DANMARKS NATIONALBANK

# DANMARKS NATIONALBANK WORKING PAPERS 2002 • 9

**Allan Bødskov Andersen**

Danmarks Nationalbank

**Tom Wagener**

Danmarks Nationalbank

# Extracting risk neutral probability densities by fitting implied volatility smiles: Some methodological points and an application to the 3M Euribor futures option prices

December 2002

ISSN (trykt/print) 1602-1185

ISSN (online) 1602-1193

# Extracting risk neutral probability densities by fitting implied volatility smiles: Some methodological points and an application to the 3M Euribor futures option prices[1]

Allan Bødskov Andersen[2]
Danmarks Nationalbank
Havnegade 5
DK-1093 Copenhagen K
aban@nationalbanken.dk

Tom Wagener
Danmarks Nationalbank
Havnegade 5
DK-1093 Copenhagen K
twa@nationalbanken.dk

19 December 2002

[1]The authors wish to thank seminar participants at Danmarks Nationalbank and the Aarhus School of Business, and participants in the workshop "The measures and determinants of financial market uncertainty" at the European Central Bank premises 27-28 May 2002, for helpful discussion and suggestions. Niels Christian Beier, Peter E. Storgaard, Nicki Sørensen, Christian Ølgaard, and an anonymous referee deserve special thank for comments to earlier drafts of the paper. The usual disclamer applies. The views in this paper are those of the authors, and do not necessarily reflect those of Danmarks Nationalbank.

[2]Corresponding author.

**Abstract**

Following Shimko (1993), a large amount of research has evolved around the problem of extracting risk neutral densities from options prices by interpolating the Black-Scholes implied volatility smile. Some of the methods recently proposed use variants of the cubic spline. These methods produce non-differentiable probability densities. We argue that this is an undesirable feature, and suggest circumventing the problem by fitting a smoothing spline of higher order polynomials. We apply this technique to the LIFFE three-month Euribor futures option prices. Summary statistics from constant horizon risk neutral densities are calculated and used to assess market uncertainty on a day-by-day basis. Finally, we analyse the impact of the 11 September attacks on the expectation of future Euribor interest rates.

**Resumé**

Med udgangspunkt i Shimko (1993) har der udviklet sig en litteratur omkring estimering af risikoneutrale tæthedsfunktioner ved at interpolere det implicitte Black-Scholes volatilitetssmil. Nogle af de senere bidrag benytter varianter af cubic splines til interpolationen. Denne metode leder til ikke-differentiable tæthedsfunktioner. Vi argumenterer for, at dette er en uheldig egenskab og foreslår at løse problemet ved at tilpasse en spline af højere ordens polynomier. Teknikken anvendes på priser på optioner på LIFFEs tremåneders Euribor future. Simple summariske mål fra risikoneutrale tætheder på konstant tidshorisont beregnes og anvendes til at vurdere markedsusikkerheden dag-til-dag. Endelig analyseres hvordan terror-angrebene 11. september påvirkede forventningen til den korte rente i Euro-området.

*Key words:* Implied volatilty, risk neutral density estimation, interest rate expectations

*JEL Classification:* C14, F33, G15

# 1 Introduction

In this paper we develop and apply a modified version of Shimko's (1993) Black-Scholes smile interpolation technique to extract risk neutral probability densities for the three-month Euribor interest rate from option prices on interest rate futures.

The forward interest rate expresses the market's average risk neutral expectation of the future short-term interest rate. The forward interest rate can be derived implicitly from the yield curve, or it can be computed directly from the prices of different interest rate derivatives. From option prices it is possible to extract the whole risk neutral distribution (henceforth RND) for the expected future interest rate. This allows a more detailed picture of the expectation. In particular, it allows an assessment of the uncertainty attached to any given level of the forward rate.

The market participants have an obvious interest in the RND, since it gives an expression of the market's average position taking. In principle, an investor can compare the RND with his/her own (subjective) probability assessment of the future market outcome, and act accordingly. Policy makers, especially central banks, have also shown great interest in RNDs, because they can be used to reveal a more detailed picture of the market expectation of future monetary policy decisions. For example, the central banks can extract information about the extend to which the market will be surprised by a decision to change the interest rate (or leave it unchanged, for that matter.) Also, while we believe that the use of RNDs is still in its infancy, future research might be able to reveal typical reaction patterns of the market, and central banks may potentially use such information as a part of the large information set underlying actual policy making.

A number of methods to extract RNDs from option prices have been derived, building on seminal work by Ross (1976), Cox and Ross (1976) and Breeden and Litzenberger (1978). In this paper we explore one of the more promising approaches, which consists of interpolating the implied Black-Scholes volatility smile. This approach was first presented by Shimko (1993).

