

Mizrach, Bruce

Working Paper

Recovering probabilistic information from options prices and the underlying

Working Paper, No. 2007-02

Provided in Cooperation with:

Department of Economics, Rutgers University

Suggested Citation: Mizrach, Bruce (2008) : Recovering probabilistic information from options prices and the underlying, Working Paper, No. 2007-02, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:

<https://hdl.handle.net/10419/31297>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Recovering Probabilistic Information From Options Prices and the Underlying

Bruce Mizrach*
Rutgers University

First Draft: January 16, 2007

This Draft: February 10, 2008

Abstract

This paper examines a variety of methods for extracting implied probability distributions from option prices and the underlying. The paper first explores non-parametric procedures for reconstructing densities directly from options market data. I then consider local volatility functions, both through implied volatility trees and volatility interpolation. I then turn to alternative specifications of the stochastic process for the underlying. I estimate a mixture of log normals model, apply it to exchange rate data, and illustrate how to conduct forecast comparisons. I finally turn to the estimation of jump risk by extracting bipower variation.

Keywords: options; implied probability densities; volatility smile; jump risk; bipower variation.

JEL Classification: G12, G14, F31.

* Mizrach is with Department of Economics, Rutgers University, e-mail: mizrach@econ.rutgers.edu, (732) 932-8261 (voice) and (732) 932-7416 (fax). <http://snde.rutgers.edu>. I would like to thank seminar participants at the OCC, Federal Reserve Bank of Philadelphia, and the FDIC Derivative Securities and Risk Management Conference for helpful comments.

1. Introduction

Traders, economists and central bankers all have interest in assessing market beliefs. The problem is that markets speak to us in the language of prices, and extracting their views can often be a difficult inference problem.

This paper surveys and occasionally extends a variety of methods for obtaining not just market expectations, but the full range of predictive densities. I examine and implement procedures for both options and the underlying. A mastery of these approaches can help banks, trading firms and their regulators avoid excessive risks.

The paper begins by explaining the theoretical connection between option prices and probability measures. The relationship linking the option price and the risk neutral density is established, under very general conditions, using the Feynman-Kac lemma.

I then establish empirically that the log normality restrictions implied by Black-Scholes are too severe. Our evidence, drawn from an exchange rate example, illustrates the volatility smile and skew.

The paper then moves beyond the log-normal and extracts densities using histogram estimators. These approaches perform fairly poorly, but I introduce a well-behaved method that can be solved in a spreadsheet.

A closely related approach using Rubinstein's (1994) implied binomial trees follows. We use the modifications of Barle and Cakici (1998) and Cakici and Foster (2002) in a numerical example.

The most popular approach in the practitioner literature, local volatility modeling, is discussed next. We follow Shimko (1993) and Dumas, Fleming and Whaley (1998) and interpolate the volatility surface. These techniques are computationally straightforward, but are limited by potential arbitrage violations in the tails, precisely the region of greatest interest to policy authorities.

I next turn to alternative stochastic processes for the underlying. I parameterize the spot price process as a mixture of log normals, as in Ritchey (1990) and Melick and Thomas (1997), and fit the model to options prices. This model has proven useful in analyzing exchange rate crises in Haas, Mittnik and Mizrach (2006) and, in Mizrach (2006), predicting Enron's bankruptcy risk prior to its collapse.

Econometric issues are considered in the stochastic process section. These include the choice of metric. The paper endorses the view that matching option volatility is the preferred distance measure. Hypothesis testing against simpler alternatives is not straightforward due to nuisance

parameters, and a bootstrap approach is discussed. Comparing value-at-risk over a specific interval is done using the approach of Christoffersen (1998), and comparison of the entire predictive density is also done using Berkowitz (2001).

I then examine jump processes and introduce the crucial recent results of Barndorff-Neilsen and Shephard (2006). These enable a straightforward extraction of the jump risk component in the underlying. I find that, in exchange rates, jump risk does not move in lockstep with stochastic volatility.

I now begin in the obvious place, with the Black-Scholes model as a baseline.

2. Black Scholes Baseline

I specify the price process and the hedge portfolio in the first part. The link between the risk neutral measure and the solution to the portfolio differential equation is in the second part.

2.1 The B-S differential equation

We assume that the spot price S_t follows a geometric Brownian motion,

$$S_t = S_0 \exp(\alpha t + \sigma W_t) \quad (1)$$

where S_0 is an initial condition, α and σ are constants and W_t is a standard Brownian motion. This process is log normal because $\log(S_t) = \log(S_0) + \alpha t + \sigma W_t$ is normally distributed. Ito's lemma implies that

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (2)$$

with drift $\mu = (\alpha + \sigma^2/2)$ and deterministic volatility σ .

Suppose that Y_t is the price of a derivative security contingent on S and t . From Ito's lemma, we obtain,

$$dY_t = \left(\frac{\partial Y}{\partial S} \mu S_t + \frac{\partial Y}{\partial t} + \frac{1}{2} \frac{\partial^2 Y}{\partial S^2} \sigma^2 S_t^2 \right) dt + \frac{\partial Y}{\partial S} \sigma S_t dW_t. \quad (3)$$

Consider next a portfolio, Π , short one derivative security and long an amount $\partial Y/\partial S$ of the underlying,

$$\Pi_t = -Y_t + \frac{\partial Y}{\partial S} S_t. \quad (4)$$

The change in value of this portfolio in a short interval is

$$d\Pi_t = -dY_t + \frac{\partial Y}{\partial S} dS_t. \quad (5)$$

Substituting (2) and (3) into (5) yields,

$$d\Pi_t = - \left(\frac{\partial Y}{\partial t} + \frac{1}{2} \frac{\partial^2 Y}{\partial S^2} \sigma^2 S_t^2 \right) dt. \quad (6)$$

Since this equation does not involve dW_t , the portfolio is riskless during the interval dt . Under a no arbitrage assumption, this portfolio can only earn the riskless rate of interest r ,

$$d\Pi_t = r\Pi_t dt \quad (7)$$

Substituting from (4) and (6), this becomes

$$\left(\frac{\partial Y}{\partial t} + \frac{1}{2} \frac{\partial^2 Y}{\partial S^2} \sigma^2 S_t^2 \right) dt = r \left(Y_t - \frac{\partial Y}{\partial S} S_t \right) dt. \quad (8)$$

This can be expressed as a partial differential equation (PDE),

$$\frac{\partial Y}{\partial t} + \frac{\partial Y}{\partial S} r S_t + \frac{1}{2} \frac{\partial^2 Y}{\partial S^2} \sigma^2 S_t^2 = r Y_t. \quad (9)$$

This equation was initially proposed and solved by Black and Scholes (1973) in their seminal paper. A number of derivative securities can be priced using the solution to (9) along with the appropriate boundary conditions,

$$Y_T = \exp(-r\tau)\psi(S_T) \quad (10)$$

where T is the date of expiry, and $\tau = T - t$. For a European call with strike price K , the boundary condition would be

$$\psi(S_t) = \exp(-r\tau) \max[S_T - K, 0]. \quad (11)$$

We then obtain the well-known Black-Scholes formula,

$$BS^C(S_t, K, T, r, \sigma) = S_t \Phi(d_1) - K \exp(-r\tau) \Phi(d_2), \quad (12)$$

with

$$\begin{aligned} d_1 &= \frac{\ln(S_t/K) + (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}, \\ d_2 &= d_1 - \sigma\sqrt{\tau}, \end{aligned}$$

and $\Phi(\cdot)$ denotes the cumulative standard normal distribution.

2.2 A probability density approach

The link between the typically numerical PDE approach and stochastic processes is provided by the Feynman-Kac analysis.

Consider again a spot price process that solves (1), and a derivative security that follows the PDE,

$$\frac{\partial Y}{\partial S} \mu S_t + \frac{\partial Y}{\partial t} + \frac{1}{2} \frac{\partial^2 Y}{\partial S^2} \sigma^2 S_t^2 = 0, \quad (13)$$

and satisfies the boundary condition (11). Applying Ito's lemma to (13), we again obtain (3).

The first term in (3) drops out from (13), and then integrating on both sides,

$$\int_t^T dY_t = Y_T - Y_t = \int_t^T \frac{\partial Y}{\partial S} \sigma S_t dW_t. \quad (14)$$

Taking expectations, the right hand side is zero, leaving

$$Y_t = E[Y_T] = \exp(-r\tau) E_Q[\psi(S_T) | S_t = s] \quad (15)$$

Note that we have replaced the solution of the PDE with a conditional expectation taken under the risk neutral measure Q . Regularity conditions are established in Karatzas and Shreve (1991).

