-3.278
std of ln(price)	-2.46383	-2.844	-2.9823	-0.72
production	-0.06472	-6.86	-1.44701	-3.865
age	0.033441	2.403	0.045638	2.205
age^2	-0.001	-2.497	-0.00107	-1.863
depth	-0.11158	-7.288	0.18284	2.528
m_{t-t}	2.438686	11.672	5.445022	8.629
m_{t-t}*[std of ln(price)]	2.360712	1.391	5.053622	1.151
m_{t-t} *depth	0.240605	6.785	-0.07713	-0.964
s_{t-1}	-	-	7.89235	12.555
s_{t-1}*[std of ln(price)]	-	-	1.012697	0.232
s_{t-1}*depth	-	-	-0.16062	-2.06

Logit with 3 choices, "produce" is the base category

Table 7- Logit Estimates of a Five Choices,Model

Choice	Produce and increase capacity		Produce and decrease capacity		Mothball		Shutdown	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t
const.	-8.44689	-13.125	-4.18161	-4.11	0.64898	1.486	-4.52151	-5.386
ln(price)	1.94192	11.059	-0.73558	-2.621	-1.01179	-7.708	-0.47044	-2.685
std of ln(price)	-0.52557	-0.402	-3.35781	-1.917	-2.48957	-2.878	-2.91524	-0.71
production	0.000615	1.15	0.00027	0.305	-0.06488	-6.87	-1.45003	-3.874
age	-0.11137	-5.82	0.123424	3.58	0.029492	2.116	0.03819	1.842
age^2	0.00347	6.658	-0.00085	-1.037	-0.00085	-2.105	-0.00081	-1.404
depth	-0.25403	-10.203	0.085016	2.355	-0.11461	-7.476	0.180958	2.503
m_{t-t}	1.527581	1.322	-3.21858	-1.688	2.436925	11.519	5.451773	8.651
m_{t-t}*[std of ln(price)]	-19.7538	-1.029	15.65138	1.888	2.353582	1.373	4.969751	1.139
m_{t-t} *depth	-0.01471	-0.063	0.289602	1.205	0.241854	6.754	-0.07637	-0.953
s_{t-1}	0.82207	1.317	-0.62496	-0.298	-	-	7.963819	12.651
s_{t-1}*[std of ln(price)]	4.050372	0.884	2.07896	0.132	-	-	1.140224	0.262
s_{t-1}*depth	0.025219	0.206	0.046384	0.153	-	-	-0.17211	-2.2

Logit with 5 choices, "produce and don't change capacity" is the base category