Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis

Author(s): William R. Melick and Charles P. Thomas

Source: *The Journal of Financial and Quantitative Analysis*, Mar., 1997, Vol. 32, No. 1 (Mar., 1997), pp. 91-115

Published by: Cambridge University Press on behalf of the University of Washington School of Business Administration

Stable URL: https://www.jstor.org/stable/2331318

JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS VOL. 32, NO. 1, MARCH 1997

Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis

William R. Melick and Charles P. Thomas*

Abstract

We develop a general method for estimating the implied, martingale equivalent, probability density function (PDF) for futures prices from American options prices. The early exercise feature of American options precludes expressing the price of the option in terms of the PDF. There exist tight bounds for the price of American options in terms of the PDF. We demonstrate how these bounds, together with observed option prices, can be used to estimate the parameters of the PDF. We estimate the distribution for crude oil during the Persian Gulf crisis and find the distribution differs significantly from that recovered using standard techniques.

I. Introduction

Option prices contain information about market participants' perceptions of the distribution of the underlying asset. To recover this information, analysts typically assume a stochastic process for the underlying asset, such as Brownian motion; use observed option prices to recover the parameters of the assumed process; and then integrate the process to recover the distribution of the underlying asset. In certain instances, however, it is more natural to begin with an assumption about the future distribution of the underlying asset, rather than the stochastic process by which it evolves, and to use option prices to directly recover the parameters of that distribution.[1] This paper develops a method to directly estimate

*The authors are staff economists in the Division of International Finance, Board of Governors of the Federal Reserve System, Washington, DC 20551. This paper represents the views of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or other members of its staff. We would like to especially thank David Bates (the referee), Jon Faust, Greg Duffee, Christian Gilles, and Ed Green for their extensive comments. We also thank Allan Brunner, Neil Ericsson, Jeff Fuhrer, Ludger Hentschel, and George Moore, as well as seminar participants at Vanderbilt's Owen School of Management, the Ohio State University Economics Department, the American Finance Association winter 1995 meetings, the Bank for International Settlements, the University of Neuchatel, the Deutsche Bundesbank, the European Monetary Institute, and the Federal Reserve Board International Finance Division. Elizabeth Vrankovich and Dara Akbarian provided valuable research assistance.

[1]Other studies that focus on the asset's terminal PDF include Breeden and Litzenberger (1978), Jarrow and Rudd (1982), Shimko (1991), Neuhaus (1995), and Malz (1997).

such a distribution from American options and applies it to the crude oil market during the Persian Gulf crisis. We also compare our estimated distributions to those recovered by more standard methods.

We find that the estimated distributions are consistent with the market commentary at the time, in that they reflect a significant probability of a major disruption in the oil markets. We also find that if policy makers or analysts had used the standard Black-Scholes (1973) model, they would have generally overestimated the market's assessment of the probability of a major disruption and underestimated the impact on prices of such a disruption.

Recovering information from option prices is complicated by two factors. First, option prices incorporate preferences towards risk as well as beliefs about outcomes. Short of modeling these preferences or assuming that oil prices are unrelated to other determinants of investor wealth and, thus, the oil-price risk is unpriced, the estimated parameters of the stochastic processes or the implicit distributions can only represent the risk-neutral (martingale equivalent) parameters rather than the true (actuarial) parameters.[2] Second, the early exercise feature of American options makes it difficult to derive a closed form expression for the price of an option. To surmount this problem, we express the option price as a weighted average of upper and lower bounds. These bounds, which are quite tight, are in terms of the risk-neutral distribution of the underlying asset, allowing us to proceed with the estimation of the parameters of the assumed distribution.

Our methodology requires us to place some structure on the form of the implied distribution. This structure is similar in spirit to the assumption of a particular stochastic process made in other studies (for example, the jump-diffusion assumption of Bates (1991)).[3] Given the wide range of possible outcomes during the Persian Gulf crisis, we assume that market participants expected oil prices to be drawn from a mixture of three lognormal distributions. This assumption is discussed more fully in Section III.

Placing structure on the terminal distribution rather than the stochastic process has both costs and benefits. With regard to costs, the recovered distribution is silent about the evolution of the asset price prior to expiration. This means that the technique will provide no guidance for constructing dynamic hedges or replication strategies for the option. In addition, the technique does not allow the time series properties of the underlying asset to be used in ex post evaluation of the model.[4]

The benefits of the technique arise from its flexibility, generality, and directness. A reasonably flexible functional form for the terminal distribution, such as that used here, can easily accommodate a wide variety of shapes for the terminal distribution. Manipulating equally flexible processes can be quite difficult. More importantly, starting with the terminal distribution is a more general approach since

[2]Evidence that the risk-neutral and true parameters are very different would include a finding that the oil futures price did not follow a martingale process. Available evidence (see Dominguez (1989), Kumar (1992), and Deaves and Krinsky (1992)) indicates that the oil futures price does, in fact, martingale. This at least allows for the possibility that our recovered parameters are quite close to the actuarial parameters.

[3]Both Aït-Sahalia and Lo (1995) and Jackwerth and Rubinstein (1996) estimate distributions nonparametrically, using intra-day quotes for European options.

[4]However, the method as a whole, rather than a particular estimate, can be evaluated ex post using EDF tests. See Fackler and King (1990) and Silva and Kahl (1993) for examples.

a given terminal distribution encompasses many stochastic process, where a given process is consistent with only one terminal distribution. Finally, as shown by the bounds, for a given terminal distribution over the relevant horizons, the prices of American options are determined almost entirely by the terminal distribution regardless of the stochastic process generating that terminal distribution. Thus, most of the information in the American option prices pertains to the terminal distribution rather than the particular process. Our technique allows the options data to speak directly to the shape of the distribution rather than forcing them to pass through the potentially distortionary filter of a misspecified process. This is particularly important in situations, such as the Persian Gulf Crisis, where interest is more naturally focused on possible asset price outcomes rather than on the asset price process.

The rest of the paper is organized as follows: Section II presents our method for estimating an asset's PDF from American options prices. Section III discusses the particulars of an application to the oil market, and Section IV presents the results of that application. A summary and concluding remarks are found in Section V.

II. A Method for Recovering an Asset's PDF from American Option Prices

In this section, we describe our method for estimating an asset's PDF from American options prices, present a "standard" or benchmark model, and discuss some of the inherent limitations of options data. The first subsection presents bounds on American option prices in terms of the PDF of the underlying asset and discusses why these bounds are quite close together. The second subsection discusses how we weight the bounds and place some structure on the terminal PDF to arrive at a system that can be estimated by non-linear least squares. The third subsection presents the Barone-Adesi and Whaley (1987) model, which we take as a benchmark method. The fourth subsection discusses why the information content of available options is limited and how this limitation interacts with the functional form assumed for the distribution.

A. Expressing American Options Prices in Terms of the Terminal PDF

With European style options, the relationship between the distribution of futures prices and the option price is very direct. For calls (puts), the value of the option is simply the value of the portion of the distribution above (below) the strike discounted back to the present using an appropriate interest rate.[5] For American style options, the relationship between the distribution and the option price is less direct, owing to the early exercise premium. In general, the option's value will depend on the entire stochastic process for futures prices, not just the distribution for futures prices at the option's expiration. To deal with this early exercise premium, we use recently derived bounds[6] for the maximum and minimum value of an option, given that the futures price is taken from a particular distribution at the

[5] See Cox and Ross (1976) for a discussion of the risk-neutral valuation technique.

[6] The lower bound is well known. The upper bound was first published in Chaudhury and Wei (1994). Melick and Thomas (1992), (1996) derive a discrete time analogue of this bound.

option's expiration. That is, for all stochastic processes that imply a given distribution for the futures at the option's expiration, there are bounds for the option's value that can be expressed in terms of that distribution alone.

The bounds can be written as follows:

$$C_t^u = E_t\left[\max\left[0, f_0 - X\right]\right]; \tag{1}$$

$$C_t^l = \max\left[E_t\left[f_0\right] - X,\quad e^{-rt}\cdot E_t\left[\max\left[0, f_0 - X\right]\right]\right]; \tag{2}$$

$$P_t^u = E_t\left[\max\left[0, X - f_0\right]\right]; \tag{3}$$

$$P_t^l = \max\left[X - E_t\left[f_0\right],\quad e^{-rt}\cdot E_t\left[\max\left[0, X - f_0\right]\right]\right]; \tag{4}$$

where f_0 denotes the (random) price of the underlying asset at the expiration of the option and X denotes the option's strike price. We index time by periods prior to the option's expiration and the total time to expiration is t. $E_t[\,]$ denotes expectations taken with respect to the risk-neutral distribution t periods prior to expiration; r is the risk-free interest rate; and e^{-rt} is the discount factor to the option's expiration. The bounds are predicated on the assumption that the underlying asset martingales with respect to the probabilities used in the $E_t[\,]$.[7]

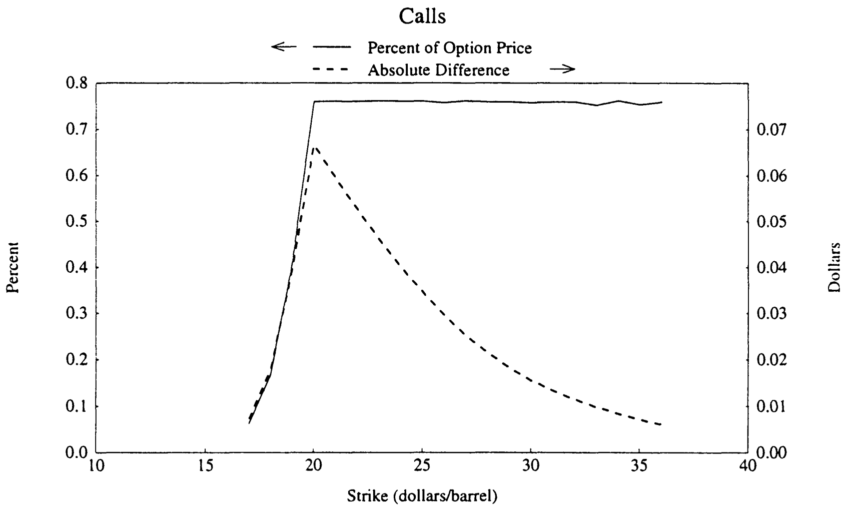
For reasonable discount factors, the bounds are remarkably close together. Note that the upper bound is the undiscounted European value and that, at a minimum, the lower bound is equal to the European value. Thus, at a maximum, the ratio of the upper bound to the lower bound is e^{rt}. For deeply in-the-money options, when the lower bound is determined by the value of exercising today, the ratio is even smaller. Figure 1 illustrates how close together the bounds are using the estimated distribution for a typical day. The two panels plot the distance between the bounds both in absolute terms and as a percent of the actual option price.

