American Economic Review 2014, 104(6): 1698–1734
http://dx.doi.org/10.1257/aer.104.6.1698

The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling†

By Ryan Kellogg*

This paper estimates the response of investment to changes in uncertainty using data on oil drilling in Texas and the expected volatility of the future price of oil. Using a dynamic model of firms' investment problem, I find that: (i) the response of drilling activity to changes in price volatility has a magnitude consistent with the optimal response prescribed by theory, (ii) the cost of failing to respond to volatility shocks is economically significant, and (iii) implied volatility data derived from futures options prices yields a better fit to firms' investment behavior than backward-looking volatility measures such as GARCH. (*JEL* C58, D92, G13, G31, L71, Q31)

The real options literature, beginning with Marschak (1949) and Arrow (1968) and developed in Bernanke (1983), Pindyck (1991), and Dixit and Pindyck (1994), explains how firms should make decisions about investments that involve sunk costs. Real options theory views such investments as options in that, at any point in time, a firm may choose to either invest immediately or delay and observe the evolution of the investment's payoff. A key insight is that the option to delay has value when future states of the world with positive returns to investing and states with negative returns are both possible, even holding the expected future return constant at its present level. Thus, in the presence of irreversibility and uncertainty, a naïve investment timing rule—proceed with an investment if its expected benefit even slightly exceeds its cost—is suboptimal because it does not account for the value of continuing to hold the option. Instead, firms should delay irreversible investments until a significant gap develops between the investments' expected benefits and costs. Moreover, as uncertainty increases, real options theory tells us that the incentive to delay should grow stronger and the gap between the expected benefit and cost necessary to trigger investment should widen.

While real options theory therefore prescribes how firms *should* carry out irreversible investments in uncertain environments, it is not empirically well-known

* Department of Economics, University of Michigan, 238 Lorch Hall, 611 Tappan Street, Ann Arbor, MI 48109-1220, and NBER (e-mail: kelloggr@umich.edu). I am grateful for financial support from the OpenLink Fund at the Coleman Fung Risk Management Center. I thank Maximilian Auffhammer, Ruediger Bachmann, Robert Barsky, Nicholas Bloom, Severin Borenstein, Jun Ishii, Lutz Kilian, Kai-Uwe Kühn, Jeffrey Perloff, Robert Pindyck, Matthew Shapiro, Catherine Wolfram, and two anonymous referees for helpful comments. I am also grateful for suggestions from numerous seminar and conference participants. I thank Reid Dorsey-Palmateer, Tay Feder, Eric Lewis, and Haili Pang for excellent research assistance. I declare that I have no relevant or material financial interests that relate to the research described in this paper.
† Go to http://dx.doi.org/10.1257/aer.104.6.1698 to visit the article page for additional materials and author disclosure statement(s).

how firms *actually proceed* in such situations. In particular, the theory's central prediction that firms should be more likely to delay investment if uncertainty increases, all else equal, has received only limited empirical scrutiny. The primary aim of this paper is therefore to assess the extent to which firms' responses to changes in uncertainty align with the theory, using data on oil drilling activity in Texas coupled with market-based expectations of the volatility of the future price of oil.

I first show, in a descriptive analysis, that increases in the expected volatility of the future price of oil are associated with decreases in drilling activity, conditional on the expected future oil price level. The core of this paper is then aimed at assessing whether the magnitude of the empirical relationship between drilling and uncertainty is aligned with the prediction from a real options model of investment. I construct and estimate a dynamic, econometric model of firms' optimal drilling timing that is based on Rust's (1987) nested fixed point approach but allows the volatility of the process governing future oil prices to vary over time. The use of this model allows me to do more than carry out a simple "yes/no" test of whether or not firms respond to changes in uncertainty: I can also ask whether the magnitude of firms' responses in the data agrees with the magnitude prescribed by the model. To the best of my knowledge, this paper is the first in the real options literature to empirically address this question.

I find that the response of drilling investment to changes in uncertainty is broadly consistent with optimal decision-making. That is, when the expected volatility of the future price of oil increases, drilling activity decreases by a magnitude that aligns with that predicted by the real options model. The close adherence of firms' drilling decisions to the theory is underscored by a related finding that firms have a substantial economic incentive to time their investments optimally: ignoring within-sample variation in oil price volatility can reduce the value of a drilling prospect by more than 25 percent.

These results provide a micro-empirical foundation for a large number of applications of real options theory that implicitly assume that firms optimally make decisions in the presence of time-varying uncertainty. In industrial organization, for instance, Pakes (1986), Dixit (1989), Grenadier (2002), Aguerrevere (2003), and Collard-Wexler (2013) model the implications of uncertainty and sunk costs for investment, entry, and research and development in several settings and under various forms of competition. The general dynamic oligopoly model of Ericson and Pakes (1995) is built on a framework in which firms treat many decisions as options. In macroeconomics, Bernanke (1983); Hassler (1996); Bloom (2009); Bloom, Bond, and Van Reenen (2007); and Bloom et al. (2012) construct models that emphasize the importance of changes in economy-wide uncertainty in determining the level of aggregate investment. In international trade, Handley (2012) and Handley and Limão (2012) model the effect of trade policy uncertainty on exporters' investment and entry decisions. Finally, in the environmental and resource economics literature, Arrow and Fisher (1974), among others, discuss the role of uncertainty in dictating when "green" investments should be undertaken.

I conduct my analysis using a detailed dataset of well-level drilling activity in Texas obtained from the Texas Railroad Commission. I combine the drilling data with information on the expected future oil price and price volatility from the New York Mercantile Exchange (NYMEX). I derive my measure of expected price volatility from the NYMEX futures options market, in which volatility is implicitly traded and priced. Under a hypothesis that the market is an efficient aggregator of information, the implied volatility from futures options will incorporate more information about the distribution of future prices than backward-looking volatility measures derived from price histories alone. Consistent with this hypothesis, I find that when I measure expected price volatility using historic volatility (either directly or via a GARCH model) rather than implied volatility, the model does a relatively poor job of fitting the data, and the estimated response of investment to changes in volatility is attenuated and imprecise. These results complement research in the finance literature that finds that, across many commodity and financial markets, implied volatility tends to be a better predictor of future volatility better than backward-looking measures (Poon and Granger 2003; Szakmary et al. 2003).

There exist previous studies that have empirically examined whether investments respond to changes in uncertainty, though without linking the magnitudes of the estimated effects to theory. Several of these studies, like this one, focus on natural resource industries. Hurn and Wright (1994); Moel and Tufano (2002); and Dunne and Mu (2010) examine the impact of resource price volatility on offshore oil field investments, gold mine openings and closings, and refinery investments, respectively. None of these papers uses implied volatility to measure expected price volatility—the uncertainty measure is the historic realized variance of commodity prices—and they collectively find mixed evidence on whether increases in volatility reduce investment. Other micro-empirical work includes Guiso and Parigi (1999), which finds evidence from a cross-sectional survey that Italian firms whose managers subjectively report high levels of expected demand uncertainty tend to have relatively low levels of investment. List and Haigh (2010) meanwhile provides experimental evidence that investment timing decisions of agents (drawn from student and professional trader subject pools) are generally responsive to changes in payoff uncertainty.

Another set of papers in the macroeconomics literature measures the response of aggregate output and investment to changes in economy-wide uncertainty, as measured by the volatility of stock market returns or interest rates (Hassler 2001; Alexopoulos and Cohen 2009; Bloom 2009; and Fernández-Villaverde et al. 2011). A related work is Leahy and Whited (1996), which examines firm-level investment and stock return volatilities. These papers generally find that increases in volatility are associated with decreases in output or investment. However, factors that influence the expected level of investments' payoffs are difficult to proxy for in this literature, so that a negative correlation between first and second moment shocks (a possibility suggested by Bachmann, Elstner, and Sims 2013) may cause these estimates to be downward-biased (away from zero). Leahy and Whited (1996) also notes that fluctuations in stock returns likely reflect the volatility of factors beyond those impacting the future revenues associated with new, marginal investment opportunities.

This paper's focus on the Texas onshore drilling industry as its object of study, combined with econometric modeling of the firms' investment timing problem, confers valuable advantages relative to previous work. First, I possess data at the level of each individual investment—the drilling of each well—and need not rely on aggregate data or accounting data. Second, the NYMEX futures and futures options markets provide measures of the expected level and volatility of each investment's expected return that, in principle, incorporate all available information at the time of the investment. Such measures are not available in most industry settings, and they also allow for a separation of first and second moment shocks. Finally, I take advantage of the fact that oil production is a highly competitive industry, with no one firm able to influence the price of oil, and I focus my analysis on oil fields in which common pool issues are not a concern. I am therefore able to treat each firm's investment decision as a single-agent dynamic investment problem. This approach, which would be questionable in most other industries, allows me to measure the magnitude of firms' response to uncertainty relative to the theoretical optimum, going beyond a simple test of whether or not firms respond to uncertainty shocks at all.

In what follows, I first discuss relevant institutional details of the Texas onshore drilling industry and the datasets I use. Section II follows with a descriptive analysis of the data. The remainder of the paper focuses on the construction and estimation of a structural model of the drilling investment problem with time-varying uncertainty: Section III presents the model, Section IV discusses the estimation procedure, and Section V follows with the estimation results. Section VI provides concluding remarks.

I. Institutional Setting and Data

A. *Drilling Description*, *Types of Wells Used in This Study*, *and Drilling Data*

Oil and gas reserves are found in geologic formations known as fields that lie beneath the earth's surface, and the mission of an oil production company is to extract these reserves for processing and sale. To recover the reserves, the firm needs to drill wells into the field. Drilling is an up-front investment in future production; if a drilled well is successful in finding reserves, it will then produce oil for a period of several years, requiring relatively small operating expenses for maintenance and pumping. The firm does not know in advance how much oil will be produced (if any) from a newly drilled well, though it will form an expectation of this quantity based on available information, such as seismic surveys and the production outcomes of previously drilled wells. The price that the firm will receive for the produced oil is also not known with certainty at the time of drilling. Conversations with industry participants have indicated that some, though not all, firms use the NYMEX market to hedge at least some of their price risk. This use of the NYMEX indicates that risk aversion over future oil prices is unlikely to influence drilling decisions, since any firm that is risk averse can hedge the price risk away.

Drilling costs range from a few hundred thousand dollars for a relatively shallow well that is a few thousand feet deep to millions of dollars for a very deep well (as much as 20,000 feet deep). Once drilled, these costs are almost completely sunk: the labor and drilling rig rental costs expended during drilling cannot be recovered, nor can the expensive steel well casing and cement that run down the length of the hole. Drilling can therefore be modeled as a fully irreversible investment.

Wells can be one of three types: exploratory, development, or infill. Exploratory wells are drilled into new prospective fields, and if successful they can not only be productive themselves but also lead to additional drilling activity. Development wells are those that follow the exploratory well: they increase the number of penetrations into a recently discovered field in order to drain its reserves. Finally, infill wells are drilled late in a field's life to enhance an oil field's production by "filling in" areas of the reservoir that have not been fully exploited by the pre-existing well stock.

In this paper, I exclude exploratory and development wells and analyze only the subset of data corresponding to infill wells. This exclusion facilitates this study in two important ways. First, examining only infill wells constrains the set of available drilling options to those that exist within a finite, known set of fields. Thus, I need not be concerned with the creation of new options through new field discoveries or leasing activity. Second, the majority of production from a typical infill well takes place within the first year or two of the well's life: because infill wells tap only small isolated pools of oil that have been left behind by older wells in a field, their productive life is quite short. Thus, I may rely on liquid near-term futures to provide expected prices and volatilities that are relevant for these wells rather than less liquid long-term futures.

I also distinguish wells drilled in fields operated by a single firm from wells drilled in fields operated by multiple firms. The process by which production companies acquire leases—rights to drill on particular plots of land—often leads to situations in which several firms have the right to drill in and produce from a single field (see Wiggins and Libecap 1985). This division of operating rights leads to a common pool problem to the extent that each firm's actions lead to informational and extraction externalities for its neighbors, suggesting that in such situations a dynamic game is needed to model firms' drilling problem. This paper avoids this substantial complication by focusing exclusively on wells drilled in sole-operated fields, for which a single-agent model is sufficient to model drilling behavior.[1]

I obtained drilling data from the Texas Railroad Commission (TRRC) "Drilling Permit Master and Trailer" database, yielding information regarding every well drilled in Texas from 1977 through 2003.[2] These data identify when each well was drilled, which field it was drilled in, whether it was drilled for oil or for gas, and the identity of the production company that drilled it. During the 1993–2003 period for which I also observe data on drilling costs and expected oil prices, I observe a total of 23,279 oil wells.[3] Of these, 17,456 are infill wells and 1,150 are infill wells drilled in sole-operated fields.[4]

[1] Industry participants have suggested that the degree of strategic interaction amongst firms drilling infill wells in common pool fields may be limited in practice because infill drilling targets tend to be small pools that are geologically isolated from other parts of the field. In addition, the TRRC regulates the minimum distance from a neighbor's lease at which a well may be drilled. Correspondingly, the time series of infill drilling in all fields, including common pools, is very similar to that for sole-operated fields (see online Appendix Figure A1). I nevertheless focus my analysis on sole-operated fields to be conservative, though estimating the model using the full sample of infill wells yields very similar results to those presented in Section V.

[2] While drilling data exist beyond 2003, industry participants have indicated that the dramatic increase in oil and natural gas prices that began in 2004 increased drilling activity to the extent that the rig market became extremely tight. Long wait lists developed when large production companies locked up rigs on long-term contracts so that the spot rental market could not allocate rigs based on price. Because these unobservable wait lists disconnect drilling decisions from observed drilling, I only use data through 2003.

[3] I define an oil well as a well that is marked as a well for oil (rather than for "gas" or "both") on its TRRC drilling permit and is drilled into a field for which average oil production exceeds average natural gas production on an energy equivalence basis (1 barrel of oil is equivalent to 5.8 thousand cubic feet of gas).

[4] I define infill wells as those that are drilled into fields discovered prior to January 1, 1990. I define a sole-operated field as one for which, in every year from 1993 to 2003, only a single firm is listed as a leaseholder in the field's

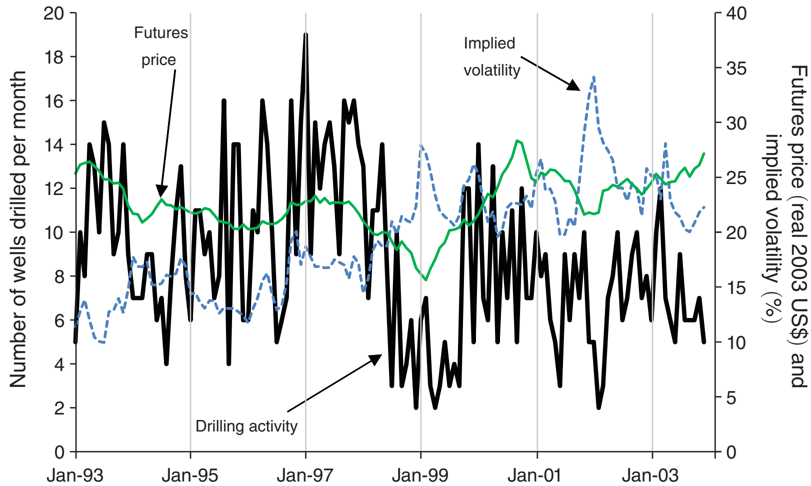


Figure 1. Time Series of Monthly Drilling Activity, Oil Futures Prices, and Implied Volatility from Futures Options Prices

Notes: Oil futures prices are 18-month-ahead prices from the NYMEX. Implied volatility is calculated from futures options prices per the discussion in Section ID. Drilling activity corresponds only to infill oil wells drilled in sole-operated fields.

The time series of Texas-wide drilling activity is depicted in Figure 1 as the number of wells drilled per month. These data appear to be noisy because they are integer count data ranging from 2 to 19 wells per month. The time series of drilling activity in a larger sample that includes wells drilled in common pool fields, provided in the online Appendix as Figure A1, does not exhibit this noisiness, confirming that it is due to the integer count nature of the data rather than a systematic feature of the industry.[5]

The drilled wells are spread over 663 sole-operated fields and 453 firms. The mean number of wells per field is 1.73, and I observe only one well drilled in the majority of fields in the data. The maximum number of wells I observe in any field is 31. In addition to the 663 fields in which I observe drilling, I also observe 6,637 sole-operated oil fields in which no infill wells are drilled. The median number of wells per firm is 1, the mean is 2.54, and the maximum is 31. Thus, the majority of wells in the dataset can be characterized as having been drilled by small firms in relatively small, old fields with few remaining drilling opportunities.

annual production data. This definition allows a field to be traded from one firm to another but disallows fields in which several firms operate simultaneously on different leases.

[5] I have also estimated a model using quarterly data rather than monthly data. Though the quarterly aggregation does substantially reduce the noise in drilling activity, it also loses important variation in prices and volatility. The estimate of firms' sensitivity to volatility in a quarterly model is therefore quite noisy: the point estimate of β from the dynamic model is 1.429 with a standard error of 1.281.