Recently, new variants of the smile interpolation technique have emerged. They differ from the original Shimko (1993) formulation by allowing more flexible shapes of the fitted smiles, thereby allowing more flexible shapes of the RNDs, too. Here we take a critical view on one of the more popular recent methods, which consists of fitting the smile with a cubic spline. The cubic spline is indeed a formulation that allows great flexibility. However, it has important drawbacks in terms of the implied RND, since the RND in general will exhibit 'kinks' (i.e. non-differentiable points corresponding to the knot points of the cubic spline.) The possibility of kinks in the RND is often mentioned as being a problem with other methods, such as fitting double lognormal distributions or in applications of Shimko's (1993) original formulation (see below). The reason is twofold. First, from an economic point of view the kinks do not occur as an implication of neither theory nor data, but are imposed by the chosen method of deriving the RND. Second, a RND infected with kinks is not easy interpretable as an underlying distribution of the market expectation of the development in the price of the asset underlying the option contract.

This lead us to suggest a new variant of the smile interpolating technique, which consists of splicing higher order polynomials. This ensures flexibility of the fitted smiles while still producing smooth RNDs. We also suggest using a more appropriate measure of curvature, which in the estimation process is applied in order to ensure smooth fitted smiles.

The paper is organized as follows. Section 2 describes the common theoretical foundation behind all the techniques to extract RNDs. Then, section 3 gives a short and non-exhaustive survey of existing techniques, and in section 4 we develop a new variant of Shimko's (1993) implied Black-Scholes volatility smile interpolation technique. In section 5 two applications of the technique are described. First, we produce constant horizon RNDs and analyze market expectations of the short-term Euribor interest rate, which under normal circumstances follows ECBs minimum bid rate rather closely. Second, the impact of the 11 September attacks on the expectation of the three month Euribor interest rate is analyzed. Section 6 contains our concluding remarks.

## 2 The common foundation of the methods

The fundamental building block in modern finance theory is contingent claims. A contingent claim is an asset or security whose return depends on the state of nature at some future point in time. An option, i.e. the right but not the obligation to buy or sell an 'underlying' asset in the future, is such a contingent claim. The value of the option, and therefore its price, will typically depend on the price of the underlying asset at some future date.[1] The price of the option therefore contains information about the investors' probability assessment of the outcome of the underlying asset price at that future date. This is the basic idea behind extraction of RNDs from option prices.

A particular simple and important example of a contingent claim is the Arrow-Debreu asset, which pays one unit if one particular state occurs, and zero otherwise. The prices of Arrow-Debreu assets are called state-prices, and they are for each state directly proportional with the risk neutral probability for the occurrence of that state (see e.g. Cochrane, 2001.) In a continuum of states the state prices are proportional with the risk neutral density function.

Unfortunately, Arrow-Debreu assets are not traded on any organized market. However, Ross (1976) shows that a portfolio of call-options of the European type can be used to construct synthetic Arrow-Debreu assets, thereby establishing a relation between option prices and the RND. By ruling out arbitrage possibilities, Cox and Ross (1976) showed that options in general, independently of investors' risk preferences, can be priced as if investors were risk neutral. The only requirement to investors preferences is local non-satiation, i.e. that more is preferred to less. The price of a call option of the European type can then be computed as the discounted value of the option's

[1] We only consider options of the European type, i.e. options that can be exercised at the options' maturity date only.

expected return under risk neutrality,

$$c = e^{-rT}\widehat{E}\left[\max(S_T - X; 0)\right] = e^{-rT}\int_0^{\infty} \max(S_T - X; 0) f(S_T) dS_T, \qquad (1)$$

where $c$ is the option price, $S_T$ is the price of the underlying asset at time $T$, the expiry date of the option, $X$ is the strike price, $r$ is the risk-free rate of interest, and $\widehat{E}[\cdot]$ denotes the expectations operator with respect to the risk neutral density of the future asset price, $f(S_T)$.

Conditional on $f(S_T)$, equation (1) can be used to price the call option. However, the form of $f(S_T)$ is never known empirically. Breeden and Litzenberger (1978) noticed that from a functional expression for $c$, $f(S_T)$ can be obtained (up to a discount factor) by differentiating (1) twice with respect to the strike price:[2]

$$\frac{\partial^2 c(X)}{\partial X^2} = e^{-rT} f(S_T). \qquad (2)$$

This is the fundamental result we will explore below, where the strategy will be to obtain an empirical relevant expression for $c$ without assuming prior knowledge of the form of $f(S_T)$.