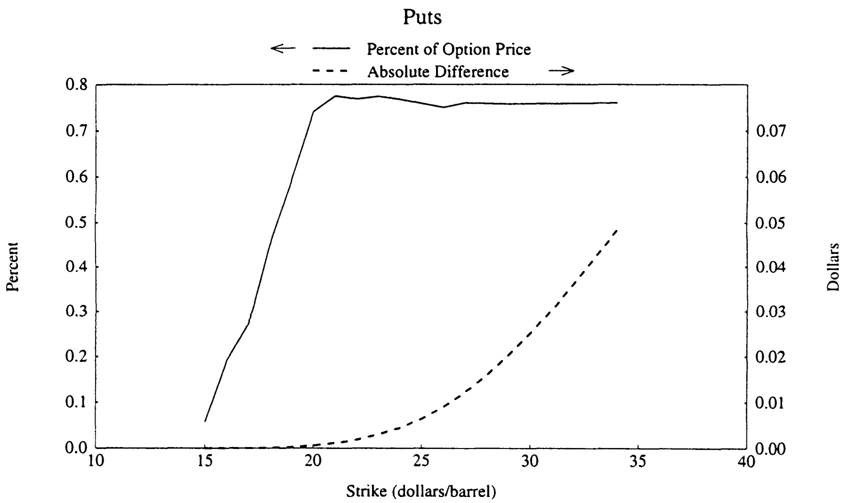
For the day plotted, the options had 38 days to expiration and the relevant T-bill rate was about 7%. Thus, the reciprocal of the discount factor in the formula for the lower bound was about 1.008. On this day, the futures price was about \$29. For the calls with strikes at or above \$20, the difference between the bounds was about 0.8% of the option price. This difference declines in absolute terms as the strike increases and the options become less in-the-money and worth less. For the deeply in-the-money calls (with strikes between \$17 and \$20), the lower bound was determined by the value of exercising today (rather than its European value), and the difference between the bounds was less than 0.8%. The pattern for the puts would generally be the mirror image of that for the calls. For the day plotted, this correspondence is masked for two reasons. First, none of the puts was sufficiently in-the-money for the lower bound to be determined by the value of exercising today, so we never reach the region where the absolute difference falls as the puts move deeply in-the-money. Second, on this day, for deeply out-of-the-money puts, the actual option price was above both estimated bounds, thus,

[7]The first term in the lower bound is usually written as $f_t - X$ rather than $E_t[f_0] - X$. We use the latter formulation to highlight that the bounds can be written in terms of the terminal distribution alone, without reference to the current futures price. This formulation also avoids the problems associated with limit days when trades at f_t may, in fact, not be possible.

FIGURE 1

Width of Bounds on Option Prices
Upper Bound–Lower Bound





the width of the bounds, although still determined by the discount factor, falls to less than 0.8% of the actual option price.[8]

B. Constructing an Estimator

To apply standard estimation techniques requires a point estimate for the option price in terms of a relatively small set of unknown parameters. To generate such a point estimate, we weight the upper and lower bounds given above and assume a specific functional form for the terminal distribution.[9]

Numerous studies have documented that option pricing errors from a variety of option pricing models vary systematically with the degree to which an option is in- or out-of-the-money (see Whaley (1986) and Cakici, Chatterjee, and Wolf (1993) for applications to options on futures). Accordingly, we use two weights, one for options in-the-money and one for options out-of-the-money.

In choosing a functional form for the terminal distribution, we tried to balance flexibility, parsimony, and ease of interpretation. For reasons explained in Section III, we specify that the futures price at the option's expiration is drawn from a mixture of three lognormal distributions (hereafter, MLN).[10] More formally, the distribution function for futures prices, $g[\,]$, is given by

$$g\left[f_0\right] = \pi_1 g_1\left[f_0\right] + \pi_2 g_2\left[f_0\right] + \pi_3 g_3\left[f_0\right], \tag{5}$$

where

$$g_i\left[f_0\right] = \left(\frac{1}{\sqrt{2\pi}\sigma_i \cdot f_0}\right) \cdot \exp\left[\left(\frac{\ln(f_0) - \mu_i}{\sigma_i}\right)^2 \Big/ 2\right]. \tag{6}$$

Let θ denote a vector of the nine parameters of $g[\,]((\pi_i, \mu_i, \sigma_i)\ i = 1,2,3)$ and let (w_1, w_2) denote the weights that describe where the actual option price falls between the bounds. Combining equations (1)–(6) and weighting the bounds yields the following pricing equations in terms of 11 estimated parameters $(\widehat{\theta}, \widehat{w}_1, \widehat{w}_2)$, two observables (X, e^{-rT}), and an error term,

$$C_t[X] = \widehat{w}_{it} \cdot C_t^u\left[X; \widehat{\theta}_t\right] + \left(1 - \widehat{w}_{it}\right) \cdot C_t^l\left[X; \widehat{\theta}_t\right] + \widehat{\epsilon}_{ct}[X], \tag{7}$$

$$P_t[X] = \widehat{w}_{it} \cdot P_t^u\left[X; \widehat{\theta}_t\right] + \left(1 - \widehat{w}_{it}\right) \cdot P_t^l\left[X; \widehat{\theta}_t\right] + \widehat{\epsilon}_{pt}[X], \tag{8}$$

$$\text{where} \quad i = \left(\begin{array}{ll} 1 & \text{if} \left(\begin{array}{l} \text{call and } X < E_t\left[f_0; \widehat{\theta}\right] \\ \text{put and } X > E_t\left[f_0; \widehat{\theta}\right] \end{array}\right) \\ 2 & \text{otherwise} \end{array}\right).$$

[8]The bounds illustrated in Figure 1 are derived from a terminal distribution estimated by the mixture technique described below.

[9]There are many bounds in the option literature but, to our knowledge, this is the first time bounds on the price of American options have been used to estimate the risk-neutral terminal distribution. Perrakis and Ryan (1984) and Levy (1985) derive bounds for European options in discrete time in terms of the true distribution. Grundy (1991) derives bounds for the true distribution from the moments of the risk-neutral distribution. Lo (1987) derives bounds on European options in terms of the first two moments of the risk-neutral distribution.

[10]This mixture assumption was also used by Ritchey (1990), which derives European options prices when the PDF is a mixture of lognormals.

Details of the derivation are relegated to the Appendix. In brief, expressions involving expectations, conditional expectations, and probabilities implicit in equations (1)–(4) are replaced with the relevant functions of the nine parameters in θ, under the restriction that the futures price martingales.

The error, $\widehat{\epsilon}$, will be the result of any error in estimating the weights or the parameters of the distribution plus any noise in the system, two examples of which are immediately obvious. First, as the same weights are applied across all options for a given contract/day, there will be a pricing error induced by weighting the two bounds. Second, actual option prices are rounded to the nearest penny, also creating errors in the equation.

The parameters of the model, exemplified by equations (7) and (8), are estimated by minimizing the sum of squared errors for all options on a given contract/day, imposing the following constraints,

$$\sum_{i=1}^{3} \pi_i = 1; \quad 0 \leq w_i \leq 1, \quad i = 1,2. \tag{9}$$

The restriction on the sum of the πs reduces the number of free parameters from 11 to 10. An additional restriction could be imposed by constraining the vector $\widehat{\theta}$ such that, under $g[\cdot;\widehat{\theta}]$, $E_t[f_0]$ equals the recorded settle price of the futures. This would reduce the number of estimated parameters to nine. However, as discussed below, price limits on the futures were binding for about 80 contract/days during the period. On these days, the futures settle price did not reflect its expectation, and this parameter restriction would not be valid. In addition, by not imposing this restriction, we were able to use the futures price as a measure of the goodness-of-fit for the estimated distribution. Details on the data and estimation are found in Section III.

C. Standard Model as Benchmark

In order to gauge the results from our model (MLN), the distribution for futures prices was also recovered using a "standard" option pricing model. The most common assumption in the option pricing literature is that the underlying commodity price follows a geometric Brownian process, which implies that the futures price at expiration will be drawn from a single lognormal distribution. This is the assumption behind Black's (1976) model for pricing European options on futures. Several approximations have been developed to price American options under this "standard" assumption, with the quadratic approximation of Barone-Adesi and Whaley (1987) (hereafter, BAW) being the most easily calculated and most commonly used.[11] We constructed a "standard" model by assuming that prices will be drawn from a single lognormal distribution, and used the BAW approximation to generate option pricing equations. We recovered the two parameters (μ_b and σ_b) of the single lognormal (hereafter SLN) distribution by minimizing the sum of squared deviations predicted from actual option prices. (See the Appendix for details of the BAW approximation and SLN estimation.)