Hambly (2006) shows how we can use the Feynman-Kac to construct probability densities for the spot price. Set $\psi(s) = I_\theta(s)$, the indicator function of a set θ . The solution to the PDE (13) with this boundary condition is

$$E[I_\theta(S_T) | S_t = s] = \Pr[S_T \in \theta | S_t = s] \quad (16)$$

The probability density f obeys the Kolmogorov (or Fokker-Planck) forward equation,

$$\frac{\partial f}{\partial S} \mu S_t + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S_t^2 = 0 \quad (17)$$

where we impose a terminal delta function. This equation describes the probability density of being at S_T given that we are at S_t .

In the case where the spot price process follows a Brownian process with drift r ,

$$dS_t = rS_t dt + \sigma S_t dW_t, \quad (18)$$

we calculate that the transition density is the log-normal,

$$f(S_T) = \frac{1}{S_T \sqrt{2\pi\sigma^2\tau}} \exp \left[-\frac{[\ln(\frac{S_T}{S_t}) - (r - \sigma^2/2)\tau]^2}{2\sigma^2\tau} \right].$$

3. Empirical Departures from Black Scholes

Data on option prices appear to be inconsistent with the Black-Scholes assumptions. In particular, volatility seems to vary across strike prices – often with a parabolic shape called the volatility “smile.” The smile is often present on only one part of the distribution giving rise to a “smirk.” Tompkins (2001) shows that these patterns are quite typical in a wide range of options markets, including stock indices, Treasury bonds and exchanges rates for Germany, the U.K. and the U.S.

The motivating examples in this section are drawn from the foreign exchange market and draw upon the work of Haas, Mittnik and Mizrach (2006). I analyze the US Dollar/British Pound

(US\$/BP) option which trades on the Philadelphia exchange. Both American and European options are traded. The BP options are for 31,250 Pounds. I use daily closing option prices that are quoted in cents. Spot exchange rates are expressed as US\$ per unit of foreign currency and are recorded contemporaneously with the closing trade. Foreign currency appreciation (depreciation) will increase the moneyness of a call (put) option. Interest rates are the Eurodeposit rates closest in maturity to the term of the option.

To obtain a rough idea about the implied volatility pattern in the currency options, I look at sample averages. I sort the data into bins based on the strike/spot ratio, S/K , and compute implied volatilities using the Black-Scholes formula. In Figure 1, I plot the data for all of 1992 and 1993, for the British Pound. It displays the characteristic pattern, with the minima of the implied volatility at the money, and with higher implied volatilities in the two tails.

[Insert Figure 1 Here]

Black-Scholes cannot account for this pattern in implied volatility. The next three sections discuss how to move beyond the restrictive Black-Scholes assumptions.

4. Beyond Black Scholes

I extend the Black-Scholes model by allowing volatility to vary with the strike price. The formal link is established using the forward Kolmogorov equation.

4.0.1 How Volatility Varies with the Strike

The Feynman-Kac analysis enables us to define a risk neutral probability in which we can price options. Let $f(S_T)$ denote the terminal risk neutral (Q -measure) probability density at time T , and let $F(S_T)$ denote the cumulative probability. A European call option at time t , expiring at T , with strike price K , is priced

$$C(K, \tau) = \exp(-r\tau) \int_K^\infty (S_T - K) f(S_T) dS_T, \quad (19)$$

In the case where $f(\cdot)$ is the log-normal density and volatility σ is constant with respect to K , this yields the Black-Scholes formula. In this benchmark case, implied volatility is a sufficient statistic for the entire implied probability density which is centered at the risk-free rate.

In the baseline case, practitioners often try to invert (2) to estimate “implied volatility”. The

academic literature has focused on whether volatility in the physical measure is better predicted by implied volatility, GARCH or stochastic volatility.

Under basic no-arbitrage restrictions, we can consider more general densities than the log-normal for the underlying. Breeden and Litzenberger (1978) show that the first derivative is a function of the cumulative distribution,

$$\partial C / \partial K |_{K=S_T} = -e^{-r\tau} (1 - F(S_T)). \quad (20)$$

The second derivative then extracts the density,

$$\partial^2 C / \partial K^2 |_{K=S_T} = e^{-r\tau} f(S_T). \quad (21)$$

4.0.2 The Kolmogorov equation

Dupire (1994) clarifies the isomorphism between the approaches that specify the density and those that specify price process. He shows that for driftless diffusions, there is a unique stochastic process corresponding to a given implied probability density.

From the Kolmogorov equation (17) for this density, I rearrange

$$\frac{\partial C}{\partial T} = \frac{1}{2} \sigma^2 K^2 \frac{\partial^2 C}{\partial K^2} - rK \frac{\partial C}{\partial K} \quad (22)$$

and obtain the local volatility formula,

$$\sigma^2(K, t) = \frac{\frac{\partial C}{\partial T} + rK \frac{\partial C}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}}. \quad (23)$$

The principal problem in estimating f is that one does not observe a continuous function of option prices and strikes. My first attempt to estimate f utilizes a histogram method to try and trace out the density.

5. Histogram Estimators

This section develops two nonparametric estimators of the risk neutral density.

5.1 A crude histogram estimator

This section describes a method first proposed by Longstaff (1990) and then modified by Rubinstein (1994). Let S be the current price of the underlying asset. Let C_1 to C_n be the prices of n associated call options with striking prices $K_1 < K_2 < \dots < K_n$ all with the same time to expiration. Denote by S_T the price of the asset at expiration, and let r be the riskless rate of interest, and δ be the

payout rate on the underlying asset (e.g. a dividend for a stock or in the case of a foreign exchange option, the interest differential).

As S_T goes from K_n to K_{n-1} the payoff goes from 0 to $K_n - K_{n-1}$ with an average payoff of $0.5(K_n - K_{n-1})$. It follows that $C_{n-1} = 0.5p_n(K_n - K_{n-1})/(1+r)$ with p_n the risk neutral probability associated with the histogram bin n .

Consider now the price of the call with the next lowest striking price. There are three outcomes one must consider. If $S_T < K_{n-2}$ then the payoff is zero. In the interval between K_{n-1} and K_{n-2} the average payoff is $0.5(K_{n-1} - K_{n-2})$. Above K_{n-1} the payoff is $p_n[(K_{n-1} - K_{n-2}) + 1/2(K_n - K_{n-1})]$. The latter part is easily seen as just $(1+r)C_{n-1}$. Adding up these components, gives us the recursion,

$$C_{n-j} = [0.5p_{n-j+1} + \sum_{k=2}^j p_{n-j+k}](K_{n-j+1} - K_{n-j})/(1+r) + C_{n-j}. \quad (24)$$

I can view the underlying asset as a call with strike $K_0 = 0$. As before, I have three outcomes to consider. If $S_T < 0$ then the payoff is zero. For $0 \leq S_T < K_1$, then the payoff is $1/2(K_1 - 0)p_1$, and for $S_T > K_1$,

$$(K_1 - 0)[p_2 + \dots + p_n]/(1+r) + C_1. \quad (25)$$

Adding up these payoff components gives us,

$$S/(1+\delta) = K_1[0.5p_1 + \sum_{k=2}^n p_k]/(1+r) + C_1. \quad (26)$$

Exploiting the fact that the probabilities sum to one,

$$S/(1+\delta) = K_1[0.5p_1 + (1-p)_1]/(1+r) + C_1, \quad (27)$$

I can rearrange (27) and solve for P_1 ,

$$p_1 = 2[1 - (1+r)(S/(1+\delta) - C_1)K_1^{-1}]. \quad (28)$$

More generally, I can write $p_j = 2[1 - \sum_{k=1}^{j-1} p_k - (1+r)(C_{j-1} - C_j)(K_j - K_{j-1})^{-1}]$.

I implement this histogram method in Figure 2. The baseline is a normal distribution for which I have priced 11 Black-Scholes calls for strikes from 75 to 125, with a spot price of 100. For this example, the Longstaff-Rubinstein estimator does not provide even a good rough approximation. The probabilities whipsaw back and forth; every other bin turns negative. After some trial and error, it was apparent this property was not uncommon. I turn next to an improved estimator.