# 3 Extracting RNDs – methodologies

The literature on the extraction of RNDs is still somewhat unsettled, and no consensus has emerged as to which technique is best. Cooper (1999) and Jondeau and Rockinger (2000) have compared a number of methods, but were unable to draw unambiguous conclusions. Below is a non-exhaustive review of the most common methods.

The most simple approach is to compute histograms, which can never be more than a rough approximation of the underlying RND. This approach builds on the discrete time approximation of equation (2).

[2] Further, Harrison and Kreps (1979) showed that the RND is exactly the distribution under which the underlying asset follows a martingale. Therefore, the RND is sometimes referred to as the equivalent martingale measure. See Campell, Lo and MacKinlay (1997) for further discussion.

One possibility to compute the RND is to assume a specific parametric stochastic process of the price dynamics of the underlying asset. Such a process will imply a specific form of the density function of the price of the underlying asset at the termination date, i.e. imply the RND. Conditional on the parameters of the process an option pricing formula can be derived using equation (1). The parameters can then be determined by minimizing the squared distance between the observed option prices and the model-implied prices. Black-Scholes is the obvious example of this approach: The underlying asset price dynamics are assumed to evolve according to the simple diffusion process (the geometric Brownian motion),

$$dS = \mu S dt + \sigma S dW, \tag{3}$$

where $\mu, \sigma$ are the parameters to be estimated, and $dW \sim N(0, 1)$. This specification implies a lognormal RND,

$$\ln(S_T) \sim N(\ln(S) + \mu T - \frac{1}{2}\sigma^2 T, \sigma^2 T), \tag{4}$$

where $S$ is the current (known) price of the underlying asset. A slightly more complicated example is given in Malz (1996). He assumes a jump diffusion model for the underlying price dynamics and shows that the RND is a double lognormal, which can repair some of the well known empirical shortcomings of the Black-Scholes model.

The problem with assuming specific parametric processes for the underlying asset's price dynamics is that only in a few cases is the implied RND known in closed form. Hence, for purposes of extracting RNDs the analyst is rather constrained in the choice of price dynamics. This inherent inflexibility means that the RND may not be able to capture all relevant features of the market price data sufficiently well.

A closely related and slightly more general approach is to assume a parametric density function for the RND itself. Note that in order to price options from (1), the stochastic process of the underlying asset price is redundant in the sense that only the implied RND of this process is needed. Hence, it is possible to proceed as described above, except that the estimated parameters are not directly interpretable in the context of the underlying asset price dynamics. Often a mixture of two or three lognormals are used for the RND, but in principle far more complicated specifications could be used. Melick and Thomas (1997) use a double lognormal, and this method is also preferred by the Bank of England, see Bahra (1997). Its merits include first and foremost its simplicity. However, this comes at the cost of being rather inflexible and can give rise to RNDs with undesirable features like 'spikes' and 'kinks.'

A non-parametric approach is developed by Aït-Sahalia and Lo (1998) and Aït-Sahalia, Wang and Yared (2001). By heavy use of Kernel regressions they (almost) avoid imposing parametric restrictions on neither the stochastic process for the underlying nor the RND. However, their technique is computational very demanding, and it differs from the other methods by requiring time series price data, and therefore it essentially requires a stable stochastic process governing the evolution of the underlying asset price.

It is worth stressing that the presence of risk aversion generally causes the RND to be different from the "true" statistical probability of the movement of the underlying asset. By comparison of the RND and the corresponding estimated probability density based on historical returns it is possible to extract an implicit assessment of the underlying risk aversion. A complete treatment of this issue is beyond the scope of this paper; see Aït-Sahalia and Lo (2000) and Jackwerth (2000) for further discussion.

# 4 Fitting the implied Black-Scholes volatility smile

We will develop and apply a modified version of Shimko's (1993) technique, which is based on interpolation of the implied Black-Scholes volatility smile. Our basic motivation for using this approach is its flexibility, in the sense that no parametric assumptions as to neither the stochastic process for the underlying price dynamics nor the RND itself are necessary. The range of possible shapes of the RND is only limited by the flexibility of the functional form of the fitted implied volatility smile.

The starting point of the smile interpolation approach is equation (2), and the idea behind the technique is to obtain an empirically relevant expression for the call price function. From (2) it is obvious that in order to produce any RND the call price function needs to be twice differentiable with respect to the strike price. However, since the RND is a probability density function for the future outcome of a continuous variable (in the application below this variable is the short money market interest rate) it is desirable that the RND itself also is smooth (where smooth means differentiable.) As already mentioned, kinks in any given RND are hard to interpret, and since the kinks are not an implication of theory or empirical regularities we believe that these are good arguments for developing a technique that ensures smooth RNDs.