[11]Overdahl and Matthews (1988), when studying a more tranquil period in the oil market, use the BAW approximation.

Thus, there are two differences between the SLN and MLN models. First, the models account for the early exercise premium in different ways—SLN uses the BAW quadratic approximation to price an American option, while MLN uses weighted upper and lower bounds. Second, SLN assumes that the futures price at expiration will be drawn from a single lognormal distribution, while MLN assumes that the futures price will be drawn from a mixture of lognormal distributions. SLN is almost nested within MLN, except that SLN uses a different technique to account for the early exercise premium. This non-nesting will become important when the two models are statistically compared.

D. Data Limitations vis-à-vis the Distribution

Before proceeding to the estimation, it is useful to note how data limitations and the assumed functional form for the distribution interact. These comments apply to all methods that extract PDFs from options prices. The way in which option prices inform us about the higher moments of a distribution is through terms of the form $E_t[\max[0, f_0 - X]]$ and $E_t[\max[0, X - f_0]]$, where the expectations are again taken over the risk-neutral distribution. These two terms can be written as $(E_t[f_0|f_0 \geq X] - X) \cdot \text{Prob}_t[f_0 \geq X]$ and $(X - E_t[f_0|f_0 < X]) \cdot \text{Prob}_t[f_0 < X]$. It is clear that even if there were no errors in the pricing relations, the fact that strikes are at discrete intervals and, more importantly, that they do not span the entire support of futures prices, places an important limitation on what the option prices can reveal about the distribution.[12] The recorded option prices only contain information about the conditional expectation and probability mass in the following segments of the support: i) the segment below the lowest strike, ii) the segments between each strike, and iii) the segment above the highest strike. In particular, if X_L and X_H are the lowest and highest strikes, then all the information revealed by the options will be in terms of the following,

$$E_t\left[f_0|f_0 < X_L\right] \cdot \text{Prob}_t\left[f_0 < X_L\right], \tag{10}$$

$$E_t\left[f_0|X_i < f_0 < X_j\right] \cdot \text{Prob}_t\left[X_i < f_0 < X_j\right] \quad X_L \leq X_i < X_j \leq X_H, \tag{11}$$

$$E_t\left[f_0|f_0 \geq X_H\right] \cdot \text{Prob}_t\left[f_0 \geq X_H\right]. \tag{12}$$

Any number of distributions could generate the same results for the conditional expectations and probabilities in (10)–(12). For example, for any given distribution we can construct a second distribution out of a series of non-overlapping uniform densities that will be observationally equivalent to the given distribution relative to the data described by (10)–(12).

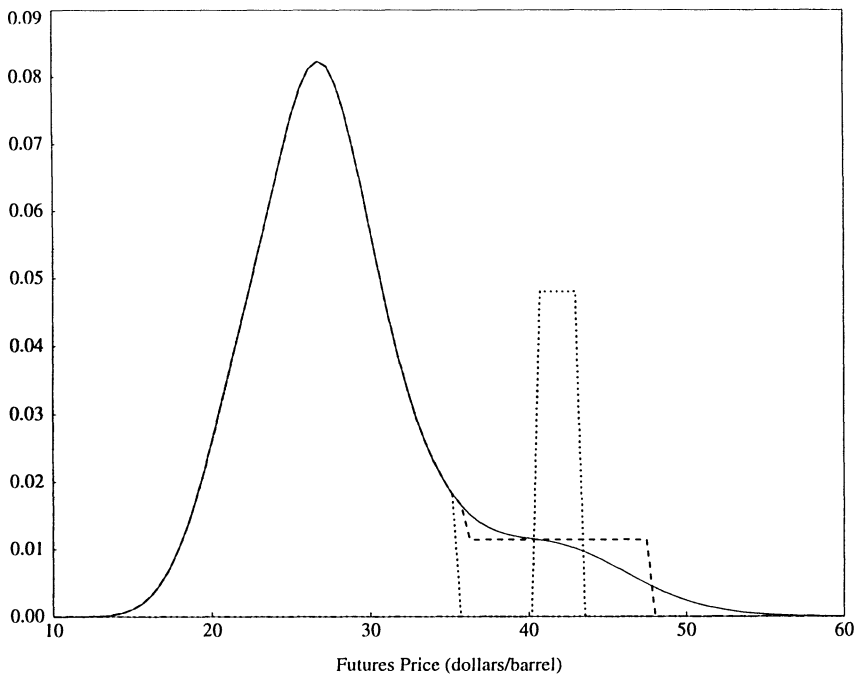
Thus, it is clear that any estimated distribution requires careful interpretation, especially in the regions below the lowest strike and above the highest strike. For crude oil, strikes are almost always \$1.00 apart (in a few instances \$5.00), allowing a fine demarcation of the distribution within the range of strikes. In the tails beyond the strikes, however, we have information only on the conditional

[12] See Rubenstein (1994) for an extensive discussion of this point.

expectations and the probabilities. Thus, the shape of the distribution in the tails will depend importantly on the functional form assumed for the distribution. Figure 2 illustrates this point with three observationally equivalent distributions. The solid line is a mixture of three lognormals, while the dashed lines replace the upper tail with uniform densities that yield the same results for (12).

FIGURE 2

Observationally Equivalent Density Functions



III. Application to the Oil Market

A. Data Sources

Data on settle prices for all crude oil options on futures for all trading days over the period July 2, 1990, through March 30, 1991, were purchased from NYMEX. We used the settle price as the value of the option in equations (7) and (8). The settle price is determined at the end of each day by a settlement committee made up of roughly 20 options market participants. The committee frequently relies on the average of bid and ask prices during the last minutes of trading as starting points for the settlement prices. Heavily traded options are priced first, with put-call parity used to price low volume options at the same strike when the futures market has settled. Price limits were in place over the sample on the futures contract but not the options. As mentioned above, price limits for the futures were hit on roughly 80 of the 642 days in our sample.[13] In this event, put-call parity cannot be used to arrive at settle prices for the thinly traded options, forcing the settlement committee to rely on the futures price from the unconstrained nearby contract and spread trading.[14] Aside from these limit days, using settle option prices avoids the problems associated with asynchronous quotes inherent in transactions data.

During July 1990 through March 1991, trading was concentrated in seven contracts. For each contract/day, all option prices that were recorded with no open interest, no volume, and no exercises were excluded from the data set. In addition, trading days within five working days of the contract's expiration were also excluded from the data set.[15] Table 1 lists summary information for each of the contracts after the exclusions.

TABLE 1

Options on Crude Oil Futures—Contract Summary Statistics

Contract	Calendar Range Used in Estimation	Total Days	Total Options	Number of Options per Day		Strikes Traded ($)		Futures Prices ($)	
				Min	Max	Min	Max	Min	Max
Oct. 1990	7/2/90–8/29/90	41	1254	17	42	14	37	17.74	31.93
Nov. 1990	7/2/90–10/4/90	66	2335	15	57	15	45	18.11	40.42
Dec. 1990	7/2/90–11/1/90	84	3150	13	55	15	44	18.38	38.80
Jan. 1991	7/2/90–11/29/90	104	3889	13	48	16	42	18.55	37.30
Feb. 1991	8/2/90–1/3/91	104	3866	11	43	5	51	23.27	35.95
Mar. 1991	9/10/90–1/31/91	99	3588	12	51	10	50	18.99	34.53
Apr. 1991	8/1/90–2/28/91	144	4827	11	44	10	45	17.91	33.21

This table reports summary statistics for the seven New York Mercantile Exchange contracts used in estimation.

Daily prices for the seven Treasury bills that matured as close as possible after the options contracts expired were used to calculate the discount factors. For each contract/day, there are N (number of options) equations similar to (7) and (8) that form a constrained, non-linear minimization problem. Among the seven contracts, there are a total of 642 contract/days; each contract/day was treated separately, therefore, 642 minimizations were performed. Each day yielded two sets of parameter estimates, the set of 11 parameters from MLN and the set of two parameters from SLN.

B. Estimation

Throughout the Persian Gulf crisis, market commentary focused on three distinct outcomes: i) a return to pre-crisis conditions, e.g., Iraq would peacefully withdraw from Kuwait; ii) a severe disruption to Persian Gulf oil supplies, e.g., damage to Saudi Arabian facilities during a war; and iii) a continuation of unsettled

[13] In December of 1990, the limits on crude oil futures price movements were widened substantially.

[14] We are grateful to NYMEX Board of Directors member, Jim Zamora, of ZAHR Trading and former NYMEX employee, Brad Horne, for their descriptions of the settlement prices.

[15] One day's worth of data for the December contract was also excluded due to an obvious error in data entry on the part of NYMEX.

conditions over the relevant horizon, e.g., a prolonged stalemate in which outcomes 1 or 2 might eventually occur. Given these three possibilities, we chose a mixture of three lognormals as the form of the distribution to be estimated. If, in fact, market participants felt that prices were likely to be drawn from a tri-modal distribution, this could be easily captured by the mixture. Moreover, the mixture could also easily accommodate a single lognormal distribution if that would best fit the data (e.g., $\pi_1 = \pi_2 = 0$). Ex ante, we expected that as news hit the market, the relative weighting of the three lognormals might change, as well as the parameters of each of the three lognormals. For example, news of an Iraqi rocket attack on a Saudi Arabian oil field might increase the weighting on the lognormal distribution with the highest mode, as well as increase the relevant range encompassed by this lognormal distribution. Section IV presents estimated distributions for selected events during the Persian Gulf crisis.