[Insert Figure 2 Here]

5.2 An improved histogram estimator

Let $b_i > 0$, $i = 0, \dots, n+1$ be the cardinality of a set uniformly distributed with mean $K_{i+1} - K_i$ for $i = 1$ to $n - 1$, and \bar{K}_i at the end points. To transform b_i into a probability measure, define

$$p_j = b_j / \sum_{j=1}^{n+1} b_j \quad (29)$$

The improved estimator minimizes the percentage deviation from the simulated and observed call prices,

$$\min_{b_0, \dots, b_{n+1}} \sum_{j=1}^n (\hat{C}_j - C_j) / C_j, \quad (30)$$

where the \hat{C}'_j s can be determined using the recursion (24).

I used this optimized histogram estimator to fit the same 11 Black-Scholes calls from the prior example. For a small number of calls, this algorithm can be solved in a spreadsheet. I plot the risk neutral probabilities in Figure 2.

The optimized histogram is a substantial improvement over the crude estimator. There are no negative probabilities (by construction) and with the exception of the long bin from 0 to 75, it provides a good approximation to the underlying normal. This method is quite reliable for back of the envelope calculations, but for more complicated processes, one needs more powerful techniques.

6. Tree Methods

6.1 A standard tree

I begin the discussion by defining a notation for a standard binomial tree. Let $S_{0,0}$ denote the current spot price. Any element of the spot price tree is a sequence of recombining up and down moves u, d :

$$S_{i,j} = S_{0,0} \times u^j d^{i-j}, \quad i = 0, \dots, N, \quad j = 0, \dots, i, \quad (31)$$

where

$$u = \exp^{\sigma \sqrt{T/N}}, \quad d = 1/u,$$

N is the number of steps in the tree, T is the time to expiration, and σ is the annualized volatility. The risk neutral probability of an up move is

$$p = \frac{\exp^{rT/N} - d}{u - d} \quad (32)$$

where r is the risk free rate of return.

The main advantage of the standard approach is simplicity. One of the important limitations

for our purposes is the assumption that the probabilities are constant throughout the tree. This implies that the *nodal probability* of reaching any particular terminal point $S_{N,j}$ is just the binomial probability

$$\begin{aligned} p_{N,j} &= \Pr[S_T = S_{N,j}] = \Phi(j, N, p), \\ &= \frac{j!(N-j)!}{N!} \times p^j(1-p)^{N-j}. \end{aligned} \quad (33)$$

6.2 Implied binomial trees

The implied binomial tree approach of Rubinstein (1994) and Jackwerth and Rubinstein (1996) relaxes the assumption of constant probabilities as in (32). Our discussion here follows the developments in Barle and Cakici (1998) and Cakici and Foster (2002). The first key modification is to center the tree at the forward price $F_{i,j}$,

$$F_{i,j} = p_{i,j}S_{i,j+1} + (1-p_{i,j})S_{i,j}. \quad (34)$$

The other change is to value options using Black-Scholes with interpolated volatilities rather than the Arrow-Debreu prices $\lambda_{i,j}$,

$$BS_{i,j}^C(S_{0,0}, K, T, \sigma) = \sum_{j=1}^n \lambda_{i,j} \exp(-r\Delta t) \max(S_{i,j} - K, 0), \quad (35)$$

which is a call with strike price K , expiration Δt , and volatility σ .

To grow the tree from the initial spot price (or any odd number step in the tree), we have two even nodes. The upper central node is

$$S_{i+1,j+1} = F_{i,j} \frac{\lambda_{i,j}F_{i,j} + \Delta_{i,j}^C}{\lambda_{i,j}F_{i,j} - \Delta_{i,j}^C} \quad (36)$$

where

$$\Delta_{i,j}^C = \exp(r\Delta t)C_{i,j}(F_{i,j}, \Delta t) - \sum_{k=j+1}^i \lambda_{i,k}(F_{i,k} - F_{i,j}), \quad (37)$$

The lower central node must then satisfy $S_{i+1,j}S_{i+1,j+1} = F_{i,j}^2$.

We then move up from the central node using the recursions,

$$S_{i,j+1} = \frac{\Delta_{i-1,j}^C S_{i,j} - \lambda_{i,j} F_{i-1,j} (F_{i-1,j} - S_{i,j})}{\Delta_{i-1,j}^C - \lambda_{i,j} (F_{i-1,j} - S_{i,j})}. \quad (38)$$

Moves down from the central node are given by

$$S_{i,j} = \frac{\lambda_{i-1,j} F_{i-1,j} (S_{i,j+1} - F_{i-1,j}) - \Delta_{i-1,j}^P S_{i,j+1}}{\lambda_{i-1,j} (S_{i,j+1} - F_{i-1,j}) - \Delta_{i-1,j}^P} \quad (39)$$

where

$$\Delta_{i,j}^P = \exp(r\Delta t)P_{i,j}(F_{i,j}, \Delta t) - \sum_{k=1}^{j-1} \lambda_{i,k}(F_{i,k} - F_{i,j}) \quad (40)$$

The Arrow-Debreu probabilities are then updated,

$$\lambda_{i+1,j} \exp(r\Delta t) = \left\{ \begin{array}{ll} p_{i,j} \lambda_{i,j} & \text{for } j = n + 1 \\ p_{i,j-1} \lambda_{i,j-1} + (1 - p_{i,j}) \lambda_{i,j} & \text{for } 2 \leq j \leq n \\ (1 - p_{i,1}) \lambda_{i,1} & \text{for } j = 1 \end{array} \right\}. \quad (41)$$

6.3 Numerical example

We illustrate the implied binomial trees using a numerical example. Consider an 2-step tree, six months from expiration, which implies $\Delta t = 0.25$. The initial price $S_{0,0} = 100$. Volatility is a smirk, rising when the strike price is above 100, $K = \max(20, 20 \times F_{i,j}/S_{0,0})$. The continuously compounded interest rate is $r = \ln(4.8790) = 5\%$.

Starting from $S_{0,0}$, we compute the forward price, $F_{0,0} = 100 \exp(0.04879 \times 0.25) = 101.2272$. The Arrow-Debreu price at the origin is $\lambda_{0,0} = 1.0$. The Black-Scholes call struck at the futures price, expiring next period and volatility $\sigma = 20 \times 101.2272/100 = 20.2454$ is

$$BS_{i,j}^C(S = S_{0,0}, K = F_{0,0}, T = 0.25, \sigma = 20.2454) = 4.0367,$$

which implies

$$\Delta_{i,j}^C = \exp(0.04879 \times 0.25) 4.0367 = 4.0862.$$

We then use (36) to find

$$S_{1,1} = 101.2272 \frac{(1.0)(101.2272) + 4.0862}{(1.0)(101.2272) - 4.0862} = 109.7434.$$

We find the lower central node by equating spot and futures prices,

$$S_{1,0} = F_{0,0}^2 / S_{1,1} = 93.3719.$$

Having found the new spot prices, we can update the transition probabilities,

$$p_{0,0} = \frac{(F_{0,0} - S_{1,0})}{(S_{1,1} - S_{1,0})} = 0.4798,$$

and the Arrow-Debreu prices,

$$\lambda_{1,1} = \lambda_{0,0} p_{0,0} \exp(-r\Delta t) = 0.4740,$$

$$\lambda_{1,0} = \lambda_{0,0} (1 - p_{0,0}) \exp(-r\Delta t) = 0.5139.$$

The next central node is at the 2-step ahead forward price,

$$S_{2,1} = 100 \exp(0.04879 \times 0.25 \times 2) = 102.4695.$$

We work up from there using (38). This requires the forward prices at the prior node,

$$F_{1,1} = \exp(r\Delta t) S_{1,1} = 111.0902,$$

and the call prices. We compute the new local volatility $\sigma = 22.2180$, and price the call

$$BS_{i,j}^C(S_{0,0}, K = F_{1,1}, T = 0.5, \sigma = 22.2180) = 3.1597.$$

Since this call is at expiry, we don't need to discount, $\Delta_{i,j}^C = BS_{i,j}^C$. Substituting into (38),

$$\begin{aligned} S_{2,2} &= \frac{(3.1597)(102.4695) - (0.4740)(111.0902)(111.0902 - 102.4695)}{3.1597 - (0.4740)(111.0902 - 102.4695)} \\ &= 140.4882. \end{aligned}$$

Moving down from the central node requires (39). We use the futures price

$$F_{1,0} = \exp(r\Delta t)S_{1,0} = 94.5178,$$

and the volatility $\sigma = 20$, to price the put,

$$BS_{i,j}^P(S_{0,0}, K = F_{1,0}, T = 0.5, \sigma = 20) = 2.3970.$$

Substituting into (39),

$$\begin{aligned} S_{2,0} &= \frac{(0.5139)(94.5178)(102.4695 - 94.5178) - (2.3970)(102.4695)}{(0.5139)(102.4695 - 94.5178) - 2.3970} \\ &= 83.2339. \end{aligned}$$

This completes this simple example.

