

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 $\ln(F_{s,\tau} / F_{s-1,\tau})$ 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

$$\ln \sigma_{t,\tau} = \eta_{\tau} + \delta_t + \varepsilon_{\tau,t} \tag{A1.1}$$

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 $A_t \exp(\eta_\tau)$, 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_t). 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 P , D , and σ 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 $\log \zeta$ 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 $\log \zeta$ 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 $\log \zeta$ 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.

$$\Sigma = R_2^{-1} + R_2^{-1} R_3' R_1^{-1} R_3 R_2^{-1} \quad (\text{A2.1})$$

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

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

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:

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

The parameters α_1 and γ_1 allow the estimated decline curve to steepen or flatten for more productive wells.⁹ I estimate that $\alpha_0 = 0.337$, $\alpha_1 = -0.0269$, $\beta = 0.144$ (t is measured in months since drilling), $\gamma_0 = 3.97$, and $\gamma_1 = -0.319$. 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.

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

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

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

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:

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

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:

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

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):

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

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:

$$\tau_0: -9.257 (5.492)$$

$$\tau_1: 2.997 (5.520)$$

$$\sigma_p: 1.765 (0.068)$$

$$\lambda_0: 1.712 (5.492)$$

$$\lambda_1: 5.847 (3.472)$$

$$\lambda_2: -2.061 (2.560)$$

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 $-2.629 + \sigma_p$ 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

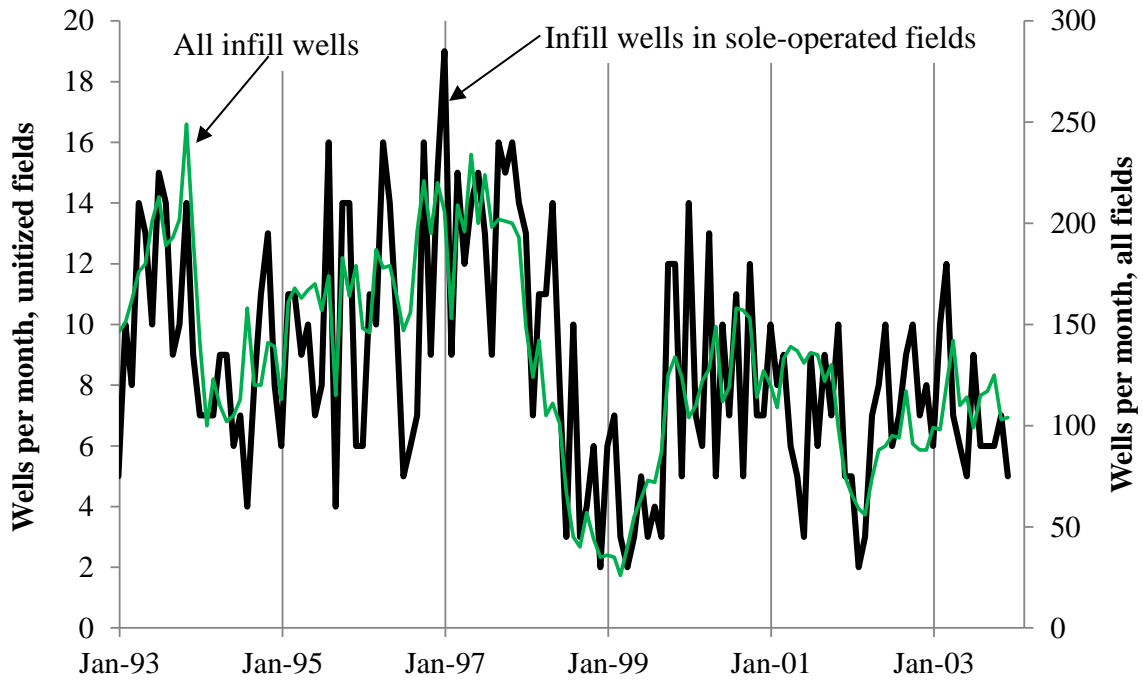
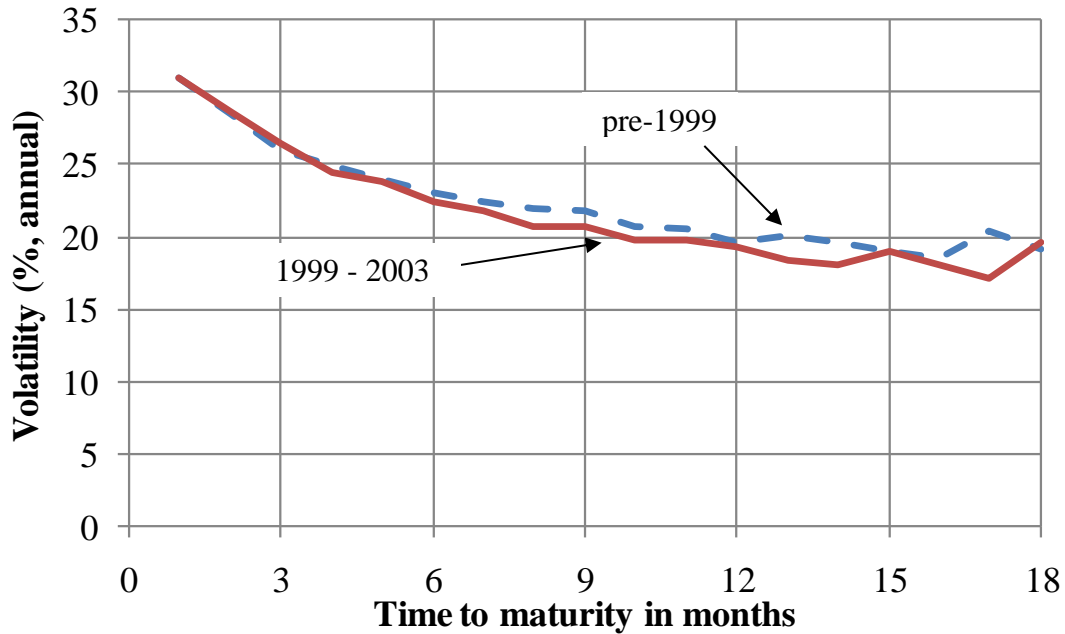


Figure A2: Estimated term structures of futures price volatility



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.