Estimation of MLN and SLN was performed with the Numerical Algorithms Group (NAG) FORTRAN algorithm E04UPF on an IBM RS-6000. Bounds for the parameters were set so that $0 < \widehat{\mu}_i < \infty$, $0.0001 < \widehat{\sigma}_i < \infty$.[16] Analytic derivatives were provided for both estimations. The derivatives were calculated using Mathematica and they were numerically verified within the E04UPF algorithm prior to estimation. (Details of the estimated equations are relegated to the Appendix.)

The estimation procedures are illustrated in Figure 3. The top panel plots the estimated density function using both the MLN and SLN models. Given the density from the single lognormal, the BAW formulae give predicted values for the option prices. The triangles in the lower panels plot the difference between these SLN predictions and the actual prices.

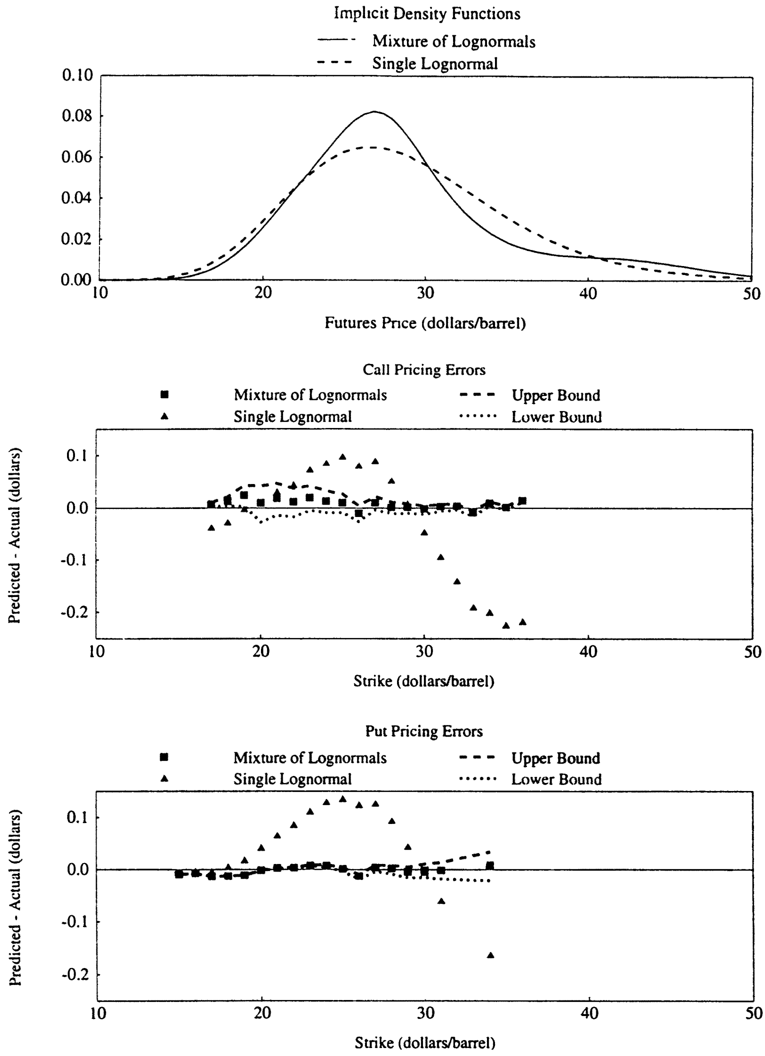
Given the MLN estimated distribution, we can compute the upper and lower bounds for the option prices. The dashed (dotted) line in the lower panels plots the difference between the upper (lower) bound for the option price and the actual option price. The predicted option price is a weighted average of these bounds, where (as noted above) the weights are determined in the minimization routine. The squares plot the difference between the option prices predicted from the MLN distribution and the actual prices. We note that this same MLN distribution was used to draw the plots in Figure 1.[17]

[16] In addition, each (μ, σ) pair was restricted such that the probability of the futures price reaching \$150 per barrel was less than 5%, under each of the lognormal distributions. These restrictions prevented the algorithm from taking unreasonable first steps.

[17] The bounds and errors plotted in Figures 1 and 3 differ slightly from those used by the minimization routine for the following technical reason: the minimization routine behaves significantly better if the objective function is differentiable and if analytic derivatives are supplied. The derived formulae for the bounds are not differentiable since they include the max operator. In estimation, we used a logit weighting scheme to construct a differentiable approximation to the max operator where the weights on the two items in the max move to zero and one as the items move farther apart. The data plotted in the figures were constructed using the actual formulae for the bounds, rather than the differentiable approximation, together with the estimated distribution and weights.

FIGURE 3

Representative Results for a Single Day



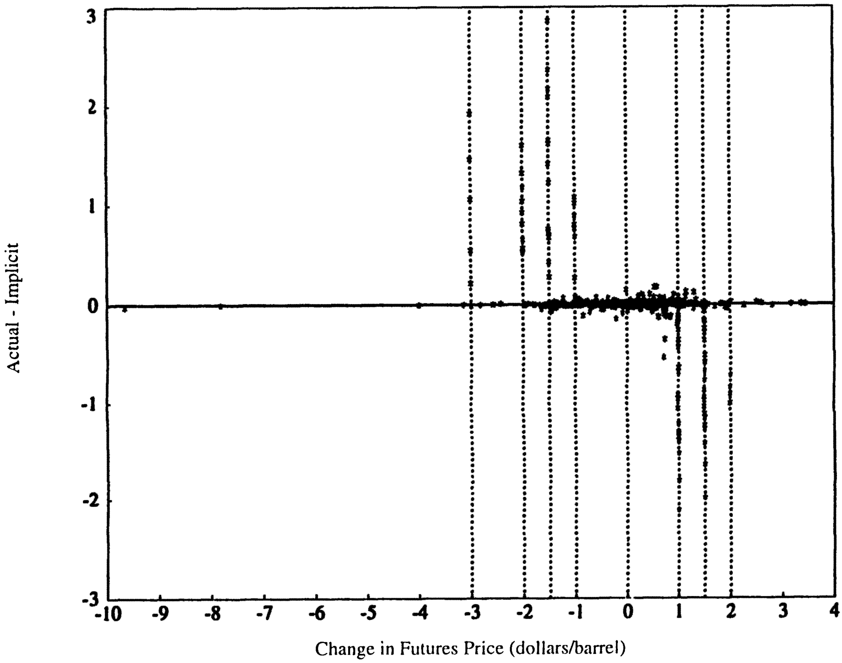
IV. Results

A. Summary Measures

The means of the estimated distributions from MLN and SLN were very similar and were extremely close to the actual futures price. (The actual futures price can be viewed as an independent estimate of the mean of the distribution.) The percentage mean absolute difference (PMAD) between the mean from MLN and the mean from SLN was 0.1%. The PMAD between the mean from MLN and the actual futures price across all contract/days amounted to 0.45%, while the PMAD between the mean from SLN and the actual futures price amounted to 0.51%. Even though they are small, these latter measures significantly overstate the difference between the estimated mean and (under the martingale assumption) the true mean, because they include days for which the futures price was subject to price limits. This can be seen in Figure 4, which plots the difference between the MLN estimated mean and the futures price against the daily movement in the futures price. Note that the contract/days for which there was a substantial discrepancy between MLN's mean and the futures price (points off the horizontal line through zero) were the contract/days for which actual futures prices moved exactly \$1.00, \$1.50, \$2.00, \$3.00, or \$4.00, that is, contract/days for which there was a limit move on the futures contract. As discussed above, there were no limits in the options market, hence, the mean from MLN on these days should not be expected to equal the futures price.

FIGURE 4

Actual-Implicit Futures Price vs. Change in Actual Futures Price (All Contracts)



The similarity between the MLN and SLN distributions does not carry over to higher moments. Figure 5 depicts representative probability density functions taken from the two models; the top panel uses estimates from the October contract on July 10th (three weeks before the crisis), and the bottom panel uses estimates from the January contract on October 10th (in the midst of the crisis). Prior to the outbreak of the crisis, there is little qualitative difference in the two estimates, while during the crisis, the estimates from SLN cannot as easily accommodate the significant probability mass above \$50 per barrel without over-weighting the distribution between \$40 and \$50 per barrel.