A simpler approach which interpolates the implied volatility surface is my next topic.

7. Local Volatility Functions

The most popular practitioners method is the use of *local volatility functions*. The seminal references are Shimko (1993) and Dumas, Fleming and Whaley (1998).

Shimko (1993) proposed to fit a polynomial to the smile so as to infer how volatility varied with the striking price. He chose to use a quadratic,

$$\hat{\sigma}(K, t) = a_0 + a_1K + a_2K^2. \quad (42)$$

In the case where volatility varies in this fashion, the density $f(S_T)$ can be written,

$$N'(d_2) [d_{2K} - (a_1 + 2a_2K)(1 - d_2d_{2K}) - 2a_2K]. \quad (43)$$

Shimko's method has several advantages. It can be used to generate call prices for a continuum of strikes. In this way, I can fill in the histogram proposed by Longstaff (1990) and Rubinstein (1994). The downside is that it can easily generate either negative probabilities or pricing inconsistencies. I can illustrate this in an empirical example now.

The difficulties with Shimko's method arise when $\partial\sigma/\partial K$ is too steep. At some point, the

volatility may become so high that it implies a more deeply out of the money option has a higher price than those close to being in the money.

I utilize data from the Philadelphia Exchange for June 25, 1992 DM/\$ puts and calls. I fit a quadratic to the smile, $0.4035 \times (\%M)^2 + 0.288 \times (\%M) + 10.362$ in Figure 3(a).

[Insert Figure 3(a) and (b) Here]

At -4% out of the money, the implied volatility rises to 15.67 and at -6% , 23.16. Once I translate these volatilities into prices in Figure 3(b), one can see that an arbitrage exists. The -6% call is more valuable than the -4% , 0.16 versus 0.10.

A further difficulty is that even when this arbitrage is not a problem, there is no guarantee that the bracketed term in (58) will remain positive for all choices of a . I also conjecture that one set of restrictions does not necessarily imply the other, making this a fairly complicated constrained optimization. Brunner and Hafner (2003) is a promising step in this direction.

Bliss and Panigirtzoglou (2002) endorse the use of local volatility functions¹ over the mixture-of-normals approach in the next section, but this conclusion does not seem to hold when we look just at the tails of the distribution.

8. PDF Approaches

8.1 Mixture-of-Log-Normals Specification

I assume that the stock price process is a draw from a mixture of three (non-standard) normal distributions, $\Phi(\mu_j, \sigma_j)$, $j = 1, 2, 3$, with $\mu_3 \geq \mu_2 \geq \mu_1$. Three additional parameters λ_1 , λ_2 and λ_3 define the probabilities of drawing from each normal. To nest the Black-Scholes, we restrict the mean to equal the interest differential, $\mu_2 = i_d - i_f$. Risk neutral pricing then implies restrictions on either the other means or the probabilities. I chose to let μ_1 , λ_1 and λ_3 vary, which implies

$$\mu_3 = \frac{\mu_1 \lambda_1}{\lambda_3}, \tag{44}$$

and

$$\lambda_2 = 1 - \lambda_1 - \lambda_3. \tag{45}$$

¹ They use natural splines in the implied volatility-delta space. When fitting the mixture of normals, they use the call price-strike metric.

For estimation purposes, this leaves six free parameters $\theta = (\theta_1, \theta_2, \dots, \theta_6)$. To keep the estimates positive, I embed the parameters into the exponential function. The left-hand tail distribution is given by

$$\Phi(\mu_1, \sigma_1) = \Phi(i_d - i_f - e^{\theta_1}, 100 \times e^{\theta_2}). \quad (46)$$

The only free parameter of the middle normal density is the standard deviation,

$$\Phi(\mu_2, \sigma_2) = \Phi(i_d - i_f, 100 \times e^{\theta_3}). \quad (47)$$

I use the logistic function for the probabilities to bound them on $[0, 1]$,

$$\lambda_1 = \frac{e^{\theta_4}}{1 + e^{\theta_4}}, \quad (48)$$

$$\lambda_3 = \frac{e^{\theta_5}}{1 + e^{\theta_5}}. \quad (49)$$

The probability specification implies the following mean restrictions on the third normal,

$$\Phi(\mu_3, \sigma_3) = \Phi\left((i_d - i_f + e^{\theta_1}) \times \frac{e^{\theta_4}/(1 + e^{\theta_4})}{e^{\theta_5}/(1 + e^{\theta_5})}, 100 \times e^{\theta_6}\right). \quad (50)$$

Mizrach (2006) shows that this data generating mechanism can match a wide range of shapes for the volatility smile.

8.2 Data and Estimation Results

8.2.1 Data

I focus on a subset of the results from Haas, Mittnik and Mizrach (2006). I utilize the British Pound/US\$ option data described in Section 3. For estimation purposes, I excluded options that were less than 5 or more than 75 days to maturity, more than 10% in or out of the money, and with volumes less than 5 contracts. This seemed to eliminate most data points with unreasonably high implied volatilities.

8.2.2 Implied Density Estimation

There are two key issues in fitting a stochastic process to the options data. The first is to extend the analysis to American options which can be exercised before expiration. The second is choosing the loss function for estimation.

I approximate American puts and calls using the Bjerksund and Stensland (1993) approach. Hoffman (2000) shows that the Bjerksund-Stensland algorithm compares favorably in accuracy and

computational efficiency to the Barone-Adesi and Whaley (1987) quadratic approximation. Our estimates were also quite similar using implied binomial trees.²

Let

$$\{d_{j,t}\}_{j=1}^n = [c(\tau_1, K_1), \dots, c(\tau_m, K_m), p(\tau_{m+1}, K_{m+1}), \dots, p(\tau_n, K_n)]$$

denote a sample of size n of the calls c and puts p traded at time t , with strike price K_j and expiring in τ_j years, and denote the pricing estimates from the model by $\{d_{j,t}(\theta)\}$.

In estimation, Christoffersen and Jacobs (2004) emphasize that the choice of loss function is important. Bakshi, Cao and Chen (1997), for example, match the model to data using option prices. This can lead to substantial errors among the low priced options though. Since these options are associated with tail probability events, this is not the best metric for our exercise. We obtained the best fit overall using the implied Bjerksund-Stensland implied volatility,

$$\sigma_{j,t} = BJST^{-1}(d_{j,t}, S_t, i_t). \quad (51)$$

Let the estimated volatility be denoted by

$$\sigma_{j,t}(\theta) = BJST^{-1}(d_{j,t}(\theta), S_t, i_t). \quad (52)$$

We then minimize the sum of squared deviations from the implied volatility in the data,

$$\min_{\theta} \sum_{j=1}^n (\sigma_{j,t}(\theta) - \sigma_{j,t})^2. \quad (53)$$

As Christoffersen and Jacobs note, this is just a weighted least squares problem that, with the monotonicity of the option price in θ , satisfies the usual regularity conditions.

[Insert Table 1 Here]

I report estimates for two sample days around the ERM crisis in Table 1. Haas, Mittnik and Mizrach (2006) establish that the model fits well both before and after the crisis. Inferences drawn from the BP/US\$ options are analyzed in the next section.

9. Inferences from the Mixture Model

There is often a belief that the standard model is sufficient in all but the most extreme circumstances. It turns out that a formal test of this kind in a mixture model is not at all straightforward.

² We can also price exotics in this framework, pricing options using Monte Carlo or other numerical procedures. Conditions under which these simulated moments models are appropriate is considered in Mizrach (2006).

The forecast comparisons discussed in the second and third parts may make more sense for practitioners.

9.1 Tests of the adequacy of Black-Scholes

Hansen (1996) considers the case of additive nonlinear models like our mixture model,

$$\begin{aligned}\sigma_{j,t} &= (1 - \lambda_1 - \lambda_3)\Phi(i_d - i_f, \sigma_2(\theta_3)) + \lambda_1\Phi(\mu_1(\theta_1), \sigma_1(\theta_2)) \\ &\quad + \lambda_3\Phi(\mu_3, \sigma_3\sigma_1(\theta_6))\end{aligned}$$

Under Black-Scholes, $\lambda_1 = \lambda_3 = 0$, and it is tempting to compare the models using the likelihood ratio. Unfortunately, under the Black-Scholes alternative, the parameters in the two tail log normals, θ_1, θ_2 and θ_6 , are nuisance parameters. Formally, the derivative of the likelihood function is flat with respect to these parameters. Their non-identification invalidates the distribution theory for the standard LR test. Hansen (1996) reports severe size distortions in several cases for the standard test.