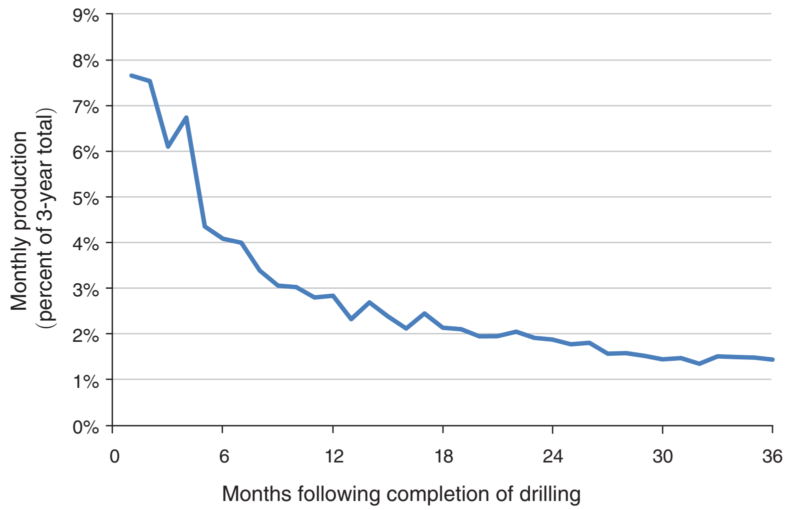


Figure 2. Average Monthly Production Profile from a Drilled Well

Notes: Production data are from the subset of observed drilled wells that are the only active producing well on their respective lease for the first 36 months subsequent to drilling. This subset amounts to 160 of the observed 1,150 drilled wells from 1993 to 2003.

B. *Oil Production*

I acquired oil production data from the TRRC's "Final Oil and Gas Annuals" dataset to assess the production that resulted from the observed drilling activity. Unfortunately, drilling data can only be matched to production records for a fraction of drilled wells. There are two impediments to merging these data. First, drilling data may only be linked to production data using the name of the lease, which is not uniform across the drilling and production databases.[6] Even after making a number of corrections to lease names (such as changing all instances of "and" to "&" and removing all periods), I am only able to match 527 of the 1,150 drilled wells to a lease in the production data. Second, the TRRC records monthly oil production at the lease level, not the well level, because individual wells are not flow-metered. I am therefore only able to identify production from wells that are drilled on leases on which there exist no other producing wells and there is no subsequent drilling: this is the case for 160 of the 527 matched wells. For these 160 wells, I tabulate the total production of each for the three years subsequent to drilling: the median well produces 8,417 barrels (bbl), and the mean produces 15,395 bbl. Seven (4.4 percent) of the wells are dry holes that produce nothing; the maximum production is 164,544 bbl.

Figure 2 displays the average monthly production profile of a drilled well in the sample. Production begins immediately subsequent to drilling, and depletion of the oil pool results in a fairly steep production decline so that a typical well's monthly production falls to one-half of its initial level only seven months into the well's

[6]The lease names in the drilling and production datasets come from different forms that are filed by producers at potentially very different times. According to TRRC staff, the two datasets were never intended to be merged together.

life. In addition, Anderson, Kellogg, and Salant (2014) show that firms do not alter production rates or delay production due to oil price changes: the shape of the production profile is consistent throughout the data, including the 1998–1999 period when the price of oil was very low. This profile is consistent with a production technology in which production rates are constrained by geologic characteristics of the oil reservoir such as its pressure, the remaining volume of oil near the well, and rock permeability. It is also consistent with low operating expenses, so that the probability that the oil price will fall below the point at which revenues equal operating costs is extremely low. Thus, the option value represented by the ability to adjust a well's production rate in response to price changes is negligible, implying that drilling and production do not need to be modeled as a compound option.

C. *Expected Oil Prices*

I measure expected oil prices using the prices of NYMEX crude oil futures contracts, obtained from Price-Data. With risk neutral traders and efficient aggregation of information by the market, the futures price is in theory the best predictor of the future price of oil. In practice, while futures prices have been found to provide slightly more precise predictions than the current spot price (i.e., a no-change forecast) during the 1993–2003 period I study here (Chernenko, Schwarz, and Wright 2004), the improvement is not statistically significant. Moreover, when data through 2007 are used, spot prices actually slightly outperform futures prices, though again the difference is not statistically significant (Alquist and Kilian 2010). Given the slightly superior performance of NYMEX futures during the sample period of this paper and the fact that a majority of producers claim to use futures prices in making their own price projections (Society of Petroleum Evaluation Engineers 1995), I will use futures prices as the measure of firms' expected price of oil. In a secondary specification, I explore how the use of spot prices impacts the results.

I focus on the prices of futures contracts with 18 months to maturity.[7] This maturity is the longest time horizon for which NYMEX futures are traded regularly (on 84 percent of all possible trading days over 1993–2003). In addition, the typical production profile of drilled infill wells suggests that 18 months might be a reasonable forecast horizon for a firm to use when evaluating a drilling prospect, since approximately one-half of the well's total expected discount production is likely to be exhausted at this time.[8]

Futures prices are consistent with mean-reverting expectations about the future price of oil, as shown in Figure 3. When the front-month (nearest delivery month) oil price exceeds approximately \$20/bbl (real 2003 US\$), the price of an 18-month futures contract tends to be lower than the front-month price, and the reverse holds when the front-month price is below \$20/bbl.

[7] In reality, it is rare that a NYMEX futures contract has a time to maturity of exactly 18 months (548 days), since the available contracts that can be traded have maturities that are either one full month or one full quarter apart. On any given trading date, I therefore treat contracts with a time to maturity that is within 46 days of 18 months as having a maturity of 18 months. When more than one such contract is traded on any given trading date, I average the prices across the contracts.

[8] This half-life is derived by fitting a hyperbolic curve to the average production data (Figure 2) and extrapolating production beyond three years. Based on this curve and a 9.9 percent real discount rate (see Section IVA), half of a typical well's expected discounted production is exhausted in 19.1 months.

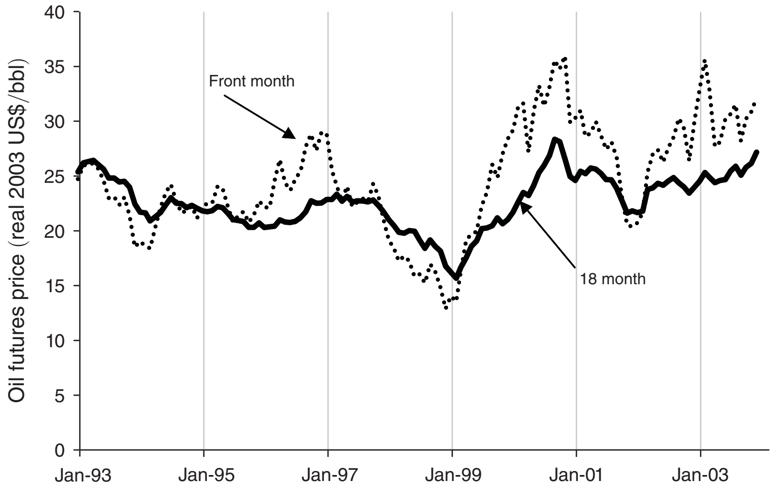


Figure 3. NYMEX 18-Month and Front-Month Oil Futures Prices

Source: Monthly averages of NYMEX futures prices obtained from PriceData

D. *Expected Oil Price Volatility*

I derive my primary measure of firms' expected future price volatility from the volatility implied by NYMEX futures options prices.[9] Across numerous commodity and financial contracts, implied volatility has been found to be a better predictor of future volatility than measures based on historic price volatility, including GARCH models (Poon and Granger 2003; Szakmary et al. 2003). Intuitively, if markets are efficient then options prices incorporate up-to-date information beyond that available from price histories alone, improving their predictive power.

The classic formula for the value of a commodity option contract is based on the Black-Scholes model (1973) and given by Black (1976). Given the price of an option, its time to maturity and strike price, the price of the underlying futures contract, and the riskless rate of interest, Black's formula can be inverted to yield the expected volatility implied by the option.[10] An important assumption of Black (1976) is that, on any given trade date, the volatilities of prices across all times to maturity are equal.[11] However, it is apparent in Figure 3 that front-month futures prices are, on average, more volatile than 18-month futures prices, violating this assumption. Hilliard and Reis (1998) shows that, in this case, applying Black (1976) to 18-month futures options yields the average volatility of futures price contracts with maturities between the front-month and 18 months. The empirical analysis below requires the volatility of 18-month futures prices rather than this average. In online Appendix 1, I discuss how I correct the Black (1976) implied volatilities to address this issue. The resulting time series of implied 18-month futures price

[9] I obtained data on daily crude futures options prices from Commodity Systems Inc.

[10] I use the interest rate on treasury bills (obtained from HSH Associates) to measure the riskless rate of interest.

[11] The Black (1976) formula also assumes that the options are European and that volatility is not stochastic. As discussed in online Appendix 1, however, these assumptions are likely to be of minor importance in this setting.

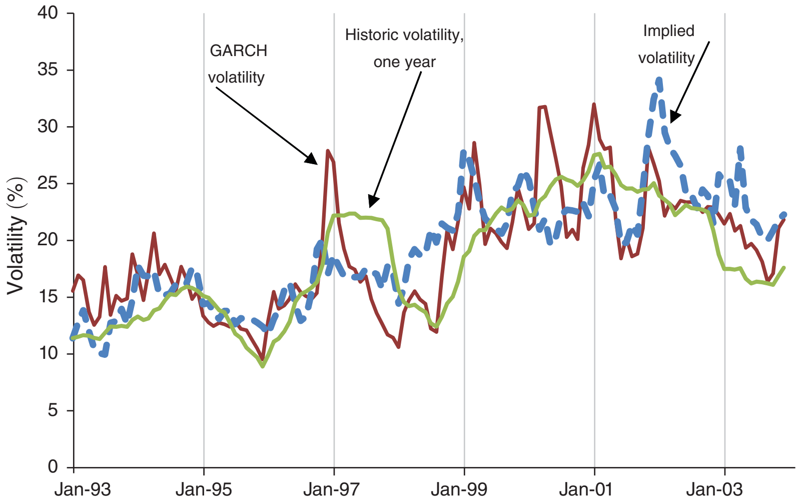


Figure 4. Comparison of Implied Volatility to the One-Year Historic Volatility of the Future Price and a GARCH(1, 1) Forecast

Notes: Implied volatility is calculated from futures options prices as described in Section ID. Historic volatility at any point in time is the standard deviation of the return on the 18-month futures price within a 1-year rolling window. The GARCH(1, 1) model is estimated at each date using a 4-year rolling window of 18-month futures prices.

volatilities is given in Figure 1 alongside the time series of 18-month futures prices (both series are averages of daily observations within each month).

In secondary empirical specifications, I construct volatility forecasts using historic futures price volatility rather than implied volatility derived from futures options. These specifications address the possibility that oil production firms' volatility forecasts differ from those of the market. One possible forecast is a no-change forecast; that is, the expected future volatility of the NYMEX futures price is its recent historic volatility. Figure 4 compares the historic volatility of the futures price, measured over a rolling window of one year, to the implied volatility series. Historic volatility sometimes deviates substantially from implied volatility: it is relatively high in 1997 and low in 1998, and it does not reflect the implied volatility spikes in 1999 and September 2001.

I have also forecast volatility using a GARCH(1, 1) model. For each date in the dataset, I estimate the GARCH parameters using a four-year rolling window of daily 18-month futures prices.[12] At each date, I then use the estimated GARCH model to forecast volatility over the upcoming month. The average forecasted volatility over this month is then used as the measure of firms' expected price volatility. Figure 4 plots the series of GARCH volatility forecasts. These GARCH forecasts align more closely with the implied volatilities than do simple historic volatilities, though the

[12] In the GARCH model, the mean price equation is a seventh-order autoregression; this number of lags is necessary to eliminate serial correlation in the price residuals. A GARCH(1, 1) process is then sufficient to eliminate conditional heteroscedasticity in the residuals (the *p*-value for rejecting a null hypothesis of no conditional heteroscedasticity is 0.423).

GARCH and implied volatility time series still differ substantially at various points, most notably 1997–1998 and late 2001.

E. *Drilling Costs*

The primary source for information on drilling costs is RigData, a firm that collects data on daily rental rates ("dayrates") for drilling rigs from surveys of drilling companies and publishes these data in its *Day Rate Report* (1990–2005).[13] Rig rental comprises the single largest line-item in the overall cost of a well, and industry sources have suggested that at typical dayrates rig rental accounts for one-third of a well's total cost.[14] Because I observe dayrates but not other components of drilling costs, I assume that non-rig costs are constant in real terms and equal to twice the rig rental cost at the average sample dayrate. This constant cost assumption seems reasonable over the 1993–2003 sample. Prices for steel, which factor into prices for drill pipe, bits, and well casing, were fairly stable over this time, nominally increasing by an average of 1.8 percent per year according to data from the Bureau of Labor Statistics. Other substantial components of cost, such as site preparation, construction, and general equipment rental (pumps, for example), should be based primarily on prices for non-specialized labor and capital inputs and therefore also be stable in real terms.[15] As for the assumption that these non-rig costs constitute two-thirds of total drilling costs on average, I explore the use of alternative ratios as robustness tests when estimating the model.

Because drilling rigs are pieces of capital that are specific to the oil and gas industry, rig rental rates are positively correlated with oil and gas prices and, accordingly, vary over the sample frame. For a well of average depth (5,825 feet in the sample), the dayrate ranges from \$5,327 to \$10,805, with an average of \$6,710.[16] Given an average drilling time of 19.2 days, the average rig rental cost for a well is therefore \$128,834, and average non-rig costs, estimated to be twice this amount, are \$257,667 (all figures in real US\$(2003)).

For each month in the sample, I calculate the total drilling cost of an average well as the sum of 19.2 days times the prevailing dayrate for that month (in real terms) with average non-rig costs. The time series of drilling costs for an average well is

[13] The oil production firms that hold leases, make drilling decisions, and are the focus of this study do not actually own the drilling rigs that physically drill their wells. Rigs are instead owned by independent drilling companies that contract out their drilling services. See Kellogg (2011) for further information regarding the contracting process between production firms and drilling companies.

[14] This one-third figure was suggested by RigData and substantiated by information from the Petroleum Services Association of Canada's (PSAC's) Well Cost Study (summers of 2000 through 2004). This study provides a breakout of the costs of drilling representative wells across Canada during the summer months. For the non-Arctic, non-offshore areas that most closely resemble conditions in Texas, rig rental costs averaged 35.2 percent of total costs.

[15] Evidence in support of this claim is available from the 2002, 2003, and 2004 PSAC Well Cost Studies, during which time the specifications for the representative wells were essentially unchanged. These data indicate that non-rig drilling costs changed, on average, by only −0.2 percent in 2003 and +3.1 percent in 2004. Rig-related drilling costs, however, increased by 9.8 percent in 2003 and 30.9 percent in 2004, following increases in the price of oil.

[16] RigData reports dayrates separately for rigs drilling wells between 0 and 5,999 feet deep and for rigs drilling wells between 6,000 and 9,999 feet deep. The dayrates used in this study are the average of these two depth classes for the Gulf Coast/South Texas region. The RigData dataset is quarterly and continuously reported from 1993 onward. Because I conduct my analysis at a monthly level, I generate monthly dayrate data by assigning each quarterly reported dayrate to the central month of each quarter and then linearly interpolating dayrates for the intervening months. The alternative approach of simply treating dayrates as constant within each quarter has only a minor effect on the estimated results.

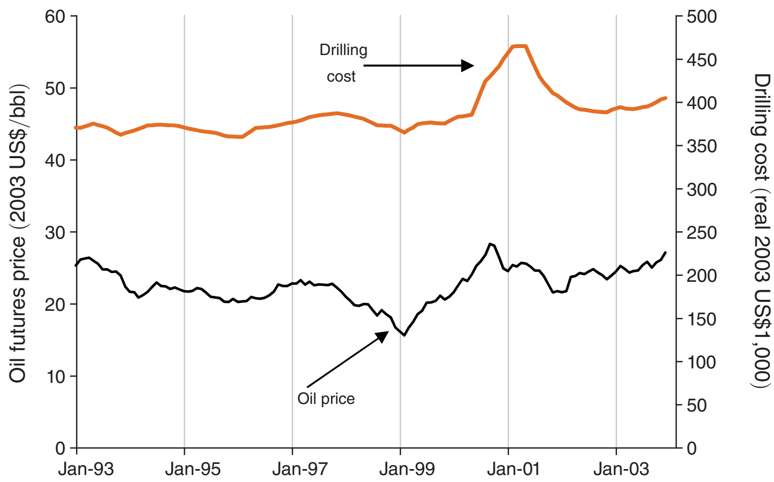


Figure 5. Drilling Costs and Oil Futures Prices

Notes: Oil futures prices are 18-month-ahead prices from the NYMEX. Drilling costs are those for an average well that requires 19.2 days of drilling time. These costs are based on daily rig rental rates obtained from RigData, as discussed in Section IE.

plotted alongside oil futures prices in Figure 5. The positive correlation between these two series is readily apparent.

II. Descriptive Results

Figure 1 plots the three time series of primary data: drilling activity, oil futures prices, and implied oil price volatility from futures options. Several features of the plot are worth noting. First, drilling activity rises and falls with the oil price. In particular, the oil price crash of 1998–1999 that was driven by the Asian financial crisis (Kilian 2009) is associated with a sharp reduction in drilling activity. Second, following the 1998–1999 price crash, oil prices rapidly recovered and by the beginning of 2000 actually surpassed their pre-1998 levels. However, oil drilling did not enjoy a similar recovery: activity did increase once prices began to rise in the summer of 1999 but recovered only to approximately two-thirds of its pre-1998 level. Why did drilling activity not reach its earlier level despite such a high oil price? The third line on the graph—implied volatility—suggests that an increase in uncertainty following the 1998 price crash may have caused producers to delay the exercise of their drilling options. Implied volatility increases sharply in 1998 and remains at an elevated level for the remainder of the sample; this high level of volatility is associated with the period in which expected oil prices were high yet drilling activity was low. Moreover, several positive shocks to volatility subsequent to 1999, such as the volatility spike following September 11, 2001, appear to be associated with reductions in drilling activity.