FIGURE 5

Implicit Density Functions

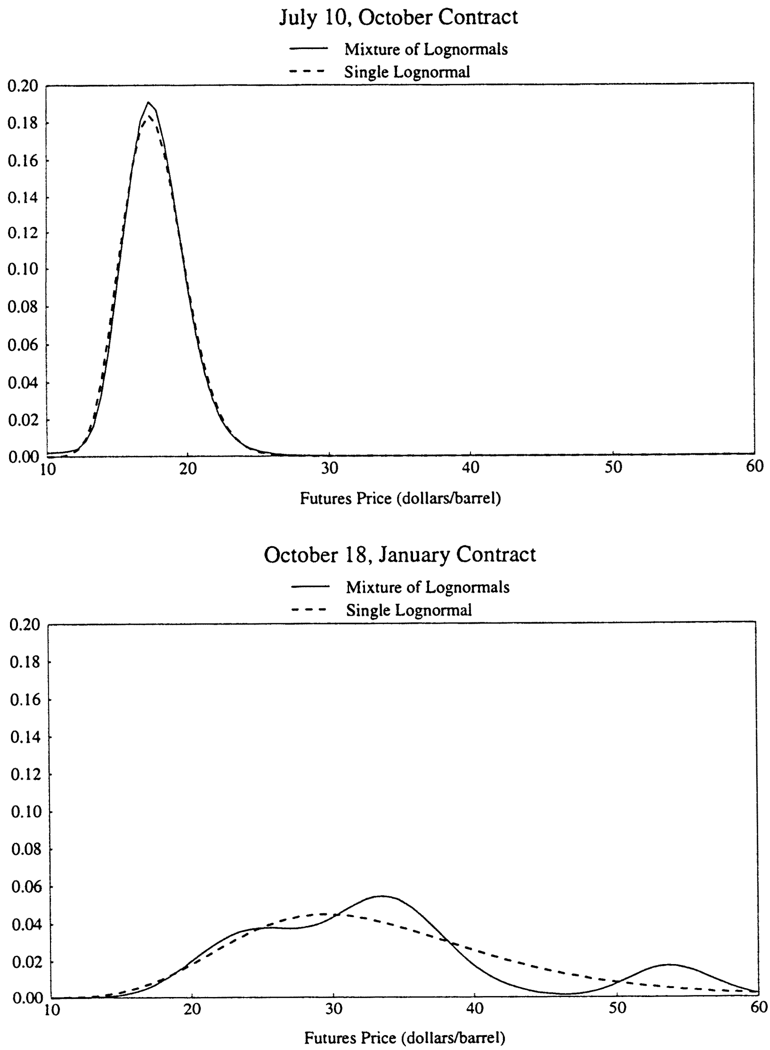


Figure 6 attempts to shed some light on the differences between the right-hand tails of the two estimates, using the April contract as an example. The top panel plots $1.25 \cdot f_t$ along with $\dot{E}_t[f_0 | f_0 > 1.25 \cdot f_t]$ from MLN and SLN. The bottom panel plots $\text{Prob}_t[f_0 > 1.25 \cdot f_t]$ from the two models. The figure shows that the conditional expectation from MLN is generally above that of SLN, while the probability that the futures price will rise by 25% is generally lower in MLN than in SLN. The reason for this result is visually apparent in the bottom panel of Figure 5. The large σ_b, estimated via SLN, forces relatively more of the probability mass to the right but, since the distribution must remain unimodal, leaves the bulk of the right-hand mass nearer the unconditional mean. These results hold across all the contracts. For 574 out of the 642 contract/days, the conditional expectation from MLN is above that of SLN. For 476 out of the 642 contract/days, the probability of being above $1.25 \cdot f_t$ from MLN is below that of SLN. To make this more concrete: if a policy maker or analyst were using the SLN estimates when the MLN were closer to the truth, she would tend to overestimate the market's assessment of the probability of a major disruption while underestimating the impact on prices of such a disruption.

These differences in the right-hand tails of the distributions are also apparent when examining the pricing errors generated by SLN and MLN. The right-hand tail of the distribution will be more important for pricing out-of-the-money calls and in-the-money puts. In Table 2, for out-of-the-money calls across all contract/days, SLN had a mean error (actual−predicted) of \$0.0865, compared to \$0.0005 for MLN. For in-the-money puts, SLN had a mean error of \$0.0445, compared to −\$0.0004 for MLN. For these options, SLN, on average, underpredicted the prices, indicating that SLN did not allocate enough probability mass to the right-hand tail. As might be expected, SLN tended to overpredict the prices for in-the-money calls (mean error of −\$0.0430) and out-of-the-money puts (mean error of −\$0.0388), an overallocation of probability mass to the left-hand tail of the distribution.

TABLE 2

Mean Model Option Pricing Errors
(Actual−Predicted, \$)

	In-the-Money		Out-of-the-Money	
	Calls ($X < f_t$)	Puts ($X > f_t$)	Calls ($X > f_t$)	Puts ($X < f_t$)
SLN	−0.0430	0.0445	0.0865	−0.0388
MLN	−0.0001	−0.0004	0.0005	0.0013

This table reports the average pricing errors for both the single lognormal (SLN) and mixture of lognormal (MLN) models, averaging across all contracts and all trading days. Average pricing errors are categorized according to the type of option (call or put) and moneyness, where X denotes the strike price and f_t denotes the futures price at time t.

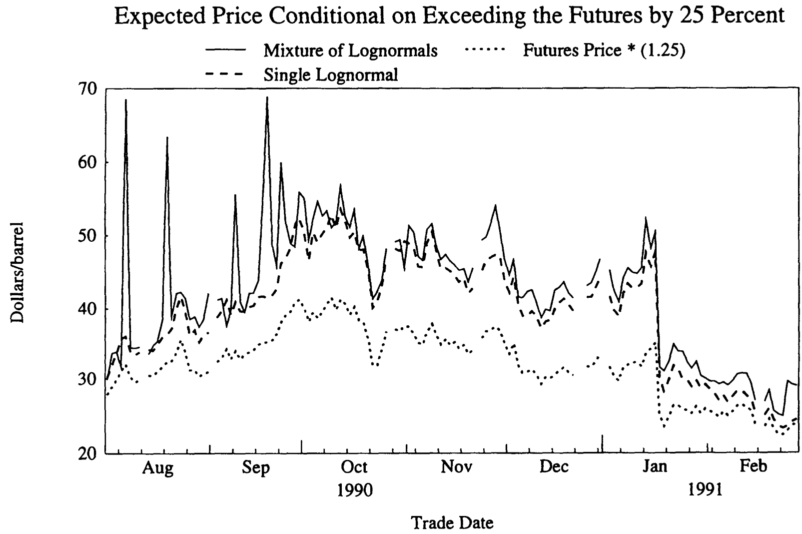
B. Statistical Model Comparison

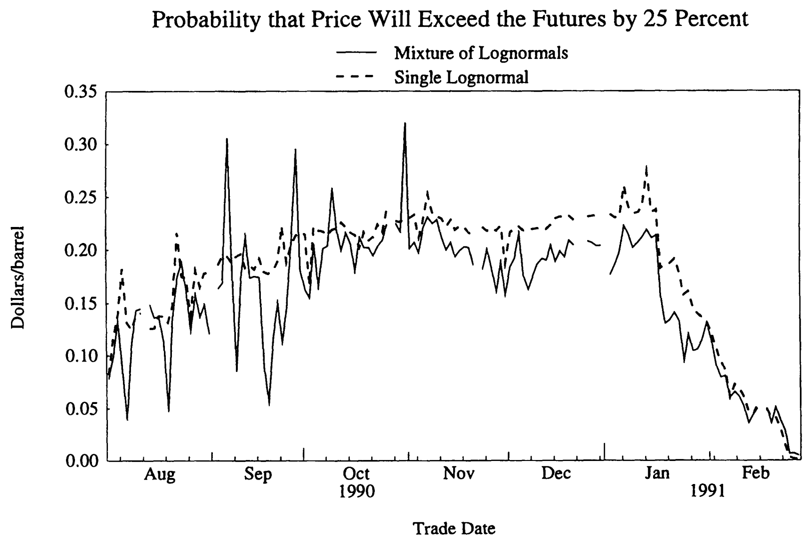
Although the differences between the estimates from MLN and those from SLN are apparent, it may be the case that these differences are not significant in a statistical sense. This issue is complicated since the SLN model cannot be nested within MLN.[18] Since the models are not nested, the asymptotic, chi-square

[18]The nesting issue is as follows: we have two competing non-linear models that explain a vector of option prices (y) on any given day. Denote them by MLN: $y = g[\pi_1, \pi_2, \pi_3, \mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3, w_1, w_2|Z]$ and SLN: $y = h[\mu, \sigma|Z]$, where Z is a data matrix containing strikes and interest

FIGURE 6

Right-Hand Tail Characteristics: Implicit Density Functions
April Contract





assumption cannot be used when forming a likelihood ratio test (or its F-test analog). The standard tests for formally comparing such non-linear, non-nested

rates, $g[\,]$ involves the weighted bounds, and $h[\,]$ involves the BAW approximation. If $g[\,]$ and $h[\,]$ did not represent different functional forms, SLN could be nested within MLN with the restrictions, ($\pi_1 = \pi_2 = 0$, or $\mu_1 = \mu_2 = \mu_3$ and $\sigma_1 = \sigma_2 = \sigma_3$).

models are the J- and P- (or JA and PA) tests.[19] We use the P-test, which compares the two models with the following regression,

$$\widehat{\epsilon}_{1,j} \quad = \quad \widehat{d}_{1,j}\,\beta + \alpha \cdot \left(\widehat{p}_{2,j} - \widehat{p}_{1,j}\right) + \text{residual}, \tag{13}$$

where $\widehat{\epsilon}_{1,j}$ is the pricing error for option "j" from Model 1; $\widehat{d}_{1,j}$ are the derivatives of Model 1 with respect to the parameters of Model 1 for option "j" evaluated at the estimates for the Model 1 parameters; $\widehat{p}_{1,j}$ and $\widehat{p}_{2,j}$ are the fitted values for option "j" from the two models; β is a vector of coefficients with length equal to the number of parameters in Model 1; and α is a single coefficient. The P-test is simply the t-statistic for α. Intuitively, the difference in the fitted values from the two models should not help explain the errors of Model 1. Hence, Model 1 is rejected in the event of a significant t-statistic for α. Obviously, SLN and MLN can both serve as either Model 1 or Model 2. Table 3 presents the t-statistics for α, after equation (13) has been pooled across options and trading days for each contract. The first two columns of the table treat SLN as Model 1 (with degrees of freedom denoted by DF), while the last two columns treat MLN as Model 1.