In order to obtain a differentiable RND we can observe from (2) that $\frac{\partial^2 c}{\partial X^2}$ needs to be differentiable, and in effect the call price function itself must be (at least) three times differentiable with respect to the strike price in order to produce a smooth RND. This seems trivial, but in the literature it is nevertheless an apparently overlooked aspect of RND extraction. We will come back to this important point below.

How can we obtain a sufficiently smooth empirical relevant expression for $c$? It cannot be observed directly from the market, since (exchange traded) options are traded with a fixed strike price interval. Thus interpolation is needed. One possibility would be to interpolate observed call prices directly, however, this has proved to be very difficult indeed, since even very small changes in the interpolation can give rise to very large changes in the second order partial derivative and thereby the RND. Shimko (1993) solved this problem rather ingeniously, namely by fitting the implied Black-Scholes volatility smile instead.

The Black-Scholes call option price formula for an interest rate future is given by (see e.g. Hull (1989))

$$c = e^{-rT}[F\Phi(d_1) - X\Phi(d_2)], \tag{5}$$

where

$$d_1 = \frac{\ln(\frac{F}{X}) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \qquad d_2 = d_1 - \sigma\sqrt{T}, \tag{6}$$

$\Phi(\cdot)$ is the cumulative normal distribution function, and $F$ is the price of the underlying asset, here the forward interest rate. Even though the smile interpolation technique builds on the Black-Scholes option pricing formula, it is not assumed that the Black-Scholes model holds true. The Black-Scholes model is merely used to convert a set of empirical option prices with given strike prices and time to maturity into a set of strike prices and corresponding volatilities implied by the Black-Scholes model. In other words, given an observed option price $c^{obs}$, the implicit volatility is the value of $\sigma$ which makes the Black-Scholes formula produce the observed market price.

It is clear from equation (5) that the Black-Scholes model assumes constant implicit volatility. However, it is by now widely accepted that market data shows higher implicit volatilities for options that are far in-the-money or far out-of-the-money. In a plot of the implicit volatilities against the strike price with the at-the-money strike roughly as a midpoint the relation describes a 'smile'. This is shown in figure 1, albeit here the implicit volatilities are plotted against the option's '$\delta$', which is the partial derivative of the call price function with respect to the forward rate,

$$\delta = \frac{\partial c}{\partial F} = \Phi(d_1). \tag{7}$$

The latter equation follows from (5), and using that

$$e^{-rT}[F\phi(d_1) - X\phi(d_2)] = e^{-rT}[F\phi(d_1) - F\phi(d_1)] = 0, \tag{8}$$

where $\phi(\cdot)$ denotes the normal probability density function. The full derivation of (8) is given in appendix A. We will return to the rationale for using $\delta$ below. To see why the implicit volatility function needs to be three times differentiable in order to produce smooth RNDs, we will for now stay with a formulation in which the implicit volatility function is dependent on the strike price (e.g. $\sigma = \sigma(X)$), since only in that case are closed-form formulae available. This is done purely for simplicity, as the arguments carry over to a formulation in $\delta$.

The first derivative of (5) is

$$\frac{\partial c}{\partial X} = -e^{-rT}\Phi(d_2) + \left(e^{-rT}X\sqrt{T}\phi(d_2)\right)\frac{\partial \sigma}{\partial X}, \tag{9}$$

from which the second derivative can be derived as

$$\frac{\partial^2 c}{\partial X^2} = e^{-rT}\phi(d_2)\times$$
$$\left[\frac{1}{\sigma X\sqrt{T}} + \left(\frac{2d_1}{\sigma}\right)\frac{\partial \sigma}{\partial X} + \left(\frac{d_1 d_2 X\sqrt{T}}{\sigma}\right)\left(\frac{\partial \sigma}{\partial X}\right)^2 + \left(X\sqrt{T}\right)\frac{\partial^2 \sigma}{\partial X^2}\right], \tag{10}$$

where (8) has been used again. First, note that in the Black-Scholes case with constant implicit volatility, all terms in the squared bracket in (10) except the first are zero. Hence, this confirms that in the Black-Scholes case the RND is lognormal. Secondly, it is immediately apparent that in order to produce a smooth RND the critical term is the fourth one in the square brackets. For this to be differentiable the implied volatility function $\sigma(X)$ needs to be three times differentiable and this should indeed be a guiding factor in the choice of the specification of $\sigma(X)$.