A descriptive statistical analysis using a hazard model confirms that the negative relationship between drilling and expected price volatility that is apparent in Figure 1 is in fact statistically significant. The unit of observation in this analysis is an individual drilling prospect, and I model 7,787 such prospects: the 1,150 observed infill wells plus one prospect for each of the 6,637 sole-operated fields in which I observe no drilling activity. In doing so, I treat prospects that exist within the same field as independent of one another. While this treatment does not allow for the modeling of factors that might cause wells within the same field to be drilled at nearly the same time, the fact that most fields have zero or one well suggests that the impact of modeling all drilling decisions independently of one another may be minor.

I choose a hazard model, rather than an OLS regression of drilling investment on expected price and volatility that would be more typical of both the macro and micro real options literatures, to capture the idea that drilling activity should decline over time as the set of available options is gradually reduced through drilling.[17] In the simplest possible model, I model the hazard rate $\gamma(t)$ as an exponential function of the expected future price level and expected price volatility per (1) below.

$$(1) \qquad \gamma(t) = \exp(\beta_0 + \beta_p \cdot Price_{t-3} + \beta_v \cdot Vol_{t-3}).$$

In estimating both this model and the structural model described below, I lag all covariates by three months, as industry participants have indicated that the engineering, permitting and rig contracting processes generally drive a three month wedge between the decision to drill and the commencement of drilling. For inference, I use a "sandwich" variance-covariance matrix estimator that allows arbitrary within-field correlation of the likelihood scores (Wooldridge 2002).[18] In practice, this estimator increases the estimated standard errors by about 25 percent, on average, relative to the standard BHHH estimator.

The results of estimating (1) are presented in column 1 of Table 1. A $1 increase in the expected future price of oil is associated with an increase in the likelihood of drilling of 4.1 percent, and a one percentage point increase in expected price volatility is associated with a decrease in the likelihood of drilling of 3.0 percent. Both of these point estimates are statistically significant at the 1 percent level. Column 2 includes the cost of drilling as an additional covariate and finds that drilling costs are negatively associated with drilling, as expected (though the relationship is not statistically significant). Oil prices and volatility continue to be positively and negatively, respectively, associated with drilling in this specification. Columns 3 and 4 show that these correlations are robust to allowing for unobserved prospect-specific heterogeneity and a time trend.

Because these descriptive results, in the absence of an economic model, cannot speak to the optimality of firm decision-making or welfare, the remainder of this

[17] Nonetheless, results from an OLS regression are similar to those from the hazard model. When I regress the log of the number of wells drilled each month on the oil price, volatility, and drilling cost, I obtain estimated coefficients of 0.082, −0.035, and −0.398, respectively. These coefficients are all similar, statistically and economically, to those from the analogous hazard model regression in column 2 of Table 2 (in which percent impacts are measured away from one rather than zero).

[18] Wooldridge (2002) shows that this approach, which is analogous to clustering in linear regression models, still produces consistent estimates of the parameters even though serial and cross-well correlation within each field is not explicitly accounted for in the likelihood function. I also use this approach when estimating the structural model, discussed in Sections III through V. I have also estimated these models while clustering the standard errors on time rather than field to account for cross-sectional correlation of the likelihood scores that might arise from technological or macroeconomic shocks. These estimated standard errors are generally similar to those obtained from the standard BHHH estimator.

TABLE 1—HAZARD MODEL RESULTS FOR THE PROBABILITY OF DRILLING

Coefficient on covariate	Basic exponential hazard (1)	Include drilling cost (2)	Prospect-specific heterogeneity (3)	Drilling cost and time trend (4)
Oil futures price ($/bbl)	1.041***	1.056***	1.056***	1.055***
	(0.016)	(0.021)	(0.021)	(0.021)
Implied volatility of future price (percent)	0.969***	0.976**	0.976**	0.967**
	(0.008)	(0.010)	(0.010)	(0.013)
Drilling cost ($100,000)	—	0.754	0.754	0.716
	—	(0.175)	(0.175)	(0.170)
Linear time trend (in years)	—	—	—	1.019
	—	—	—	(0.021)
Unobserved heterogeneity (inverse Gaussian distribution)	No	No	Yes	No
log likelihood	−3,979.5	−3,978.5	−3,978.5	−3,977.9

Notes: Reported coefficients are hazard ratios: the multiplicative effect on the hazard rate of a one unit increase in the covariate. All estimates use prices of futures and options that are 18 months from maturity. All covariates are lagged by three months. Standard errors are estimated using a sandwich estimator that allows for correlation of the likelihood scores across wells within the same field, thereby accounting for spatial and serial correlation. Significance levels are for a two-tailed test that the coefficient is different from one.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

paper focuses on formulating and estimating a model of the infill drilling problem faced by oil production companies in Texas. The primary goal of this model is to relate firms' observed responses to changes in uncertainty to the theoretically optimal response.

III. A Model of Drilling Investment Under Time-Varying Uncertainty

A. *Model Setup*

Consider a risk-neutral, price-taking oil production firm that is deciding whether to drill some prospective well i at date t. Using geologic and engineering estimates, the firm generates an expectation regarding the monthly oil production from the well should it be drilled. The present value of the well's expected revenue is then equal to the sum, over the months of the well's productive life, of the product of the well's expected monthly production with the expected oil price each month, net of taxes and royalties, and discounted at the firm's discount factor δ. Rather than model this discounted sum explicitly, I model it instead as simply the product $r_i P_t$. Here, r_i represents the sum of the well's expected monthly production, net of taxes and royalties, and discounted so that it is in present value terms.[19] P_t represents the "average"

[19] A narrow view of r_i suggests that I am assuming that the ongoing production from any previously drilled wells in the same field as well i is unaffected by the drilling of well i. This assumption is incorrect if the new well is, at least to some extent, only accelerating the recovery of reserves from the field rather than exploiting new reserves that the existing well stock did not reach. However, the model can handle wells drilled with the purpose of acceleration by interpreting the expected productivity r_i as the expected production of the new well net of its expected impact on the production from the existing well stock (if any).

oil price that will prevail over all barrels of oil expected to be produced by the well, so that the product $r_i P_t$ is equal to the original discounted sum of monthly revenue. In the estimation that follows, I will use the 18-month futures price of oil as P_t. This simplification allows me to model the price level using only the single state variable P_t rather than a vector of state variables for the expected price in each month of the well's productive life.

I emphasize that $r_i P_t$ is the firm's expectation of the value that will be obtained from drilling. Realized value may differ substantially from $r_i P_t$ because the realized oil price may differ from P_t (though the firm could hedge this risk away) and because realized production may differ from r_i. Recall that some of the wells observed in the sample yielded zero oil production. Clearly, a dry hole was not the firms' expected outcome for these wells.

In month t, the well's drilling cost is equal to the sum of non-rig costs c_i with the product of the dayrate D_t and the number of days d_i required to drill the well.[20] Then, given an expected oil price P_t and a dayrate D_t, the expected profits π_{it} from drilling the well are given by the function π_i:

$$\pi_{it} = \pi_i(P_t, D_t) = r_i P_t - c_i - d_i D_t. \tag{2}$$

It will be useful for estimation to rearrange (2), defining the expected productivity of a well as the ratio of its expected production r_i to its drilling cost at the average dayrate. Denote this cost by $\overline{C}_i = c_i + d_i \overline{D}$ and let this ratio be denoted by x_i. Further, let $\tilde{c}$ denote $c_i/\overline{C}_i$ and let $\tilde{d}$ denote $d_i/\overline{C}_i$. Assuming that the ratio of non-rig costs to total costs at the average dayrate is constant across wells implies that both $\tilde{c}$ and $\tilde{d}$ are constant across wells (in the reference case model, I set $\tilde{c} = 2/3$ and $\tilde{d}\overline{D} = 1/3$ per the discussion in Section IE). Then, expected profits π_{it} can be re-written as (3) below, in which all cross-well productivity heterogeneity relevant to the drilling timing decision is collapsed into the single variable x_i.

$$\pi_{it} = \pi_i(P_t, D_t) = \overline{C}_i(x_i P_t - \tilde{c} - \tilde{d} D_t). \tag{3}$$

I treat all firms as price takers, in the sense that they believe that their decisions do not impact P_t or D_t. This assumption almost certainly holds institutionally. The market for oil is global, and Texas as a whole constitutes only 1.3 percent of world oil production. With respect to oil producers' monopsony power in the market for drilling services, the largest firm in the dataset is responsible for only 2.2 percent of all wells drilled in Texas during the sample period, a quantity that seems insufficient for exertion of substantial market power.

Let the processes by which firms believe the price of oil and rig dayrates evolve be first-order Markov and given by (4) and (5) below. P_t denotes the oil price (18-month

[20] I assume that d_i does not vary over time. Learning-by-doing could cause d_i to decrease as more wells are drilled in the field (Kellogg 2011); however, since most of the observed sole-operated fields have only a few new wells during the sample, this effect is likely to be negligible. Technological progress might also decrease d_i over time; this possibility is part of the motivation for allowing for a time trend in an alternative specification.

NYMEX future) in the current month t, and P_{t+1} is the price in month $t+1$. D_t and D_{t+1} represent the current and next month's dayrates.[21]

$$\ln P_{t+1} = \ln P_t + \mu(P_t, \sigma_t^2) - \sigma_t^2/2 + \sigma_t \varepsilon_{t+1} \tag{4}$$

$$\ln D_{t+1} = \ln D_t + \hat{\mu}(D_t, \hat{\sigma}_t^2) - \hat{\sigma}_t^2/2 + \hat{\sigma}_t \hat{\varepsilon}_{t+1}. \tag{5}$$

The firm's current expectation of the volatility of the oil price is denoted by σ_t, and the price shock ε_{t+1} is an i.i.d. standard normal random variable that is realized subsequent to the firm's drilling decision in the current period. Because I do not observe expectations of dayrate volatility $\hat{\sigma}_t$, I assume this volatility is a scalar multiple of the oil price volatility so that $\hat{\sigma}_t = \alpha\sigma_t$. The cost shock $\hat{\varepsilon}_{t+1}$ is drawn from a standard normal that has a correlation of ρ with ε_{t+1}.

$\mu(P_t, \sigma_t^2)$ and $\hat{\mu}(D_t, \hat{\sigma}_t^2)$ denote the expected price and drilling cost drifts as stationary functions of the current expected level and volatility of the oil price and dayrate. Dependence of these drifts on the price and dayrate levels allows for the mean reverting expectations exhibited by NYMEX futures prices (Figure 3). I also allow the drifts to depend on volatility because, as pointed out by Pindyck (2004), an increase in volatility may increase the marginal value of storage and therefore raise near-term prices. In addition, a volatility increase may also affect investments related to oil production and consumption (via the real options mechanism considered here, for example), affecting expectations of future prices. The specification and estimation of $\mu(P_t, \sigma_t^2)$ and $\hat{\mu}(D_t, \hat{\sigma}_t^2)$ is discussed in Section IVA, where I also discuss the estimation of the correlation of oil price shocks ε_{t+1} with dayrate shocks $\hat{\varepsilon}_{t+1}$.

B. *Optimal Drilling with Time-Varying Volatility*

The firm's problem at a given time t is to maximize the present value of the well V_{it} by optimally choosing the time at which to drill it. This optimal stopping problem is given by (6) below, in which Ω denotes a decision rule specifying whether the well should be drilled in each period $\tau \geq t$ as a function of P_τ and D_τ (conditional on the well not having been drilled already). I_τ denotes a binary variable indicating the outcome of this decision rule each period, and δ denotes the firm's real discount factor.

$$V_{it} = \max_{\Omega} E\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} I_\tau \pi_i(P_\tau, D_\tau)\right]. \tag{6}$$

In formulating (6), I assume that firms holding multiple drilling options treat them independently of one another. Given that I only observe zero or one well drilled in most fields in the sample, this assumption does not seem particularly strong. In those cases in which a firm holds multiple drilling options within the same field, it may be that the outcome from drilling one well may convey information regarding other prospects. That is, if the first well drilled by a firm in a field turns out to be highly

[21] These transition functions are the discrete time analogue to geometric Brownian motion with drift (see Dixit and Pindyck 1994). Volatility is assumed to be constant within each time step.

productive, the firm may increase its estimate of x_i for its remaining prospects.[22] This contingent reevaluation will result in temporal clustering of drilling activity in multi-well fields relative to what would be predicted by (6) alone.

Because drilling a well is irreversible and because future prices and costs are uncertain, the decision rule for maximization of (6) is not simply to invest in the first period in which $\pi_{it} \geq 0$. The firm must trade off the value of drilling immediately against the option value of postponing the investment to a later date, at which time the expected oil price may be higher or the drilling cost lower. This trade-off is captured by restating the optimal stopping problem as the Bellman equation (7) below, in which V_i represents the current maximized value of the drilling option as a function of the state variables P, D, and σ (from which I now remove the subscript t). P', D', and σ' denote next period's state.

$$V_i(P, D, \sigma) = \max\{\pi_i(P, D), \delta \cdot E[V_i(P', D', \sigma')]\}. \tag{7}$$

Equation (7) includes the firm's expected oil price volatility σ as a state variable even though it does not appear in the profit function $\pi_i(\cdot)$. Volatility impacts drilling decisions through its impact on the distribution of next period's expected oil price P' given the current expected price P. An increase in σ increases the variance of P' conditional on P, thereby increasing the value of holding the drilling option relative to the value of drilling immediately.

Intuition suggests that the solution to (7) will be governed by the following "trigger rule:" at any given P, D, and σ, there will exist a unique $x^*(P, D, \sigma)$ such that it will be optimal to drill prospect i if and only if $x_i \geq x^*(P, D, \sigma)$. Furthermore, x^* will be strictly decreasing in P and strictly increasing in D and σ. The following conditions on the stochastic processes governing the evolution of P, D, and σ (none of which is rejected by the data) are sufficient for this trigger rule to hold. S denotes the state space.

(i) $\delta E[P' | P, D, \sigma] < P \;\; \forall P, D, \sigma \in S$ (oil prices cannot be expected to rise more quickly than the rate of interest).

(ii) $\frac{\partial E[P' | P, D, \sigma]}{\partial P} < \frac{1}{\delta}$, with the same holding for D and σ, $\forall P, D, \sigma \in S$ (the expected rates of change of each state variable cannot increase too quickly with the current state).

(iii) $\rho < 1$ (oil price shocks and dayrate shocks are not perfectly correlated).

(iv) The distribution of P' is stochastically increasing in P, with the same holding for D and σ.

[22] The process by which firms learn about the quality of fields through drilling is examined by Levitt (2011), which develops and estimates a dynamic learning model. That paper's approach cannot be used here because it requires data on oil production outcomes for all drilled wells and because the separate identification of learning effects and location-specific heterogeneity requires observations of different firms drilling wells in the same field (as well as an assumption of no cross-firm information spillovers).

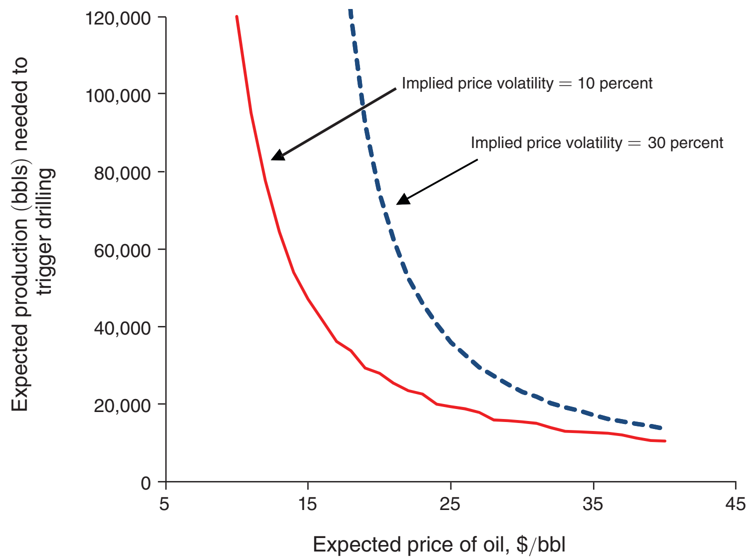


Figure 6. The Impact of Expected Oil Price Volatility on the Expected Production Required to Trigger Drilling

Notes: The relationships shown are for a well of average cost facing an average dayrate, so that the drilling cost is $386,501. The model used to generate these curves uses the state transition parameters estimated in Section IVA.

(v) $\delta E[\pi(P', D', \sigma')|P, D, \sigma] < \pi(P, D, \sigma)\ \forall P, D, \sigma \in S$ (the Hotelling condition necessary for drilling to be optimal: expected profits cannot rise more quickly than the rate of interest).

It is straightforward to show that conditions (i)–(iii) imply that $\pi(s) - E[\pi(s'|s)]$ is strictly increasing in P and x_i, and strictly decreasing in D and σ. Given this result and conditions (iv) and (v), a fixed point contraction mapping argument given in Dixit and Pindyck (1994) proves that the trigger $x^*(P, D, \sigma)$ must exist. There must also exist similar triggers $P^*(D, \sigma, x_i)$, $D^*(P, \sigma, x_i)$, and $\sigma^*(P, D, x_i)$, representing the minimum price, maximum drilling cost, and maximum volatility at which drilling is optimal as functions of the other variables. The existence of all four triggers implies that $x^*(P, D, \sigma)$ must be strictly decreasing in P and strictly increasing in D and σ.

Thus, an increase in expected volatility σ will cause a fully optimizing firm to increase the productivity trigger x^* necessary to justify investment, holding the expected price and dayrate constant. Consider such a firm for which the price volatility expectation σ is equal to the volatility implied by the futures options market, which I denote by σ^m. Figure 6 illustrates how the firm's critical productivity x^* will vary with P and σ^m for a well with an average drilling cost at the average sample dayrate. The relationship between x^* and P is shown at both low (10 percent) and high (30 percent) levels of expected price volatility σ^m. At both volatility levels, x^* decreases with price so that less productive wells may be drilled in relatively high price environments. Holding price constant, x^* is greater in the high volatility case than the low volatility case.