TABLE 3

P-Tests of Model Specification

	SLN as Model 1			MLN as Model 1		
Contract	*t*-Statistic	DF	*P*-Value	*t*-Statistic	DF	*P*-Value
Oct. 1990	133.98	590	0.00	−0.29	581	0.61
Nov. 1990	172.98	1329	0.00	0.27	1320	0.39
Dec. 1990	130.71	1958	0.00	0.83	1949	0.20
Jan. 1991	122.75	2504	0.00	−0.12	2495	0.55
Feb. 1991	164.54	2504	0.00	−1.28	2495	0.90
Mar. 1991	219.26	2242	0.00	−1.08	2233	0.86
Apr. 1991	260.51	3020	0.00	−0.44	3011	0.67

This table reports *t*-statistics, degrees of freedom, and probability values (*P*-values) for *P*-tests of model specification hypotheses concerning two non-nested, non-linear models. Informally, a significant test statistic indicates that Model 2 is useful in explaining the errors of Model 1, casting doubt on the relative performance of Model 1. On the left side of the table, the single lognormal model (SLN) is treated as Model 1, while on the right side of the table, the mixture of lognormals model (MLN) is treated as Model 1.

Every t-statistic in the first column is significant at the 5% level, while no t-statistic in the fourth column is significant at the 5% level. Clearly, the difference in the fitted values from MLN and SLN helps to explain the pricing errors of SLN, while the converse is not true. Thus, available evidence (the data and MLN) can reject SLN, but MLN cannot be rejected.

MacKinnon (1992) notes that the P-test can have poor finite sample properties, especially when Model 2 has a large number of parameters. For example, in the limiting situation where a completely over parameterized Model 2 exactly matches the observed option prices, the regression in (13) collapses with an infinite t-statistic for α. Given that MLN has 11 parameters and that the average trading day has roughly 35 option prices, there may be reason to question the results of the P-test.

[19] See MacKinnon (1992) for a discussion of these tests.

To ensure that this is not the case, we ran the following Gauss Newton Regression (GNR) described in MacKinnon (1992),

$$\widehat{\epsilon}_{1,j} \quad = \quad \widehat{d}_{1,j}\,\beta + \widetilde{d}_{2,j}\,\gamma + \text{residual}, \tag{14}$$

where $\widehat{d}_{2,j}$ are the derivatives of Model 2 with respect to the parameters of Model 2 for option "j" evaluated at a point chosen independently of the data used to estimate Model 1. These derivatives will not be correlated with $\widehat{\epsilon}_1$ in small samples, and if Model 1 is correct, they will not help explain the errors from Model 1. Table 4 presents the F-tests for $\gamma = 0$. The conclusion is the same as that of the P-test: SLN is rejected while MLN is not.[20]

TABLE 4

Gauss Newton Regression Tests of Model Specification

	SLN as Model 1			MLN as Model 1		
Contract	F-Test	DF (n, d)	P-Value	F-Test	DF (n, d)	P-Value
Oct. 1990	94.77	11, 1031	0.00	0.02	2, 1031	0.98
Nov. 1990	197.32	11, 2045	0.00	0.01	2, 2045	0.99
Dec. 1990	275.64	11, 2872	0.00	0.54	2, 2872	0.58
Jan. 1991	264.02	11, 3638	0.00	0.86	2, 3638	0.42
Feb. 1991	297.17	11, 3638	0.00	0.06	2, 3638	0.94
Mar. 1991	428.72	11, 3321	0.00	2.50	2, 3321	0.08
Apr. 1991	462.05	11, 4594	0.00	0.89	2, 4594	0.41

This table reports F-statistics, degrees of freedom, and probability values (P-values) for Gauss Newton Regression tests of model specification hypotheses concerning two non-nested, non-linear models. Informally, a significant test statistic indicates that Model 2 is useful in explaining the errors of Model 1, casting doubt on the relative performance of Model 1. On the left side of the table, the single lognormal model (SLN) is treated as Model 1, while on the right side of the table, the mixture of lognormals model (MLN) is treated as Model 1.

C. Selected Events

Throughout the Persian Gulf crisis, the oil market often experienced large price movements as "news" hit the market and participants revised their expectations concerning likely outcomes to the crisis. Comparing the estimated PDFs from the two models immediately before and after the receipt of news allows us to infer how the market interpreted the news and further highlights the differences between the MLN and SLN models.

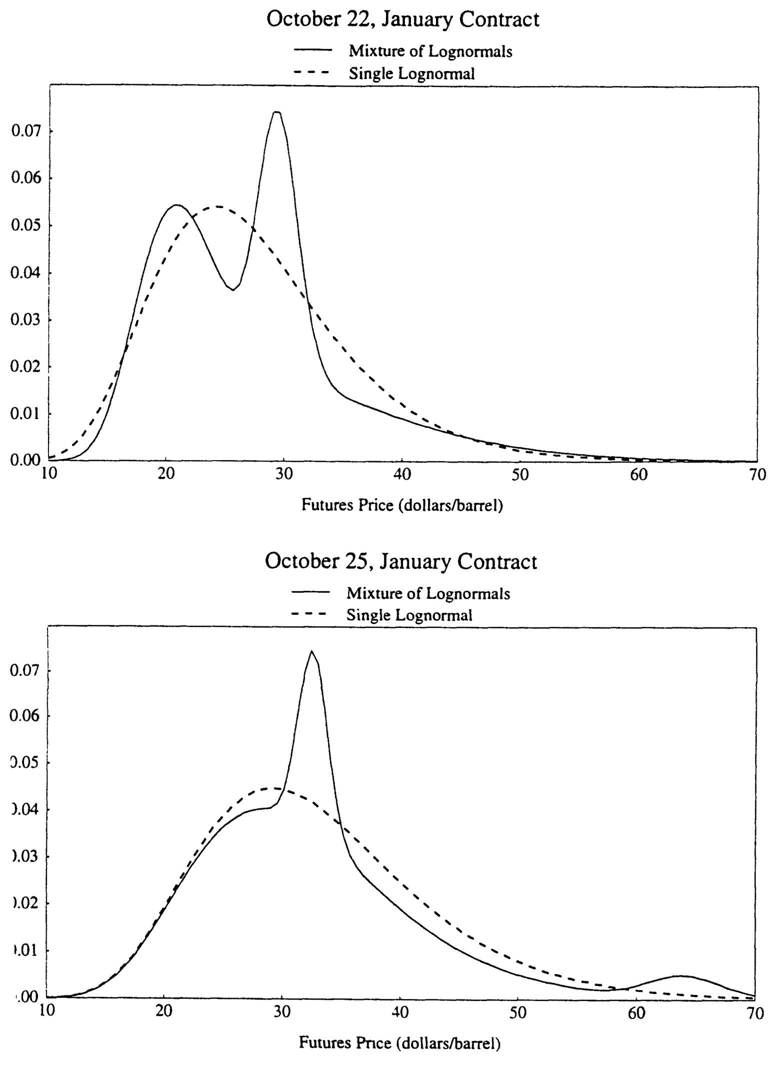
On Thursday, October 25, 1990, the London Financial Times reported that Iraqi forces had attached explosives to 300 of Kuwait's 1000 oil wells, quoting a

[20]When SLN is treated as Model 1 in equation (14), the derivatives of MLN were evaluated at $\sigma_1 = \sigma_2 = \sigma_3 = 0.2, w_1 = w_2 = 0.5, \mu_2 = \ln(f_t) - \sigma_1^2/2, \mu_1 = 0.95 \cdot \mu_2, \mu_3 = 1.05 \cdot \mu_2, \pi_1 = \pi_3 = 0.1$, and $\pi_2 = 0.8$. Given that the futures price for a given trading day was not used in the estimation of the parameters for either MLN or SLN, these derivatives, assuming SLN is correct, will not be correlated with the errors from SLN. When MLN is treated as Model 1 in equation (14), the derivatives of SLN were evaluated at the estimated values of μ and σ from SLN. This is a conservative approach that increases the likelihood of a significant F-statistic.

senior Kuwaiti engineer who had left Kuwait one week earlier. This revelation pushed oil prices up sharply, with the futures contract nearest to expiration (December) rising $3.17 per barrel. Figure 7 plots the PDFs from MLN and SLN for October 22 (top panel) and October 25 (bottom panel) using the January contract. On October 22, market expectations for futures prices were centered quite tightly around $24 per barrel. The news of the mining widened each model's distribution significantly, with MLN allowing for a sizeable probability mass between $60 and $70 per barrel.

FIGURE 7

Mining of Wells: Implicit Density Functions



The largest one-day change in oil prices in NYMEX history occurred on Thursday, January 17, 1991, when i) several governments announced a coordinated release of oil from their emergency inventories, and ii) it became clear that the coalition forces had total air supremacy. On that day, the futures price for the March contract fell \$9.66 while the futures price for the April contract fell \$7.82. The six panels of Figure 8 trace the evolution of expected PDFs on the days surrounding January 17. Prior to the first air strike (as can be seen in the first two panels), the market was still expecting a fairly significant chance of a major oil market disruption (perhaps Iraqi damage of Saudi Arabian oil facilities) that could push prices to the \$40–\$60 per barrel range. On January 17th, these PDFs tightened dramatically and, on ensuing days, the PDF generated from MLN moved closer and closer to that from SLN. By January 23, there was little difference between the two PDFs, as the market returned to almost a pre-crisis distribution (compare right-hand panels with the top panel of Figure 5).