Fortunately, Hansen also offers a constructive alternative. In our present setting, we can take minimizing the squared residuals,

$$\varepsilon_t^2 = (\sigma_{j,t}(\theta) - \sigma_{j,t})^2$$

as the objective function. The restricted alternative is given by

$$\varepsilon_{R,t}^2 = (\sigma_{j,t}(\theta_3) - \sigma_{j,t})^2$$

Construct the F -test,

$$F_n = n \sum_{j=1}^n (\varepsilon_{R,t}^2 - \varepsilon_t^2) / \sum_{j=1}^n \varepsilon_{R,t}^2. \quad (54)$$

Hansen then shows that

$$F_n = \sup_{\theta \in \Theta} F_n(\theta).$$

whose distribution is *not* asymptotically F and must be obtained by bootstrap. Since coefficients are likely to change daily, this is potentially tedious procedure.

9.2 Forecast comparison

9.2.1 Hypothesis tests on the forecast intervals

To compare the value-at-risk on the 2 days at both confidence levels. I adapt the framework of Christoffersen (1998). In each of the three cases, I construct the test of the null hypothesis from a

sequence of Bernoulli trials,

$$I_t^\alpha = \begin{cases} 1, & \Pr [(S_T - S_t)/S_t < -\%VaR^\alpha] > \alpha \\ 0 & \Pr [(S_T - S_t)/S_t < -\%VaR^\alpha] < \alpha \end{cases}, \quad (55)$$

where S_t is the current spot price, $T - t$ is four weeks, and α is the critical level of the VaR . I take the VaR loss intervals from August 20 as the null, and compute $-\%VaR^{0.05} = -4.84\%$ and $-\%VaR^{0.01} = -7.14\%$

Since the VaR is assessing tail risk, we are concerned with the coverage of the forecast interval from $(-\infty, -\%VaR^\alpha)$. Under what Christoffersen calls unconditional coverage, I test

$$H_0 : E[I_t^\alpha] = \alpha,$$

using the likelihood ratio

$$LR_\alpha = -2 \ln[\alpha^{n_0}(1 - \alpha)^{n_1} / (\hat{\pi}_\alpha^{n_0}(1 - \hat{\pi}_\alpha)^{n_1})], \quad (56)$$

where

$$\hat{\pi}_\alpha = n_0 / (n_1 + n_0),$$

is the maximum likelihood estimator of α . Under H_0 , LR_α is distributed $\chi^2(1)$.

As one might expect, the VaR rose substantially with the higher volatility during the crisis. Values at risk rises to $-\%VaR^{0.05} = -6.11\%$ and $-\%VaR^{0.01} = -8.73\%$. To assess the statistical significance of this, I implement the test in (56) in Table 1. I simulate $N = 250$ forecasts from the September 17 parameterization. Let n_0 be the number of times that $(S_T - S_t)/S_t$ is less than $-\%VaR^\alpha$. I find n_0 under the August loss coverage rejects too often, 26 times, rather than the expected $0.05 \times 250 = 12.5$ times under the null. The likelihood ratio is 11.87 which rejects the null at better than 1% significance, At the 99% level, I find $n_0 = 3$, and the test, with $LR_{0.01} = 0.094$, is not nearly powerful enough to reject that the forecasts are different.

9.2.2 Comparing the entire density

Next, we evaluate the forecast densities produced across their entire support. The approach we take is the one originally proposed by Berkowitz (2001). He notes that the probability integral transform

$$\hat{F}(s_t) = \int_{-\infty}^{s_t} f(u) du.$$

generates uniform, independent and identically distributed estimates under fairly weak assumptions.

Testing for an independent uniform density in small samples can be problematic, so Berkowitz

suggests transforming the data into normal random variates,

$$z_t = \Phi^{-1}(\widehat{F}(s_t)),$$

A simple test of the null hypothesis that the transformed forecast statistics, z_t , have mean zero can be performed using the likelihood ratio,

$$LR = \sum_{t=1}^T \left(\frac{z_t^2}{\hat{\sigma}^2} - 1 \right), \quad (57)$$

where $\hat{\sigma}$ is the forecast standard deviation. LR is then approximately distributed $\chi^2(1)$.

I graph the forecast density for August 20, 1992 in Figure 4.

[Insert Figure 4 Here]

The realized 4-week returns are then plotted in comparison to the density. To the naked eye, the left-tail, associated with BP depreciation, is substantially longer. This was indeed detected in our VAR exercise in the prior section.

When looking across the entire density, the power of the test is substantially weaker. The average forecast is 0.9756 or approximately a 2.5% decline. The average z -value is -0.4261 , and $\hat{\sigma}^2 = 2.4195$. I compute $LR = 1.5008$ which has a p -value of only 0.2206. Clearly, when the risks are one-sided, you want to exploit this information.

10. Jump Processes

I now move back to the case of a single process for the underlying. The key step is to introduce discontinuities in the price process through jumps. I begin with the Merton (1976) model as a baseline, and then allow for stochastic volatility in the second part.

10.1 Merton model

Merton (1976) has proposed a *jump diffusion model*

$$dS_t = (\mu - \lambda k)S_t dt + \sigma S_t dW_t + dq_t, \quad (58)$$

where dW_t is a Wiener process, dq_t is the Poisson process generating the jumps, and σ is the volatility. dW and dq are considered independent. This assumption is important because we cannot apply risk-neutral valuation to situations where the jump size is systematic.

λ is the rate at which jumps happen, μ is the expected return, and k is the average jump size.

This model gives rise to fatter left and right tails than Black-Scholes. If we assume that the log of k is normal with standard deviation δ , the European call option price is

$$C = \sum_{n=0}^{\infty} \frac{e^{-\lambda'T} (\lambda'\tau)^n}{n!} f_n, \quad (59)$$

where $\lambda' = \lambda(1+k)$ and f_n is the Black-Scholes option price when the variance is

$$\sigma^2 + n\delta^2/T,$$

and the risk free rate is

$$r - \lambda k + n \ln(1+k)/T.$$

Carr and Wu (2004) consider the generalization of the jump diffusion model to Levy processes. Bates (1991) has estimated this model to infer the risk in options prior to the 1987 stock market crash.

10.2 Bipower Variation

I follow Andersen, Bollerslev and Diebold (2007) and consider a stochastic volatility model with jumps,

$$dp_t = \mu_t dt + \sigma_t dW_t + \kappa_t dq_t \quad (60)$$

where $p_t = \ln(S_t)$, q_t is a counting process with intensity λ_t , and κ_t is the jump size with mean μ_κ and standard deviation σ_κ . The *quadratic variation* for the cumulative return process, $r_t = p_t - p_0$ is then

$$[r, r]_t = \int_0^t \sigma_s^2 ds + \sum_{0 < s \leq t} \kappa^2(s). \quad (61)$$

Estimation of the quadratic variation proceeds with discrete sampling from the log price process. Denote the Δ -period returns by $r_{t+1, \Delta} = p_{t+1} - p_{t+1-\Delta}$. The *realized volatility* is

$$RV_{t+1, \Delta} = \sum_{j=1}^{1/\Delta} r_{t+j \cdot \Delta, \Delta}^2. \quad (62)$$

In the now vast literature on the stochastic volatility model, $\kappa = 0$, researchers have employed realized volatility as an estimator of the integrated volatility, $\int_0^t \sigma_s^2 ds$.