Now, however, suppose that firms have time-varying expectations about future volatility that coincide with those of NYMEX market participants but do not take these expectations into account when making drilling decisions, so that in terms of the model σ is effectively constant over time. In this case, the two lines in Figure 6 will coincide. It is this difference in investment behavior between firms that respond to σ^m and those that do not that will provide identification in the empirical exercise described below. Note, however, that an observed lack of response to σ^m could also reflect the possibility that, while firms properly take expected volatility into account when making investment decisions, they hold a belief that volatility σ is constant over time rather than equal to the time-varying σ^m. Thus, to the extent that the data imply differences between σ and σ^m, I will not be able to identify whether the differences are due to suboptimal investment decision-making or to differences between firms' beliefs and those of the broader market.

I capture the extent to which firms optimally respond to the market's implied volatility σ^m by parameterizing firms' beliefs through a behavioral parameter β. First, I define $\overline{\ln \sigma}$ to be the average log of the market volatility over the first year of the sample (12.8 percent) and let $\ln \sigma^d$ be the deviation of $\ln \sigma^m$ from $\overline{\ln \sigma}$. That is,

$$\ln \sigma^m = \overline{\ln \sigma} + \ln \sigma^d. \tag{8}$$

I then relate the firms' expected volatility σ to σ^d via (9):

$$\ln \sigma = \overline{\ln \sigma} + \beta \ln \sigma^d. \tag{9}$$

Through this formulation, the behavioral parameter β regulates the extent to which firms respond to changes in σ^m. A firm that behaves according to $\beta = 1$ is a firm that shares the market's beliefs regarding future price volatility and correctly optimizes its investment decisions according to those beliefs. Conversely, a firm with $\beta = 0$ does not respond to changes in σ^m because it either has beliefs that are orthogonal to σ^m or does not optimize its investment decisions correctly. The primary objective of the empirical work is to obtain an estimate of β and test whether the estimate is consistent with investment behavior that is an optimal response to beliefs that coincide with those of the market.

The final component of the model is the process by which firms believe volatility itself evolves over time. Like the price and dayrate processes, I assume this process is first-order Markov per equation (10):

$$\ln \sigma^{m\prime} = \ln \sigma^m + \mu_\sigma(\sigma^m) - \gamma^2/2 + \gamma \eta'. \tag{10}$$

In (10), γ denotes the volatility of the volatility process, and η' is an i.i.d. standard normal random variable. $\mu_\sigma(\sigma^m)$ denotes the expected change in volatility as a function of the current volatility. This function can be set to zero to allow for random walk beliefs or it can be specified to allow for mean reversion, though in practice the distinction will not substantially impact the estimation results. I discuss the specification of $\mu_\sigma(\sigma^m)$ and the estimation of γ in Section IVA below. Note that, by specifying the volatility evolution process on market volatility σ^m, firms' beliefs about the evolution of their expected volatility σ will be scaled by β.[23]

IV. Empirical Model and Estimation

The parameter of primary interest is β, the behavioral parameter that reflects firms' sensitivity to the implied volatility of the price of oil. To obtain an estimate of β, I must also estimate the parameters α, ρ, and γ that govern the state transition processes as well as the oil price and dayrate drift functions $\mu(P_t, \sigma_t^2)$ and $\hat{\mu}(D_t, \hat{\sigma}_t^2)$. An estimate of the discount factor δ is also required. In what follows, I first discuss how I estimate these "secondary" parameters independently of the full model before turning to the estimation of β via a procedure based on the nested fixed point approach of Rust (1987).

A. *Estimates of the Discount Factor and State Transition Processes*

While the firms' discount factor δ can in principle be estimated as part of the nested fixed point routine, obtaining precise inference in practice is challenging. I adopt the standard approach in the literature by setting δ in advance. According to a 1995 survey by the Society of Petroleum Evaluation Engineers, the median nominal discount rate applied by firms to cash flows is 12.5 percent. Given average inflation over 1993–2003 of 2.36 percent, I set δ equal to the quotient 1.0236/1.125, approximately 0.910.

I assume that $\mu(P_t, \sigma_t^2)$, the expected drift of the log oil futures price, is the stationary linear function given by (11):

$$\mu(P_t, \sigma_t^2) = \kappa_{p0} + \kappa_{p1} P_t + \kappa_{p2} \sigma_t^2. \tag{11}$$

Per equation (4), consistent estimates of κ_{p0}, κ_{p1}, and κ_{p2} may be obtained via an OLS regression of $E[\ln P_{t+1}] - \ln P_t + \sigma_t^2/2$ on P_t and σ_t^2. Because the reference case specification uses 18-month futures prices for P_t, I use 19-month futures prices to measure $\mathrm{E}[\ln P_{t+1}]$ in this regression. I estimate that $\kappa_{p0} = 0.0094$, $\kappa_{p1} = -0.00054$, and $\kappa_{p2} = 0.401$. These values are consistent with mean reversion to an oil price of \$19.51 per barrel at the sample average volatility of 19.4 percent.

I similarly assume that $\hat{\mu}(D_t, \hat{\sigma}_t^2)$, the expected dayrate drift, is a linear function of the current dayrate, so that $\hat{\mu}(D_t, \hat{\sigma}_t^2) = \kappa_{d0} + \kappa_{d1} D_t + \kappa_{d2} \hat{\sigma}_t^2$. There does not exist a futures market for rig dayrates to facilitate estimation of the κ_d. Rather than attempt to estimate these parameters from a short time series of quarterly drilling cost observations, I instead assume that the parameters κ_{d0}, κ_{d1}, and κ_{d2} match those from the oil price drift equation, with κ_{d1} rescaled by the ratio of the average dayrate to the average oil price.

I next estimate the scalar parameter α, the ratio of dayrate volatility $\hat{\sigma}_t$ to oil price volatility σ_t. To do so, I first calculate $\xi_t = \ln P_t - \ln P_{t-1}$ and $\hat{\xi}_t = \ln D_t - \ln D_{t-1}$

[23] Specifically, firms' belief about the volatility of volatility will be $\beta\gamma$, and their belief about the expected drift of volatility will be $\beta\mu_\sigma(\sigma^m)$.

in each period. α is then estimated by the ratio of the standard deviation of $\hat{\xi}_t$ to the standard deviation of ξ_t. I then estimate ρ, the parameter governing correlation between dayrate and oil price shocks, to be the correlation coefficient between $\hat{\xi}_t$ and ξ_t. The estimate of α is 1.16, and that for ρ is 0.413.

Finally, I specify and estimate equation (10), which dictates how firms believe current implied volatility σ^m will evolve into next month's implied volatility. There do not exist markets that allow for direct measurement of either the expected change in volatility $\mu_\sigma(\sigma^m)$ or the volatility of volatility γ. I therefore rely on the time series of volatility realizations. For $\mu_\sigma(\sigma^m)$, I make use of two specifications: the first is a random walk volatility forecast ($\mu_\sigma(\sigma^m) = 0$), and the second is a mean-reverting expectation. Firms might reasonably have a random walk forecast because a unit root cannot be rejected in the implied volatility data.[24] Alternatively, the GARCH analysis of historic volatility realizations discussed in Section ID indicates modest mean reversion in expected future volatility. When the month-ahead GARCH forecast (the forecast used in generating the time series in Figure 4) is relatively high, the two-month ahead forecast predicts a fall in volatility, and the reverse holds when the month-ahead forecast is relatively low. Therefore, as an alternative to the random walk, I also estimate $\mu_\sigma(\sigma^m)$ using expected mean reversion rates derived from the GARCH forecasts. I estimate that, in a simple linear mean reversion model, the expected change in logged volatility is given by 0.0326 minus 0.0010 times the current volatility (in annualized percent).[25] This estimate implies that if the current volatility is 10 percent, then the expected volatility next month is 10.22 percent. In contrast, if the current volatility is 35 percent, then the expected volatility next month is 34.86 percent.

In the random walk specification, I estimate γ, the volatility of the volatility process, to be the standard deviation of $\ln \sigma_{t+1}^m - \ln \sigma_t^m$. This value is 0.119. Under the mean-reverting specification, I first correct $\ln \sigma_{t+1}^m - \ln \sigma_t^m$ for the predicted change in volatility between the two months before taking the standard deviation. This value is 0.118.

B. *Primary Empirical Model and Estimation*

Given the state transition functions estimated above, the remaining unknowns in the econometric model are the behavioral parameter β and the unobserved expected productivity of each drilling prospect, the x_i. Given a value for β and the realized oil prices, dayrates, and oil price volatilities, the solution to the model determines the productivity cutoffs x_t^* each period. Because all firms face the same price, volatility, and dayrate processes, x_t^* will be the same for all prospects in the data at any given time. If x_i is modeled as identical across prospects, then all firms would make the decision to drill at the same time, a prediction that conflicts with the spread of drilling activity over time evident in Figure 1. Clearly, there must exist a distribution of x_i across prospects.

[24] In an augmented Dickey-Fuller test, the p-value for rejecting the null of a unit-root process with 12 lags is 0.2417.

[25] To obtain this estimate, I regress, for each month for which four years of data are available to estimate the GARCH model (July 1994–December 2003), the difference between the two-month and one-month GARCH volatility forecasts on the one-month GARCH forecast. I then correct the constant term by γ squared divided by two.

It is therefore tempting, at first, to estimate a model in which expected productivity x_i varies across prospects but for each individual prospect is constant over time. However, this model is also incapable of rationalizing the data. Given the trigger rule described in Section III, such a model implies that in each period t all wells with productivity $x_i > x_t^*$ will be drilled. Now consider what would happen should x^* rise in period $t+1$, perhaps because the oil price fell or because volatility increased. In this case, only prospects with $x_i \geq x_{t+1}^*$ will be drilled. However, all such prospects will already have been drilled in period t since $x_{t+1}^* > x_t^*$. Thus, an implication of a model in which x_i does not vary over time is that there cannot be any drilling activity following an increase or no change in x^*. Such a model is clearly inconsistent with the drilling data. In 1999, for example, the expected price is considerably lower than it was in 1998 and the expected volatility is higher; however, drilling activity does not go to zero.

To fit the data, the model requires changes in the x_i over time or some other mechanism to smooth out drilling activity. That said, the true productivity of each prospect is constant, since the underlying geology is time-invariant. Moreover, new information that would cause a firm to update its expectation of a prospect's productivity does not arrive exogenously on its own, since the prospective reservoir is buried thousands of feet below the surface. Nonetheless, there exist several possible mechanisms capable of explaining the drilling data. First, a firm's evaluation of any given prospect's productivity is likely not stable over time, even if the information set does not change. The process by which geologists and engineers develop an estimate of a prospective well's production is inherently challenging and error-prone. They must make inferences about a buried oil reservoir using only limited information from seismic surveys, production outcomes from previously drilled wells, and electromagnetic "logs" of the rock characteristics at nearby wells. Any individual geologist or engineer may change his or her views regarding a prospect as more time is spent studying the information, and different personnel may draw different conclusions from the same set of information (much like different econometricians may draw different inferences from the same data). Such re-evaluations of prospects, which are anecdotally common in the industry, particularly if there is personnel turnover, can drive substantial variation in a prospect's x_i over time. Second, firms may sometimes "discover" new prospects in old fields in their analyses of their data. Observationally, such discoveries are equivalent to an increase in the x_i of what had been a low-quality prospect. Third, firms can engage in costly gathering of information by, for example, taking a seismic survey of their field, thereby triggering a revision of expected productivity. Similarly, the results from the drilling of one well may yield information about the quality of another prospect. Finally, variance in the lag between the decision to drill and the actual commencement of drilling may arise due to delays in engineering design, management approval, permitting, or drilling contracting. These stochastic lags will lead to drilling at times not predicted by the model.

Rather than separately model each of the mechanisms above, I instead account for them jointly by allowing each well's expected productivity x_i to vary over time. In the absence of data on firms' engineering estimates, their use of seismic surveys, or well-specific delays in drilling, separate identification of each source of time variation would require strong functional form assumptions and a substantially more complex model than that given here.[26] Allowing the x_i to vary over time is closest in spirit to the prospect reevaluation mechanism and is sufficient for the model to predict drilling activity following an increase in the productivity cutoff x^*.

In the reference case empirical specification, I treat the log of expected productivity as an i.i.d. normal variable across both prospects i and time t, with a mean μ and variance ζ that are to be estimated in the main estimation procedure (along with the behavioral parameter β). Given μ and ζ, the number of wells that the model predicts will be drilled at any time t is given by the expected number of prospects for which x_{it} exceeds x_t^*.

Despite the emphasis of the above discussion on time variance in x_{it}, there may exist some persistent cross-sectional heterogeneity in the expected productivity of each prospect. I therefore also consider a model in which log x_{it} is the sum of a time-invariant normally distributed random variable φ_i, with mean and standard deviation given by μ_1 and ζ_1, and an i.i.d. normal variable ν_{it} with a zero mean and standard deviation ζ_2. In this specification, I estimate μ_1, ζ_1, and ζ_2 in addition to β.

Note that I do not explicitly model expected productivity as a stochastic state variable when solving the Bellman equation (7). That is, I do not model firms as anticipating or waiting for future changes in x_i, so there is no option value in this dimension. I do so because the mechanisms discussed above that drive changes in the x_i over time do not reflect exogenous inputs of new information or actual changes in the underlying geology, so that firms do not anticipate and wait for exogenous productivity shocks (unlike, for example, the Rust (1987) model in which Harold Zurcher receives new information about the state of each bus when it arrives in his shop).[27] This modeling choice is supported by the data. If I instead structure the model so that firms do anticipate productivity shocks, the model's ability to match actual drilling behavior is substantially reduced because firms have a very strong incentive to wait for a large positive shock and then drill as soon as one is realized. This incentive effectively mutes firms' incentive to respond to uncertainty about future oil prices, so that the model fails to predict the response of drilling to oil price volatility that is apparent in the data.[28]

Given the state transition processes discussed in Section IVA, the parameters governing the distribution of the x_{it}, the behavioral parameter β, and the realized monthly time series of futures prices, rig dayrates, and implied volatilities (denoted by $\mathbf{P}$, $\mathbf{D}$, and $\boldsymbol{\sigma}$, respectively), the model's solution yields the likelihood that a given prospect

[26] For example, a firm that undertakes a seismic survey is, in reality, making an endogenous investment that should in principle be modeled dynamically in conjunction with the drilling model. The present model can, however, accommodate costly information gathering to the extent that drilling a well can be viewed as a compound investment: when prices rise or volatility falls so that the firm contemplates drilling, it undertakes a seismic survey before drilling the well. I also continue to model each prospect independently, abstracting away from the process by which the drilling of a well in a field can influence the firm's beliefs about other prospects in the same field. This abstraction may result in unmodeled correlation of drilling activity in fields with multiple wells drilled, motivating the use of a clustered variance-covariance estimator (Wooldridge 2002).

[27] In the context of firms' revisions of a prospect's expected productivity based on a reanalysis of preexisting information, thinking of firms as anticipating and waiting for future shocks to expected productivity seems problematic because neither the underlying true productivity of the prospect nor the available information will actually change. The possibility that an engineer might, in the future, positively update the evaluation of the prospect doesn't actually increase the prospect's value, so firms shouldn't wait for such reevaluations.

[28] The estimated variance over time in expected productivity is sufficiently large (in order to rationalize drilling during low price periods) that it swamps oil price volatility. When I model firms as anticipating this variation, increasing β to even 2.0 has essentially no effect on simulated drilling.

will be drilled in any given month t conditional on not having been drilled already. This likelihood is simply the probability that x_{it} exceeds the trigger productivity x_t^*. Starting from the initial period of January 1993, these conditional probabilities yield the probability that any given prospect will be drilled in each month t as well as the probability that the prospect will not be drilled by the end of the sample.[29] These probabilities form the basis for the likelihood function. Let I_{it} denote an indicator variable that takes on a value of one if prospect i is drilled in month t and zero otherwise, let T denote the final month of the sample, let N_t denote the number of wells actually drilled at t, and let N_0 denote the number of prospects not drilled ($N_0 = 6{,}637$, the number of undrilled sole-operated fields).[30] The log-likelihood function is therefore:

$$(12) \qquad \begin{aligned} &\ell((N_1, N_2, \ldots, N_T), N_0 \,|\, \mathbf{P}, \mathbf{D}, \boldsymbol{\sigma}; \beta, \mu, \zeta) \\ &\quad = \sum_{t=1}^{T} N_t \log \Pr(I_{it} = 1 \,|\, \mathbf{P}, \mathbf{D}, \boldsymbol{\sigma}; \beta, \mu, \zeta) \\ &\qquad + N_0 \log \Pr(I_{it} = 0 \;\; \forall t \,|\, \mathbf{P}, \mathbf{D}, \boldsymbol{\sigma}; \beta, \mu, \zeta). \end{aligned}$$

Estimation of β, μ, and ζ is carried out by maximizing this likelihood function using a nested fixed point routine. The outer loop searches over the unknown parameters while the inner loop solves the model and calculates the likelihood function at each guess. Details regarding this simulation and estimation procedure, such as the discretization of the state space used to numerically solve the model, are provided in online Appendix 2. The specification with cross-sectional heterogeneity proceeds by integrating the likelihood over the distribution of φ_i.