V. Conclusion

This paper develops a method for using option prices to estimate the market's probability distribution for the underlying asset's price. The method is quite general, allowing the standard lognormal distribution to be replaced by any distribution from within a wide class. The particular assumption of a mixture of three lognormal distributions used here was driven by conditions in the oil market during the Persian Gulf crisis. As the focus is only on the asset's probability distribution, minimal structure is placed on the stochastic process governing movements in the asset price over time. This lack of structure is appealing since we generally have little a priori information about the stochastic process that market participants have assumed. Our methodology should be useful to researchers who wish to impose a minimum of structure and are i) examining other markets during unsettled times, or ii) investigating asset price distributions that are not adequately described by the lognormal distribution (e.g., leptokurtotic distributions). For example, in the foreign exchange market, our methodology can be used to shed light on the "Peso problem," or, in the interest rate market, it can be used to infer the market's assessment of possible changes in monetary policy.[21]

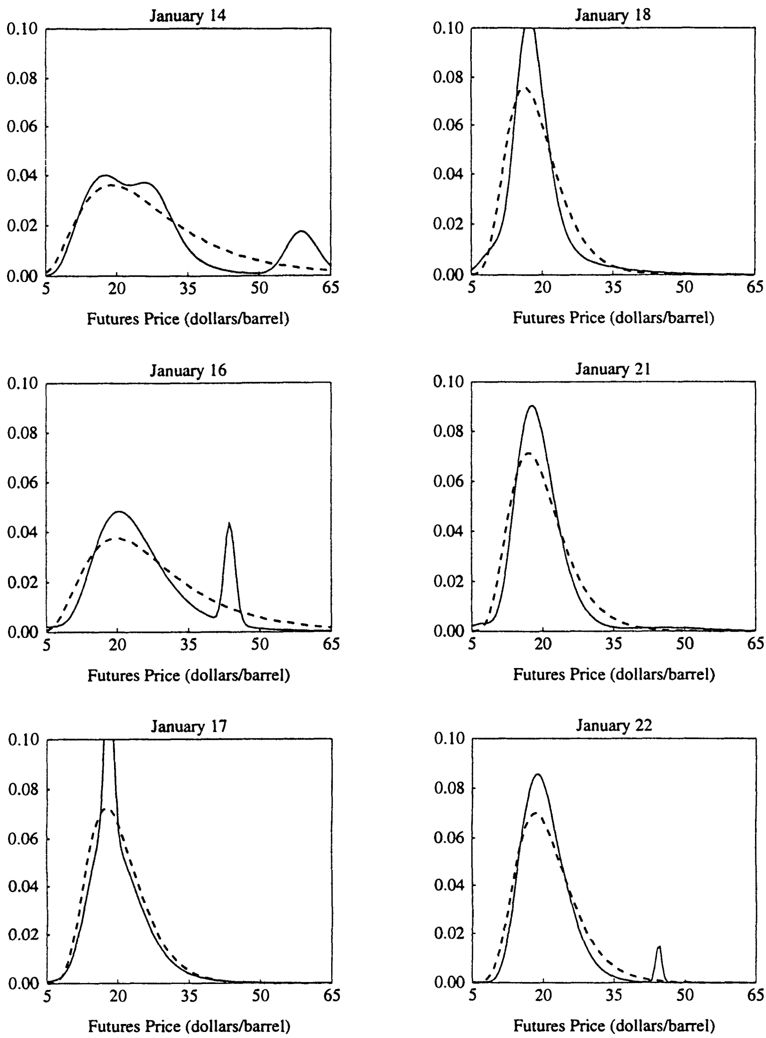
In the application to the oil market, we find that the options markets were consistent with the market commentary at the time, in that they reflected a significant probability of a major disruption in oil prices. We find that the estimated price of oil conditional on a major disruption was often in the \$50–\$60 per barrel range, which is also consistent with market commentary. We also find that the standard lognormal assumption did a poorer job of characterizing the data than did our method. In particular, we find that if policy makers or analysts had used the distribution from the lognormal model where our model was closer to the truth, they would have generally overestimated the market's assessment of the probability of a major disruption and underestimated the impact on prices of such a disruption.

[21]See Leahy and Thomas (1996) and McCauley and Melick (1996). Deutsche Bundesbank (1995) provides examples using related techniques.

FIGURE 8

Start of Air War: Implicit Density Functions

Mixture of Lognormals(———) Single Lognormal (— — —), April Contract



Finally, examination of particular days confirmed the large shift in market expectations that occurred when significant crisis-related news reached the oil market.

Appendix: Estimation Details

A. MLN Model

Equations (7) and (8) in the text give the pricing equations for puts and calls in terms of the bounds. Substituting for the bounds using equations (1)–(4), and noting that $E[\max[0, z]] = E[z|z \geq 0] \cdot \Pr[z \geq 0]$, yields pricing equations written in terms of the terminal distribution,

$$\text{(A-1)} \qquad \begin{aligned} C_t[X] &= \widehat{w}_{it} \cdot \left[\left(\widehat{E}_t\left[f_0|f_0 \geq X\right] - X\right) \cdot \widehat{\Pr}\left[f_0 \geq X\right]\right] + \left(1 - \widehat{w}_{it}\right) \\ &\quad \cdot \max\left[\widehat{E}_t\left[f_0\right] - X, e^{-rt} \cdot \left(\widehat{E}_t\left[f_0|f_0 \geq X\right] - X\right) \cdot \widehat{\Pr}\left[f_0 \geq X\right]\right] \\ &\quad + \widehat{\epsilon}_{ct}[X], \end{aligned}$$

$$\text{(A-2)} \qquad \begin{aligned} P_t[X] &= \widehat{w}_{it} \cdot \left[\left(X - \widehat{E}_t\left[f_0|f_0 \leq X\right]\right) \cdot \widehat{\Pr}\left[f_0 \leq X\right]\right] + \left(1 - \widehat{w}_{it}\right) \\ &\quad \cdot \max\left[X - \widehat{E}_t\left[f_0\right], e^{-rt} \cdot \left(X - \widehat{E}_t\left[f_0|f_0 \leq X\right]\right) \cdot \widehat{\Pr}\left[f_0 \leq X\right]\right] \\ &\quad + \widehat{\epsilon}_{pt}[X], \end{aligned}$$

$$\text{where} \qquad i = \left(\begin{array}{l} 1 \text{ if} \left(\begin{array}{l} \text{call and } X < \widehat{E}_t\left[f_0\right] \\ \text{put and } X > \widehat{E}_t\left[f_0\right] \end{array}\right) \\ 2 \text{ otherwise} \end{array}\right).$$

It is assumed that the distribution for futures prices (f_0) is a mixture of lognormals, i.e.,

$$\text{(A-3)} \qquad g\left[f_0\right] = \pi_1 g_1\left[f_0\right] + \pi_2 g_2\left[f_0\right] + \pi_3 g_3\left[f_0\right],$$

where

$$g_i\left[f_0\right] = \left(\frac{1}{\sqrt{2\pi}\sigma_i \cdot f_0}\right) \cdot \exp\left[\left(\frac{\ln(f_0) - \mu_i}{\sigma_i}\right)^2 \Big/ 2\right].$$

Using the properties of the lognormal distribution, the terms involving the expectations operator in equations (A-1) and (A-2) may be substituted for as follows,

$$\text{(A-4)} \qquad E\left[f_0\right] = \sum_{i=1}^{3} \pi_i \cdot \exp\left[\mu_i + \frac{\sigma_i^2}{2}\right],$$

$$\text{(A-5)} \qquad \Pr\left[f_0 \geq X\right] = \sum_{i=1}^{3} \pi_i \cdot \left(1 - \Phi\left[\frac{\ln[X] - \mu_i}{\sigma_i}\right]\right),$$

$$\text{(A-6)} \qquad \Pr\left[f_0 \leq X\right] = \sum_{i=1}^{3} \pi_i \cdot \Phi\left[\frac{\ln[X] - \mu_i}{\sigma_i}\right],$$

$$\text{(A-7)} \qquad E\left[f_0|f_0 \geq X\right] = \sum_{i=1}^{3} \pi_i \cdot \frac{\exp\left[\frac{\sigma_i^2 + 2\mu_i}{2}\right]\left(\Phi\left[\frac{\ln[X] - \mu_i - \sigma_i^2}{\sigma_i}\right] - \frac{1}{2}\right)}{\Pr\left[f_0 \geq X\right]},$$

$$\text{(A-8)} \qquad E\left[f_0 | f_0 \leq X\right] = E\left[f_0\right] - E\left[f_0 | f_0 \geq X\right],$$

where Φ represents the cumulative normal distribution function.[22] The pricing equations are now written in terms of the parameters of the model. A final difficulty is presented by the non-differentiability of the max operator in equations (A-1) and (A-2). Constrained, non-linear minimization is greatly enhanced when analytic derivatives are provided. Therefore, we replaced the max operator with a logistic approximation as follows,

$$\text{(A-9)} \qquad \text{logitmax}[x, y] \equiv \frac{1}{1 + \exp[-5 \cdot (x - y)]},$$

$$\text{(A-10)} \qquad \max[x, y] \approx \text{logitmax}[x, y] \cdot x + (1 - \text{logitmax}[x, y]) \cdot y.$$

A similar technique was used to handle the non-differentiability associated with jumping from one weight to the other when the estimated mean crosses through an option's strike price.

