Implied risk-neutral probability density functions from option prices: theory and application

Bhupinder Bahra

Bank of England, Threadneedle Street, London, EC2R 8AH

The views expressed in this paper are those of the author and do not necessarily reflect those of the Bank of England. The author would like to thank Creon Butler, Paul Doran and Jim Steeley for their helpful suggestions and technical advice. He is also grateful for many helpful discussions with colleagues, and with Silio Aparacio and Stewart Hodges (University of Warwick, Financial Options Research Centre), Cedric Kohler (Union Bank of Switzerland), Mark Laycock (formerly at the Bank of England, now at Deutsche Morgan Grenfell), Allan Malz (Federal Reserve Bank of New York), Holger Neuhaus (formerly at the Deutsche Bundesbank, now at the European Monetary Institute) and Charles Thomas (Federal Reserve Board). Kevin Barry and Gareth James provided valuable research assistance.

Issued by the Bank of England, London, EC2R 8AH to which requests for individual copies should be addressed: envelopes should be marked for the attention of the Publications Group. (Telephone 0171 601 4030).

Bank of England 1997
ISSN 1368-5562

Contents

Abstract

Derivative markets provide monetary authorities with a rich source of information for gauging market sentiment. For example, a futures price is the market's expectation of the future value of an asset. More interestingly, it is possible to derive the higher moments of future asset values from the market prices of options. These can be extracted in the form of a risk-neutral probability distribution of the underlying asset price at the maturity date of the options. In this paper we develop various techniques for estimating the market's *ex ante* risk-neutral probability density function of an underlying asset price from the prices of options on that asset. We then illustrate the potential value of this type of information to the policy-maker in assessing monetary conditions, monetary credibility, the timing and effectiveness of monetary operations, and in identifying anomalous market prices.

1 Introduction

Many monetary authorities routinely use the information that is embedded in financial asset prices to help in formulating and implementing monetary policy. In this context, derivative markets provide them with a rich source of information for gauging market sentiment; due to their forward-looking nature, futures and options prices efficiently encapsulate market perceptions about underlying asset prices in the future. Although the information that is embedded in futures prices can be derived from cash market instruments, options prices do reveal genuinely new information about underlying price processes.

For example, the variance that is implied by an option's price is the market's *ex ante* estimate of the underlying asset's return volatility over the remaining life of the option. More interestingly, it is possible to derive the higher moments of future asset values from the market prices of *European* options.[1] These can be extracted in the form of an *ex ante* risk-neutral probability distribution of the underlying price at the maturity date (or terminal date) of the options.

In this paper we develop various techniques for estimating the market's implied terminal risk-neutral density (RND) function of an underlying asset price from the prices of options on that asset. We then illustrate the potential value of this type of information to the policy-maker in assessing monetary conditions, monetary credibility, the timing and effectiveness of monetary operations, and in identifying anomalous market prices. In Section 2 we look at the theoretical relationship between option prices and RND functions. In Section 3 we describe and apply some techniques for estimating RND functions, and consider some of the advantages and disadvantages of each approach. We then describe how our preferred approach can be applied to LIFFE equity and interest rate options, and to Philadelphia Stock Exchange currency options. Section 4 illustrates the potential value of implied RND functions to the policy-maker in terms of the information they provide that is additional to mean estimates of future asset prices. In Section 5 we describe some data limitations. We conclude in Section 6 and provide a relevant technical derivation in a mathematical appendix.

[1] A European call (put) option on a given underlying asset is a contract that gives the holder the right, but not the obligation, to buy (sell) that asset at a certain date in the future at a predetermined price. The predetermined price at which the underlying asset is bought or sold, which is stipulated in an option contract, is known as the *exercise price* or *strike price*. The date at which an option expires is known as the *maturity date* or *exercise date*. Options that can be exercised at any time up to and including the maturity date are known as *American* options.

2 The relationship between option prices and RND functions

European call options on the same underlying asset, and with the same time to maturity, but with different exercise prices, can be combined to mimic other state-contingent claims, that is, securities whose returns are dependent on the 'state' of the economy at a particular time, T, in the future. The prices of such state-contingent securities reflect investors' assessments of the probabilities of particular states occurring in the future.

Intuitively, this can be seen by noting that the difference in the price of two call options with adjacent exercise prices reflects the value attached to the ability to exercise the options when the price of the underlying asset lies between their exercise prices. This clearly depends on the probability of the underlying asset price lying in this interval. In this way, the prices of European call options of a given maturity, but with a range of different exercise prices, are related to the weights attached by the representative risk-neutral agent to the possible outcomes for the terminal price of the underlying security.[2]

An important example of a state-contingent claim is the *elementary claim*. First introduced in the time-state preference model of Arrow (1964) and Debreu (1959), it is the fundamental building block from which we have derived much of our current understanding of the theory of finance under uncertainty. An elementary claim, also known as an 'Arrow-Debreu' security, is a derivative security that pays £1 at future time T if the underlying asset (or portfolio of assets) takes a particular value, or 'state', S_T, at that time, and zero otherwise. The prices of Arrow-Debreu securities, known as state prices, at each possible state are directly proportional to the risk-neutral probabilities of each of the states occurring.[3]

Given its enormous informational value, it is unfortunate that the Arrow-Debreu security is not a traded commodity on any exchange, and hence its price is not directly observable. However, as indicated above, such a security can be replicated by investing in a suitable combination of European call options, known as a *butterfly spread*. The state price at any given state is the cost of the butterfly spread centred on that particular state.[4]

[2] See Cox and Ross (1976) for the pricing of options under risk neutrality.
[3] The constant of proportionality is the present value of a zero-coupon bond that pays £1 at time T, with the discount rate being the risk-free rate of interest.
[4] See Banz and Miller (1978), Breeden and Litzenberger (1978) and Ross (1976).

Ross (1976) first demonstrated how to relate call option prices to state prices (and hence to risk-neutral densities). Breeden and Litzenberger (1978) showed that if the underlying price at time T has a continuous probability distribution, then the state price at state S_T is determined by the second partial derivative of the European call option pricing function for the underlying asset(s) with respect to the exercise price, $\partial^2 c/\partial X^2$, evaluated at an exercise price of $X=S_T$.[5] When applied across the continuum of states, $\partial^2 c/\partial X^2$ equals the 'state pricing function.' It follows that $\partial^2 c/\partial X^2$ is directly proportional to the risk-neutral probability density function of S_T. All of the techniques for estimating terminal RND functions from options prices can be related to this result.

2.1 Pricing elementary claims from option prices

The Breeden and Litzenberger (1978) approach, which was developed within a time-state preference framework, provides the most general approach to pricing state-contingent claims.

The one-unit payoff to an elementary claim at a given future state, $S_T=X$, can be achieved by selling two call options, each with exercise price $X=S_T$, and buying two call options, one with exercise price $S_T-\Delta S_T$ and one with exercise price $S_T+\Delta S_T$, where ΔS_T is the step size between adjacent calls. This portfolio of four call options is a butterfly spread centred on state $S_T=X$. The payoff to the butterfly spread, evaluated at the exercise price $X=S_T$ is given by:

$$\left.\frac{[\bar{c}(S_T+\Delta S_T,\tau)-\bar{c}(S_T,\tau)]-[\bar{c}(S_T,\tau)-\bar{c}(S_T-\Delta S_T,\tau)]}{\Delta S_T}\right|_{X=S_T}=1 \qquad \textbf{(1)}$$

where $\bar{c}(X,\tau)$ denotes the payoff to a European call option with strike price X and time-to-maturity τ. As ΔS_T tends to zero, the payoff function of the butterfly tends to a Dirac delta function with its mass at $X=S_T$, that is, in the limit the butterfly becomes an Arrow-Debreu security paying £1 if $S_T=X$ and zero for all other states.

For example, consider a butterfly spread centred on state $S_T=3$ with unit step size between adjacent options. This consists of two short (written) call options, each with exercise price $X=3$, and two bought calls, one with exercise price 3-1=2 and one with exercise price 3+1=4. The payoffs, for

[5] The call pricing function relates the call price to the exercise price for options on the same underlying instrument and with the same time-to-maturity.

different possible future values of the underlying asset, of each of these call options are given below in Table A.

Table A
Payoff to a butterfly spread centred on state S_T=3 with unit step size between adjacent options

Possible values of S_T	Payoff to a long call with exercise price X=2	Payoff to a short call with exercise price X=3	Payoff to a long call with exercise price X=4	Payoff to butterfly - calculated using equation **(1)**
1	0	0	0	0
2	0	0	0	0
3	1	0	0	1
4	2	1	0	0
5	3	2	1	0
6	4	3	2	0

Table A shows that the butterfly pays one unit only if the terminal value of the underlying asset is 3. All other integer values of S_T result in a zero payoff.

If we denote $P(S_T,\tau;\Delta S_T)$ as the time-t (current) price of an elementary claim (or butterfly spread) centred on state $S_T=X$, then $P(S_T,\tau;\Delta S_T)$ divided by the step size between adjacent calls, ΔS_T, may be written in terms of the prices of the constituent call options as the following second-order difference quotient:

$$\frac{P(S_T,\tau;\Delta S_T)}{\Delta S_T} = \frac{[c(S_T+\Delta S_T,\tau)-c(S_T,\tau)]-[c(S_T,\tau)-c(S_T-\Delta S_T,\tau)]}{(\Delta S_T)^2} \qquad \textbf{(2)}$$

where $c(X,\tau)$ denotes the price of a European call option with strike price X and time-to-maturity $\tau=T-t$. In the limit, as the step size tends to zero, the price of the butterfly spread at state $S_T=X$ tends to the second derivative of the call pricing function with respect to the exercise price, evaluated at $X=S_T$;

$$\lim_{\Delta S_T \to 0} \frac{P(S_T,\tau;\Delta S_T)}{\Delta S_T} = \left.\frac{\partial^2 c(X,\tau)}{\partial X^2}\right|_{X=S_T} \qquad \textbf{(3)}$$

It can be seen that if we could price butterfly spreads across the full continuum of states, each with infinitely small step sizes between exercise prices, then we would have the complete state pricing function.

The price of an Arrow-Debreu security can also be expressed as an expected future payoff, that is, the present value of £1 multiplied by the risk-neutral probability of the state that gives rise to that payoff, $S_T=X$, occurring. Equated with equation **(3)** and applied across the continuum of possible values of S_T, this gives the result that the second derivative of the call pricing function with respect to the exercise price is equal to the discounted RND function of S_T conditioned on the underlying price at time t, S; i.e.

$$\frac{\partial^2 c(X,\tau)}{\partial X^2} = e^{-r\tau} q(S_T) \qquad \textbf{(4)}$$

where r is the (annualised) risk-free rate of interest over the time period $\tau=T-t$, and $q(S_T)$ is the RND function of S_T.

In the absence of arbitrage, $c(X,\tau)$ is convex and monotonic decreasing in exercise price, which implies that all butterfly spreads that can be formed along the continuum of states have a positive price. This results in a positive RND function. If arbitrage opportunities do exist at some states, then $c(X,\tau)$ will not be monotonic decreasing and convex in exercise price and the values of $q(S_T)$ will be negative at those states. Also note that with techniques which do not specify a distribution for $q(S_T)$ it may be necessary to impose the condition that $q(0)=q(\infty)=0$.

The derivation of the Breeden and Litzenberger result makes no assumptions about the underlying asset price dynamics. Aside from the assumption that markets are perfect,[6] the only requirement to be able to estimate $q(S_T)$ is that $c(X,\tau)$ be twice differentiable; even this is not necessary for calculating a discretised state pricing function using equation **(2)**. Agents' preferences and beliefs have not been restricted since option prices are risk neutral with respect to the underlying risky asset.

2.2 The Black-Scholes (1973) formula and its RND function

We now review the assumptions of the classic Black-Scholes (1973) option pricing model and show how they relate to a lognormal implied terminal RND function. We will then show how the model is modified in practice, and how these modifications to the theoretical Black-Scholes prices result in non-lognormal implied terminal RND functions.

In order to calculate an option's price, one has to make an assumption about how the price of the underlying asset evolves over the life of the option, and therefore what its RND function, conditioned on S, is at the maturity date of the option. The Black-Scholes (1973) model assumes that the price of the

[6] Breeden and Litzenberger assume perfect markets, that is, there are no restrictions on short sales, there are no transactions costs or taxes, and investors may borrow at the risk-free rate of interest.

underlying asset evolves according to a stochastic process called *geometric Brownian motion* (GBM) with an instantaneous expected drift rate of μS and an instantaneous variance rate of $\sigma^2 S^2$:

$$dS = \mu S dt + \sigma S dw \qquad \textbf{(5)}$$

where μ and σ are assumed to be constant and dw are increments from a *Wiener process*. Applying Ito's Lemma to equation **(5)** yields the result:[7]

$$\ln S_T \sim \varphi\left[\ln S + \left(\mu - \tfrac{1}{2}\sigma^2\right)\tau, \sigma\sqrt{\tau}\right] \qquad \textbf{(6)}$$

where $\varphi(\alpha, \beta)$ denotes a normal distribution with mean α and standard deviation β. Therefore, the Black-Scholes GBM assumption implies that the RND function of S_T, $q(S_T)$, is lognormal with parameters α and β (or, alternatively, that the RND function of underlying *returns* is *normal* with parameters μ and σ). The lognormal density function is given by:

$$q(S_T) = \frac{1}{S_T \beta\sqrt{2\pi}} e^{\{-(\ln S_T - \alpha)^2\}/2\beta^2} \qquad \textbf{(7)}$$

Like Cox and Ross (1976), Black and Scholes (1973) show that options can be priced as if investors are risk neutral by setting the expected rate of return on the underlying asset, μ, equal to the risk-free interest rate, r. The formula that Black and Scholes (1973) derived for pricing European call options is as follows:

$$c(X, \tau) = SN(d_1) - e^{-r\tau} XN(d_2) \qquad \textbf{(8)}$$

where

$$d_1 = \frac{\ln(S/X) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$$

$$d_2 = \frac{\ln(S/X) + (r - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} = d_1 - \sigma\sqrt{\tau}$$

and $N(x)$ is the cumulative probability distribution function for a standardised normal variable; i.e. it is the probability that such a variable will be less than x.