V. Estimation Results and Discussion

A. *Reference Case Estimation Results*

I begin by estimating the version of the model in which $\log x_{it}$ is assumed to be i.i.d. across prospects i and time t. As a baseline, column 1 of Table 2 provides the estimation results when I impose the restriction that $\beta = 0$; that is, firms do not respond to changes in implied volatility.[31] I find that a broad distribution of expected productivity x_{it} is needed to sufficiently smooth the model's simulated drilling activity such that it rationalizes the data. The estimated mean μ and standard deviation ζ of $\log x_{it}$ are -0.431 and 3.023, respectively. Here, and throughout the presentation of the results, x_{it} is given in barrels of expected discounted production per \$100,000 of drilling cost at the average rig dayrate. These estimates together imply that, in

[29] For example, the probability that the prospect will be drilled in February 1993 is the conditional probability that it is drilled in February 1993 multiplied by the probability that it was not drilled in January 1993. The probability that it is drilled in March 1993 is then the conditional probability that it is drilled in March 1993 multiplied by the probability that it was not drilled in February 1993 or earlier, and so on.

[30] Throughout this section, I use "drilled" as shorthand for the drilling decision. As with the descriptive hazard model, I allow for a three-month lag between the drilling decision and the actual start of drilling. Thus, for example, the model's drilling probability for January 1993 is matched with drilling activity for April 1993. The final period of the sample is September 2003, which is matched with drilling activity for December 2003.

[31] With $\beta = 0$, the random walk and mean reverting specifications for future volatility are equivalent.

Table 2—Results from Estimation of the Dynamic Model

Parameter	Beta restricted to zero (1)	Reference case, random walk volatility forecast (2)	Reference case, mean reverting volatility (3)	Small firm subsample (4)	Large firm subsample (5)
β (sensitivity to volatility)	—	1.118	1.191	1.136	1.085
	—	(0.141)	(0.188)	(0.127)	(0.212)
μ (mean of $\log(x_{it})$)	−0.431	−12.045	−12.616	−18.991	−6.998
	(2.668)	(7.993)	(8.183)	(15.025)	(9.068)
ζ (SD of $\log(x_{it})$)	3.023	6.961	7.212	9.268	5.282
	(0.887)	(2.664)	(2.733)	(4.991)	(3.039)
log likelihood	−8,670.5	−8,661.2	−8,661.3	−4,350.9	−4,309.7

Notes: All estimates use prices of futures and options that are 18 months from maturity. x_{it} is expressed as expected oil production (in bbl) divided by the cost of drilling (in \$100,000) at the average sample dayrate. Drilling data are matched to drilling likelihoods with a three month lag. Columns 4 and 5 assume that firms have a random walk forecast for future volatility. Standard errors are estimated using a sandwich estimator that allows for correlation of the likelihood scores across wells within the same field, thereby accounting for spatial and serial correlation. Standard errors also account for the sampling error in the estimated function for the expected drift of future oil prices.

the model, the average prospect at any point in time is expected to produce only 63 barrels of oil per \$100,000 of cost, well below the productivity necessary to justify investment at any reasonable oil price.[32] This estimate reflects the presence of a large number of fields in the data (6,637) in which no drilling occurs. A large estimate of the variance ζ is therefore necessary to rationalize the observed drilling. For example, a prospect with average costs and a log x_{it} 3.5 standard deviations greater than the mean will be expected to produce 25,578 barrels of oil, sufficient to trigger drilling over a range of prices and implied volatilities in the sample.

In column 2, I allow β to be a free parameter and model firms as having random walk beliefs about future volatility. I obtain a point estimate of β of 1.118. This value is very close to one in both an economic and statistical sense (the standard error is 0.141),[33] consistent with an optimal response of investment to volatility expectations that match the implied volatility of NYMEX futures options.[34] Moreover, a likelihood ratio test strongly rejects, with a p-value less than 0.001, a null hypothesis that firms do not respond at all to implied volatility ($\beta = 0$).[35] The time series of predicted drilling under models I and II are given in Figure 7, alongside actual drilling activity.[36] The prediction from model II, allowing for a response to volatility,

[32] The 63 barrel per \$100,000 figure is equal to $\exp(-0.431 + 3.023^2/2)$.

[33] Standard errors are clustered on field (thereby accounting for within-field spatial and serial correlation) and take into account the standard errors of the estimated parameters in the price drift function (11). Clustering the standard errors on month-of-sample, which would account for broader cross-sectional correlation (perhaps associated with technological or macroeconomic shocks), generally results in smaller standard errors than those that are not clustered at all. See online Appendix 2 for details on the standard error calculations.

[34] Note that, in column 2, the distribution of log x_{it} is estimated to have a lower mean and higher variance than in column 1. This shift in parameters is necessary to rationalize non-zero drilling activity in early 1999 when oil prices were low and implied volatility was high: the increased variance allows simulated prospects to have an expected quality x_{it} greater than the high drilling trigger productivity x_t^* during this period.

[35] Rejection of the restricted estimate with a test size of 5 percent requires a difference in log likelihoods of 1.92. A likelihood ratio test does not take clustering of the likelihood scores on field into account so will therefore underestimate the true p-value.

[36] To connect the structural model to the descriptive hazard model from Section II, I have taken the simulated drilling data and used it to reestimate specification II from Table 1 (technical note: to deal with the fact that the

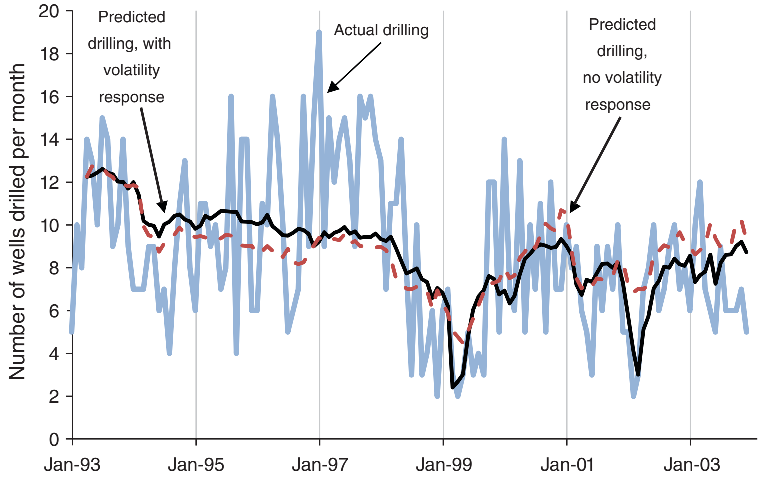


FIGURE 7. PREDICTED DRILLING FROM THE ESTIMATED MODEL VERSUS ACTUAL DRILLING

Notes: Predicted drilling with no volatility response corresponds with the estimates of column 1 of Table 2, in which the behavioral parameter β is restricted to zero. Predicted drilling with a volatility response refers to column 2 of Table 2, in which β is estimated to be 1.118.

yields a better fit to the data, particularly during the 1999 low price period and the volatility spike following September 11, 2001. More broadly, the model that does not allow a response to time-varying volatility underpredicts drilling in the early part of the sample and overpredicts drilling in the latter part. Allowing for a volatility response largely corrects these mispredictions, though there remain sections of the time series, such as early 1997, that the model does not fit well (and, of course, the model smoothes over the month-to-month noise in the actual drilling data).[37]

In column 3 of Table 2, I model firms as having mean-reverting beliefs about future volatility. When endowed with these beliefs, firms believe that changes in volatility will not be persistent, and they will therefore not respond as strongly to such changes as they would with a random walk forecast. Thus, the column 3 estimate indicates a higher value of β of 1.191 (s.e. $=$ 0.188) because the value of 1.118 from column 2 will not yield sufficient sensitivity to volatility to match the data.[38] The random walk and mean-reverting models yield nearly identical log likelihoods ($-$8,661.2 versus $-$8,661.3), so they cannot be distinguished by the data. In the presentation of the results from alternative specifications and robustness tests below, I will focus on the results from the random walk specification and simply footnote the

number of simulated wells each month is not an integer, I multiply the number of wells by 10 and then round to the nearest integer before estimating (I also multiply the number of undrilled wells by 10)). The estimated coefficients on the oil price, volatility, and drilling cost are 1.034, 0.966, and 0.949. Each of these numbers is well within the 95 percent confidence interval of the estimates obtained from the actual data as presented in column 2 of Table 1.

[37] Estimating the model on the sample of all infill wells, rather than the reference case sample of infills in sole-operated fields, yields essentially the same estimate of β as in the reference case (the point estimate is 1.117). The reduction in the noise in the drilling data (see online Appendix Figure A1 for a plot of drilling data from all infill wells) also improves the fit of the model: when I regress actual drilling on simulated drilling, I obtain an R^2 of 0.260 with the reference case data and 0.471 with the data from all infills.

[38] To be clear, this estimate uses implied volatility, not GARCH volatility, as the measure of expected volatility over the current month. The GARCH model is used only to generate a forecast of expected mean reversion.

mean reverting results, which are consistently qualitatively similar to those from the random walk specifications.

When a model allowing for time-invariant prospect-specific heterogeneity is estimated, the log likelihood is maximized when this heterogeneity (ζ_1) is zero and the model's other parameters match the Table 2, column 2 or column 3 estimates discussed above (depending on whether the random walk or mean-reverting volatility specification is used). Persistent prospect-level heterogeneity would be manifest in the data as a steady decrease in the rate of drilling activity over time as the best prospects are gradually removed from the pool.[39] However, such a decrease is not a prominent feature of the data. An explanation for the lack of empirical support for prospect-level heterogeneity is likely to be that all drilling prospects were marginal at the start of the dataset in 1993. Any particularly promising prospects were likely to have already been "skimmed off" before this year, especially given that there were periods of very high oil prices in the early 1980s and during 1990–1991. Thus, at the beginning of my dataset in 1993, there were no ex ante "best" prospects in the pool to drill first, meaning that overall prospect quality—and therefore the rate of drilling—would not decline substantially over time.[40]

B. *Firms' Incentive to Respond Optimally*

Why might the estimate of drilling activity's response to changes in expected volatility accord so well with theory? Given the small size of many of the firms in the data, it seems unlikely that they are formally solving Bellman equations. However, they may have developed decision heuristics that roughly mimic an optimal decision-making process. Moreover, the firms have a strong financial incentive to get their decision-making at least approximately right. Consider a firm that has a drilling prospect of average cost that is expected to produce 17,000 bbl and faces an average dayrate (so that the drilling cost is \$386,501). The value of the prospect to the firm, over a range of prices and for several expected price volatilities, is given in Figure 8. Suppose that the firm is somewhat myopic, acting as if volatility were 15 percent when volatility is actually 30 percent (both of these values are well within the range of in-sample realizations). In this case, the firm will incorrectly choose to drill when the oil price is between \$29 and \$35/bbl, losing as much as \$29,000 in value. Put another way, behaving optimally rather than myopically in this example can increase the prospect's value by 27 percent. Expanding the range of volatilities, and therefore the extent to which the firm can be incorrect, naturally increases the potential loss from suboptimal behavior. In the extreme, ignoring volatility—and therefore option value—altogether can cause a firm to drill a prospect with an expected profit of nearly zero despite the fact that holding the prospect can have substantial value. In the example above, a firm that completely ignored price

[39] Even without time-invariant prospect-level heterogeneity, the model would still predict a steady decrease in drilling over time because the total number of prospects available (7,787) is finite. Adding time-invariant heterogeneity steepens the rate of decrease.

[40] It may also be that the effects of any decline in prospect quality over time caused by heterogeneity have been masked by technological improvements, such as adoption of 3-D seismic imaging, that have pushed prospects' expected productivity upward.

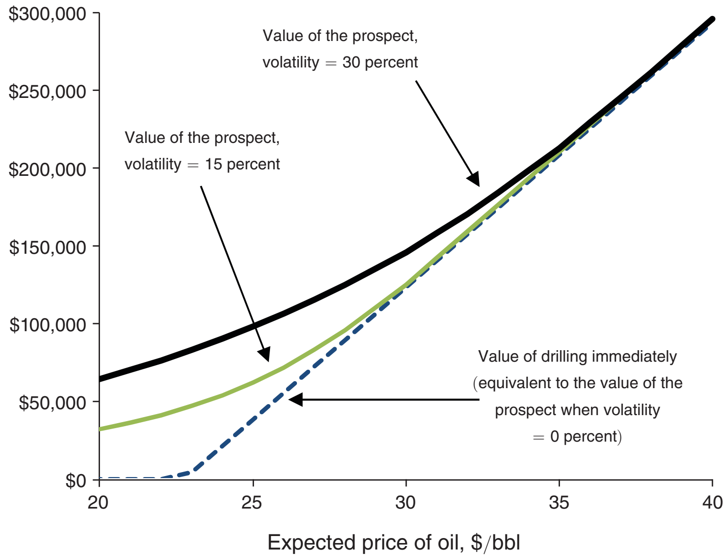


FIGURE 8. COMPARISON OF THE VALUE OF HOLDING A DRILLING PROSPECT TO THE PROFITS FROM DRILLING

Notes: The relationships shown are for a drilling prospect with an expected production of 17,000 bbl and an average depth, facing an average dayrate, so that the drilling cost is $386,501. With expected oil price volatilities of 0 percent, 15 percent, and 30 percent, drilling is triggered at oil prices of $23, $29, and $36/bbl, respectively. The model used to generate these curves has a behavioral parameter $\beta = 1$ and uses the state transition parameters estimated in Section IVA of the text (with a random walk volatility forecast). Values shown are taken directly from the model's value function.

volatility of 30 percent would drill at a price of $23/bbl, achieving an expected profit of $4,500, even though holding the prospect has a value of $135,000.

Do different types of firms respond differently to volatility shocks? In Figure 9 and columns 4 and 5 of Table 2, I examine whether there is a difference in drilling behavior between small and large firms. I classify each firm as small or large based on the total number of wells (including non-infill wells) the firm drills during the 1993–2003 sample period, splitting the sample at the median number of wells per firm.[41] Figure 9 plots the time series of drilling activity by each type of firm against futures prices and implied volatility. Though the time series are noisy, the drilling rates of small and large firms overlay each other fairly closely, suggesting that the two types respond similarly to price and volatility signals. When I estimate the model separately for each type, the estimates are statistically and economically indistinguishable: the estimated β for small firms is 1.136 while that for large firms is 1.085 (columns 4 and 5 of Table 2).[42] Thus, it is not the case that small firms in this industry are "unsophisticated" and cannot respond properly to market signals. Instead, small firms respond to volatility shocks as optimally as do large firms,

[41] In the reference case sample (1,150 infill wells in sole-operated fields), the median well is drilled by a firm that drilled 49 wells in total over 1993–2003.

[42] Under a hypothesis that firms believe volatility is mean reverting, the point estimates (standard errors) for β in columns 4 and 5 are 1.212 (0.187) and 1.181 (0.201), respectively. The log likelihoods are −4,350.7 and −4,310.0.

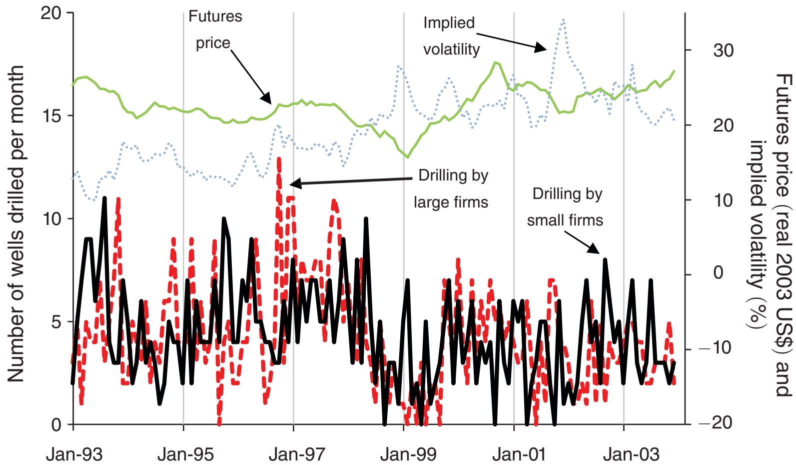


Figure 9. Time Series of Monthly Drilling Activity Split by Small versus Large Firms, with Oil Futures Prices and Implied Volatility from Futures Options Prices

Notes: Drilling activity is broken into small versus large firms at the median firm size (measured by total number of wells drilled) in the dataset. Oil futures prices are 18-month ahead prices from the NYMEX. Implied volatility is calculated from futures options prices per the discussion in Section ID. Drilling activity corresponds only to infill oil wells drilled in sole-operated fields.

perhaps reflecting the possibility that it is difficult for a firm to survive in this highly competitive environment if its decision-making is poor.

C. *Sources of Identification*

The above results indicate that, in periods of high expected oil price volatility, drilling activity falls in a way that is commensurate with the predictions of real options theory. This section examines which aspects of the data drive the identification of this result and also studies whether realized production data can offer additional identifying variation.

A prominent feature of the data is that volatility is greater at the end of the sample than at the beginning, and drilling activity is lower at the end of the sample then at the beginning. This feature raises the question of the extent to which identification is coming from an overall trend in the data. I examine this issue by including in the model a time trend in wells' expected productivity x_{it}. The results from estimating this model are presented in column 2 of Table 3. I find that the estimated time trend is effectively zero, with a point estimate of a productivity increase of about 0.1 percent per year that is not statistically significant. Moreover, the estimate of β is virtually unchanged.[43] This result is related to the lack of evidence for prospect-specific heterogeneity, which, like a time trend, would be manifest as a steady decrease in drilling activity over time. Identification of β does not appear to arise from such a trend.

I next examine the extent to which the reference case results are driven by the sharp increase in volatility and commensurate decrease in drilling activity that

[43] Under a hypothesis that firms believe volatility is mean reverting, the point estimate (standard error) for β in column 2 of Table 3 is 1.194 (0.189), with a log likelihood of −8,661.3.