When the smile depends on $\delta$ instead of the strike prices $X$ there is no closed-form solution for the RND like (10). Instead the second order derivative of the call price function must be computed numerically.

## 4.1 Specifying the implied volatility function

We confine the discussion to possible parametric specifications of $\sigma(X)$, i.e. we assume $\sigma(X) = \sigma(X;\theta)$ where $\theta$ is a vector of parameters to be estimated.

In Shimko's (1993) formulation, $\sigma(X;\theta)$ was chosen to be a simple parabolic function. The second order derivative in this case is simply a constant and the RNDs will therefore in general be smooth. However, the parabolic function is not very flexible. Consequently, to ensure a reasonable fit within the range of available (and reliable) strike prices, another model has to be used for the tails of the RND, where in general no empirical observations are available. Note that it is common to all methods that the modelling of the tails of the RND outside the region where empirical observations are available is arbitrary. Shimko (1993) simply pasted lognormal tails on the RND, creating 'kinks' at the pasting points.

In order to improve the overall fit and flexibility of the fitted smile, Campa, Chang and Reider (1997) used the cubic spline for $\sigma(X;\theta)$. Similar in spirit, Clews, Panigirtzoglou and Proudman (2000) used a natural smoothing cubic spline. In cubic spline formulations, $\sigma(X;\theta)$ consists of piecewise cubic functions, which are spliced at the knot points under the condition that the spline is twice differentiable at those points. Increasing the number of knot points will improve the fit, but at the cost of increasing the dimension of $\theta$.

However, the cubic spline formulation does not comply with the requirement that the implied volatility function must be three times differentiable. The second derivative of a cubic spline is a continuous piecewise linear function, which is non-differentiable at the knot points and which will therefore produce kinks in the RND corresponding to the knot points of the spline. Applying various schemes which penalize curvature when undertaking the numerical estimation of $\theta$ will make the kinks less visible in a plot of the RND, but this cannot remove the problem. The RND will only be smooth if the penalty on curvature is given so much weight in the estimation procedure that it essentially makes the third derivative of the spline uniformly linear over the whole support of the RND. However, this is a rather uninteresting special case since it is equivalent to applying a single cubic function over the whole support for the RND. Consequently, the splining would not add anything in terms of flexibility, which was what the use of the cubic spline aimed for.

Here we suggest a method which circumvents the problem. The goal is to find a formulation which has a smooth second order derivative while preserving some flexibility in it. We also require a rather parsimonious model because typically the number of observations per smile is rather limited, cf. figure 1. These considerations lead us to suggest using a spline of fourth order polynomials with the condition that the spline is three times differentiable at the knot points. In this case the second order derivative is a differentiable spline of parabolic functions, a considerable more flexible function than in Shimko's (1993) formulation, in which the corresponding expression was a constant. As pointed out by Lancaster and Salkauskas (1986) splining higher order polynomials often results in 'overshooting', i.e. an exaggerated zigzag pattern of the function. This problem can be practically eliminated by reducing the number of knot points. In our empirical study below we found that just one knot point (corresponding to $\delta = 0.5$) was sufficient to ensure a reasonable fit. In comparison, the natural smoothing cubic spline has a knot point for every observation.

Our preferred specification with one knot point only will generally not be as flexible as the natural cubic spline. In particular, the possibility of obtaining a perfect fit is sometimes sacrificed with our specification, depending on the irregularity of the data. However, although a perfect fit might be desirable in other settings, such as fitting term structures of interest rates, we are often not interested in fitting implied volatilities perfectly. This is due to the fact that trade spreads, unequal timing of intra-day quotes, lack of market deepness etc. are likely to cause arbitrage violations in the price data, giving rise to e.g. negative probability densities if a perfect fit of the implied volatilities forms the basis of the RND. Thus, the goodness-of-fit of the implied volatility smile is generally not an appropriate metric to guide neither the functional form, nor the number of knot points. Of course, it would be nice if quoted option prices were more consistent with finance theory, but this seems too much to hope for in most empirical settings. Should the option price data comply with no-arbitrage conditions, our specification of the spline should be flexible enough to ensuring a fit very close perfect. Still, it is important to stress that improving the fit is not the main motivation for our proposed method, cf. the discussion above.

The trade-off between goodness-of-fit and curvature is controlled by a parameter $\lambda$, see below. The choice of $\lambda$ is very much similar to the choice of the bandwidth in applications of kernel smoothing. Here, various cross-validation schemes have been developed, but the specification choice often ultimately relies on aesthetics. This is also recognized in the literature, see e.g. Wand and Jones (1995), where an 'eye-balling' procedure for choosing the appropriate bandwidth is provided. This is essentially the procedure we have followed in the paper.