[7] See Hull (1993), chapter 10.

Since the price of an option does not depend upon μ, the expected rate of return on the underlying asset, except through *S*, a distribution recovered from option prices will not be the true distribution unless universal risk-neutrality holds, in which case $\mu = r$, the risk-free rate of interest.

2.3 The implied volatility smile curve

Of the parameters that determine the price of an option, the only one that is unobservable at time *t* is the underlying asset's return volatility over the remaining life of the option, σ. However, an estimate of this can be inferred from the prices of options traded in the market: given an option price, one can solve an appropriate option pricing model for σ to obtain a market estimate of the future volatility of the underlying asset returns. This type of estimate of σ is known as *implied volatility*.

Under the Black-Scholes assumption that the price of the underlying asset evolves according to GBM, the implied volatility ought to be the same across all exercise prices of options on the same underlying asset and with the same maturity date. However, implied volatility is usually observed in the market as a convex function of exercise price which is commonly referred to as the smile curve (illustrated in Figure 1). In other words, market participants price options with strikes which are less than *S*, and those with strikes greater than *S*, with higher volatilities than options with strikes that are equal to *S*.[8]

The existence of the volatility smile curve indicates that market participants make more complex assumptions than GBM about the path of the underlying asset price. And as a result, they attach different probabilities to terminal values of the underlying asset price than those that are consistent with a lognormal distribution.[9] The extent of the convexity of the smile curve indicates the degree to which the market RND function differs from the Black-Scholes (lognormal) RND function. In particular, the more convex the smile curve, the greater the probability the market attaches to extreme outcomes for S_T. This causes the market RND function to have 'fatter tails' than are consistent with a lognormal density function.[10] In addition, the direction in which the smile curve slopes reflects the skew of the market RND function: a positively (negatively) sloped implied volatility smile curve

---

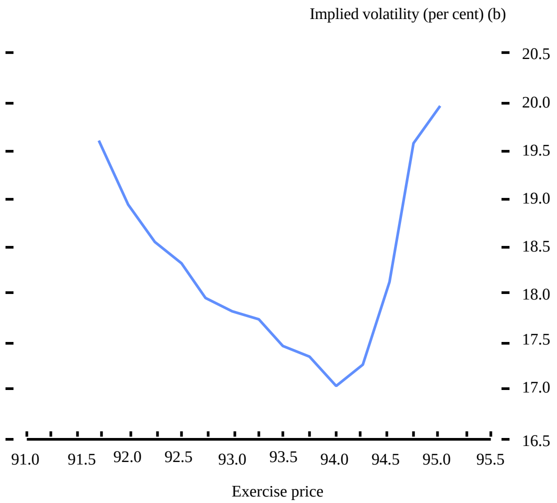
[8] Call options with strike prices which are less than *S* are said to be *in-the-money* (ITM), whilst call options with strikes which are greater than *S* are *out-of-the-money* (OTM), and those with strikes that are equal to *S* are *at-the-money* (ATM).

[9] Early empirical studies that document the differences between theoretical Black-Scholes prices and observed market prices include Black (1975), MacBeth and Merville (1980), Rubinstein (1985), and Whaley (1982).

[10] Evidence of fat tails, or leptokurtosis, in stock prices was first noted by Fama (1965).

results in an RND function that is more (less) positively skewed than the lognormal RND function that would result form a flat smile curve.

Figure 1
Implied volatility smile curve for LIFFE December 1996 options on the short sterling future[(a)]



(a) As at 16 April 1996. These options expire on 18 December 1996.
(b) Implied volatility is an annualised estimate of the instantaneous standard deviation of the return on the underlying asset over the remaining life of the option.

Any variations in the shape of the smile curve are mirrored by corresponding changes in the slope and convexity of the call pricing function. The slope and convexity of the smile curve, or of the call pricing function, can be translated into probability space to reveal the market's (non-lognormal) implied RND function for S_T. In the next section, we review techniques for undertaking this translation.

3 Some techniques for estimating implied terminal RND functions

In this section, we outline and implement various approaches for estimating RND functions from options prices. Four related approaches have been used in the literature: **(i)** assumptions are made about the stochastic process that governs the price of the underlying asset and the RND function is inferred from it;[11] **(ii)** a parametric assumption is made about the RND function itself and its parameters are recovered by minimising the distance between the observed option prices and those that are generated by the assumed functional form;[12] **(iii)** the RND function is derived directly from some parametric specification of the call pricing function (or of the implied volatility smile curve);[13] and **(iv)** the RND function is estimated *nonparametrically*, that is, with no parametric restrictions on either of the underlying asset price dynamics, the call pricing function, or the terminal RND function.[14]

Implementation of the Breeden and Litzenberger (1978) result, which underlies all of the techniques, requires that a continuum of European options with the same time-to-maturity exist on a single underlying asset spanning strike prices from zero to infinity. Unfortunately, since option contracts are only traded at discretely spaced strike price levels, and for a very limited range either side of the at-the-money (ATM) strike, there are many RND functions that can fit their market prices. Hence, all of the procedures for estimating RND functions essentially amount to interpolating between observed strike prices and extrapolating outside of their range to model the tail probabilities. Before describing some of these procedures, we begin by outlining a simple way of approximating the implied RND function via an implied risk-neutral histogram.

3.1 A simple approach: risk-neutral histograms

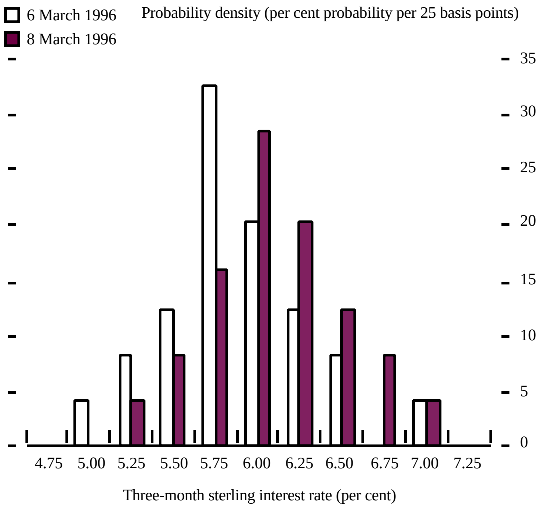
As shown in Section 2.1, the discrete valuation equation, **(2)**, gives the value of a butterfly spread centred on a given state, $S_T=X$. This value can be compounded at the risk-free interest rate to give an approximation to the risk-neutral probability of the underlying asset price lying at state $S_T=X$ at time T; i.e.

[11] See Bates (1991, 1995), and Malz (1995b).
[12] See Jackwerth and Rubinstein (1995), Melick and Thomas (1994), and Rubinstein (1994).
[13] See Bates (1991), Jarrow and Rudd (1982), Longstaff (1992, 1995), Malz (1995a) and Shimko (1993).
[14] See Aït-Sahalia and Lo (1995).

$$q(S_T) = e^{r\tau} \left. \frac{[c(S_T + \Delta S_T, \tau) - c(S_T, \tau)] - [c(S_T, \tau) - c(S_T - \Delta S_T, \tau)]}{\Delta S_T} \right|_{X = S_T} \quad \textbf{(9)}$$

Applying equation **(9)** to call prices observed across a range of exercise prices (states) results in the implied terminal risk-neutral histogram of the underlying asset price.[15] Figure 2 shows how the implied histogram for the three-month sterling interest rate on 19 June 1996 (as implied by the June short sterling futures price) changed between 6 March and 8 March 1996, a period which included a cut of 25 basis points in official UK interest rates and the publication of stronger-than-expected US non-farm payrolls data.[16]

Figure 2
Implied risk-neutral histograms for the three-mont sterling interest rate in June 1996[(a)]



(a) Derived using LIFFE June 1996 options on the short sterling future, as at 6 March and 8 March 1996. These options expire on 19 June 1996.

---

[15] Neuhaus (1995) prefers to derive the histogram indirectly by first computing the discretised implied *cumulative* distribution of the underlying price. This way the implied histogram shows the probabilities of the underlying asset price lying *between* two adjacent strike prices rather than in a fixed interval *around* each strike. His procedure also obviates the need to subjectively allocate any residual probability between the tails: because he derives the cumulative distribution the mass in each tail is automatically determined.

[16] The histograms were calculated using data for the LIFFE June 1996 option on the short sterling future. We used LIFFE settlement prices to avoid the problems associated with asynchronous data.

One of the drawbacks of this method of calculating the risk-neutral histogram is that it relies on options being traded at equally spaced strikes. Also, there is no systematic way of modelling the tails of the histogram, which may not be observable due to the limited range of exercise prices traded in the market.

Furthermore, nothing in this procedure can adjust for badly behaved call pricing functions and the existence of arbitrage opportunities. Observed prices sometimes exhibit small but sudden changes in convexity across strikes as well as small degrees of concavity in exercise price, which result in large variations in the probabilities over adjoining strike intervals, and negative probabilities respectively. These irregularities may be due, in cases where bid-ask spreads are observed instead of actual traded prices, to measurement errors arising from using middle prices. Irregular call pricing functions may also result if the option price data are not synchronous across exercise prices. The bias due to asynchronous data can be reduced significantly by using exchange settlement prices rather than intra-day quotes. Despite these shortcomings, this method provides a useful first approximation of implied RND functions.

Alternatively, sensible RND functions can be obtained by smoothing the call pricing function in a way that places less weight on data irregularities while preserving its overall form. The procedures described below attempt to do this by applying various interpolation and extrapolation techniques to model the complete RND function under the assumption of no arbitrage, thereby ensuring continuity, monotonicity and convexity of the call pricing function in exercise price.

3.2 Interpolating the call option pricing function directly

At first glance it would seem that the most obvious way of estimating the implied RND function is by direct application of the Breeden and Litzenberger (1978) result to the call option pricing function. This requires an interpolated call pricing function, $c(X,\tau)$, that is consistent with the monotonicity and convexity conditions, and that can be differentiated twice. This can be achieved either parametrically, by imposing a particular parametric functional form directly on the observed call prices and estimating its parameters by solving a (nonlinear) least squares problem, or nonparametrically, by applying a statistical technique called nonparametric kernel regression.[17]

[17] See Härdle (1991), chapter 5. Broadly this involves locally fitting polymonials along the call pricing function.

Bates (1991) interpolates the set of observed call prices directly by fitting a cubic spline to the observed data subject to convexity, monotonicity and level constraints. Because the call pricing function takes a fairly complex functional form, a relatively large number of degrees of freedom are required in order to infer it accurately in this way. Aït-Sahalia and Lo (1995) take a nonparametric approach and apply the ‘Nadaraya-Watson’ kernel estimator to estimate the entire call pricing function. They undertake the ambitious task of applying the kernel estimator to a time series of option prices across strikes in order to estimate all of the underlying determinants, namely, S, X, r, σ and τ.

The fact that the nonparametric regression approach involves a large number of regressors, coupled with the necessity to compute the second-order derivative of $c(X,\tau)$, makes nonparametric estimation of the implied RND function particularly data-intensive. Various assumptions can be made to reduce the dimensionality of the problem and to force the asymptotic convergence of the higher-order derivatives in small samples, but many of these are precisely the type of constraints that one wishes to avoid in the first place by employing a nonparametric estimator. Since we are working with a fairly limited number of prices quoted across strikes at any one point in time, this approach is practically unimplementable for most of the option markets with which we are concerned.

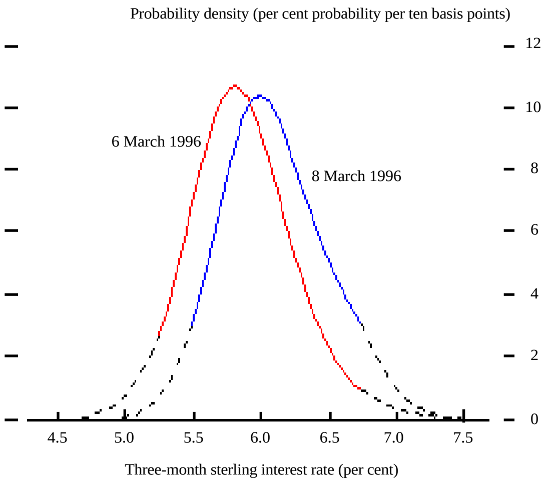
3.3 Interpolating the implied volatility smile curve

Shimko (1993) proposes an alternative methodology for interpolating the call pricing function and for modelling the tail probabilities. Arguing that Black-Scholes implied volatilities are more smooth than the option prices themselves, he proposes interpolating in the implied volatility domain instead of the call price domain, and he assumes that implied volatility is a quadratic function of exercise price for every exercise price within the traded range. He then uses the Black-Scholes formula to invert the interpolated smile curve, solving for the call price as a continuous function of the strike price. Note that Shimko’s use of the Black-Scholes formula to transfer between the call price and implied volatility domains does not require it to be true. He merely uses the formula as a translation device that allows him to interpolate implied volatilities rather than the observed option prices themselves. The indirectly interpolated call pricing function can then be differentiated twice to determine the implied RND function between the lowest and the highest strike options.