TABLE 3—ALTERNATIVE SPECIFICATIONS: SOURCES OF IDENTIFICATION

Parameter	Reference case model (1)	Time trend (2)	July 1998 dummy (3)	July 1998 dummy, alternative local maximum (4)	Model includes production data (5)
β (sensitivity to volatility)	1.118 (0.141)	1.117 (0.138)	1.137 (0.127)	0.453 (0.280)	0.923 (0.169)
μ (mean of $\log(x_{it})$)	−12.045 (7.993)	−11.922 (9.107)	−20.913 (15.667)	−3.113 (4.160)	−3.709 (4.428)
ζ (SD of $\log(x_{it})$)	6.961 (2.664)	6.920 (3.065)	9.959 (5.261)	3.964 (1.400)	4.166 (1.487)
Time trend (years)		0.001 (0.038)			
Dummy for date $\geq$ July 1998			0.433 (0.501)	0.250 (0.169)	
log likelihood	−8,661.2	−8,661.2	−8,659.5	−8,659.9	−9,457.9

Notes: All estimates use prices of futures and options that are 18 months from maturity. x_{it} is expressed as expected oil production (in bbl) divided by the cost of drilling (in \$100,000) at the average sample dayrate. Drilling data are matched to drilling likelihoods with a three month lag. All estimates assume that firms have a random walk forecast for future volatility. Column 5 involves estimates of parameters not shown; full estimates are given in online Appendix 3. Standard errors are estimated using a sandwich estimator that allows for correlation of the likelihood scores across wells within the same field, thereby accounting for spatial and serial correlation. Standard errors also account for the sampling error in the estimated function for the expected drift of future oil prices.

occurred in the summer of 1998. To do so, I model a permanent shock to the x_{it} that begins in July 1998 and is common to all prospects, thereby removing this source of identifying variation. The results from estimating the model while including this shock are given in column 3 of Table 3. The estimated shock is large, decreasing $\log x_{it}$ by 0.433, although it is not statistically significant. The point estimate of β is not significantly affected, changing from 1.118 in the reference case to 1.137 here.[44] This result suggests that identification does not strongly rely on the substantial changes that occurred in the summer of 1998—other variation, such as that occurring around the events of September 11, 2001, is also important.[45] However, with the inclusion of the post-July 1998 productivity dummy, the likelihood function exhibits strong nonconcavities, and the data cannot reject an alternative local maximum at $\beta = 0.453$, shown in column 4 of Table 3 (the difference in log likelihoods between columns 3 and 4 is only 0.4). Thus, statistical power is coming in no small part from the sharp changes in volatility and drilling that occurred in 1998.[46]

[44] Under a hypothesis that firms believe volatility is mean reverting, the point estimate (standard error) for β in column 3 of Table 3 is 1.208 (0.198), with a log likelihood of −8,659.1. There is an alternative local maximum with an estimated β of 0.385 and a log likelihood of −8,660.4.

[45] I have also examined the importance of 9/11 to the results by estimating a specification in which there is a common shock across all prospects for September 2001 through January 2002. The log likelihood for this specification is maximized at $\beta = 1.101$, though again there is another local maximum at 0.631. The difference between these estimates' log likelihoods is larger than that of the post-July 1998 shock model, however. The log likelihood at the estimate of 1.101 is −8,661.0, while that at the 0.631 estimate is −8,662.4. A likelihood ratio test rejects the local maximum at the 10 percent level, with the caveat that this test does not take the clustering of standard errors into account.

[46] It seems unlikely that there was an actual sudden (and potentially confounding) drop in wells' productivity in 1998, since productivity is a function of geology and there is no obvious reason why firms' beliefs about the geologic characteristics of oil reservoirs would simultaneously decrease. In addition, the statistical insignificance of the estimated coefficient on the post-July 1998 dummy in column 3 is consistent with the absence of a true shock.

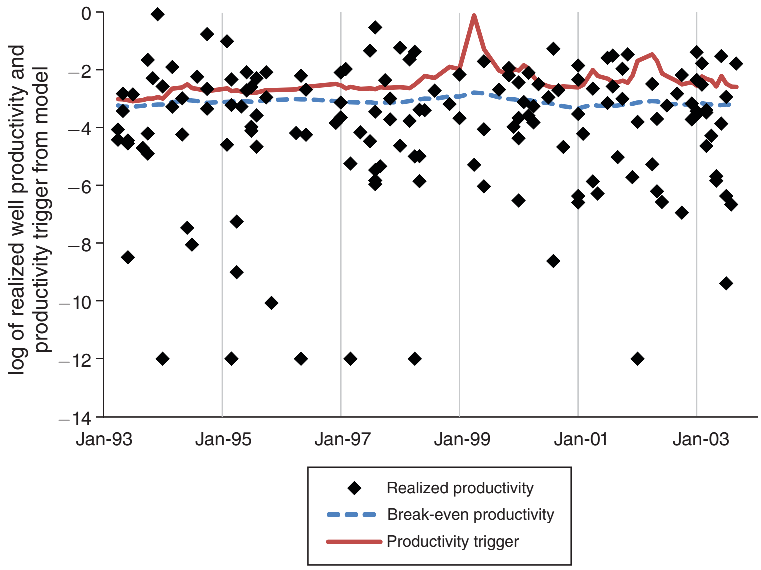


Figure 10. Realized Productivity versus Productivity Triggers x^* from the Reference Case Model

Notes: Production data are from the subset of observed drilled wells that are the only active producing well on their respective lease for the first 36 months subsequent to drilling (these are 160 of the observed 1,150 drilled wells from 1993 to 2003). Realized productivity for each well is the estimated lifetime discounted production (in barrels) divided by the cost of drilling (in real 2003 US$). Dry holes are plotted as having a log(productivity) of −12.

Finally, I explore the extent to which the data on realized production—for the subset of 160 wells for which production is observable—can provide additional information for identification of β. In principle, the realizations of well productivity (discounted lifetime production divided by the cost of drilling) should be related to variation in the trigger productivities x^* over time and should therefore be informative for β. Below, I summarize my use of the productivity data in this regard; details can be found in online Appendix 3.

To make the production data comparable to the x^* trigger productivities, I first transform them from production over the first three years of each well's life to total discounted production over each well's lifetime, estimating hyperbolic decline curves to carry out the extrapolation. I then divide each well's discounted lifetime production by its estimated drilling cost at the average dayrate to obtain its realized productivity (in barrels per $ of drilling cost). Figure 10 plots (in logs) each well's thusly calculated realized productivity against the x^* productivity triggers generated by the reference case model. This figure reveals two obstacles to using the realized productivity data in the model. First, these data are extremely noisy, reflecting the substantial geologic uncertainty involved in drilling for oil and masking any obvious relationship between realized productivity and x^*. Second, in this sample of wells average productivity falls substantially below the average x^*. This shortfall is consistent with selection: the wells for which productivity is observable are those drilled in leases in which no other wells were subsequently drilled or operated, and this lack of follow-up activity may have been due to poor results from the initial well. Data from the 367 drilled wells for which I have production data but there also exist other active wells provide support for this selection mechanism, in that they exhibit substantially higher productivity on average.[47] Figure 10 also provides the logged "break-even" productivity each period. This productivity level falls below x^* due to option value, and the figure shows that the median well for which production is observed performs slightly worse than break-even. In levels, however, the right-tail of highly productive wells implies that these 160 wells earn strictly positive profits on average, consistent with real options theory's prescription that expected profits from irreversible investments should be strictly positive.[48]

With the above issues in mind, I augment the model and likelihood function to incorporate the productivity realization data. As discussed in detail in online Appendix 3, the augmented likelihood function takes into account both the probability of each productivity realization (which will depend on x^* each period and on the estimated variance of productivity realizations about their expectation) and the probability that production is observable for each drilled well (which will depend on the well's productivity). I find that making use of the production data does not substantially influence the estimate of β. As shown in column 5 of Table 3, the estimate of β in this specification is 0.923 with a standard error of 0.169, still consistent with an optimal response by firms to changes in oil price volatility.[49] The production data do shift the estimated distribution of well productivities, raising the mean μ and lowering the standard deviation ζ, though the estimates still lie well within the 95 percent confidence interval from the reference case.[50] The full set of estimated parameters from this model is given in online Appendix 3.

D. *Alternative Specifications*

Alternative Measures of Expected Volatility.—The analysis thus far has used implied volatility from the NYMEX futures options market as the measure of firms' oil price volatility expectations. Column 2 of Table 4 reports results in which expected volatility is instead measured by the historic volatility of futures prices over a one year rolling window. The use of historic volatility yields a worse fit to the drilling data than does implied volatility, evidenced by the substantial decrease in the log likelihood relative to the implied volatility results in column 1 of Table 4. Moreover, the estimate of β is only 0.348 and not statistically significant, indicating that firms do not respond as strongly to historic volatility as they do to volatility signals that are reflected in the NYMEX futures options market.[51]

[47] For these 367 wells, the noise in their leases' monthly production data makes it impractical to estimate each well's productivity. However, I have estimated the average production across these wells. They are substantially more productive on average than the 160 drilled wells that are the only active wells on their lease, consistent with selection. The estimated average log realized productivity in the 367 well sample is −2.482, while that for the 160 well sample is −4.109. Consistent with the selection story, the majority of these wells' leases exhibit production from wells that come online after the well in question was drilled.

[48] The average estimated discounted profit for the 160 well sample is \$243,720, relative to an average drilling cost of \$403,697. The median well, however, loses \$121,469.

[49] Under a hypothesis that firms believe volatility is mean reverting, the point estimate (standard error) for β in column 5 of Table 3 is 0.938 (0.172), with a log likelihood of −9,457.7.

[50] These changes have a net effect of reducing expected production conditional on drilling, thereby better matching the production data.

[51] Under a hypothesis that firms believe volatility is mean reverting, the point estimates (standard errors) for β in columns 2, 3, and 4 are 0.292 (0.323), 0.790 (0.281), and 1.690 (0.447), with log likelihoods of −8,669.7, −8,666.1, and −8,663.8, respectively.

Table 4—Alternative Specifications of Price and Volatility Beliefs

Parameter	Reference case model (Table 2, column 2) (1)	Historic volatility of futures prices, one year window (2)	GARCH volatility (3)	Front-month futures and implied volatility (4)
β (sensitivity to volatility)	1.118	0.348	0.587	1.679
	(0.141)	(0.338)	(0.194)	(0.572)
μ (mean of $\log(x_{it})$)	−12.045	−0.331	−2.481	−3.462
	(7.993)	(2.587)	(3.552)	(2.972)
ζ (SD of $\log(x_{it})$)	6.961	2.996	3.745	4.146
	(2.664)	(0.860)	(1.189)	(1.000)
log likelihood	−8,661.2	−8,670.4	−8,666.0	−8,662.5

Notes: x_{it} as expressed in expected oil production (in bbl) divided by the cost of drilling (in \$100,000) at the average sample dayrate. Drilling data are matched to drilling likelihoods with a three month lag. All estimates assume that firms have a random walk forecast for future volatility. Standard errors are estimated using a sandwich estimator that allows for correlation of the likelihood scores across wells within the same field, thereby accounting for spatial and serial correlation. Standard errors also account for the sampling error in the estimated function for the expected drift of future oil prices.

Column 3 of Table 4 uses the GARCH(1, 1) model to forecast future volatility. This model yields an estimate of β of 0.587 that is statistically significant at the 1 percent level, though the fit of the model is still substantially worse than when implied volatility is used (the decrease in the log likelihood is equal to 4.8). The reduced fit reflects the fact that, while GARCH provides a closer match to implied volatility than does historic volatility, the GARCH and implied volatility series still diverge substantially at several points in time (Figure 4). This result, as well as that obtained from the direct use of historic volatility, suggests an explanation for why some previous empirical studies (Hurn and Wright 1994; Moel and Tufano 2002) have not found strong evidence that time-varying volatility significantly affects investment. These studies measure firms' volatility expectations using historic volatility, which may only be a noisy measure of firms' true beliefs because it does not reflect up-to-date information regarding volatility shocks.

Front-Month Rather than Futures Prices.—Column 4 of Table 4 considers a model in which firms respond to the front-month price and volatility of oil rather than 18-month futures and volatilities. I replace the price series P_t with the NYMEX front-month futures contract, and I replace the market's implied 18-month price volatility σ^m with that of front-month futures options. Because firms' use of current prices as expected prices is consistent with a no-change forecast for the price of oil, I set the price and cost drift functions $\mu(\cdot)$ and $\hat{\mu}(\cdot)$ to zero. The estimate of β from this model is 1.679, with a relatively large standard error of 0.572. The front month model also yields a weaker fit to the data than does the reference case model, with a log likelihood of $-8{,}662.5$ rather than $-8{,}661.2$. The increase in β, imprecision, and decreased fit with the front month model likely reflects the zero expected price drift assumption inherent in the use of front-month prices. A relatively high volatility state in this model is not associated with an expectation that prices will increase in the future, as was the case in the reference case model. In addition, the lack of mean reversion in the price forecast means that firms would not expect an increase

TABLE 5—ALTERNATIVE SPECIFICATIONS: DISCOUNT RATES AND DRILLING COSTS

Parameter	Reference case model (Table 2, column 2) (1)	14.5 percent nominal discount rate (2)	10.5 percent nominal discount rate (3)	Rig costs average 20 percent of total drilling cost (4)	Rig costs average 50 percent of total drilling cost (5)
β (sensitivity to volatility)	1.118	1.189	0.973	1.122	1.127
	(0.141)	(0.176)	(0.174)	(0.143)	(0.133)
μ (mean of $\log(x_{it})$)	−12.045	−10.164	−11.470	−12.871	−13.163
	(7.993)	(7.599)	(8.045)	(8.545)	(7.672)
ζ (SD of $\log(x_{it})$)	6.961	6.325	6.776	7.237	7.336
	(2.664)	(2.538)	(2.684)	(2.847)	(2.555)
log likelihood	−8,661.2	−8,661.5	−8,661.3	−8,661.9	−8,661.0

Notes: All estimates use prices of futures and options that are 18 months from maturity. x_{it} is expressed as expected oil production (in bbl) divided by the cost of drilling (in \$100,000) at the average sample dayrate. Drilling data are matched to drilling likelihoods with a three month lag. All estimates assume that firms have a random walk forecast for future volatility. Standard errors are estimated using a sandwich estimator that allows for correlation of the likelihood scores across wells within the same field, thereby accounting for spatial and serial correlation. Standard errors also account for the sampling error in the estimated function for the expected drift of future oil prices.

in the oil price during periods such as 1998–1999 when the front-month price was low (and expected volatility was high). Because expectations of higher prices in the future dampen the incentive to drill today, a higher estimate of β is required in order to offset the use of a no-change price forecast and fit the data.

Alternative Discount Rate and Drilling Cost Assumptions.—The estimates heretofore have been based on an assumed 12.5 percent nominal discount rate, taken from a 1995 survey by the Society of Petroleum Evaluation Engineers. Columns 2 and 3 of Table 5 examine the use of alternative discount rates. A 14.5 percent discount rate yields an estimate of β of 1.189 while a 10.5 percent discount rate yields $\beta = 0.973$.[52] Neither estimate is statistically distinct from one. These changes to the estimated β are in line with real options theory's predictions. As the discount rate increases, firms value the future less, option value decreases, and firms become less responsive to changes in expected volatility. Thus, to fit the empirical volatility response, the volatility sensitivity parameter β must increase when the assumed discount rate increases.

Finally, columns 4 and 5 of Table 5 examine the estimates' sensitivity to the assumption that rig costs constitute one-third of total drilling costs on average. Assuming a value of 20 percent or 50 percent does not substantially alter the estimate of β.

VI. Conclusions

The importance of irreversibility and uncertainty in investment decision-making has been recognized since Marschak (1949) and Arrow (1968). Theoretical work

[52] Under a hypothesis that firms believe volatility is mean reverting, the point estimates (standard errors) for β in columns 2 through 5 of Table 5 are, in order, 1.281 (0.139), 1.027 (0.197), 1.194 (0.189), and 1.194 (0.191). The log likelihoods are −8,661.5, −8,661.8, −8,662.0, and −8,660.9.

has since derived optimal timing rules for irreversible investments and demonstrated that firms should defer projects when uncertainty is relatively high. These concepts have taken a prominent role in industrial organization and the macroeconomic modeling of aggregate investment. However, there has been a shortage of empirical evidence regarding the extent to which firms actually take option value into account when making irreversible investments.

This paper tests the sensitivity of firms' investment decisions to changes in the uncertainty of their economic environment by assembling a new, detailed dataset that combines information on well-level oil drilling with expected oil price volatility data from the NYMEX futures options market. I build and estimate a dynamic model of firms' drilling investment timing problem, taking advantage of industry features that make a single-agent approach appropriate. I find not only that firms reduce their drilling activity when expected volatility rises but also that the magnitude of this reduction is consistent with the optimal response prescribed by theory. This result provides micro-empirical support for the frequent use of real options models in economic research. It is also consistent with the existence of a strong incentive for firms to behave optimally. I find that the cost of failing to respond to changes in volatility can be substantial, potentially exceeding 25 percent of a drilling prospect's value at in-sample oil price and volatility realizations.

I also show that a forward-looking measure of expected price volatility derived from futures options is a more powerful determinant of drilling behavior than are backward-looking measures based on historic volatility. The relative strength of the implied volatility measure is consistent with the hypothesis that participants in the NYMEX commodity market and physical industry participants share common beliefs about future price uncertainty. This result thereby provides support for the use of data from financial markets as measures of firms' expectations in applied work. It is also well-aligned with other research regarding the predictive power of option-based implied volatility and supports the intuition that options prices incorporate up-to-date information about uncertainty shocks that cannot be conveyed by price histories alone.

REFERENCES

Aguerrevere, Felipe L. 2003. "Equilibrium Investment Strategies and Output Price Behavior: A Real-Options Approach." *Review of Financial Studies* 16 (4): 1239–72.