In the case of discontinuous price paths, Barndorff-Nielsen and Shephard (2006) show that the realized volatility will also include the jump component,

$$p \lim_{\Delta \rightarrow 0} RV_{t+1, \Delta} = \int_t^{t+1} \sigma_s^2 ds + \sum_{t < s \leq t+1} \kappa^2(s). \quad (63)$$

To extract the integrated volatility from (60), Barndorff-Nielsen and Shephard have also introduced the *realized bi-power variation*,

$$BV_{t+1, \Delta} = \mu_1^{-2} \sum_{j=2}^{1/\Delta} |r_{t+j \cdot \Delta, \Delta}| |r_{t+(j-1) \cdot \Delta, \Delta}| \quad (64)$$

where $\mu_1 = \sqrt{2/\pi}$. It is then possible to show

$$p \lim_{\Delta \rightarrow 0} BV_{t+1,\Delta} = \int_t^{t+1} \sigma_s^2 ds. \quad (65)$$

By comparing (63) and (65), we have the estimate of just the jump portion of the process,

$$p \lim_{\Delta \rightarrow 0} (RV_{t+1,\Delta} - BV_{t+1,\Delta}) = \sum_{t < s \leq t+1} \kappa^2(s). \quad (66)$$

I next turn to an application of this analysis in the ERM exchange rate context

10.3 An application

I will now take the sampling interval Δ to be daily changes, and compute 50-period rolling sample estimates of realized volatility,

$$RV_t^{1/2} = (\sum_{j=1}^{50} r_{t-j}^2 / 50)^{1/2} \quad (67)$$

and bipower variation,

$$BV_t = (\pi/2) \sum_{j=1}^{50} |r_{t-j}| |r_{t-j-1}| / 50. \quad (68)$$

We constrain the jump risk to be positive,

$$\kappa_t = (\max[RV_t - BV_t, 0])^{1/2} \quad (69)$$

I again use the British Pound/US\$ exchange rate sample for January to October 1992, and plot the realized standard deviation and jump risk in Figure 5.

[Insert Figure 5 Here]

The figure has several features worth noting. The first is that while realized volatility for the British Pound was quite high leading into the Spring of 1992, jump risk is zero until April 24. Jump risk rises rather steadily through the rest of the spring, peaking at 0.2836% on June 1. Jump risk then falls through the summer, returning to zero from July 24 to August 24.

On August 25, the jump risk spikes to 0.2188% and then falls back until September 14th, when it spikes again to 0.4155%. Following a third local maximum of 0.3850% on Britain's September 17th ERM withdrawal, jump risk falls back to zero by September 22. This is true even though the realized volatility remains as high as it was during the crisis.

Although not shown in the figure, both the volatility and jump risks remain high into 1993. At the end of 1992, more than 3 months after the crisis, the realized volatility is at 0.9007%. The jump risk spikes twice above 0.30%, October 19 and December 2.

11. Conclusion

It is important to stress what this paper has not discussed: the preservation of moments into the change to the risk neutral measure. While straightforward in the Black-Scholes case, the results do not hold in general. This paper has only shown, empirically, that there is useful information in the risk neutral measure. The stochastic volatility model was not addressed in great detail; my emphasis on tail behavior gave priority to jumps. I also have not discussed approaches like Bates (2006) that rely on the characteristic function.

The paper has covered ground in three important areas: (1) nonparametric approaches estimating implied densities directly from option prices; (2) parametric modeling of the local volatility surface; (3) generalizations of the PDEs for the underlying process.

I also evaluate and implement tools for hypothesis testing and evaluating forecasts from alternative models. The nuisance parameter issue is addressed with a bootstrap. Value-at-risk and forecast comparison relies on the formal statistical procedures introduced by Christoffersen (1998) and Berkowitz (2001).

I last considered the very recent work of Barndorff-Neilsen and Shephard on bipower variation. This enables researchers to easily extract the jump component from the underlying. These discontinuities are unhedgeable risks that may be systemic.

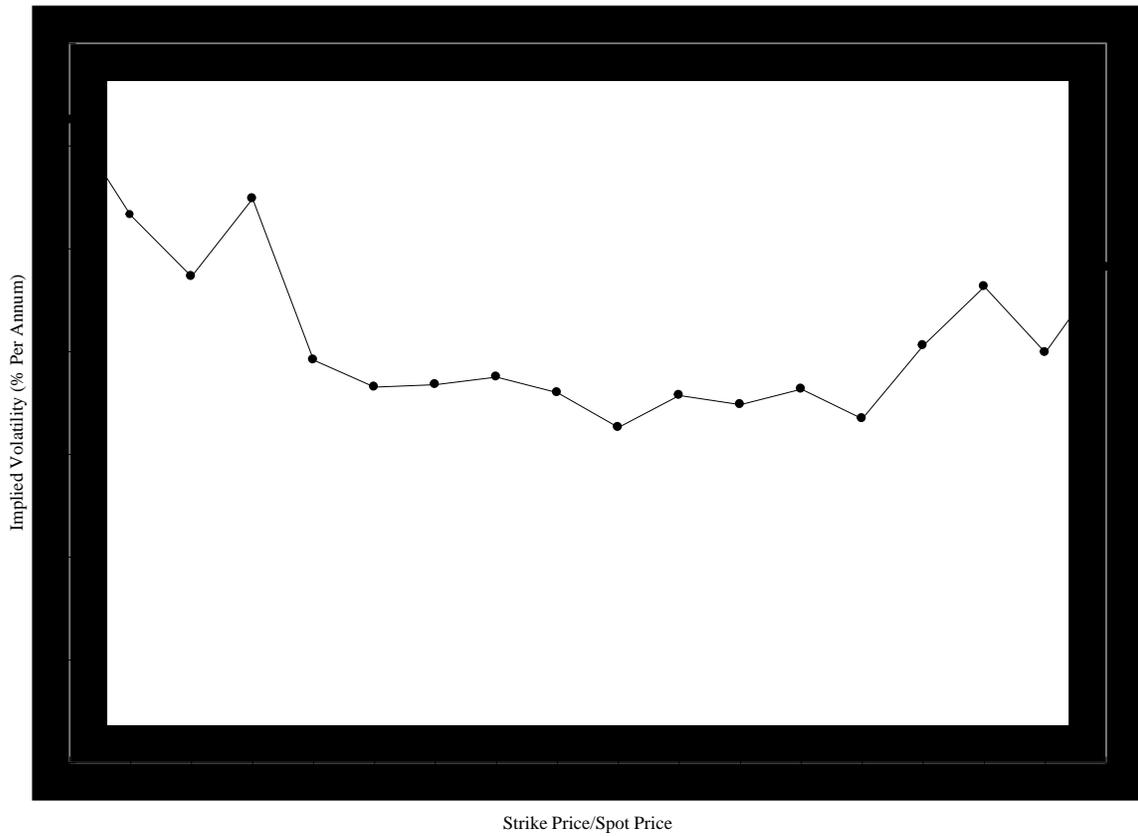
Policy makers may find these tools and inference worthwhile in a variety of contexts. Their subjective weights between type I and type II errors should not only be tested ex-post but incorporated directly in the estimation. Both Skouras (2007) and Christoffersen and Jacobs (2004) have made progress along these lines.

References

- Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2007), "Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility" *Review of Economics and Statistics* 89, 701-20.
- Bakshi, C., Cao, C. and Z. Chen (1997), "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance* 52, 2003-2049.
- Barle, Stanko and Nusret Cakici (1998). "How to Grow a Smiling Tree," *Journal of Financial Engineering* 7, 127-46.
- Barndorff-Nielsen. O. and N. Shephard (2006), "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Financial Econometrics* 4, 1-30.
- Barone-Adesi, G. and R.E. Whaley, (1987), "Efficient Analytic Approximation of American Option Values," *Journal of Finance* 42, 301-20.
- Bates, D. (1991), "The Crash of '87: Was it Expected? The Evidence from Options Markets," *Journal of Finance* 46, 1009-44.
- Bates, D. (2006), "Maximum Likelihood Estimation of Latent Affine Processes," *Review of Financial Studies*, forthcoming.
- Berkowitz, J. (2001), "Testing Density Forecasts with Applications to Risk Management," *Journal of Business and Economic Statistics* 19, 466-74.
- Bliss, R. and N. Panigirtzoglou, (2002), "Testing the Stability of Implied Probability Density Functions," *Journal of Banking and Finance* 26, 381-422.
- Bjerksund, P. and G. Stensland, (1993), "Closed-Form Approximation of American Options," *Scandinavian Journal of Management* 9, 87-99.
- Breeden, D. and R. Litzenberger (1978), "State Contingent Prices Implicit in Option Prices," *Journal of Business* 51, 621-51.
- Brunner, B. and R. Hafner (2003). "Arbitrage-free estimation of the risk-neutral density from the implied volatility smile," *Journal of Computational Finance* 7, 76-106.
- Cakici, Nusret, and Kevin R. Foster (2002). "Risk-Neutralized At-the-Money Consistent Historical Distributions in Currency Options Pricing," *Journal of Computational Finance* 6, 25-47.
- Carr, P. and L. Wu (2004), "Time-changed Levy processes and option pricing," *Journal of Financial Economics* 71, 113-41.
- Christoffersen, P., (1998), "Evaluating Interval Forecasts," *International Economic Review* 39, 841-862.
- Christoffersen, P. and K. Jacobs (2004), "The Importance of the Loss Function in Option Valuation," *Journal of Financial Economics* 72, 291-318.
- Dumas, B., J. Fleming, and R. Whaley, (1998), "Implied Volatility Functions: Empirical Tests," *Journal of Finance* 53, 2059-2106.