Shimko extrapolates beyond the traded strike range by grafting lognormal tails onto each of the endpoints of the observable density such that the total cumulative probability is one. This is done by matching the frequency and cumulative frequency of the implied RND with a lognormal distribution in each tail. This ensures continuity at the endpoints of both the (observable) density function and the (observable) cumulative distribution.

We experimented with various interpolating functions: piecewise-linear, hyperbolic, parabolic, the best-fit polynomial, and various quadratic and cubic spline structures with different numbers of knot points. In most cases a cubic spline with two knot points fitted the data better than any of the other functional forms. In particular, we find Shimko's assumption that the smile can be represented by a quadratic function to be somewhat restrictive. Actual implied volatilities, especially in the equity market, tend not to follow a parabolic form at far away-from-the-money strikes. Figure 3 illustrates the use of the Shimko technique, interpolating with a cubic spline, with LIFFE options on the short sterling future. It shows how the implied RND function for the implied three-month interest rate on 19 June 1996 changed between 6 March and 8 March 1996. Compare this with Figure 2 which shows the risk-neutral histograms on the same trade dates and for the same maturity.

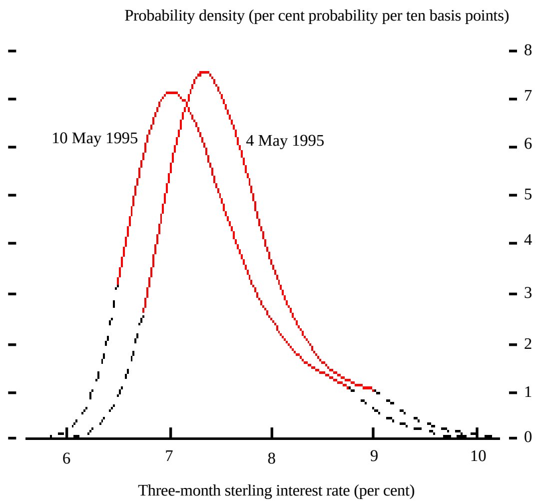
Figure 3
Implied RND functions for the three-month sterling interest rate in June 1996(a)



(a) Derived using LIFFE June 1996 options on the short sterling future, as at 6 March and 8 March 1996. These options expire on 19 June 1996. These graphs illustrate the result when using the Shimko technique. In this case the smile curve was interpolated using a cubic spline with two knot points. The tails, which are shown as dotted lines, are lognormal.

The problem with Shimko's extrapolation procedure (which grafts lognormal tails onto the observable part of the implied RND function) is that it arbitrarily assigns a constant volatility structure to the smile outside of the traded strike range. Since the final distribution is pieced together from three separate parts it is not always possible to ensure a smooth transition from the observable part of the distribution to the tails.[18] The transitions to the upper tails of the RND functions in Figure 4 are examples of cases when Shimko's approach does not produce plausible results. The alternative techniques for estimating implied RND functions that are described below ensure both continuity and smoothness at the endpoints of the observable segment of the RND function.

Figure 4
Implied RND functions for the three-month sterling interest rate in September 1995[(a)]



(a) Derived using LIFFE September 1995 options on the short sterling future, as at 4 May and 10 May 1995. These options expire on 20 September 1995. These graphs illustrate the result when using the Shimko technique. In this case the smile curve was interpolated using a cubic spline with two knot points. The tails, which are shown as dotted lines, are lognormal. Note that this extrapolation procedure does not always result in a smooth transition from the observable part of the distribution to the tails, as evidenced by the kink in the right tails.

Note that, as with the procedure for estimating implied histograms (see Section 3.1), nothing in the Shimko approach can prevent negative probabilities. Their occurrence depends on the interpolation procedure

[18] This would require continuity in (at least) the first derivative.

employed and its implications for the shape of the call pricing function. Some interpolated smile curves could conceivably result in non-asymptotic behaviour of the call pricing function, or in concavities in the option prices with respect to exercise price. For example, if the slope of the smile curve becomes too steep the volatility becomes so high as to imply that a more deeply OTM option has a higher price than those close to being ITM. This clearly implies the existence of arbitrage and would result in negative implied probabilities.

3.4 Fitting an assumed option pricing model to observed option prices

An approach which has been followed is to assume a particular stochastic process for the price of the underlying asset and to use observed option prices to recover the parameters of the assumed process. These, in turn, can be used to infer the RND function that is implied by the assumed stochastic process. Under sufficiently strong assumptions about the underlying price dynamics the RND function is obtainable in closed form. For example, in the Black-Scholes case, the assumption that the underlying price evolves according to GBM with a constant expected drift rate and constant volatility implies a lognormal RND function. Malz (1995b) assumes that exchange rates evolve according to a jump-diffusion process and uses *risk reversal* prices to recover the parameters of the model.[19] Since he uses the Bernoulli distribution version of the jump-diffusion model, which assumes that the jump size is non-stochastic and that there is either zero or one jump in the exchange rate over the life of the option, he is able to derive a closed-form solution for the terminal implied RND function.[20] Under these assumptions this turns out to be a mixture of two lognormal distributions.

---

[19] A risk reversal price, or *skewness premium*, is the difference between the price of a put option and that of a call option on the same underlying variable, with the same time to maturity and the same *delta* (the delta of an option is a metric for moneyness, that is, it provides a measure of the amount by which the option is away from the money). Under the Black-Scholes lognormality assumption the risk reversal price is zero; the probability of an OTM call being ATM at maturity is the same as that of an equally OTM put being ATM at maturity. In practice, positive risk reversal prices exist when market expectations are skewed relative to the lognormal distribution. Thus a risk reversal is a measure of the skew of an implied RND function.

[20] For further details about the Bernoulli distribution version of the jump-diffusion model see Ball and Torous (1983, 1985), and Bates (1988).

3.5 *Fitting an assumed parametric form for the implied RND function to observed option prices*

Rather than specifying the underlying asset price dynamics to infer the RND function, it is possible to make assumptions about the functional form of the RND function itself and to recover its parameters by minimising the distance between the observed option prices and those that are generated by the assumed parametric form.[21] As Melick and Thomas (1994) point out, starting with an assumption about the terminal RND function, rather than stochastic process by which the underlying price evolves, is a more general approach. This is because a given stochastic process implies a unique terminal distribution, but the converse is not true, that is, any given RND function is consistent with many different stochastic price processes.

The prices of European call and put options at time t can be written as the discounted sums of all expected future payoffs:

$$c(X,\tau) = e^{-r\tau}\int_X^{\infty} q(S_T)(S_T - X)dS_T \qquad \textbf{(10)}$$

$$p(X,\tau) = e^{-r\tau}\int_0^{X} q(S_T)(X - S_T)dS_T \qquad \textbf{(11)}$$

In theory any functional form for the density function, $q(S_T)$, can be used in equations **(10)** and **(11)**, and its parameters recovered by numerical optimisation. The problem with using (finite variance) models other than the Gaussian one is that the underlying price distribution changes as the holding period changes. In the Gaussian world we can say that if daily prices are lognormally distributed then other arbitrary length holding period price distributions must also be lognormal. No other finite variance distribution is similarly stable under addition. Under these circumstances, and given that observed financial asset price distributions are in the neighbourhood of the lognormal distribution, it seems economically plausible to employ the same framework suggested by Ritchey (1990) and to assume that $q(S_T)$ is the weighted sum of k-component lognormal density functions, that is,

[21] Rubinstein (1994) employs an optimisation method that solves for an RND function which is, in the least squares sense, closest to lognormal distribution that causes the present values of the underlying asset and all the options priced with it to fall between their respective bid and ask prices. Jackwerth and Rubinstein (1995) experiment with different distance criteria.

$$q(S_T) = \sum_{i=1}^{k} \left[\theta_i L(\alpha_i, \beta_i; S_T) \right] \qquad \textbf{(12)}$$

where $L(\alpha_i, \beta_i; S_T)$ is the i^{th} lognormal density function in the k-component mixture with parameters α_i and β_i;

$$\alpha_i = \ln S + \left(\mu_i - \tfrac{1}{2}\sigma_i^2 \right)\tau \text{ and } \beta_i = \sigma_i \sqrt{\tau} \text{ for each } i. \qquad \textbf{(13)}$$

(see equation **(7)** for the formula of the lognormal density function).

The probability weights, θ_i, satisfy the conditions

$$\sum_{i=1}^{k} \theta_i = 1, \ \theta_i > 0 \text{ for each } i. \qquad \textbf{(14)}$$

Moreover the functional form assumed for the RND function should be relatively flexible. In particular, it should be able to capture the main contributions to the smile curve, namely the skewness and the kurtosis of the underlying distribution. A weighted sum of independent lognormal density functions fits these criteria.[22] Each lognormal density function is completely defined by two parameters. The values of these parameters and the relative weighting applied to the two density functions together determine the overall shape of the mixture implied RND function.

Melick and Thomas (1994) apply this methodology to extract implied RND functions from the prices of American-style options on crude oil futures.[23] They assume that the terminal price distribution is a mixture of three independent lognormal distributions. However, given that, in many of the markets with which we are concerned, options are only traded across a relatively small range of exercise prices, there are limits to the number of distributional parameters that can be estimated from the data. Therefore, on grounds of numerical tractability, we prefer to use a two-lognormal mixture, which has only five parameters: α_1, β_1, α_2, β_2 and θ. Under this assumption the values of call and put options, given by equations **(10)** and **(11)**, can be expressed as follows:

[22] Note that this functional form implicitly ensures that the fitted call pricing function is monotonic decreasing and convex in exercise price, and is therefore consistent with the absence of arbitrage.

[23] To deal with the early exercise feature of the options that they examine, Melick and Thomas (1994) derive bounds on the option price in terms of the terminal RND function.

$$c(X,\tau) = e^{-r\tau} \int_X^{\infty} \left[\theta L(\alpha_1, \beta_1; S_T) + (1-\theta) L(\alpha_2, \beta_2; S_T)\right](S_T - X)\, dS_T \qquad \textbf{(15)}$$

$$p(X,\tau) = e^{-r\tau} \int_0^{X} \left[\theta L(\alpha_1, \beta_1; S_T) + (1-\theta) L(\alpha_2, \beta_2; S_T)\right](X - S_T)\, dS_T \qquad \textbf{(16)}$$

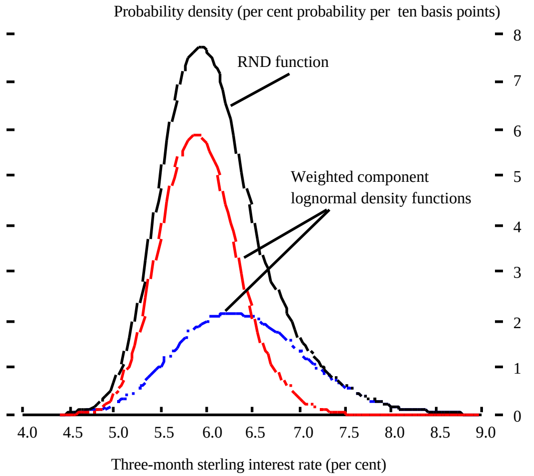
For fixed values of X and τ, and for a set of values for the five distributional parameters and r, equations **(15)** and **(16)** can be used to provide fitted values of $c(X,\tau)$ and $p(X,\tau)$ respectively. This calculation can be applied across all exercise prices to minimise the sum of squared errors, with respect to the five distributional parameters and r, between the option prices generated by the mixture distribution model and those actually observed in the market. In practice, since we can observe interest rates which closely approximate r, we use this information to fix r, and thereby reduce the dimensionality of the problem. Therefore, the minimisation is carried out with respect to the five distributional parameters only.

Since both calls and puts are priced off the same underlying distribution, we include both sets of prices in the minimisation problem. Also, in the absence of arbitrage opportunities, the mean of the implied RND function should equal the forward price of the underlying asset. In this sense we can treat the underlying asset as a zero-strike option and use the incremental information it provides by including its forward price as an additional observation in the minimisation procedure. The minimisation problem is:

$$\underset{\alpha_1,\alpha_2,\beta_1,\beta_2,\theta}{\mathrm{Min}} \sum_{i=1}^{n} \left[c(X_i,\tau) - \hat{c}_i\right]^2 + \sum_{i=1}^{n} \left[p(X_i,\tau) - \hat{p}_i\right]^2 + \left[\theta e^{\alpha_1 + \frac{1}{2}\beta_1^2} + (1-\theta) e^{\alpha_2 + \frac{1}{2}\beta_2^2} - e^{r\tau} S\right]^2 \qquad \textbf{(17)}$$

subject to $\beta_1,\beta_2>0$ and $0\leq\theta\leq1$, over the observed strike range $X_1,X_2,X_3,\ldots,X_n$. The first two exponential terms in the last bracket in equation **(17)** represent the means of the component lognormal RND functions. Their weighted sum therefore represents the mean of the mixture RND function. Figure 5 shows an example of an implied RND function derived using the two-lognormal mixture distribution approach. It also shows the (weighted) component lognormal density functions of the mixture RND function.