Alexopoulos, Michelle, and Jon Cohen. 2009. "Uncertain Times, Uncertain Measures." University of Toronto Working Paper 352.

Alquist, Ron, and Lutz Kilian. 2010. "What Do We Learn from the Price of Crude Oil Futures?" *Journal of Applied Econometrics* 25 (4): 539–73.

Anderson, Soren T., Ryan Kellogg, and Stephen W. Salant. 2014. "Hotelling under Pressure." Unpublished.

Arrow, Kenneth J. 1968. "Optimal Capital Policy with Irreversible Investment." In *Value, Capital and Growth, Essays in Honor of Sir John Hicks*, edited by James N. Wolfe, 1–20. Edinburgh, Scotland: Edinburgh University Press.

Arrow, Kenneth J., and Anthony C. Fisher. 1974. "Environmental Preservation, Uncertainty, and Irreversibility." *Quarterly Journal of Economics* 88 (2): 312–19.

Bachmann, Ruediger, Steffen Elstner, and Eric R. Sims. 2013. "Uncertainty and Economic Activity: Evidence from Business Survey Data." *American Economic Journal: Macroeconomics* 5 (2): 217–49.

Bernanke, Ben S. 1983. "Irreversibility, Uncertainty, and Cyclical Investment." *Quarterly Journal of Economics* 98 (1): 85–106.

Black, Fischer. 1976. "The Pricing of Commodity Contracts." *Journal of Financial Economics* 3 (1/2): 167–79.

Black, Fischer, and Myron S. Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81 (3): 637–54.

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3): 623–85.

Bloom, Nicholas, Stephen Bond, and John Van Reenen. 2007. "Uncertainty and Investment Dynamics." *Review of Economic Studies* 74 (2): 391–415.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2012. "Really Uncertain Business Cycles." National Bureau of Economic Research Working Paper 18245.

Chernenko, Sergey V., Krista B. Schwarz, and Jonathan H. Wright. 2004. "The Information Content of Forward and Futures Prices: Market Expectations and the Price of Risk." Board of Governors of the Federal Reserve System, International Finance Discussion Paper 808.

Collard-Wexler, Allan. 2013. "Demand Fluctuations in the Ready-Mix Concrete Industry." *Econometrica* 81 (3): 1003–37.

Commodity Systems Inc. 1986–2006. "End of Day Market Data." NYMEX Crude Oil Commodity Options. http://www.csidata.com/?page_id=10 (accessed June 5, 2006).

Dixit, Avinash K. 1989. "Entry and Exit Decisions under Uncertainty." *Journal of Political Economy* 97 (3): 620–38.

Dixit, Avinash K., and Robert S. Pindyck. 1994. *Investment Under Uncertainty*. Princeton, NJ: Princeton University Press.

Dunne, Timothy, and Xiaoyi Mu. 2010. "Investment Spikes and Uncertainty in the Petroleum Refining Industry." *Journal of Industrial Economics* 58 (1): 190–213.

Ericson, Richard, and Ariel Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies* 62 (1): 53–82.

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, Juan F. Rubio-Ramírez, and Martín Uribe. 2011. "Risk Matters: The Real Effects of Volatility Shocks." *American Economic Review* 101 (6): 2530–61.

Grenadier, Steven R. 2002. "Option Exercise Games: An Application to the Equilibrium Investment Strategies of Firms." *Review of Financial Studies* 15 (3): 691–721.

Guiso, Luigi, and Giuseppe Parigi. 1999. "Investment and Demand Uncertainty." *Quarterly Journal of Economics* 114 (1): 185–227.

Handley, Kyle. 2012. "Exporting under Trade Policy Uncertainty: Theory and Evidence." http://webuser.bus.umich.edu/handleyk/eutpu.pdf.

Handley, Kyle, and Nuno Limão. 2012. "Trade and Investment under Policy Uncertainty: Theory and Firm Evidence." National Bureau of Economic Research Working Paper 17790.

Hassler, John A. A. 1996. "Variations in Risk and Fluctuations in Demand: A Theoretical Model." *Journal of Economic Dynamics and Control* 20 (6-7): 1115–43.

Hassler, John A. A. 2001. "Uncertainty and the Timing of Automobile Purchases." *Scandanavian Journal of Economics* 103 (2): 351–66.

Hilliard, Jimmy E., and Jorge Reis. 1998. "Valuation of Commodity Futures and Options Under Stochastic Convenience Yields, Interest Rates, and Jump Diffusions in the Spot." *Journal of Financial and Quantitative Analysis* 33 (1): 61–86.

HSH Associates. 1980–2006. "Treasury Constant Maturity Series." 3-month, 6-month, 1-year, 3-year, and 5-year rates. http://www.hsh.com/indices/tsec00s.html and http://www.hsh.com/indices/3n6tsec00s.html (accessed June 23, 2006).

Hurn, A. S., and Robert E. Wright. 1994. "Geology or Economics? Testing Models of Irreversible Investment Using North Sea Oil Data." *Economic Journal* 104 (423): 363–71.

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

Kellogg, Ryan. 2014. "The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling: Dataset." *American Economic Review*. http://dx.doi.org/10.1257/aer.104.6.1698.

Kilian, Lutz. 2009. "Not All Oil Price Shocks are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review* 99 (3): 1053–69.

Leahy, John V., and Toni M. Whited. 1996. "The Effects of Uncertainty on Investment: Some Stylized Facts." *Journal of Money, Credit, and Banking* 28 (1): 64–83.

Levitt, Clinton J. 2011. "Learning Through Oil and Gas Exploration." http://clintlevitt.weebly.com/uploads/3/1/5/8/3158449/exploration.pdf (accessed May 31, 2013).

List, John A., and Michael S. Haigh. 2010. "Investment under Uncertainty: Testing the Options Model with Professional Traders." *Review of Economics and Statistics* 92 (4): 974–84.

Marschak, Jacob. 1949. "Role of Liquidity under Complete and Incomplete Information." *American Economic Review* 39 (3): 182–95.

Moel, Alberto, and Peter Tufano. 2002. "When Are Real Options Exercised? An Empirical Study of Mine Closings." *Review of Financial Studies* 15 (1): 35–64.

Pakes, Ariel S. 1986. "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." *Econometrica* 54 (4): 755–84.

Petroleum Services Association of Canada. 2000–2004. "PSAC Well Cost Study." http://www.psac.ca/resources/well-cost-study-overview/ (accessed May 13, 2009).

Pindyck, Robert S. 1991. "Irreversibility, Uncertainty, and Investment." *Journal of Economic Literature* 29 (3): 1110–48.

Pindyck, Robert S. 2004. "Volatility and Commodity Price Dynamics." *Journal of Futures Markets* 24 (11): 1029–47.

Poon, Ser-Huang, and Clive W. J. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 41 (2): 478–539.

Price-Data. 1983–2005. "End of Day Futures." NYMEX Crude Oil. http://www.grainmarketresearch.com/eod_futures.cfm (accessed December 17, 2005).

RigData. 1990–2005. *Day Rate Report* http://www.rigdata.com/news-and-analysis/day-rate-report.aspx (accessed January 17, 2006).

Rust, John. 1987. "Optimal Replacement of GMC Bus Engines; An Empirical Model of Harold Zurcher." *Econometrica* 55 (5): 999–1033.

Society of Petroleum Evaluation Engineers. 1995. "Survey of Economic Parameters Used in Property Evaluations."

Szakmary, Andrew, Evren Ors, Jin Kyoung Kim, and Wallace N. Davidson III. 2003. "The Predictive Power of Implied Volatility: Evidence from 35 Futures Markets." *Journal of Banking and Finance* 27 (11): 2151–75.

Texas Railroad Commission. 1977–2005. "Drilling Permit Master and Trailer." http://www.rrc.state.tx.us/data/datasets/DrillingPermitData.php (accessed December 16, 2005).

Texas Railroad Commission. 1990–2007. "Final Oil and Gas Annuals." http://www.rrc.state.tx.us/data/datasets/WellData.php (accessed February 22, 2007).

Wiggins, Steven N., and Gary D. Libecap. 1985. "Oil Field Unitization: Contractual Failure in the Presence of Imperfect Information." *American Economic Review* 75 (3): 368–85.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

ONLINE APPENDIX

"The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling"

by Ryan Kellogg

Appendix 1: Construction of the time series of implied futures price volatility

This appendix describes how I construct a time series of the implied volatility of 18-month NYMEX oil futures contracts. As discussed in the main text, I cannot simply use the Black (1976) formula directly because it assumes that the term structure of volatility (the function by which the volatility of the future price of oil varies as time to maturity increases) is constant. My strategy for addressing this issue proceeds in three steps. First, I use the realized volatility of futures prices to estimate the average term structure of volatility. Second, I use liquidly traded short-term futures options to generate a time series of the implied volatility of one-month futures option contracts. Because a one month time horizon is short, this time series is equivalent to the time series of the implied volatility of one-month futures price contracts. Finally, I combine the one-month futures price volatilities with the estimated term structure to generate the desired time series of the implied volatility of 18-month futures price contracts.[1] The remainder of this appendix discusses these three steps in turn.

Let $F_{t,\tau}$ denote the price of a NYMEX futures contract traded at date t with time to maturity τ measured in months.[2] For each t and τ, I calculate the realized volatility at t of the τ-month futures contract as the standard deviation of $\ln(F_{s,\tau} / F_{s\text{-}1,\tau})$ for all dates s within the 6 months prior and subsequent to t.[3] Let this volatility be denoted by $\sigma_{t,\tau}$. I then estimate the term structure of futures price volatility by regressing the log of $\sigma_{t,\tau}$ on fixed effects for each τ and t:[4]

$$\ln \sigma_{t,\tau} = \eta_\tau + \delta_t + \varepsilon_{\tau,t} \qquad \text{(A1.1)}$$

[1] An alternative procedure to that used here would use the term structure of the implied volatility of futures options directly to derive the implied volatility of 18-month futures prices. This approach would use the fact that the volatility of a τ-month futures price is equal to the volatility of a τ-month futures option plus τ times the derivative of the futures option term structure (with respect to τ) at τ. The use of the derivative implies that this approach requires a very precise estimate of the term structure of futures options' implied volatility. Thin markets for futures options beyond 6 months render this procedure impractical. For example, 18-month futures options are traded, on average, only 18 days each year from 1993-2003.

[2] Time to maturity in months is equal to the time to maturity in days divided by 365.25, multiplied by 12, and rounded to the nearest whole number.

[3] Observations $F_{s,\tau}$ for which date s - 1 is missing (for example, if s - 1 is a Sunday) are excluded.

[4] I use the log of $\sigma_{t,\tau}$ as the dependent variable rather than the level because the levels regression does not yield an estimated term structure that is stable over time. In levels, the term structure is has a steeper slope during 1999-2003 than in the earlier part of the data.

The fixed effects η_τ represent the estimated term structure while the δ_t control for the level of volatility on each date t. Given estimates of these fixed effects, the predicted volatility of a τ-month futures price on date t is given by $A_t \cdot \exp(\eta_\tau)$, where $A_t = \exp(\delta_t + v^2 / 2)$ and v^2 is the variance of the estimated residuals. Thus, for a fixed trade date t, varying τ will trace out the term structure of volatility. Figure A2 verifies that the term structure of volatility is stable over the sample by plotting two estimates of the term structure: one using data from 1999-2003 and another using data prior to 1999. The constant term A for each plotted estimate is set so that the one-month future price volatility is 31%, approximately equal to the average one-month volatility over 1993-2003. The plots overlay each other closely, indicating that the term structure of volatility is stable over the sample despite the substantial increase in the overall level of volatility after 1999.

Given the estimated term structure (the η_τ), all that is needed to compute expected 18-month futures price volatilities is a time-series of short-run (one month) expected futures price volatilities. I derive this time series from the implied volatility of short-term futures options with a time to maturity between 60 and 180 days. The implied volatility of options with a shorter time to maturity are noisy, potentially reflecting low option values and integer problems (options prices must be in whole cents), while options with a longer time to maturity are thinly traded.

For each trade date and time to maturity within the 60 to 180 day window, I use the Black (1976) model to find the implied volatilities of the call and put options that are nearest to at-the-money.[5] I then estimate the implied volatility term structure by regressing the log of each option's implied volatility on its time to maturity τ (in days), a call/put dummy, and trade date fixed effects δ_t.[6] I then use this estimated term structure (the estimated coefficient on τ) to extrapolate implied volatility back to a 30 day maturity.

As a validation check on the this procedure, I compare the average, over 1993-2003, of the estimated implied volatilities of 30-day futures options to the average realized volatility of one-month futures prices over the same timeframe. These two averages should be approximately equal given the short one month time to maturity. The former series has an average volatility of

---

[5] The Black (1976) model assumes that the options are European rather than American and that volatility is not stochastic. Neither of these assumptions holds here; however, their effects are likely to be minor and they save considerable computational complexity. Hilliard and Reis (1998) demonstrate that the American premium is no more than 2% of the European option price for volatilities similar to those considered here. Stochastic volatility acts in the opposite direction, causing the Black (1976) model to slightly over-price at-the-money options (this effect is particularly small for the relatively short maturities considered here); see Hull and White (1987), Wiggins (1987), and Poon and Granger (2003). The argument that these assumptions are of minor effect is supported by the close agreement between the average realized and average implied volatility over the 1993-2003 sample.

[6] Inspection of the residuals indicates that a linear term structure specification is appropriate. Moreover, when a squared time to maturity term is added, it is not statistically significant (p-value = 0.114).

30.83% while the average of the latter is 31.07%. The closeness of these two numbers (derived from two completely different data sets) supports the argument that implied volatilities from one-month futures options can be used as implied volatilities of one-month futures prices.

Finally, I convert the time series of implied volatilities of one-month futures prices to implied volatilities of 18-month futures prices using the estimated term structure of futures price volatility (the η_τ). This conversion amounts to multiplying the one-month volatility at each trade date t by $\exp(\eta_{18} - \eta_1)$.

References

Hull, John, and Alan White. 1987. "The Pricing of Options on Assets with Stochastic Volatilities." *Journal of Finance* 42(2): 281-300.

Wiggins, James B. 1987. "Option Values Under Stochastic Volatility." *Journal of Financial Economics* 19(2): 351-372.

Appendix 2: Numerical solution and estimation methods

A2.1 Value function iteration

I solve the value function (12) on a grid of points in (P,D,σ,x) space (in logs) using standard value function iteration. An important factor in defining the grid is that, while the price, dayrate, and volatility states that are realized in the data are bounded, the stochastic processes for these variables (equations 4, 5 and 10) imply that agents place nonzero probabilities on realizations outside of these bounds. Thus, the value function must be solved for states extending beyond the boundaries of the data. The state space I use extends from one-fifth of the lowest realized price and dayrate to five times the highest price and dayrate, and from one-half the lowest realized volatility to twice the highest volatility. With this state space, marginal reductions or extensions in size do not substantially affect the estimated parameters or the value function within the range of realized observations.

I found that a relatively dense grid was required to accurately capture the effects of stochastic volatility. The grid I use has 1,875,000 points: 50 price states by 50 dayrate states by 15 volatility states by 50 productivity states. Starting from this density, the estimated results are insensitive to increases or decreases in the number of grid points.

In the full estimation routine, the initial value function used for each guess of parameters is the value function from the previous guess. For the first parameter guess, the initial value function is zero in all states. The convergence criterion is a tolerance of 10^{-6} on the sup norm of the value function (the value function used in the computations is in units of \$386,501, the average drilling cost at the average dayrate). Increasing the tolerance to 10^{-7} has essentially no affect on the parameter estimates or value function.

With the value function solved, I can then find, for any given P, D, and σ, the critical productivity x^* such that drilling is optimal iff $x_i > x^*$. Because the P, D, and σ realizations do not coincide with the grid states used in the model, I use linear interpolation to find x^*. At each x_i grid point, I calculate the value function at the realized P, D, and σ by linearly interpolating the value function between the states immediately above and below the P, D, and σ. I then find the smallest x_i grid point such that the value of waiting exceeds the realized profits from drilling immediately and the largest x_i such that it is optimal to drill immediately (these two values of x_i will be adjacent grid points). Interpolation gives x^* as the productivity level for which the firm is indifferent: the value of waiting equals the value of drilling immediately. As described in the text, the realized time series of $\boldsymbol{P}$, $\boldsymbol{D}$, and $\boldsymbol{\sigma}$ can then be combined with a parameterized distribution on the x_{it} to yield the probability that a given prospect will be drilled each period.

In most of the estimated models, there is no initial conditions problem because the productivity shocks x_{it} are modeled as iid. An initial conditions problem is present, however, in the specification allowing for time-invariant prospect heterogeneity (though the specification ultimately finds no evidence of such heterogeneity). I address this issue by extending the simulation back to January 1992, so that by 1993, when drilling likelihoods start to be taken, an equilibrium is approximately reached. This extension requires the interpolation of missing rig dayrate data for the fourth quarter of 1992.

A2.2 Estimation

I search for the parameters β, μ, and log ζ that maximize the log-likelihood function (13) via a gradient-based search that uses the BFGS method for computing the Hessian at each step (I take the logarithm of ζ to allow for negative values in the parameter search). I accelerate the search by conducting it in two stages. First, holding β fixed, I search for the μ and log ζ that maximize the likelihood. This stage is fast because changing μ and ζ does not require re-solving the model. The outer-most loop then searches for β. The stopping criterion is a tolerance on the likelihood function (scaled down by a factor of 10,000) of 10^{-10} for the μ and ζ loop and 10^{-8} for the β loop.