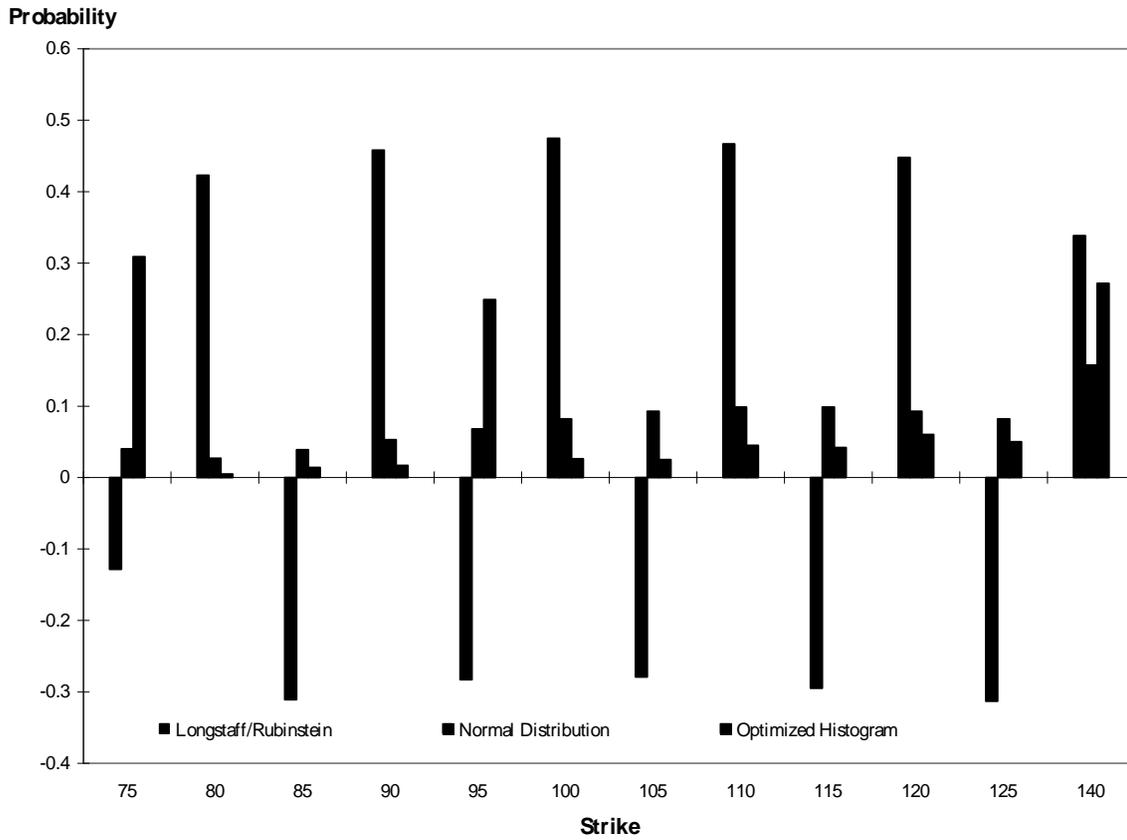
- Dupire, B. (1994), "Pricing with a Smile," *Risk* 7, 18–20.
- Haas, M., S. Mittnik, and B. Mizrach, "Assessing Central Bank Credibility During the EMS Crises: Comparing Option and Spot Market-Based Forecasts," *Journal of Financial Stability* 2, 2006, 28-54.
- Hambly, B. (2006), *Lecture Notes on Financial Derivatives*, Oxford University.
- Hoffman, C. (2000), "Valuation of American Options," Oxford University Thesis.
- Karatzas, I. and S. Shreve (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. New York: Springer-Verlag,
- Melick, W. and C. Thomas (1997), "Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil During the Gulf Crisis," *Journal of Financial and Quantitative Analysis* 32, 91-115.
- Merton, R. (1976), "Option Pricing when Underlying Stock returns are Discontinuous," *Journal of Financial Economics*, 3, 124-44.
- Mizrach, B., (1995), "Target Zone Models with Stochastic Realignment: An Econometric Evaluation." *Journal of International Money and Finance* 14, 641–657.
- Mizrach, B. (2006), "The Enron Bankruptcy: When Did The Options Market Lose Its Smirk," *Review of Quantitative Finance and Accounting* 27, 2006, 365-82.
- Ritchey, R. (1990), "Call Option Valuation for Discrete Normal Mixtures," *Journal of Financial Research* 13, 285-295.
- Shimko, D. (1993), "Bounds of Probability," *Risk* 6, 33-37.
- Skouras, S. (2007), "Decisionmetrics: A Decision-Based Approach To Econometric Modelling," *Journal of Econometrics* 127, 414-40.
- Stein, E.M. and J.C. Stein (1991), "Stock Price Distributions with Stochastic Volatility: An Analytic Approach," *Review of Financial Studies* 4, 727-52.
- Tompkins, R. (2001). "Implied Volatility Surfaces: Uncovering Regularities for Options on Financial Futures," *The European Journal of Finance* 7, 198-230.
- Wiggins, J.B. (1987), "Option Values under Stochastic Volatility: Theory and Empirical Estimates," *Journal of Financial Economics* 19, 351-72.

Figure 1
Averages of Implied Volatility BP/US\$ Options 1992-93



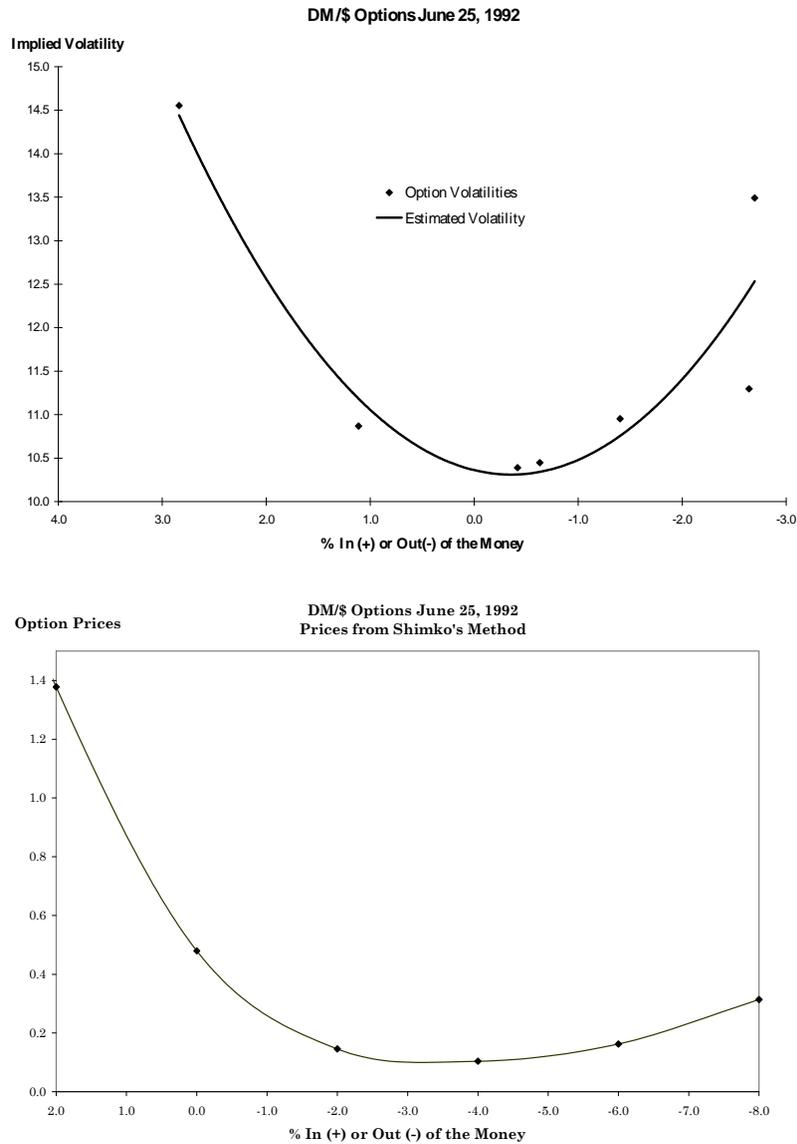
NOTES: The figure reports implied volatility for BP/US\$ options from the Philadelphia Exchange for 1992-93.