Figure 5
An implied RND function derived using the two-lognormal mixture distribution approach[(a)]



(a) Shown with its (weighted) component lognormal density functions. This RND function was derived using LIFFE December 1996 options on the short sterling future as at 10 June 1996. These options expire on 18 December 1996.

We would expect the five distributional parameters to vary over time as news changes and option prices adjust to incorporate changing beliefs about future events. The two-lognormal mixture can incorporate a wide variety of possible functional forms which, in turn, are able to accommodate a wide range of possible scenarios, including a situation in which the market has a bi-modal view about the terminal value of the underlying asset; for example, if participants are placing a high weight on an extreme move in the underlying price but are unsure of its direction.

It can be seen that although this mixture distribution methodology is similar in spirit to the approach taken by Bates (1991) and others in deriving the parameters of the underlying stochastic process, it focuses directly on possible future outcomes for the underlying asset price, thereby obviating the need to specify the underlying price dynamics.

It is important to remember that the implied density functions derived are risk neutral, that is, they are equivalent to the true market density functions only when investors are risk neutral. In reality investors are likely to be risk averse, and option prices will incorporate these preferences towards risk as well as beliefs about future outcomes. To distinguish between these two factors would require specification of the aggregate market utility function (which is unobservable) and estimation of the corresponding coefficient of risk aversion. However, even if the market does demand a premium for taking on risk, the true market implied density function may not differ very much from the RND function, at least for some markets.[24] Moreover, on the assumption that the market's aversion to risk is relatively stable over time, changes in the RND function from one day to the next should mainly reflect changes in investors' beliefs about future outcomes for the price of the underlying asset.

3.6 Application of the two-lognormal mixture approach to equity, interest rate and foreign exchange markets

We apply the two-lognormal mixture distribution approach outlined above to LIFFE equity index (European) options, short interest rate and long bond futures options, and to PHLX currency options. In this section, we begin by describing the mixture distribution model as applied to LIFFE equity index options. This is shown to be the weighted sum of two Black-Scholes solutions. We then describe some institutional features of LIFFE short interest rate and long bond futures options, and show how the general mixture distribution model used in the equity case can be modified to take account of these features. Lastly, we discuss some features of PHLX currency options.

In order to avoid the problems associated with asynchronous intra-day quotes we use exchange settlement prices.[25] Settlement prices are established at the end of each day and are used as the basis for overnight 'marking-to-market' of all open positions. Hence, they should give a fair reflection of the market at the close of business, at least for contracts on which there is open interest.[26]

LIFFE takes into account the following factors when calculating provisional settlement prices: **(i)** the final bid-offer spread at close; **(ii)** the final bid, or

[24] For example, Rubinstein (1994) converts an RND function for an equity index to a 'consensus subjective' density function under the assumption that the representative investor maximises his/her expected utility of wealth with constant relative risk aversion (CRRA). He finds that for assumed market risk premia of between 3.3 per cent and 5 per cent, the subjective distribution is only slightly shifted to the right relative to the risk-neutral distribution, and that the qualitative shapes of the two distributions are quite similar.
[25] LIFFE data are obtained directly from their database and PHLX data via *Reuters*.
[26] Settlement prices are only actual traded prices in those cases when the deal is done during the last few minutes of trading (known as the closing range). We have to accept that the ideal situation of being able to observe traded prices at all available strikes at exactly the same moment in time is not likely to occur very often.

offer, or trade at close; **(iii)** put-call parity; **(iv)** related options and futures; **(v)** intra day implied volatility and/or the previous day's implied volatility at settlement; **(vi)** advice obtained from option practitioners, floor/pit committee members and designated market-makers. Provisional settlement prices become final settlement prices approximately 30 minutes after the closing bell, subject to agreement with the London Clearing House (LCH). PHLX have a similar settlement procedure, with market makers giving closing quotes at each strike for which there is open interest by taking into account the last traded price and the movement in the underlying market since the last trade.

a) LIFFE equity index options

Consider applying equations **(15)** and **(16)** in Section 3.5, which give the values of call and put options under the assumption that the underlying asset is distributed as a mixture of two lognormal distributions, to LIFFE's FT-SE 100 index options. Although these are options on the equity index, they are normally hedged using the FT-SE 100 index future rather than a basket of stocks. They are therefore priced as though they are options on the index future; i.e. $S_T=F_T$, the terminal value of the 'implied' FT-SE 100 index future.[27] Evaluating equation **(15)** numerically results in compounded numerical errors due to the upper limit of infinity. Because of this and for computational ease, we prefer to optimise the objective function, given by equation **(17)**, using the following closed-form solutions to equations **(15)** and **(16)**:[28]

$$c(X,\tau) = e^{-r\tau}\left\{\theta\left[e^{\alpha_1+\frac{1}{2}\beta_1^2}N(d_1) - XN(d_2)\right] + (1-\theta)\left[e^{\alpha_2+\frac{1}{2}\beta_2^2}N(d_3) - XN(d_4)\right]\right\} \quad \textbf{(18)}$$

$$p(X,\tau) = e^{-r\tau}\left\{\theta\left[-e^{\alpha_1+\frac{1}{2}\beta_1^2}N(-d_1) + XN(-d_2)\right] + (1-\theta)\left[-e^{\alpha_2+\frac{1}{2}\beta_2^2}N(-d_3) + XN(-d_4)\right]\right\} \quad \textbf{(19)}$$

where

$$d_1 = \frac{-\ln X + \alpha_1 + \beta_1^2}{\beta_1}, \qquad d_2 = d_1 - \beta_1$$

$$d_3 = \frac{-\ln X + \alpha_2 + \beta_2^2}{\beta_2}, \qquad d_4 = d_3 - \beta_2$$

---

[27] The implied future value, which we use as the underlying asset price, *F*, is one that incorporates basis and dividend yield adjustments.

[28] The relevant single lognormal model is Black (1976). For the complete derivation of equations **(18)** and **(19)** see the Mathematical appendix.

This two-lognormal mixture model is the weighted sum of two Black-Scholes solutions, where θ is the weight parameter, and α_1, β_1 and α_2, β_2 are the parameters of each of the component lognormal RND functions. The *d* terms are the same as those in the Black-Scholes model, but have been reformulated here in terms of the relevant α and β parameters by applying the definitions given in equation **(6)**. Notice that the closed-form solutions involve the cumulative normal distribution function rather that the lognormal density function. This obviates the need for numerical integration since the cumulative normal distribution can be calculated to six decimal place accuracy using a polynomial approximation.[29]

Also note a subtle change in the objective function. As suggested in Section 3.5, the mean of the implied RND should equal the forward price of the underlying asset. In this case the underlying asset for pricing purposes is the equity index future, whose expected growth rate in a risk-neutral world is zero. Hence, the implied mean is the time-*t* implied futures price, that is, $e^{r\tau}S$ in equation **(17)** is replaced by *F*. We use equations **(18)** and **(19)** to estimate call and put prices and minimise the objective function to obtain estimates for the five distributional parameters, α_1, β_1, α_2, β_2 and θ.[30]

b) LIFFE options on short interest rate and long bond futures

Here we consider the following LIFFE contracts: long gilt, Bund, Euromark and short sterling futures options.[31] The long gilt, Bund and Euromark futures options are available for both quarterly and serial expiry dates. The short sterling futures option is only available for quarterly maturity dates. The exercise of a serial option gives rise to a futures contract of the associated quarterly delivery month; e.g. exercise of a February 1997 option gives rise to a March 1997 futures contract.

An institutional feature that is particular to LIFFE is that there are no carrying costs for their short interest rate and long bond futures options, that is, the buyer is not required to pay the premium up front. Instead the buyer has to deposit collateral, which remains his/her own property, and the option position is marked-to-market each day. This ensures that, by the maturity date of the option, the buyer pays the time value of the option to the seller. The seller, on the other hand, is compensated for not having the premium in hand by charging the buyer the compounded value of the usual premium. Since

[29] See Hull (1993), chapter 10.

[30] It sometimes aids optimisation if the α's and β's are expressed in terms of μ's and σ's, by applying the definitions of the lognormal parameters that are given in equation **(6)** , and the objective function is minimised with respect to $\mu_1, \mu_2, \sigma_1, \sigma_2$ and θ.

[31] LIFFE also have traded options on BTP, Eurolira and Euroswiss futures.

there is no opportunity cost to the buyer to holding an option, LIFFE's short interest rate and long bond futures options, which are all American style, are priced as European style options, that is, without the early exercise premium. This is because the option buyer can keep the position open at zero cost for as long as favourable movements in the underlying price generate positive cash flows into his/her margin account, whilst losses can be mitigated by closing out the position.

The two-lognormal mixture model for pricing LIFFE options on long bond futures is essentially the same as that given by equations **(18)** and **(19)**. The only difference is that, since there are no carrying costs for these options, the discount factor is omitted to give the time-*T* call and put pricing equations as:[32]

$$c(X,\tau) = \theta\left[e^{\alpha_1+\frac{1}{2}\beta_1^2}N(d_1) - XN(d_2)\right] + (1-\theta)\left[e^{\alpha_2+\frac{1}{2}\beta_2^2}N(d_3) - XN(d_4)\right] \quad \textbf{(20)}$$

$$p(X,\tau) = \theta\left[-e^{\alpha_1+\frac{1}{2}\beta_1^2}N(-d_1) + XN(-d_2)\right] + (1-\theta)\left[-e^{\alpha_2+\frac{1}{2}\beta_2^2}N(-d_3) + XN(-d_4)\right] \quad \textbf{(21)}$$

where d_1, d_2, d_3, and d_4 are the same as in equations **(18)** and **(19)**.

For LIFFE options on short interest rate futures further modifications are required. These take into account the fact that the instrument underlying a short-rate futures option is the interest rate that is implied by the futures price, given by one hundred minus the futures price, rather than the futures price itself. Therefore, a call (put) option on an interest rate futures price is equivalent to a put (call) option on the implied interest rate. The modified formulae are:

$$c(X,\tau) = \theta\left[-e^{\alpha_1+\frac{1}{2}\beta_1^2}N(-d_1) + (100-X)N(-d_2)\right] + (1-\theta)\left[-e^{\alpha_2+\frac{1}{2}\beta_2^2}N(-d_3) + (100-X)N(-d_4)\right] \quad \textbf{(22)}$$

$$p(X,\tau) = \theta\left[e^{\alpha_1+\frac{1}{2}\beta_1^2}N(d_1) - (100-X)N(d_2)\right] + (1-\theta)\left[e^{\alpha_2+\frac{1}{2}\beta_2^2}N(d_3) - (100-X)N(d_4)\right] \quad \textbf{(23)}$$

where

[32] An additional point to note is that LIFFE long bond futures prices are in fractions of 32 and the long bond futures option prices are in fractions of 64. All other LIFFE prices are quoted in decimal points.

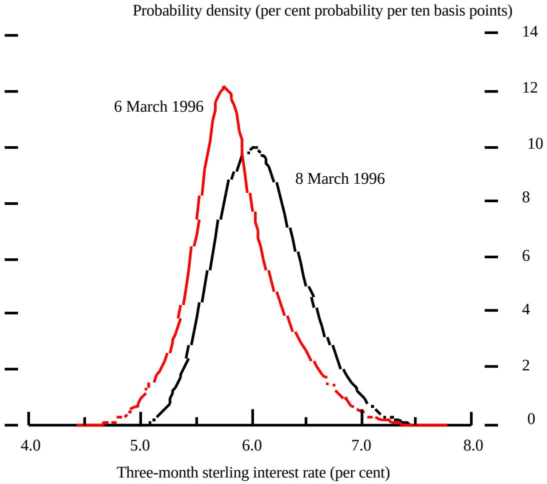
$$d_1 = \frac{-\ln(100 - X) + \alpha_1 + \beta_1^2}{\beta_1}, \qquad d_2 = d_1 - \beta_1$$

$$d_3 = \frac{-\ln(100 - X) + \alpha_2 + \beta_2^2}{\beta_2}, \qquad d_4 = d_3 - \beta_2$$

$$\alpha_i = \ln(100 - F) + \left(\mu_i - \tfrac{1}{2}\sigma_i^2\right)\tau \text{ and } \beta_i = \sigma_i\sqrt{\tau} \text{ for } i=1, 2. \qquad \textbf{(24)}$$

and the implied mean is the time-*t* implied interest rate, (100-*F*). Figure 6 illustrates the use of the two-lognormal mixture distribution approach with LIFFE options on the short sterling future. It shows how the implied RND for the implied three-month interest rate on 19 June 1996 changed between 6 March and 8 March 1996. Compare this figure with Figures 2 and 3 which show the risk-neutral histograms and the RND functions derived with the same data but using the simple (histogram) approach and the Shimko approach respectively.