As a practical illustration of the consequence of expression (10), the smiles fitted by a smoothed cubic spline and a smoothed spline of fourth order polynomials (based on the same price observations) are shown in figure 1, and the corresponding RNDs are shown in figure 2 (section 4.2 below gives the estimation details.) The option prices are from 4 April 2001, a date at which the kinks in the RND with the cubic spline formulation are particular visible. The spline of fourth order polynomials tracks the cubic spline fairly close while preserving a smooth RND.

## 4.2 Estimation of the spline

The spline of two fourth order polynomials is a very parsimonious specification. There are ten parameters from the two polynomials and four restrictions in the knot point (equating the level and the first three derivatives of the polynomials at the knot point), leaving only six free parameters to be estimated.

Following Lancaster and Salkauskas (1986) we use the following representation of the spline:

$$\sigma(\delta;\theta) = \alpha_0 + \alpha_1\delta + \alpha_2\delta^2 + \alpha_3\delta^3 + \alpha_4\delta^4 + \sum_{i=1}^{s} \alpha_{4+i} \left|\delta - k_i\right|^4 , \tag{11}$$

where $s$ is the number of knot points ($=$ 1 in our implementation), $\theta$ $=$

$(\alpha_0, ..., \alpha_5)'$ is the parameter vector to be estimated, and $k_1$ is the knot point ($= 0.5$).

We estimate the function in the following way. First the observed option put prices are converted to call prices using the put-call parity. Together with the observed call prices the implied Black-Scholes volatility for each price is then computed numerically as the value of $\sigma$ that minimizes the squared distance between the observed price and the price implied by the Black-Scholes formula (5). For each estimated implied volatility the corresponding value of delta is computed, yielding a set of connected values of deltas and implied volatilities, $\{\delta_i, \sigma_i\}_{i=1}^{N}$, where $N$ is the number of observed prices. Now, the parameters of the spline are determined by numerically minimizing the squared distance between the 'observed' $\sigma_i$'s and the fitted values $\widehat{\sigma}(\delta_i)$ given by (11), with a penalty on curvature,

$$(\alpha_0, ..., \alpha_5) = \arg\min \sum_{i=1}^{N} [\sigma_i - \widehat{\sigma}(\delta_i)]^2 + \lambda \int_0^1 \kappa(\delta)^2 d\delta, \tag{12}$$

where

$$\kappa(\delta) = \frac{\left|\widehat{\sigma}''(\delta_i)\right|}{\left[1 + \widehat{\sigma}'(\delta_i)^2\right]^{\frac{3}{2}}}. \tag{13}$$

The value of $\lambda$ governs the relative weight attributed to the penalty on curvature of the fitted smile. Our choice for $\kappa(\cdot)$ (which is to the best of our knowledge new to the literature) requires an explanation, as it is different from the more common measure of curvature, which is simply the numerical value of the second order derivative of the spline. However, measuring curvature by the simple second order derivative may be inappropriate in the context of smile fitting, since the segments of the smile potentially can become rather steep. By the second order derivative, curvature at steep segments of the smile is penalized too much relative to the measure in (13), which by the denominator takes into account the steepness of the curve. For example, consider the function $f(x) = x^2$. The second order derivative of $f$ is 2 for all $x$, although visually the curvature is much larger near the vertex $f(0)$, than in, say, $f(10)$ where indeed the curve seems close to being linear.

In this example, the measure (13) is still 2 at $f(0)$, but is only 0.0007 at $f(10)$.

The interpretation of (13) is that it measures the inverse radius of the circle curve that is tangent to the function, and that has a second order derivative of the circle line that is equal to the second order derivative of the function in the tangent point. Thus, this measure is 'visually consistent', or in other words invariant to rotation. Measured by the second order derivative alone, the smile will always tend to curve relatively more near the bottom of the smile, which is an arbitrary restriction to impose on the shape. An example of the difference between the two measures of curvature is given in figure 3. The difference is apparent in the right part of the smile, where the simple second order derivative measure of curvature makes the smile less flexible (for a given value of $\lambda$), although for this particular date the difference is not great. Note that the particular scaling of the axes in the plot makes the slope at the steep segments of the smile seem higher than it actually is. The slopes are actually quite small, and hence explain why the difference between the two measures of curvature is minor. Still, from a conceptual point of view we are more comfortable with the consistent treatment of the curvature. The associated implied RNDs for the two smiles are shown in figure 4. Given the limited difference of the smiles in figure 3 it comes at no surprise that the corresponding RNDs are relatively similar, too.