B. SLN Model

The SLN Model uses the approximation of Barone-Adesi and Whaley (1987) (hereafter, BAW) to express the option price in terms of the parameters of a single lognormal (u_b and σ_b). In our notation, the BAW approximation for the price of an American call option is written as

$$\text{(A-11)} \qquad C_t[X] = c\left[E_t\left[f_0\right], \sigma_b, X\right] + A_2\left[f^*, \sigma_b, X\right] \cdot \left(\frac{E_t\left[f_0\right]}{f^*}\right)^{q_2[\sigma_b]}$$

$$\text{when } E_t\left[f_0\right] < f^*,$$

$$\text{(A-12)} \qquad C_t[X] = E_t\left[f_0\right] - X, \quad \text{when } E_t\left[f_0\right] \geq f^*,$$

where

$$\begin{aligned}
c\left[E_t\left[f_0\right], \sigma_b, X\right] &\equiv E_t\left[f_0\right] \cdot e^{-rt} \cdot \Phi\left[d_1\left[E_t\left[f_0\right], \sigma_b, X\right]\right] \\
&\quad - X \cdot e^{-rt} \cdot \Phi\left[d_2\left[E_t\left[f_0\right], \sigma_b, X\right]\right], \\
A_2\left[f^*, \sigma_b, X\right] &\equiv \frac{\sigma_b}{q_2\left[\sigma_b\right]} \cdot \left(1 - e^{-rt} \cdot \Phi\left[d_1\left[f^*, \sigma_b, X\right]\right]\right), \\
q_2\left[\sigma_b\right] &\equiv \frac{1 + \sqrt{1 + \dfrac{8\rho_t}{\sigma_b^2 \cdot (1 - \exp[-rt])}}}{2}, \\
d_1\left[E_t\left[f_0\right], \sigma_b, X\right] &\equiv \frac{\ln\left[\dfrac{E_t\left[f_0\right]}{X}\right] + \dfrac{\sigma_b^2 \cdot t}{2}}{\sigma_b \cdot \sqrt{t}}, \\
d_2\left[E_t\left[f_0\right], \sigma_b, X\right] &= d_1\left[E_t\left[f_0\right], \sigma_b, X\right] - \sigma_b \cdot \sqrt{t}, \\
E_t\left[f_0\right] &= \exp\left[\mu_b + \frac{\sigma_b^2}{2}\right].
\end{aligned}$$

[22]The mean of the lognormal distribution is $\exp[u + \sigma^2/2]$. Calculation of (A-9) and (A-10) used integral 3.322 from Gradshteyn and Ryzhik (1980).

The term f^* is the "critical commodity price," and is solved for implicitly according to

$$\text{(A-13)} \quad f^* - X \quad = \quad c\left[f^*, \sigma_b, X\right] + \frac{\left(1 - e^{-rt} \cdot \Phi\left[d_1\left[f^*, \sigma_b, X\right]\right]\right) \cdot f^*}{q_2\left[\sigma_b\right]}.$$

The formula for the price of a put is similar to that for the call (see BAW for details). NAG algorithm C05AJF was used to solve for f^*. To improve the estimation of μ_b and σ_b, derivatives were provided to the NAG minimization algorithm E04UPF. Given that the option pricing formula (equations (A-11) and (A-12)) is not differentiable, the logitmax operator described above was used as follows:

$$\begin{aligned} \text{(A-14)} \quad C_t[X] \quad = \quad & \text{logitmax}\left[f^*, E_t\left[f_0\right]\right] \cdot (\text{RHS of (A-11)}) \\ & + \left(1 - \text{logitmax}\left[f^*, E_t\left[f_0\right]\right]\right) \cdot (\text{RHS of (A-12)}). \end{aligned}$$

References

Aït-Sahalia, Y., and A. W. Lo. "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices." Working Paper Series, No. 5351, National Bureau of Economic Research (Nov. 1995).

Barone-Adesi, G., and R. Whaley. "Efficient Analytic Approximation of American Option Values." *Journal of Finance*, 42 (June 1987), 301–320.

Bates, D. S. "The Crash of '87: Was It Expected? The Evidence from Options Markets." *Journal of Finance,* 46 (July 1991), 1009–1044.

Black, F. "The Pricing of Commodity Contracts." *Journal of Financial Economics*, 3 (Jan./March 1976), 167–179.

Black, F., and M. Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, 81 (May/June 1973), 637–654.

Breeden, D., and R. Litzenberger. "Prices of State-Contingent Claims Implicit in Options Prices." *Journal of Business*, 51 (Oct. 1978), 621–651.

Cakici, N.; S. Chatterjee; and A. Wolf. "Empirical Tests of Valuation Models for Options on T-Note and T-Bond Futures." *Journal of Futures Markets*, 13 (Feb. 1993), 1–13.

Chaudhury, M. M., and J. Wei. "Upper Bounds for American Futures Options: A Note." *Journal of Futures Markets*, 14 (Feb. 1994), 111–116.

Cox, J. C., and S. A. Ross. "The Valuation of Options for Alternative Stochastic Processes." *Journal of Financial Economics*, 3 (Jan./March 1976), 145–166.

Deaves, R., and I. Krinsky. "Risk Premiums and Efficiency in the Market for Crude Oil Futures." *Energy Journal,* 13 (1992), 93–118.

Deutsche Bundesbank. "The Information Content of Derivatives for Monetary Policy." *Monthly Report*, 47 (Nov. 1995), 17–32.

Dominguez, K. "The Volatility and Efficiency of Crude Oil Futures Contracts." In *Oil and Money: Coping with Price Risk through Financial Markets*, Harvard International Energy Studies, 5 (1989), 48–87.

Fackler, P. L., and R. P. King. "Calibration of Option-Based Probability Assessments in Agricultural Commodity Markets." *American Journal of Agricultural Economics*, 72 (Feb. 1990), 73–83.

Gradshteyn, I. S., and I. M. Ryzhik. *Tables of Integrals, Series, and Products*, New York, NY: Academic Press (1980).

Grundy, B. D. "Option Prices and the Underlying Asset's Return Distribution." *Journal of Finance*, 46 (July 1991), 1045–1069.

Jackwerth, J. C., and M. Rubinstein. "Recovering Probability Distributions from Option Prices." *Journal of Finance,* 51 (Dec. 1996), 1611–1631.

Jarrow, R., and A. Rudd. "Approximate Option Valuation for Arbitrary Stochastic Processes." *Journal of Financial Economics*, 10 (Nov. 1982), 347–369.

Kumar, M. "The Forecasting Accuracy of Crude Oil Futures Prices." *International Monetary Fund Staff Papers*, 39 (June 1992), 432–461.

Leahy, M. P., and C. P. Thomas. "The Sovereignty Option: The Quebec Referendum and Market Views on the Canadian Dollar." Board of Governors of the Federal Reserve System *International Finance Discussion Papers,* No. 555 (June 1996).

Levy, H. "Upper and Lower Bounds of Put and Call Option Value: Stochastic Dominance Approach." *Journal of Finance*, 40 (Sept. 1985), 1197–1217.

Lo, A. W. "Semi-Parametric Upper Bounds for Option Prices and Expected Payoffs." *Journal of Financial Economics*, 19 (Dec. 1987), 373–387.

MacKinnon, J. G. "Model Specification Tests and Artificial Regressions." *Journal of Economic Literature*, 30 (March 1992), 102–146.

Malz, A. M. "Option Prices and the Probability Distribution of Exchange Rates." In *Currency Options and Exchange Rate Economics,* Z. Chen, ed. River Edge, NJ: World Scientific (1997).

McCauley, R. N., and W. R. Melick. "Propensity and Density." *Risk,* 9 (Dec. 1996), 52–54.

Melick, W. R., and C. P. Thomas. "War and Peace: Recovering the Market's Probability Distribution of Crude Oil Futures Prices during the Gulf Crisis." *International Finance Discussion Papers*, No. 437, Board of Governors of the Federal Reserve System (Oct. 1992).

__________________. "Using Options Prices to Infer PDFs for Asset Prices: An Application to Oil Prices during the Gulf Crisis." *International Finance Discussion Papers*, No. 541, Board of Governors of the Federal Reserve System (Feb. 1996).

Neuhaus, H. "The Information Content of Derivatives for Monetary Policy: Implied Volatilities and Probabilities." *Economic Research Group Discussion Paper,* No. 3/95, Deutsche Bundesbank (July 1995).

Overdahl, J. A., and H. L. Matthews. "The Use of NYMEX Options to Forecast Crude Oil Prices." *Energy Journal*, 9 (Oct. 1988), 135–147.

Perrakis, S., and P. J. Ryan. "Option Pricing Bounds in Discrete Time." *Journal of Finance*, 39 (June 1984), 519–525.

Ritchey, R. J. "Call Option Valuation for Discrete Normal Mixtures." *Journal of Financial Research*, 13 (Winter 1990), 285–296.

Rubinstein, M. "Implied Binomial Trees." *Journal of Finance*, 49 (July 1994), 771–818.

Shimko, D. C. "Beyond Implied Volatility: Probability Distributions and Hedge Ratios Implied by Option Prices." Mimeo, Dept. of Finance and Business Economics, Univ. of Southern California (Nov. 1991).

Silva, E. M. S., and K. H. Kahl. "Reliability of Soybean and Corn Option-Based Probability Assessments." *Journal of Futures Markets*, 13 (Oct. 1993), 765–799.

Whaley, R. "Valuation of American Futures Options: Theory and Empirical Tests." *Journal of Finance*, 41 (March 1986), 127–150.