Figure 6
Implied RND functions for the three-month sterling interest rate in June 1996(a)



(a) Derived using LIFFE June 1996 options on the short sterling future, as at 6 March and 8 March 1996. These options expire on 19 June 1996. The graphs illustrate the result when using the two-lognormal mixture distribution approach.

c) Philadelphia Stock Exchange (PHLX) currency options

Currency options are traded on the Philadelphia Stock Exchange between 2.30 a.m. and 2.30 p.m. Philadelphia time. We are primarily concerned with PHLX's European-style mid-month options on the following currency pairs: US dollar/British pound, Japanese yen/Deutsche Mark, Deutsche Mark/British pound, Deutsche Mark/US dollar and Japanese yen/US dollar.[33] We focus on mid-month rather than month-end options because they tend to have higher open interest and are available for a wider range of expiration dates: March, June, September, and December for up to nine months into the future, and the two near-term months.

PHLX option prices are for the purchase or sale of one unit of a foreign currency with the domestic currency. For example, one call option contract (where the contract size is £31,250) on the British pound with exercise price 155 cents would give the holder the right to purchase £31,250 for US $48,437.50. In the case of the US dollar-based options the domestic currency is US dollars. With the two cross-rate options it is Japanese yen and Deutsche Mark respectively.

Currency options are valued using the Garman-Kohlhagen (1983) currency option pricing model. When the implied RND is a two-lognormal mixture, the model takes the same form as the weighted Black-Scholes model given by equations **(18)** and **(19)**, with the discount rate being the domestic risk-free interest rate, r_d. The mean of the implied RND is the forward foreign exchange rate, or the time-t spot exchange rate, S, compounded over period τ by the differential between domestic and foreign interest rates, r_d-r_f. Note that both S and X are defined in terms of the value of one unit of the foreign currency in domestic currency units. We use the nearest domestic and foreign Eurorates for r_d and r_f.

[33] Examples of Deutsche Mark/US dollar and Japanese yen/US dollar implied RND functions derived using the two-lognormal mixture distribution approach with over-the-counter data can be found on page 101 of the BIS 66th Annual Report, June 1996.

4 Using the information contained in implied RND functions

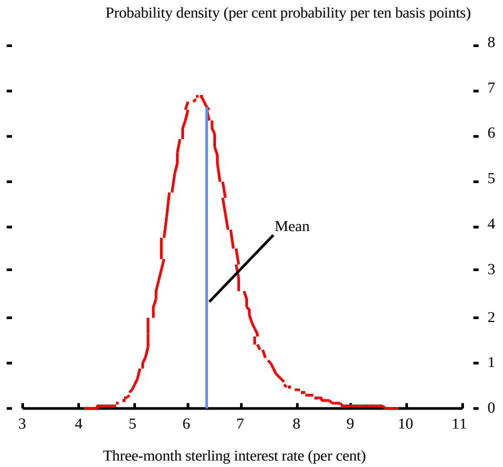
We now illustrate how the information contained in implied RND functions may be used in formulating and implementing monetary policy. We begin by describing various summary measures for density functions and then suggest a way to validate the two-lognormal mixture distribution approach. Next, we outline different ways in which implied RND functions may be used by the policy-maker. Finally, we discuss some caveats and limitations in data availability, and detail some areas for future research.

4.1 Summary statistics

Much of the information contained in RND functions can be captured through a range of summary statistics. For example, the mean is the expected future value of the underlying asset, or the average value of all possible future outcomes. Forward-looking information derived directly from futures prices and indirectly via bond yields is typically based on the mean. The median, which has 50% of the distribution on either side of it, is an alternative measure of the centre of a distribution. The mode, on the other hand, is the most likely future outcome. The standard deviation of an implied RND function is a measure of the uncertainty around the mean and is analogous to the implied volatility measure derived from options prices. An alternative dispersion statistic is the interquartile range (IQR). This gives the distance between the 25% quartile and the 75% quartile, that is, the central 50% of the distribution lies within it. Skewness characterises the distribution of probability either side of the mean. A positively skewed distribution is one for which there is less probability attached to outcomes higher than the mean than to outcomes below the mean. Kurtosis is a measure of how peaked a distribution is and/or the likelihood of extreme outcomes: the greater this likelihood, the fatter the tails of the distribution. These summary statistics provide a useful way of tracking the behaviour of RND functions over the life of a single contract and of making comparisons across contracts.

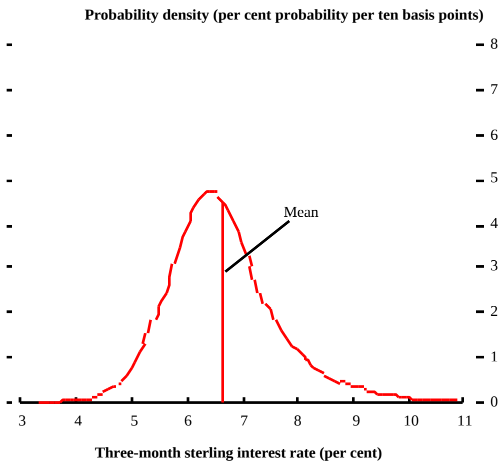
Figures 7 and 8 show the RND functions, as at 4 June 1996, for the three-month sterling interest rate in December 1996 and in March 1997. Figures 9 and 10 depict the RND functions, also as at 4 June 1996, for the three-month Deutsche Mark interest rate in the same months. Table B shows the summary statistics for these four distributions.

Figure 7
Implied RND function for the three-month sterling interest rate in December 1996[(a)]



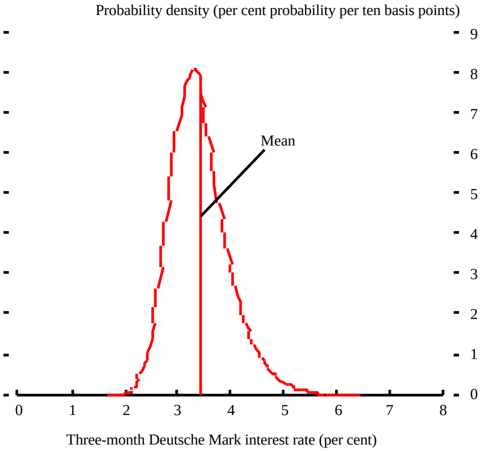
(a) Derived using LIFFE December 1996 options on the short sterling future, as at 4 June 1996. These options expire on 18 December 1996.

Figure 8
Implied RND function for the three-month sterling interest rate in March 1997[(a)]



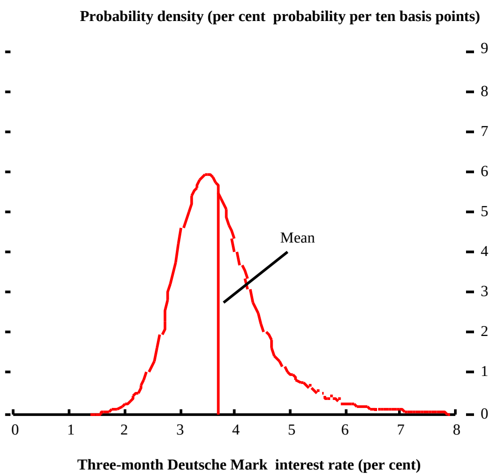
(a) Derived using LIFFE March 1997 options on the short sterling future, as at 4 June 1996. These options expire on 19 March 1997.

Figure 9
Implied RND function for the three-month Deutsche Mark interest rate in December 1996[(a)]



(a) Derived using LIFFE December 1996 options on the Euromark future, as at 4 June 1996. These options expire on 16 December 1996.

Figure 10
Implied RND function for the three-month Deutsche Mark interest rate in March 1997[(a)]



(a) Derived using LIFFE March 1997 options on the Euromark future, as at 4 June 1996. These options expire on 17 March 1997.

Table B
Summary statistics for the three-month sterling and Deutsche Mark interest rates in December 1996 and March 1997[(a)]

Sterling	December 1996	March 1997
Mean	6.33	6.66
Mode	6.18	6.43
Median	6.27	6.56
Standard deviation	0.66	1.01
Interquartile range	0.80	1.19
Skewness	0.83	0.76
Kurtosis[(b)]	4.96	4.67
Deutsche Mark		
Mean	3.45	3.73
Mode	3.29	3.47
Median	3.39	3.62
Standard deviation	0.55	0.84
Interquartile range	0.69	0.95
Skewness	0.75	1.16
Kurtosis	4.27	6.06

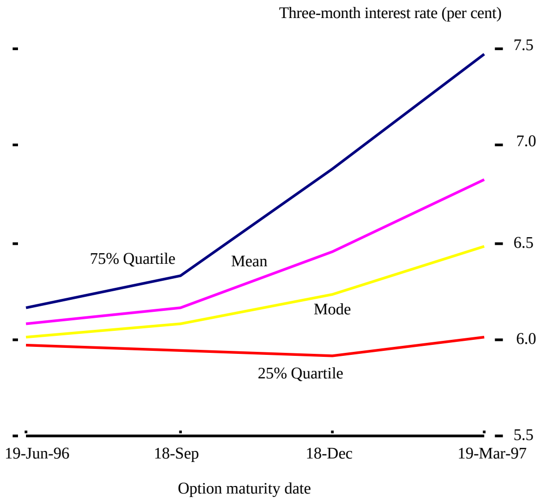
(a) Derived using LIFFE December 1996 and March 1997 options on the short sterling and Euromark futures, as at 4 June 1996.
(b) A normal distribution has a fixed kurtosis of three.

The means of the distributions are equivalent to the interest rates implied by the current prices of the relevant futures contracts, and are lower in Germany than in the United Kingdom.[34] For both countries, the dispersion statistics (standard deviation and IQR) are higher for the March 1997 contract than for the December 1996 contract. One would expect this since, over longer time horizons, there is more uncertainty about the expected outcome. Figure 11 confirms this, showing the upper and lower quartiles with the mean and the mode for the three-month sterling interest rate on four different option maturity dates as at 15 May 1996. It can be seen that the IQR is higher for contracts with longer maturities. Also, the standard deviations of the two distributions for the sterling rate are higher than the corresponding standard deviations of those for the Deutsche Mark rate, suggesting greater uncertainty about the level of future short-term rates in the United Kingdom than in Germany. Another feature of all four distributions is that they are positively skewed, indicating that there is less probability to the right of each of the means than to their left. The fact that the mode is to the left of the mean is

[34] The mean of an implied RND function should equal the forward value of the underlying asset. In this case the underlying assets are short-term interest rate futures contracts. The expected growth rate of a futures price in a risk-neutral world is zero. Hence, the means of the implied RND functions are equal to the interest rates implied by the respective current futures prices.

usually also indicative of a positive skew. This feature is discussed in greater detail below.

Figure 11
Implied RND summary statistics for the three-month sterling interest rate on four different option maturit dates(a)



(a) Derived using LIFFE options on the short sterling future, as at 15 May 1996.

4.2 Validation

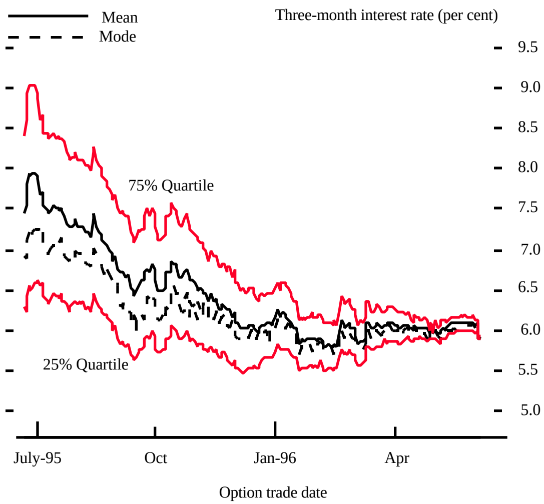
In deciding whether to place reliance on the information extracted using a new technique, one not only needs to be confident in the theory, but must also test whether in practice changes in the expectations depicted are believable in light of the news reaching the market. In the case of short-term interest rate expectations, we sought to do this by examining the way RND functions for short-term sterling interest rates change over time, and by comparing the RND functions for short-term sterling interest rates with those from Germany, a country with different macroeconomic conditions and monetary history.

4.2.1 Analysing changes in implied RND functions over time

Figures 12 and 13 show a convenient way of representing the evolution of implied RND functions over the life of a single option contract. Figure 12 shows the market's views of the three-month sterling interest rate on 19 June 1996 (as implied by the prices of LIFFE June short sterling futures options) between 22 June 1995 and 7 June 1996. Figure 13 shows the same type of information for the three-month Deutsche Mark interest rate on 17 June 1996 (as implied by the prices of LIFFE June Euromark futures options) between 20 June 1995 and 7 June 1996. Both figures depict the mean, mode, and the lower (25%) and upper (75%) quartiles of the distributions.