# 5 Empirical applications

In this section we present two applications of the technique developed above. First, following Clews et al. (2000) we use the technique to extract market expectations of the future three-month Euribor interest rate (henceforth 3M Euribor) on a constant horizon basis. We extract the RNDs from the prices of the options on the 3M Euribor future. These options are traded on London International Futures and Options Exchange (LIFFE.) Under normal circumstances the 3M Euribor rate follows the ECB's minimum bid rate rather closely. Hence, the options on the 3M Euribor can be used to extract the market's expectations about ECB's future monetary policy decisions. Extracting RNDs at a constant horizon is non-trivial since exchange traded options are contracts with fixed expiry date.

The second application is a more classic use of implied RNDs, namely an analysis of a specific event's impact on market expectations. In particular, we compare RNDs from right before and immediately after the acts of terror in the US on 11 September.

## 5.1 Application 1: Constant horizon RNDs

Exchange traded options are based on contracts with fixed expiry date. The LIFFE traded options on the 3M Euribor interest rate futures expire the third Wednesday in each of the months March, June, September and December, corresponding to the expiry dates of the underlying futures contracts. A problem with the fixed expiry dates is that the RNDs will narrow in on the forward rate the closer they are to the expiry date, simply because the market becomes more certain about the value of the underlying asset at the expiry date the closer in time we are to that date. This hampers the interpretation of changes in the RNDs as time passes by: How much can be attributed to the effect of new information to the market, and how much is merely an effect of less time to maturity? The constant horizon RND is an attempt to isolate the former effect from the latter.

In a day-to-day comparison of the risk neutral expectations, the fixed expiry date effect is a minor problem if there is long time to maturity. Whether an option has (say) four months to maturity, or four months minus one day to maturity is effectively inconsequential for the prices of the option, and hence the fixed expiry date effect can safely be ignored. Consequently, any changes in the RNDs between two consecutive days can be attributed to new information.[3] The closer in time the RNDs are, and the longer the time to maturity of the option, the smaller the effect from the fixed expiry date.

[3] Ignoring effects from liquidity changes etc.

However, we often wish to track changes in market expectations over a longer period than just a few days, and in such cases the expiry date effect cannot be ignored.

Figure 5 shows the fitted implied volatility smiles for prices from 1 March 2000 on option contracts expiring in June, September, December 2000, and March, June, and September 2001. Together these smiles make an implied volatility surface. As expected, the surface increases over time due to the simple fact that the market is more uncertain of the future short interest rate the further away in time it has to guess.

Clews et al. (2000) suggest the following technique to extract constant horizon RNDs. The idea is to interpolate between two implied volatility smiles, these being estimated with prices from the same date, but from options with different expiry dates. The principle can be seen from figure 6. With a constant horizon of, say, six months, we can construct a synthetic smile with six months to maturity by interpolating the smiles from contracts with expiry dates before and after the six months' horizon. As an example, in late April a synthetic smile at six months' sight can be estimated by interpolating the smiles from the actual expiry dates of the September contract (five months to maturity) and the December contract (eight months to maturity), weighting the smiles according to their closeness to the six months' horizon. Of course, constructing a constant horizon RND on 19 March (say) is simple because no further interpolation is needed as the September contract itself has exactly six months to maturity. The synthetic smiles can then be inserted in (10) and the constant horizon RND is computed as described above.

The smile interpolation over time is also the basic motivation for estimating the smiles over $\delta$'s instead of strikes. $\delta$ is confined to the interval between 0 and 1, while the strike price interval may change considerably over time. The time interpolation is done for constant $\delta$, and during the calculation of the the RND (by numerical differentiation) the $\delta$'s are inverted back to strike prices.

Although it is well known that implied volatility term structures can exhibit non-linearities (which is also apparent in figure 5), we nevertheless apply a simple linear interpolation in the time domain. However, as we consider a three months' period only, we have just two points in time (i.e. the maturity dates) available for interpolation, which leaves little room for anything but a linear specification.

Of course, it would be possible to estimate a more fancy implied volatility term structure by considering a greater number of contracts with maturities beyond the horizon of direct interest, but this would many-double the computational requirement, because 2,500 term structures needs to be re-estimated for each day in the sample, amounting to a total of about 150,000 term structures. Thus, the gain from taking possible non-linearities in the term structure of implied volatilities into account was judged marginal compared to the heavy increase in computational costs.