To compute the standard errors of the parameter estimates, I obtain the likelihood score of each observation (drilling prospect - month) numerically. With respect to each parameter θ_k, I calculate the derivative of the log likelihood for observation j as $\frac{\ell_j(\theta_k+\varepsilon_k)-\ell_j(\theta_k-\varepsilon_k)}{2\varepsilon_k}$. For the parameters β and μ, I use a value for ε_k of 0.001, and for log ζ I use a value of 0.0001 because the likelihood function is particularly concave in this parameter. The standard errors are robust to values of ε_k that are an order of magnitude larger or smaller.

I adjust the standard errors to account for the fact that the parameters of the expected price drift function (11) are estimated in a first stage.[7] Denoting the first-stage parameters (κ_{p0}, κ_{p1}, and κ_{p2}) and log-likelihood function by θ_1 and L_1, and denoting the second-stage parameters (β, μ, and ζ) and log-likelihood function by θ_2 and L_2, I apply the procedure of Murphy and Topel (1985) using equation (A2.1),

[7] The volatility of volatility (γ), the ratio of dayrate volatility to oil price volatility (α), and the correlation between dayrate and price shocks (ρ) are also estimated in a first stage. However, I found that these parameters contributed only negligibly to the standard errors of the main parameter estimates in the reference case model. To reduce computational burden, the results presented in the paper therefore ignore these parameters when computing Murphy and Topel two-step standard errors. In the mean-reverting volatility beliefs specifications, I also account for sampling error in the estimation of the parameters governing the volatility mean reversion function.

$$\Sigma = R_2^{-1} + R_2^{-1} R_3' R_1^{-1} R_3 R_2^{-1} \tag{A2.1}$$

where Σ denotes the corrected variance-covariance matrix for θ_2, and

$$\begin{aligned} R_1(\theta_1) &= E \frac{\partial L_1}{\partial \theta_1} \left(\frac{\partial L_1}{\partial \theta_1} \right)' = -E \frac{\partial^2 L_1}{\partial \theta_1 \partial \theta_1'} \\ R_2(\theta_2) &= E \frac{\partial L_2}{\partial \theta_2} \left(\frac{\partial L_2}{\partial \theta_2} \right)' = -E \frac{\partial^2 L_2}{\partial \theta_2 \partial \theta_2'} \\ R_3(\theta) &= E \frac{\partial L_2}{\partial \theta_1} \left(\frac{\partial L_2}{\partial \theta_2} \right)' = -E \frac{\partial^2 L_2}{\partial \theta_1 \partial \theta_2'} \end{aligned} \tag{A2.2}$$

R_1 is simply the inverse of the variance-covariance matrix from the least-squares estimate of the price drift function (11), which I compute using standard errors clustered on month-of-sample.[8] R_2 is the inverse of the unadjusted (and non-clustered) second-stage variance-covariance matrix. Calculation of R_3 requires numerical derivatives of the second-stage likelihood function with respect to the first-stage parameters. I calculate these derivatives in the same way that I calculate those with respect to the second stage parameters, as discussed above. The perturbations I use for κ_{p0}, κ_{p1}, and κ_{p2} are 10^{-5}, 10^{-6}, and 10^{-3}, respectively.

For the specifications that yield estimates of β near one, the above procedure roughly increases the estimated standard errors by a factor of 3, a magnitude similar to that found in several examples in Murphy and Topel (1985). The adjustment is not substantial for other specifications, however, as their unadjusted standard errors are already large.

References

Murphy, Kevin M., and Robert H. Topel. 1985. “Estimation and Inference in Two-Step Econometric Models.” *Journal of Business & Economic Statistics* 3(4): 370-379.

[8] Clustering on year rather than month-of-sample does not substantially affect the estimated standard errors.

Appendix 3: Estimation including productivity realizations

This appendix provides details of the process by which I use production data from the subset of wells for which production is observable to estimate an expanded version of the structural model. I first discuss how I transform the raw production data into estimates of each well's total discounted lifetime productivity. I then discuss the construction of an augmented likelihood function that incorporates these productivity data.

A3.1 Calculating discounted lifetime productivity

For 160 of the 1,150 wells in the sample, I observe the well's monthly production for the first three years of the well's life. The dynamic model presented in the main text, however, is based on the productivity of each well, defined as its discounted total lifetime production divided by its drilling cost at the average rig dayrate. To transform the three years of production data for each well into an estimate of discounted total lifetime production, I employ a decline curve analysis. The simplest possible approach would be to fit a hyperbolic curve to the average production decline data shown in figure 2 in the paper and then use this curve to extrapolate production for future years of each well's life. However, one strong feature of the production data is that decline rates are less steep for wells that are relatively productive. Therefore, I allow the parameters governing the hyperbolic decline to vary with the observed three-year production volumes.

Specifically, denoting the production from well *i* in month *t* as q_{it}, and denoting the log of well *i*'s total production in its first three years as Q_{i3}, I estimate the hyperbolic decline equation (A3.1) on the pooled monthly data from all 160 wells:

$$\frac{q_{it}}{Q_{i3}} = (\alpha_0 + \alpha_1 Q_{i3})(1+\beta t)^{-(\gamma_0+\gamma_1 Q_{i3})} \tag{A3.1}$$

The parameters α_1 and γ_1 allow the estimated decline curve to steepen or flatten for more productive wells.[9] I estimate that α_0 = 0.337, α_1 = -0.0269, β = 0.144 (*t* is measured in months since drilling), γ_0 = 3.97, and γ_1 = -0.319. The negative estimates for α_1 and γ_1 are consistent with a shallower decline rate for more productive wells.[10]

---

[9] While the β term could in principle also be interacted with total three-year production, it becomes very difficult for the estimator to converge when this interaction is included. Intuitively, allowing for this additional flexibility is unnecessary, as providing flexibility in the decline curve intercept (through α_1) and "slope" (through γ_1) is sufficient.

[10] As an alternative approach, I have also attempted to estimate decline curves well-by-well. However, the estimates are generally too noisy to be useful, especially as some wells are actually estimated to have increasing production over their first three years, which makes it impossible to project an eventual decline.

For each of the 160 wells, I use the estimate of equation (A3.1) to extrapolate future production, and I then apply the discount factor used in the model (see section IV.A of the paper) to obtain the well's discounted total lifetime production. Finally, I divide this number by the well's estimated drilling cost at the average dayrate (this cost depends on the number of days needed to drill the well, as described section I.E in the paper) to obtain its realized productivity (in barrels per $ of drilling cost). These realized productivity data are plotted in figure 10 in the main text.

A3.2 Augmenting the likelihood function

With the inclusion of the realized productivity data, the likelihood function must now incorporate the probability of each productivity realization (which will depend on x* each period and on the variance of productivity realizations about their expectation) and the probability that production is observable for each drilled well (which will depend on the well's productivity). The resulting likelihood function involves three pieces, which I now describe in turn.

The first piece of the likelihood for each drilled well is that given by equation (13) in the text: the probability that drilling would occur in the month the well was actually drilled (or not occur at all during the sample in the case of an undrilled prospect). This part of the likelihood does not change, and for undrilled prospects this is the only part of the likelihood.

The second piece of the likelihood applies only to wells for which productivity is observed (160 of the 1,150 wells). Some notation is required. Let Y_{it} denote the realized productivity of well i drilled in month t, and let y_{it} denote its log. Let Z_{it} denote the expected productivity of well i drilled in month t, and let z_{it} denote its log. This expectation is the firm's expectation of the well's productivity after it has made the decision to drill but before drilling is completed. Thus, z_{it} has a normal distribution, with mean μ and standard deviation ζ (the two parameters that govern the distribution of x_{it} as discussed in section IV.B in the text), that is left-truncated at the log of the productivity trigger at time t. In a slight abuse of notation, let x_t* now denote the logged productivity trigger. Finally, let $P_{dry}(x_t^*)$ denote the probability of a dry hole as a probit function of x_t*. Specifically, $P_{dry}(x_t^*)$ is given by equation (A3.2) below, in which τ_0 and τ_1 are parameters to be estimated.

$$P_{dry}(x_t^*) = 1 - \Phi\left(\frac{x_t^* - \tau_0}{\tau_1}\right) \tag{A3.2}$$

Given an expected productivity z_{it}, y_{it} will be -∞ with probability $P_{dry}(x_t^*)$ and will otherwise be normally distributed about z_{it} with variance σ_p^2, following (A3.3):

$$y_{it} \mid z_{it} \sim \begin{cases} -\infty & \text{with prob. } P_{dry} \\ N\left(\dfrac{z_{it}}{\log(P_{dry})} - \dfrac{\sigma_p^2}{2}, \sigma_p^2\right) & \text{with prob. } 1-P_{dry} \end{cases} \tag{A3.3}$$

It is the variance term σ_p^2 that allows for noise in the production realizations so that they can be rationalized by the model. Also note that the distribution of y_{it} is designed so that $E[Y_{it}] = Z_{it}$.

Let $f(y_{it}|z_{it})$ denote the distribution of y_{it} conditional on z_{it} and on the well not being dry (i.e., $f(y_{it}|z_{it})$ is the second part of (A3.3)). Let $g(z_{it}|x_t^*)$ denote the truncated normal distribution of z_{it}, conditional on x_t^*. The contribution of production realization y_{it} to the likelihood is then:

$$P_{dry} \text{ if } y_{it} = 0$$
$$(1-P_{dry}) \cdot \int_{x_t^*}^{\infty} f(y_{it} \mid z_{it}) g(z_{it} \mid x_t^*) dz_{it} \text{ if } y_{it} > 0$$

The third and final piece of the likelihood contribution from each drilled well is the probability that production from the well is observable in the data. There are two components of this probability, which I denote by P_{obs}. The first is the probability that the well can be matched to a lease name in the production database. I take this probability, which I denote by P_{match}, to be exogenous to the model and fix it to equal the observed match rate in the data, 527/1150. The second component addresses selection. The probability of observing a drilled well's production should increase if its realized productivity is low relative to expectations, and it should also increase if x_t^* is high, since fewer wells are drilled when the trigger productivity is high (and trigger productivities are serially correlated). I therefore specify P_{obs} per equation (A3.4) below, in which λ_0, λ_1, and λ_2 are parameters to be estimated:

$$P_{obs}(y_{it}, x_t^*) = P_{match} \cdot s(y_{it}, x_t^*) = P_{match} \cdot \left(1 - \Phi\left(\frac{(y_{it} - x_t^*) + \lambda_2 x_t^* - \lambda_0}{\lambda_1}\right)\right) \tag{A3.4}$$

Note that equation (A3.4) implies that the probability of observing a dry hole, for which $y_{it} = -\infty$, is equal to P_{match}. Thus, for all wells for which production is observed, the final component of their likelihood is given by $P_{obs}(y_{it}, x_t^*)$.

For the drilled wells for which I do not observe production, I must compute the probability that production is unobserved, conditional on the trigger productivity x_t^* at the time of drilling. This computation requires a double integral over realized productivity conditional on expected productivity and over expected productivity conditional on x_t*. The probability that production is unobserved is therefore given by equation (A3.5):

$$P_{unobs}(x_t^*) = 1 - P_{match}\left(P_{dry}(x_t^*) + \int_{x_t^*}^{\infty} \left[\int_{-\infty}^{\infty} (1 - P_{dry}(x_t^*)) s(y_{it}, x_t^*) f(y_{it} \mid z_{it}) dy_{it} \right] g(z_{it} \mid x_t^*) dz_{it} \right) \quad \text{(A3.5)}$$

For all wells for which production is unobserved, the final component of their likelihood is given by $P_{unobs}(x_t^*)$. This completes the likelihood.

Estimation involves six parameters not present in the reference case model: the parameters τ_0 and τ_1 that dictate how the dry hole probability varies with x_t*, the parameter σ_p that dictates the variance of the realized production data, and the parameters λ_0, λ_1, and λ_2 that dictate which productivity observations are likely to be observable. The estimates of these parameters, which correspond to the estimates presented in column V of table 3 in the text, are:

τ_0: -9.257 (5.492)
τ_1: 2.997 (5.520)
σ_p: 1.765 (0.068)
λ_0: 1.712 (5.492)
λ_1: 5.847 (3.472)
λ_2: -2.061 (2.560)

The estimates of τ_0 and τ_1 are imprecise but consistent with a modest decrease in the probability of a dry hole as x_t* increases. At the sample average x_t* of -2.629, the estimates imply that the probability of a dry hole is 1.4%. For comparison, I observe 7 dry holes out of the 160 observed wells and 1,150 total wells. The σ_p parameter is large and precisely estimated, consistent with the noise in realized productivity plotted in figure 10 in the paper. The λ_0, λ_1, and λ_2 estimates are consistent with the probability of observing production varying negatively with realized productivity and positively with x_t*. The parameters imply that a one standard deviation increase in realized productivity from the sample mean x_t* of -2.629 to -2.629 + σ_p reduces the probability of observing the well's production from 12.1% to 8.0%. For reference, I observe production for 13.9% (=160/1150) of the drilled wells. Thus, given the parameter estimates, observed productivity realizations must on average be lower than the sample mean x_t* in order to attain the 13.9% observation rate.

Figure A1: Infill drilling in sole operated vs. all fields

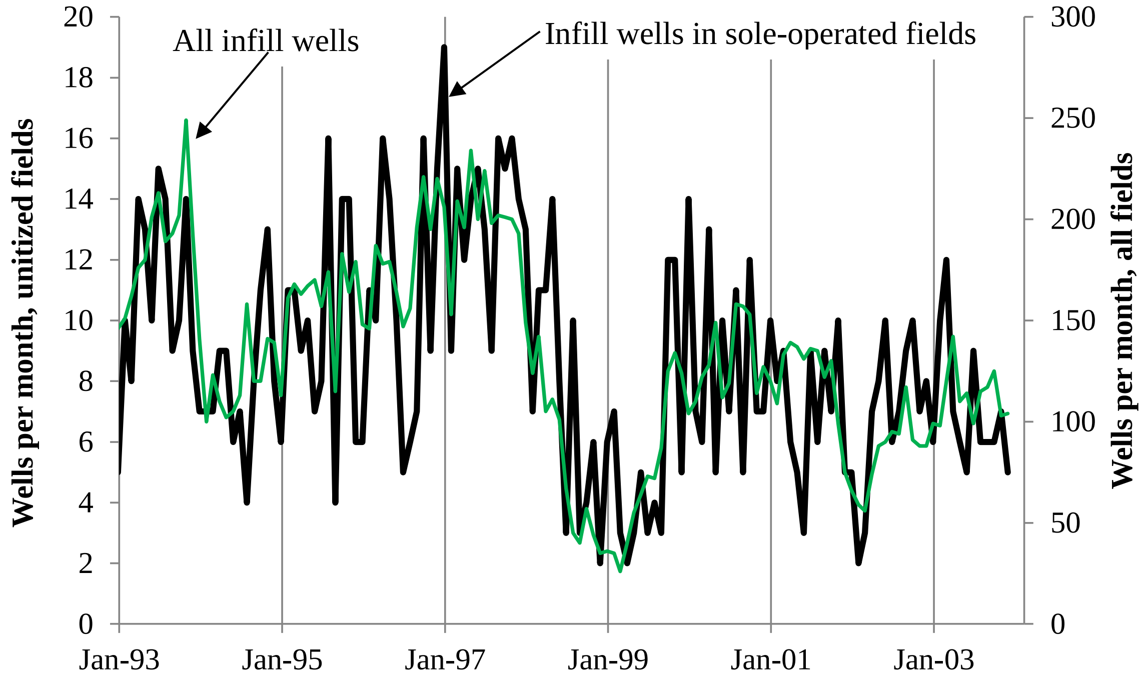
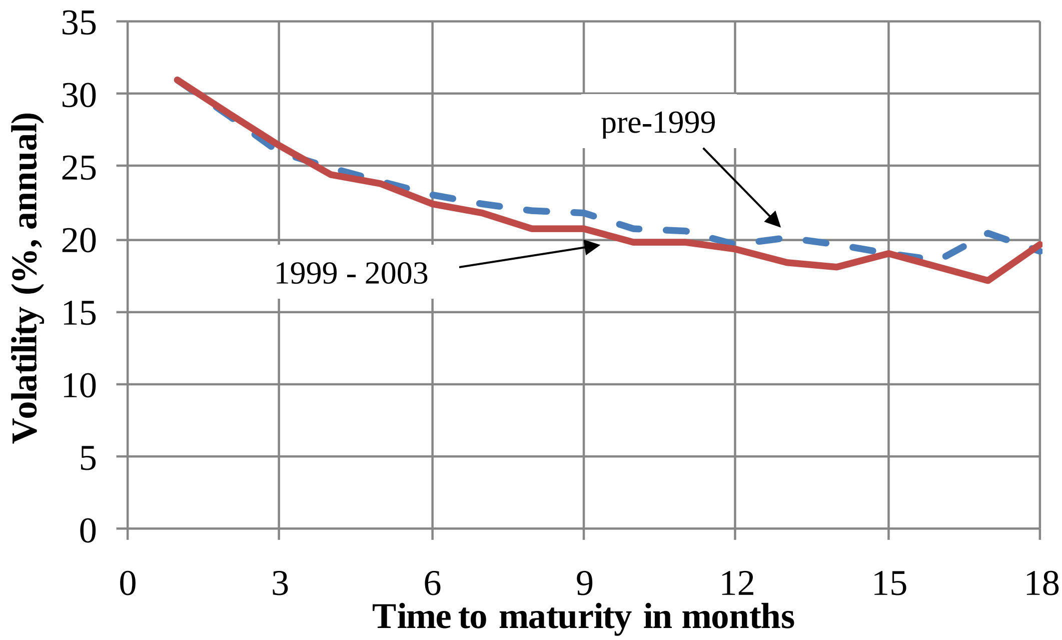


Figure A2: Estimated term structures of futures price volatility



Notes: The figure displays two term structures, one estimated using data from before 1999, the other using data from 1999-2003. Volatility of a one-month future is set to 31.0% for both term structures.