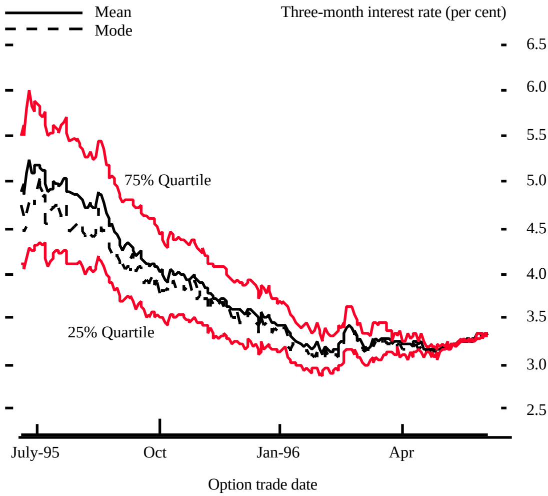
These time-series representations of implied RND functions convey how market uncertainty about the expected outcome changed over time; an increase in the distance between the lower and upper quartiles indicates that the market became more uncertain about the expected outcome. Figures 12 and 13 also convey information about changes in the skewness of the implied distributions. For example, the location of the mean relative to the lower an upper quartiles is informative of the direction and extent of the skew. Movements in the mean relative to the mode are also indicative of changes in skewness.

Figure 12
Implied RND summary statistics for the three-month sterling interest rate in June 1996[(a)]



(a) Derived using LIFFE June 1996 options on the short sterling future. These options expire on 19 June 1996.

Figure 13
Implied RND summary statistics for the three-month Deutsche Mark interest rate in June 1996(a)



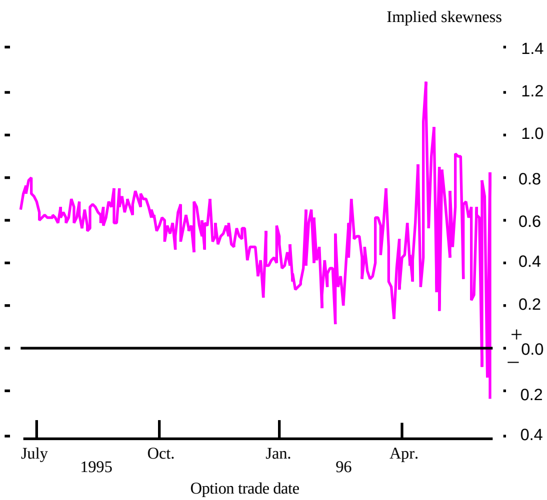
(a) Derived using LIFFE June 1996 options on the Euromark future. These options expire on 17 June 1996.

Generally, both sets of implied RND functions depict falling forward rates over the period analysed, as evidenced by the downward trend in the mean and mode statistics. At the same time, the gaps between these measures narrowed, suggesting that the distribution of market participants' expectations was becoming more symmetrical as the time horizon shortened. Figures 12 and 13 also show that as the maturity date of a contract is approached, the distributions typically become less dispersed causing the quartiles to converge upon the mean. This is because as the time horizon becomes shorter, the market, all other things being equal, becomes more certain about the terminal outcome due to the smaller likelihood of extreme events occurring. Another feature of the distributions is that the mode is persistently below the mean expectation in both countries, indicating a positive skew to expectations of future interest rates. In the United Kingdom this might be interpreted as reflecting political uncertainty, with the market attaching some probability to much higher short-term rates in the future. However, in Germany the macroeconomic and political conditions are different and yet the RND functions are also positively skewed.

One possible explanation is that the market perceives there to be a lower bound on nominal interest rates at zero. In this case, the range of possible outcomes below the current rate is restricted, whereas the range of possible outcomes above the current rate is, in principle, unlimited. If market participants are generally uncertain, that is, they attach positive probabilities to a wide range of possible outcomes, the lower bound may naturally result in the RND function having a positive skew. Moreover, the lower the current level of rates the more positive this skew may be for a given degree of uncertainty.

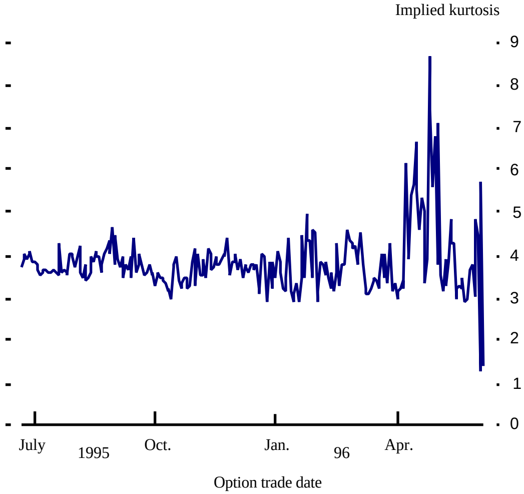
Figures 14 and 15 show how the skewness and kurtosis for the three-month sterling interest rate on 19 June 1996 changed between 22 June 1995 and 7 June 1996. It is notable that, unlike the measures of dispersion, these statistics exhibit no clear trend over their life cycles. Also, they appear to become more volatile towards the end of the contract's life.

Figure 14
Implied skewness for the three-month sterling interest rate in June 1996[(a)]



(a) Derived using LIFFE June 1996 options on the short sterling future. These options expire on 19 June 1996.

Figure 15
Implied kurtosis for the three-month sterling interest rate in June 1996(a)



(a) Derived using LIFFE June 1996 options on the short sterling future. These options expire on 19 June 1996.

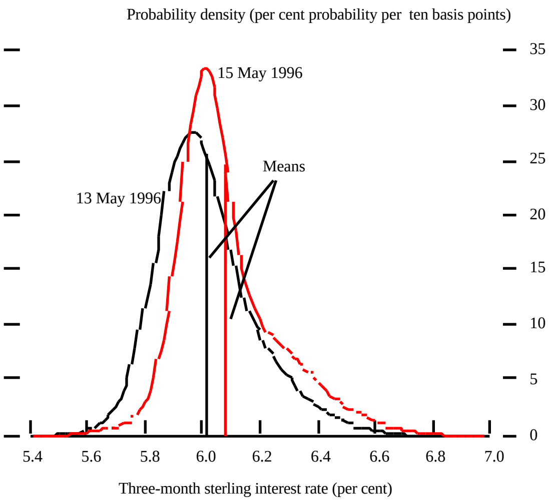
4.2.2 Analysing changes in implied RND functions around specific events

A detailed example of a change in perceptions following a particular news event is given in Figure 16 which shows the change in the shape of the implied RND function for the three-month sterling interest rate in June 1996 around the publication of the May 1996 *Inflation Report* on 14 May. The *Inflation Report* concluded that it was marginally more likely than not that inflation would be above 2.5% in two years' time were official rates to remain unchanged throughout that period. This was followed by an upward revision of the market's mean expectation for short-term interest rates between 13 May and 15 May. However, it seems that this upward move was not driven so much by a rightward *shift* in the distribution as by a change in the entire *shape* of the distribution; a reallocation of probability from outcomes between 5.6% and 5.9% to outcomes between 5.9% and 6.6% resulted in a fatter right tail which was in part responsible for the upward movement in the mean.[35]

[35] While the changes in the characteristics of the distributions are numerically distinct, they may not be statistically significantly different. Suitable tests could be designed to support the numerical results.

This type of change in the shape of implied RND functions is illustrative of how they can add value to existing measures of market expectations such as the mean.

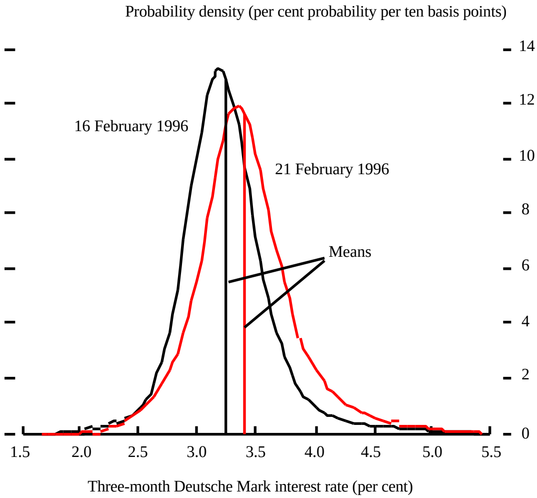
Figure 16
Change in the implied RND function for the three-month sterling interest rate in June 1996 aroun[d] the publication of the May 1996 *Inflation Report*[(a)]



(a) Derived using LIFFE June 1996 options on the short sterling future, as at 13 May and 15 May 1996. These options expire on 19 June 1996.

A similar change in market sentiment can be observed in Germany between 16 and 21 February 1996, ahead of the publication of the German M3 figure on 23 February. Figure 17 shows how the implied RND function for the three-month Deutsche Mark interest rate in June 1996 changed between these two dates. There was a significant shift in probability from outcomes between 2.5% and 3.3% to outcomes between 3.3% and 4.5%, apparently driven by market speculation ahead of the publication of the data. In particular, on 21 February the market attached a much higher probability to short-term rates being around 4% in June than it did on 16 February.

Figure 17
Change in the implied RND function for the three-month Deutsche Mark interest rate in June 1996[(a)]

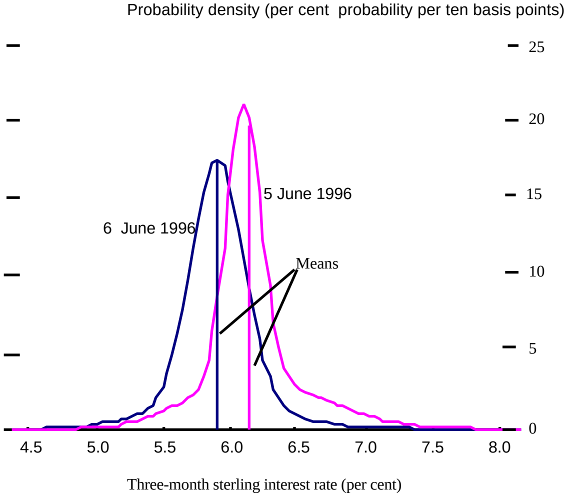


(a) Derived using LIFFE June 1996 options on the Euromark future, as at 16 February and 21 February 1996. These options expire on 17 June 1996.

The cut in UK official interest rates on 6 June 1996 provides an illustration of how market perceptions may change around a monetary policy decision. Figure 18 shows the change in the shape of the implied RND function for the three-month sterling interest rate in September 1996 between 5 and 6 June 1996. Table C shows the summary statistics for the RND functions on each of these dates.

The first point to note is that the mean moved down by 25 basis points which was the size of the interest rate cut. Secondly, the distribution on 6 June was more symmetrical (in fact the mean was almost equal to the mode) and had a higher standard deviation compared to the previous day; i.e. the market was more uncertain on 6 June than on 5 June about the short-term interest rate in September, and attached the same weight to it being above the mean as to it being below the mean. The change in the degree of skewness can also be seen by the shift in probability from outcomes between 6% and 7% to outcomes between 5.5% and 6%, resulting in a much thinner right tail and a left tail which was only slightly fatter. By comparison with other day-to-day movements, this particular change in the shape of the implied distribution

Figure 18
Change in the implied RND function for the three-month sterling interest rate in September 1996[(a)]



(a) Derived using LIFFE September 1996 options on the short sterling future, as at 5 June and 6 June 1996. These options expire on 18 September 1996.

Table C
Summary statistics for the RND functions in Figure 18

	5 June 1996	6 June 1996
Mean	6.16	5.91
Mode	6.11	5.91
Median	6.12	5.91
Mean minus mode	0.05	0.01
Standard deviation	0.30	0.35
Interquartile range	0.27	0.31
Lower quartile	6.00	5.76
Upper quartile	6.27	6.07
Skewness	0.85	0.22
Kurtosis[(a)]	6.61	7.02

(a) A normal distribution has a fixed kurtosis of three.

was quite large indicating the extent to which the market was surprised by the rate cut.

The above examples suggest that the two-lognormal mixture distribution approach is validated by recent market developments in the United Kingdom and in Germany. Although the mean expectation remains a key summary statistic, on the basis of these and other examples there is no reason to doubt that implied RND functions can add to our understanding of short-term interest rate expectations.

4.3 Use of implied RND functions by monetary authorities

We now discuss four ways in which the policy-maker may use implied RND functions.

4.3.1 Assessing monetary conditions

Assuming that financial market expectations are indicative of those in the economy as a whole, RND functions have the potential to improve the authorities' ability to assess monetary conditions on a day-to-day basis.

In principle, the whole probability distribution of future short-term interest rates is relevant to the determination of economic agents' behaviour. A lot of this information is captured in the mean of the distribution, which can already be observed directly from the yield curve or forward rates, but other summary statistics may add explanatory power. For example, suppose that agents tend to place less weight on extreme interest rate outcomes when taking investment or consumption decisions than is assumed in the mean of the interest rate probability distribution. In this case, a trimmed mean - in which the probabilities attached to extreme outcomes are ignored or given reduced weight - may reflect the information used by agents better than the standard mean, and so may provide a better indication of monetary conditions for the monetary authorities. Much of the time the standard mean and the trimmed mean may move together, but one could envisage circumstances in which the standard mean is influenced by an increase in the probabilities attached to very unlikely outcomes, while the trimmed mean is less affected. Similar issues would arise if investors or consumers placed *more* weight on extreme interest rate outcomes than allowed for in the standard mean.

At present, this kind of scenario is entirely speculative. Further empirical research is required to assess whether summary statistics such as an adjusted mean, the mode, median, interquartile range, skewness and kurtosis can add explanatory power to the standard mean interest rate in conventional economic models.