Figure 2
Histogram Estimators



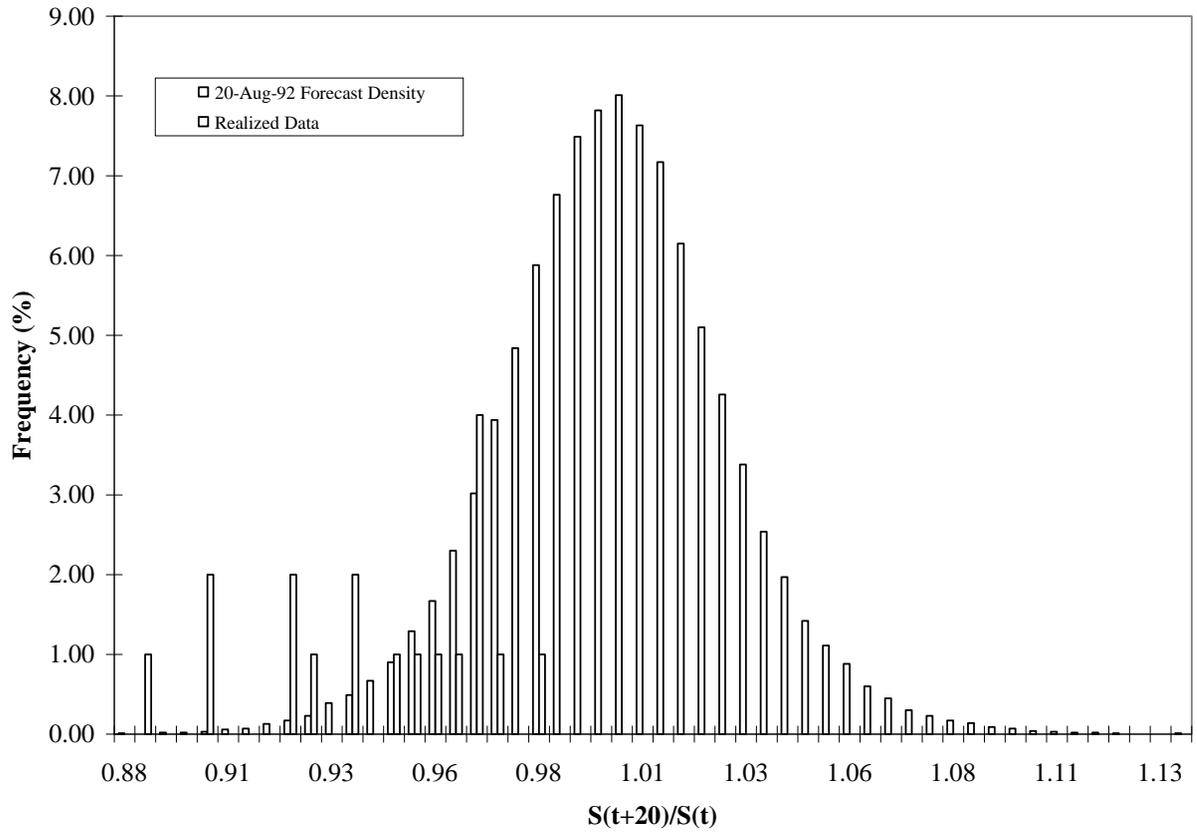
NOTES: The figure reports various different histogram estimators.

Figure 3
Local Volatility Functions and Arbitrage Violations



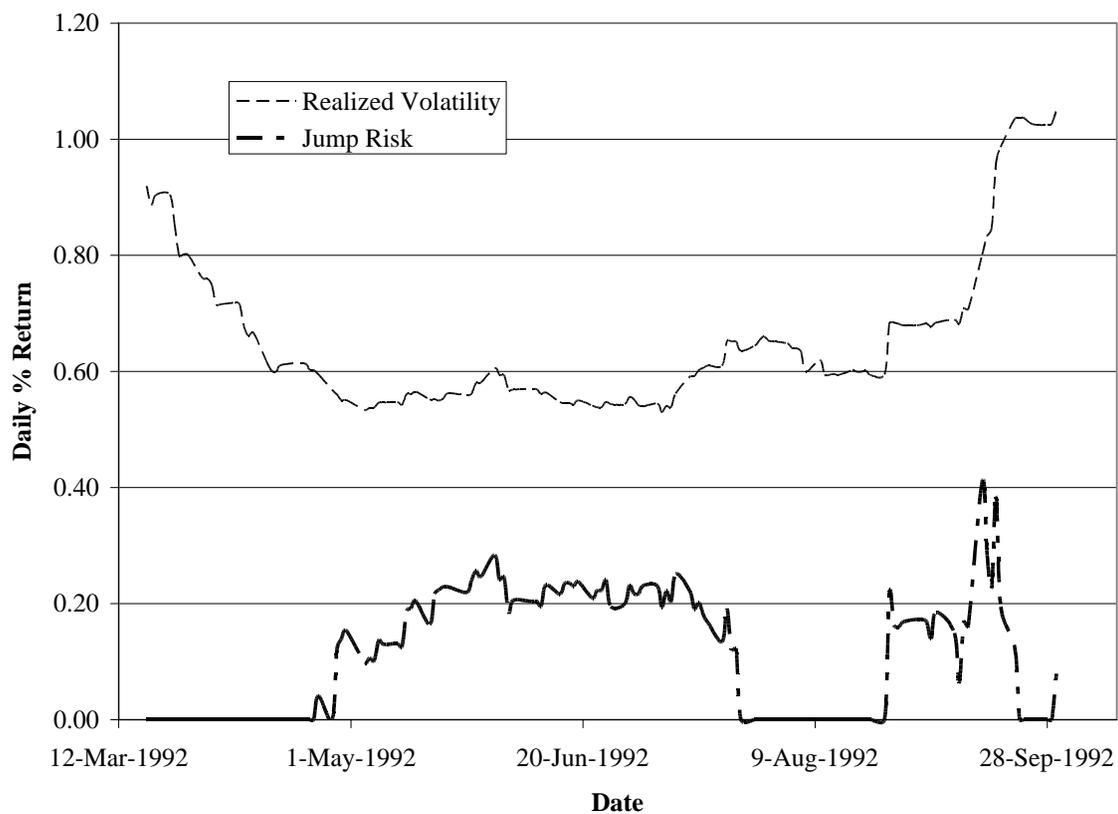
NOTES: The dark line in panel (a) is a quadratic function $0.4035 \times (\%M)^2 + 0.288 \times (\%M) + 10.362$ fit to the implied volatility of the DM/US\$ options. As we extrapolate the local volatility surface further out of the money, the model produces an arbitrage violation in panel (b) for the calls more than 5% out of the money.

Figure 4
Forecast Density Comparison



NOTES: The light bars are the forecast density for the 4-week return on the spot exchange rate using the mixture model parameters from August 20, 1992 in Table 1. The dark bars are the subsequent 20 realizations. A formal statistical comparison is conducted using (57).

Figure 5
Jump Risk Estimates



NOTES: The figure plots realized daily standard deviation (67) and jump risk (32) for the British Pound spot rate between March 18 and September 30, 1992.

Table 1
Estimates of the Mixture of Log Normals Model

Date	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	$-\%VaR^{0.05}$	$-\%VaR^{0.01}$	$LR_{0.05}$	$LR_{0.01}$
20-Aug-1992	-2.179 (0.01)	-1.908 (0.15)	-0.816 (0.01)	-1.855 (0.01)	-1.908 (0.07)	-2.440 (0.63)	-4.84%	-7.14%		
17-Sep-1992	-0.882 (0.07)	-0.943 (0.48)	-3.634 (0.48)	-3.108 (0.02)	-2.051 (0.09)	-1.933 (2.12)	-6.11%	-8.73%	11.87 (0.00)	0.09 (0.76)

NOTES: The θ' s are estimates of the model (14). t -ratios are in parentheses. $-\%VaR^{0.05}$ and $-\%VaR^{0.01}$ are the value at risk, in percent of long position in the BP over a four-week time horizon at the 95 and 99% confidence level. The LR statistic is given by (55) and compares the VaR across the two dates. The simulation size is $N = 250$ observations. The LR test, with p -values underneath, is distributed $\chi^2(1)$.