RND functions may also provide evidence of special situations influencing the formation of asset price expectations. For example, if two distinct economic or political scenarios meant that asset prices would take very different values according to which scenario occurred, then this might be revealed in bi-modal probability distributions for various asset prices.

4.3.2 Assessing monetary credibility

A monetary strategy to achieve a particular inflation target can be described as credible if the public believes that the government will carry out its plans. So, a relative measure of credibility is the difference between the market's perceived distribution of the future rate of inflation and that of the authorities.[36] Some information on this is already available in the United Kingdom in the form of implied forward inflation rates, calculated from the yields of index-linked and conventional gilts. However, this only gives us the mean of the market's probability distribution for future inflation. Even if this mean were the same as the authorities' target, this could mask a lack of credibility if the market placed higher weights on much lower and much higher inflation outcomes than the authorities.

Unfortunately, there are at present no instruments which enable the extraction of an RND function for inflation. Future research on implied probability distributions for long-term interest rates revealed by options on the long gilt future may, however, help in this respect, to the extent that most of the uncertainty over long-term interest rates - and hence news in the shape of a long gilt RND function - may plausibly be attributed to uncertainty over future inflation.

4.3.3 Assessing the timing and effectiveness of monetary operations

Implied RND functions from options on short-term interest rates indicate the probabilities the market attaches to various near-term monetary policy actions. These probabilities are in turn determined by market participants' expectations about news and their view of the authorities' reaction function.

In this context, implied RND summary statistics may help the authorities to assess the market's likely reaction to particular policy actions. For example, a decision to raise short-term interest rates may have a different impact on market perceptions of policy when the market appears to be very certain that rates will remain unchanged (as evidenced by a narrow and symmetric RND function for future interest rates) from when the mean of the probability distribution for future rates is the same, but the market already attaches non-trivial probabilities to sharply higher rates, albeit counterbalanced by higher probabilities attached to certain lower rates.

---

[36] For further explanation, see King (1995).

Equally, implied RND functions may help in the *ex post* analysis of policy actions. For example, if the shape and location of the implied RND function for short-term interest rates three months ahead remains the same following a change in base rates, this suggests, all other things being equal, that the market fully expected the change in monetary stance. By contrast a constant mean is less informative because it could disguise significant changes in skewness and kurtosis.

Implied probability distributions may also be useful for analysing market reactions to money market operations which do not involve a change in official rates, or events such as government bond auctions. These can be assessed either directly by looking at probability distributions from the markets concerned, or indirectly by looking at related markets.

4.3.4 Identifying market anomalies

All of the above uses of RND data assume that markets are perfectly competitive and that market participants are rational. However, provided one has overall confidence in the technique used, RND functions may help to identify occasional situations where one or other of these assumptions does not hold, essentially because the story being told is not believable.[37]

For example, in the face of an ‘abnormal’ asset price movement - such as a stock market crash or a sharp jump in the nominal exchange rate, which is not easily explained by news hitting the market - the information embedded in options prices for this and related assets may help the authorities to understand whether the movement in question is likely to be sustained with consequent macroeconomic effects, or whether it reflects a temporary phenomenon, possibly due to market failure. For example, if RND functions suggest that the market factored in the possibility of the very large asset price movement because it purchased insurance against the move in advance, then the amount of news required to trigger the change might reasonably be expected to be less than in the situation where there was no ‘advance knowledge’. This in turn might make it more believable that the move reflected fundamentals and hence would be sustained.

[37] The estimated distributions all fit well, as measured by the R^2 of fitted option prices computed from these distributions.

5 Data limitations

The most important limitation, from the point of view of monetary authorities, is that there are no markets that allow the authorities directly to assess uncertainty about future inflation. To learn about the market's future inflation distribution would require a market in options on inflation, for example, options on annual changes in the retail price index (RPI), or a market in options on real rates, as in index-linked bond futures. This, where necessary combined with information on nominal rates, would reveal what price agents were willing to pay to insure themselves against the risks to the inflation outturn, and hence the probabilities they attach to various future inflationary outcomes. Neither inflation options, nor options on index-linked bond futures are traded on exchanges anywhere in the world. However, such instruments could conceivably be available in the future.

Another limitation is that the two-lognormal mixture distribution approach is restricted to European options, whilst many of the more liquid exchange-traded options are often American.[38] This restriction is a feature of most of the existing techniques for deriving RND functions. Fairly complex extensions of these techniques are required to estimate terminal RND functions from the prices of American options.[39] Even then the RND function can only be derived within a bound that allows for the possibility that the options may be exercised at any time before the maturity date.

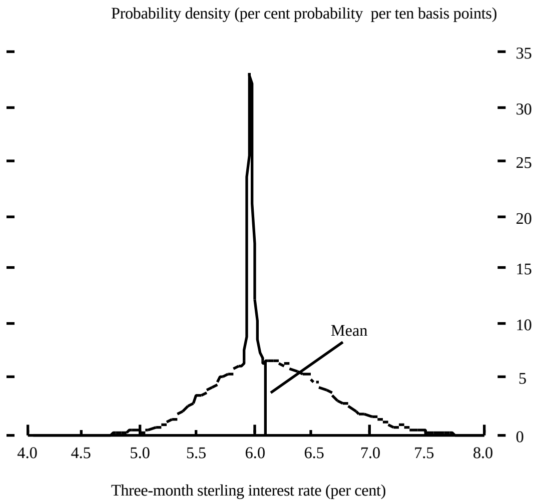
There are also limitations to the quality of the data that is available. Some option contracts are fairly illiquid, particularly at deep ITM and OTM strike prices. Options prices at these outer strikes may be less informative about market expectations, or may not be available. This data limitation sometimes results in sudden changes in the degree of convexity of the option pricing function. The two-lognormal mixture distribution approach (and other techniques) may in turn be sensitive to this. Figure 19 shows an example of the sort of (implausibly) spiked RND function that has on occasion resulted when there are relatively few data observations across strike prices.

---

[38] LIFFE options on interest rate futures, although American, can be treated as European. This is because they are margined daily, which means that the buyer is not required to pay the option premium up front. So, the buyer can keep the position open at zero cost for as long as favourable movements in the underlying price generate positive cash flows into his/her margin account, whilst losses can be mitigated by closing out the position. This means it is never optimal for the buyer to exercise such options early.

[39] See, for example, Melick and Thomas (1994).

Figure 19
Implied RND function for the three-month sterling interest rate in September 1996(a)



(a) Derived using LIFFE September 1996 options on the short sterling future, as at 8 May 1996. These options expire on 18 September 1996.

To derive implied RND functions we need options prices across the widest possible range of strike prices. To ensure that they are representative of the market's views, and that they can be estimated regularly, we use exchange-traded options contracts. However these have a limited number of fixed maturity dates, which is problematic when deriving time series of distributions and when assessing changes in market perceptions of short-term rates in the very near future. For example, if there are three months remaining until the nearest option maturity date, it is not possible to determine the market's perceptions of the short-term rate in one month's time. Also, because it is not possible with exchange-traded options to ensure that intra-day call and put prices are synchronous across exercise prices, only (end-of-day) settlement prices are usable in practice.

6 Conclusions

In this paper we have outlined the theory that relates options prices to risk-neutral probability density functions and have described five techniques for estimating such functions, discussing the relative merits and drawbacks of each technique. We have also shown how to apply our preferred approach - which assumes that the implied RND can be characterised by a weighted sum of two independent lognormal density functions - to estimate implied RND functions using LIFFE equity and interest rate options and Philadelphia Stock Exchange currency options. We have also discussed some of the institutional features of both exchanges that are relevant to the estimation routine employed.

We then went on to show how the information contained in implied RND functions can add to the type of forward-looking information available to policy-makers, particularly in assessing monetary conditions, monetary credibility, the timing and effectiveness of monetary operations, and in identifying anomalous market prices. To the extent that the distribution around the mean is observed to change in shape over time, measures such as the standard deviation, mode, interquartile range, skewness and kurtosis are useful in quantifying these changes in market perceptions. However, a good deal of further research, including event studies and the use of RND summary statistics in addition to the mean in classic economic models, is required to extract the maximum benefit from such information.

As a first step, it is important to be able to identify when a particular change in an implied probability distribution is significant by historical standards. One way of doing this is to establish suitable benchmarks. This would enable a large change in the shape of an RND function to be compared with changes in market perceptions at the time of a significant economic event in the past. In addition, RND functions could be estimated over the life cycles of many historical contracts for the same underlying asset in order to calculate average values for their summary statistics at particular points in the life cycle. These average values would identify the characteristics of a typical implied RND function during its life cycle. We plan to calculate this information for the implied RND functions of short-term sterling and Deutsche Mark interest rates. We are also in the process of implementing the technique discussed in this paper for options on long-term interest rate futures and for currency and equity options.

Mathematical appendix

This appendix shows how to derive the closed-form solutions to equations **(10)** and **(11)** under the assumption that the implied RND, $q(S_T)$, is a mixture of two lognormal distributions, $L(\alpha_1,\beta_1;S_T)$ and $L(\alpha_2,\beta_2;S_T)$, weighted by θ. Under this assumption, equations **(10)** and **(11)** can be written as follows:

$$c(X,\tau) = e^{-r\tau}\int_X^{\infty}\left[\theta L(\alpha_1,\beta_1;S_T) + (1-\theta)L(\alpha_2,\beta_2;S_T)\right](S_T - X)\,dS_T \qquad \textbf{(a1)}$$

$$p(X,\tau) = e^{-r\tau}\int_0^{X}\left[\theta L(\alpha_1,\beta_1;S_T) + (1-\theta)L(\alpha_2,\beta_2;S_T)\right](X - S_T)\,dS_T \qquad \textbf{(a2)}$$

We first derive the closed-form solution to equation **(a1)** and then use the put-call parity relationship to get the solution to equation **(a2)**.

Equation **(a1)** can be separated into two integrals:

$$\begin{aligned} c(X,\tau) &= e^{-r\tau}\int_X^{\infty} S_T\left[\theta L(\alpha_1,\beta_1;S_T) + (1-\theta)L(\alpha_2,\beta_2;S_T)\right]dS_T \\ &\quad - e^{-r\tau}X\int_X^{\infty}\left[\theta L(\alpha_1,\beta_1;S_T) + (1-\theta)L(\alpha_2,\beta_2;S_T)\right]dS_T \equiv A - B \end{aligned} \qquad \textbf{(a3)}$$

<u>Consider A</u>

Substituting in the formula for a lognormal density function, given by equation **(7)**, gives:

$$A = e^{-r\tau}\frac{1}{\sqrt{2\pi}}\int_X^{\infty}\left\{\frac{\theta}{\beta_1}e^{-[(\ln S_T-\alpha_1)^2]/2\beta_1^2} + \frac{(1-\theta)}{\beta_2}e^{-[(\ln S_T-\alpha_2)^2]/2\beta_2^2}\right\}dS_T \qquad \textbf{(a4)}$$

A change of variables allows a transformation from lognormal distributions to normal distributions. Substitute $w = \ln S_T$, where w is distributed as a mixture of two normal distributions; $w \sim \varphi(\alpha_1,\alpha_2,\beta_1,\beta_2;\theta)$, implies that,

$$S_T = e^{w} \qquad dS_T = e^{w}dw \qquad \textbf{(a5)}$$

$$A = e^{-r\tau} \frac{1}{\sqrt{2\pi}} \left\{ \frac{\theta}{\beta_1} \int_{\ln X}^{\infty} e^{w-[(w-\alpha_1)^2]/2\beta_1^2} dw + \frac{(1-\theta)}{\beta_2} \int_{\ln X}^{\infty} e^{w-[(w-\alpha_2)^2]/2\beta_2^2} dw \right\} \quad \textbf{(a6)}$$

The exponents of the exponential terms can be rewritten (by completing the square) as follows:

$$w - \frac{(w-\alpha_i)^2}{2\beta_i^2} = \frac{-(w^2 - 2\alpha_i w + \alpha_i^2 - 2\beta_i^2 w)}{2\beta_i^2}$$

$$= \frac{-[w^2 - 2w(\alpha_i + \beta_i^2) + \alpha_i^2 + 2\alpha_i\beta_i^2 + \beta_i^4] + 2\alpha_i\beta_i^2 + \beta_i^4}{2\beta_i^2} \quad \textbf{(a7)}$$

$$= \frac{-[w-(\alpha_i+\beta_i^2)]^2}{2\beta_i^2} + \alpha_i + \tfrac{1}{2}\beta_i^2$$

for $i = 1,\ 2$

$$A = e^{-r\tau} \frac{1}{\sqrt{2\pi}} \left\{ \frac{\theta}{\beta_1} e^{\alpha_1 + \frac{1}{2}\beta_1^2} \int_{\ln X}^{\infty} e^{-[w-(\alpha_1+\beta_1^2)]^2/2\beta_1^2} dw + \frac{(1-\theta)}{\beta_2} e^{\alpha_2 + \frac{1}{2}\beta_2^2} \int_{\ln X}^{\infty} e^{-[w-(\alpha_2+\beta_2^2)]^2/2\beta_2^2} dw \right\} \quad \textbf{(a8)}$$

A second change of variables allows a transformation from the normal distribution to the standard normal distribution. Substitute,

$$y_1 = [w - (\alpha_1 + \beta_1^2)]/\beta_1$$
$$y_2 = [w - (\alpha_2 + \beta_2^2)]/\beta_2 \quad \textbf{(a9)}$$
$$dw = \beta_1 dy_1 = \beta_2 dy_2$$

to get:

$$A = e^{-r\tau} \frac{1}{\sqrt{2\pi}} \left\{ \theta e^{\alpha_1 + \frac{1}{2}\beta_1^2} \int_{\ln X - (\alpha_1+\beta_1^2)]/\beta_1}^{\infty} e^{-\frac{1}{2}y_1^2} dy_1 + (1-\theta) e^{\alpha_2 + \frac{1}{2}\beta_2^2} \int_{\ln X - (\alpha_2+\beta_2^2)]/\beta_2}^{\infty} e^{-\frac{1}{2}y_2^2} dy_2 \right\} \quad \textbf{(a10)}$$

Writing this in terms of the cumulative normal distribution,

$$A = e^{-r\tau} \left\{ \theta e^{\alpha_1 + \frac{1}{2}\beta_1^2} \Pr\left(y_1 \le \frac{-\ln X + (\alpha_1+\beta_1^2)}{\beta_1} \right) + (1-\theta) e^{\alpha_2 + \frac{1}{2}\beta_2^2} \Pr\left(y_2 \le \frac{-\ln X + (\alpha_2+\beta_2^2)}{\beta_2} \right) \right\}$$

(a11)

$$A = e^{-r\tau} \left\{ \theta e^{\alpha_1 + \frac{1}{2}\beta_1^2} N\left(\frac{-\ln X + (\alpha_1+\beta_1^2)}{\beta_1} \right) + (1-\theta) e^{\alpha_2 + \frac{1}{2}\beta_2^2} N\left(\frac{-\ln X + (\alpha_2+\beta_2^2)}{\beta_2} \right) \right\} \quad \textbf{(a12)}$$

where $N(x)$ is the cumulative probability distribution function for a standardised normal variable, i.e. it is the probability that such a variable will be less than x.

<u>Consider B</u>

Substituting in the formula for a lognormal density function, given by equation **(7)**, gives:

$$B = e^{-r\tau} X \frac{1}{\sqrt{2\pi}} \int_X^{\infty} \left\{ \frac{\theta}{\beta_1} \frac{1}{S_T} e^{-[(\ln S_T - \alpha_1)^2]/2\beta_1^2} + \frac{(1-\theta)}{\beta_2} \frac{1}{S_T} e^{-[(\ln S_T - \alpha_2)^2]/2\beta_2^2} \right\} dS_T \qquad \textbf{(a13)}$$

Making the following substitution, to switch to the standard normal distribution,

$$\begin{aligned} y_1 &= (\ln S_T - \alpha_1)/\beta_1 \\ y_2 &= (\ln S_T - \alpha_2)/\beta_2 \\ dS_T &= S_T \beta_1 dy_1 = S_T \beta_2 dy_2 \end{aligned} \qquad \textbf{(a14)}$$

we get,

$$B = e^{-r\tau} X \frac{1}{\sqrt{2\pi}} \left\{ \theta \int_{(\ln X - \alpha_1)/\beta_1}^{\infty} e^{-\frac{1}{2}y_1^2} dy_1 + (1-\theta) \int_{(\ln X - \alpha_2)/\beta_2}^{\infty} e^{-\frac{1}{2}y_2^2} dy_2 \right\} \qquad \textbf{(a15)}$$

In terms of the cumulative normal distribution,

$$B = e^{-r\tau} X \left\{ \theta \Pr\left(y_1 \le \frac{-\ln X + \alpha_1}{\beta_1} \right) + (1-\theta) \Pr\left(y_2 \le \frac{-\ln X + \alpha_2}{\beta_2} \right) \right\} \qquad \textbf{(a16)}$$

$$B = e^{-r\tau} X \left\{ \theta N\left(\frac{-\ln X + \alpha_1}{\beta_1} \right) + (1-\theta) N\left(\frac{-\ln X + \alpha_2}{\beta_2} \right) \right\} \qquad \textbf{(a17)}$$

<u>Remember, $c(X,\tau) \equiv$ A-B</u>

$$\begin{aligned} c(X,\tau) = {} & e^{-r\tau}\theta \left[e^{\alpha_1 + \frac{1}{2}\beta_1^2} N\left(\frac{-\ln X + (\alpha_1 + \beta_1^2)}{\beta_1} \right) - XN\left(\frac{-\ln X + \alpha_1}{\beta_1} \right) \right] \\ & + e^{-r\tau}(1-\theta) \left[e^{\alpha_2 + \frac{1}{2}\beta_2^2} N\left(\frac{-\ln X + (\alpha_2 + \beta_2^2)}{\beta_2} \right) - XN\left(\frac{-\ln X + \alpha_2}{\beta_2} \right) \right] \end{aligned} \qquad \textbf{(a18)}$$

$$c(X,\tau) = e^{-r\tau}\left\{\theta\left[e^{\alpha_1+\frac{1}{2}\beta_1^2}N(d_1) - XN(d_2)\right] + (1-\theta)\left[e^{\alpha_2+\frac{1}{2}\beta_2^2}N(d_3) - XN(d_4)\right]\right\}$$

where

$$d_1 = \frac{-\ln X + \alpha_1 + \beta_1^2}{\beta_1} \qquad \textbf{(a19)}$$

$$d_2 = d_1 - \beta_1$$

$$d_3 = \frac{-\ln X + \alpha_2 + \beta_2^2}{\beta_2}$$

$$d_4 = d_3 - \beta_2$$

Equation **(a19)** is the closed form solution to equation **(a1)**.

The solution to equation **(a2)** can be obtained in a similar fashion, or by using the put-call parity relationship:

$$p(X,\tau) = c(X,\tau) + e^{-r\tau}X - S \qquad \textbf{(a20)}$$

Now, the forward price of the underlying asset, Se^{rt} , is equivalent to the mean of the implied RND, that is,

$$Se^{rt} = \theta e^{\alpha_1+\frac{1}{2}\beta_1^2} + (1-\theta)e^{\alpha_2+\frac{1}{2}\beta_2^2} \qquad \textbf{(a21)}$$

where the exponential terms represent the means of the component lognormal RND functions. Substituting for S in equation **(a20)** gives:

$$p(X,\tau) = e^{-r\tau}\left\{\theta\left[e^{\alpha_1+\frac{1}{2}\beta_1^2}N(d_1) - XN(d_2)\right] + (1-\theta)\left[e^{\alpha_2+\frac{1}{2}\beta_2^2}N(d_3) - XN(d_4)\right]\right\} \qquad \textbf{(a22)}$$
$$+ e^{-r\tau}X - e^{-r\tau}\theta e^{\alpha_1+\frac{1}{2}\beta_1^2} - e^{-r\tau}(1-\theta)e^{\alpha_2+\frac{1}{2}\beta_2^2}$$

Rearranging this expression gives the closed-form solution to equation **(a2)**:

$$p(X,\tau) = e^{-r\tau}\theta\left\{-e^{\alpha_1+\frac{1}{2}\beta_1^2}\left[1 - N(d_1)\right] + X\left[1 - N(d_2)\right]\right\}$$
$$+ e^{-r\tau}(1-\theta)\left\{-e^{\alpha_2+\frac{1}{2}\beta_2^2}\left[1 - N(d_3)\right] + X\left[1 - N(d_4)\right]\right\} \qquad \textbf{(a23)}$$

$$p(X,\tau) = e^{-r\tau}\left\{\theta\left[-e^{\alpha_1+\frac{1}{2}\beta_1^2}N(-d_1) + XN(-d_2)\right] + (1-\theta)\left[-e^{\alpha_2+\frac{1}{2}\beta_2^2}N(-d_3) + XN(-d_4)\right]\right\} \qquad \textbf{(a24)}$$

References

Arrow, K J (1964), 'The Role of Securities in the Optimal Allocation of Risk-Bearing', *Review of Economic Studies*, 31, No. 2 (April), pages 91-96.

Aït-Sahalia, Y and Lo, A (1995), 'Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices', *NBER, working paper No. 5351*.

Ball, C A and Torous, W N (1983), 'A Simplified Jump Process for Common Stock Returns', *Journal of Financial and Quantitative Analysis*, 18, pages 53-65.

Ball, C A and Torous, W N (1985), 'On Jumps in Common Stock Prices and Their Impact on Call Option Pricing', *Journal of Finance*, 40, pages 155-73.

Banz, R and Miller, M (1978), 'Prices for State-Contingent Claims: Some Estimates and Applications', *Journal of Business*, 51, pages 653-72.

Bates, D S (1988), 'The Crash Premium: Option Pricing Under Asymmetric Processes, With Applications to Options on Deutschemark Futures', *Rodney L. White Centre, University of Pennsylvania, working paper No. 36-88.*

Bates, D S (1991), 'The Crash of '87: Was it Expected? The Evidence from Options Markets', *Journal of Finance*, 46(3), pages 1009-44.

Bates, D S (1995), 'Post-'87 Crash Fears in S&P 500 Futures Options', *working paper, The Wharton School*.

BIS 66th Annual Report, 1 April 1995 - 31 March 1996, Basle, 10 June 1996.

Black, F (1975), 'Fact and Fantasy in the Use of Options', *Financial Analysts Journal*, pages 684-701.

Black, F (1976), 'The Pricing of Commodity Contracts', *Journal of Financial Economics*, 3, pages 167-79.

Black, F and Scholes, M (1973), 'The Pricing of Options and Corporate Liabilities', *Journal of Political Economy*, 81, pages 637-59.

Breeden, D T and Litzenberger, R H (1978), 'Prices of State-Contingent Claims Implicit in Option Prices', *Journal of Business*, Vol. 51, No. 4, pages 621-51.

Cox, J and Ross, S (1976), 'The Valuation of Options for Alternative Stochastic Processes', *Journal of Financial Economics*, 3, pages 145-66.

Debreu, G (1959), *Theory of Value*, Wiley, NY.

Fama, E (1965), 'The Behaviour of Stock Market Prices', *Journal of Business*, 38, pages 34-105.

Garman, M B and Kohlhagen, S W (1983), 'Foreign Currency Option Values', *Journal of International Money and Finance*, 2, pages 231-37.

Härdle, W (1991), *Smoothing Techniques With Implementation in S*, Springer-Verlag, NY.

Hull, J C (1993), *Options, Futures, and Other Derivative Securities*, 2nd ed., Prentice Hall International.

Jackwerth, J C and Rubinstein, M (1995), 'Implied Probability Distributions: Empirical Analysis', *Haas School of Business, University of California, working paper No. 250.*

Jarrow, R and Rudd, A (1982), 'Approximate Option Valuation for Arbitrary Stochastic Processes', *Journal of Financial Economics*, 10, pages 347-69.

King, M (1995), 'Credibility and monetary policy: theory and evidence', *Bank of England Quarterly Bulletin*, Vol 35, February, pages 84-91.

Longstaff, F (1992), 'An Empirical Examination of the Risk-Neutral Valuation Model', *working paper, College of Business, Ohio State University, and the Anderson Graduate School of Management, UCLA.*

Longstaff, F (1995), 'Option Pricing and the Martingale Restriction', *Review of Financial Studies*, Vol. 8, No. 4, pages 1091-1124.

MacBeth, J and Merville, L (1980), 'Tests of the Black-Scholes and Cox Call Option Valuation Models', *Journal of Finance*, 35, pages 285-303.

Malz, A (1995a), 'Recovering the Probability Distribution of Future Exchange Rates From Option Prices', *mimeo, Federal Reserve Bank of New York.*

Malz, A (1995b), 'Using Option Prices to Estimate Realignment Probabilities in the European Monetary System', *Federal Reserve Bank of New York Staff Reports*, No. 5.

Melick, W R and Thomas, C P (1994), ‘Recovering an Asset’s Implied PDF From Option Prices: An Application to Crude Oil During the Gulf Crisis’, *working paper, Federal Reserve Board, Washington*.

Neuhaus, H (1995), ‘The Information Content of Derivatives for Monetary Policy: Implied Volatilities and Probabilities’, *Deutsche Bundesbank Economic Research Group, discussion paper No. 3/95*.

Ritchey, R J (1990), ‘Call Option Valuation for Discrete Normal Mixtures’, *Journal of Financial Research,* Vol. XIII, No. 4, pages 285-96.

Ross, S (1976), ‘Options and Efficiency’, *Quarterly Journal of Economics*, Vol. 90, pages 75-89.

Rubinstein, M (1985), ‘Nonparametric Tests of Alternative Option Pricing Models Using all Reported Trades and Quotes on the 30 Most Active CBOE Option Classes From August 23, 1976 Through August 31, 1978, *Journal of Finance*, 40, pages 445-80.

Rubinstein, M (1994), ‘Implied Binomial Trees’, *Journal of Finance*, Vol. LXIX, No. 3, pages 771-818.

Shimko, D (1993), ‘Bounds of Probability’, *RISK*, Vol. 6, No. 4.

Whaley, R (1982), ‘Valuation of American Call Options on Dividend-Paying Stocks: Empirical Tests, *Journal of Financial Economics*, 10, pages 29-58.