Changes over time in the constant horizon RNDs can to a larger extend be attributed to changes in market expectations. However, another type of expiry date effect is induced by this technique. Note that right before a contract expires the interpolation over time is done between contracts with (a little more than) three and six months to maturity. On the other hand, right after a contract expires, the corresponding interpolation uses contracts with six and nine months to maturity. Thus, in the latter case the constant horizon RND is affected by expectations of up to nine months' sight, while in the former case it is affected by expectations of only up to six months' sight. As a consequence, the constant horizon RNDs can be unstable around maturity dates. Clews et al. (2000) test for the relevance of this effect for the short Sterling and find it significant, but numerically small. In our application we do not consider periods that cover expiry dates, why we can disregard the problem here.

### 5.1.1 Data

The option prices are drawn from Bloomberg. We use both put and call prices. The put prices are converted to call prices using the put-call parity, see e.g. Hull (1989). One obvious source of error is that the market is very thin for options that are far from at-the-money, why LIFFE's prices far from always reflect any market price. In fact, often a quoted price is simply identical to the option's internal value.

Options that are out-of-the-money or slightly in-the-money are the most liquid. Thus the gross sample consisted of all out-of-the-money contracts as well as those in-the-money contracts where the strikes were at most 0.75 percentage points from the forward rate. Outliers were removed from the gross sample, and illiquid strike prices were identified by the option's 'vega', i.e. the first derivative of the call price function with respect to $\sigma$. Observations for which Vega was below 0.025 were deleted.

### 5.1.2 Results

The constant horizon (six months' sight) RNDs are calculated day by day for a three months' period from 20 March to 21 June 2001. The RNDs are shown in figure 7. The solid curve in the zero-plane is the forward interest rate, which is the expected value of the RND by construction. The dashed line indicates 10 May 2001, the date of the ECB's interest rate reduction of 25 bp.

The resulting picture is an expectation 'mountain chain', which gives an impression of the day-to-day variation in the implied RNDs. The same picture from a bird's view (figure 8) might give a better idea of the shifting RNDs.

Statements about the certainty of market expectations are made possible by the derivation of RNDs. Normally, the market expectations are summarized by the forward rate, but figures 7 and 8 illustrate that this forward rate average can cover relatively large changes in the underlying distribution.

From an operational point of view, figures like 7 and 8 are hardly usable, at least not over longer periods. Instead, one useful approach is to summarize central features of the RNDs by plotting some key statistics derived from the RNDs. Here, the modal value (i.e. the interest rate with the highest probability), the forward rate and a measure of the underlying uncertainty of the market expectations are shown in figure 9.

When the modal is greater than the forward the underlying RND is skewed to the left, and this indicates that the market assigns relatively more probability mass to lower interest rates at six months' sight than to higher interest rates. The figure also shows a measure of the confidence level of the expectations, here calculated as the 75% quantile less the 25% quantile as a fraction of the forward rate.[4]

The dates in the figure correspond to the meeting dates of ECB's Governing Council. Uncertainty shows a degree of persistence, which is reassuring. There are some sudden changes in the confidence level, which could indicate arrival of new information.

Two incidents deserve a closer look, 11 April and 10 May. On 11 April the ECB left the interest rate unchanged. This gave an immediate upward jump in the modal and the forward rate, suggesting that the market expected a slash. The uncertainty was at a moderate level before the meeting, and took an increasing course the following ten days. Together, these findings are consistent with the interpretation that the market was surprised by the interest rate decision.

On 10 May the ECB cut its interest rate with 25 bp. In the preceding week the forward rate dropped a little, but it did not reach the low level that was seen before 11 April. The forward rate dropped considerably in response to the interest rate cut. Thus, judged from the forward rate it seems that the market was taken by surprise again. However, the modal indicates that the RNDs were skewed to the left in the days before the meeting, which

[4] This measure was chosen because the price data always cover the mid 50% of the distribution. The more standard 95% confidence level would here rely on an (always) arbitrary form of the RND in the tails.

suggests that the market did assign more probability to lower interest rate at six months' sight. Also, uncertainty rose markedly in the days before the meeting. This might indicate that the interest rate cut was to some extent foreseen by the market. However, the modal is rather volatile, why we should be cautious about hard interpretations of the changes in this statistic.

These examples illustrate how RNDs give a more detailed picture of the underlying market expectations than the simple forward rate average.

### 5.2 Application 2: The market reaction to the 11 September attacks

In this application the expiry date problem can be ignored, since the analysis consists of comparing RNDs just two days apart.

We analyze the effect on the market expectation of the 3M Euribor from the terrorist attacks on 11 September. The fitted smiles from option prices of the December 2001 contract from 10 September and 12 September are shown in figure 10. The attacks had a very significant impact. The implicit volatilities rose from a level of 15-16% on 10 September to 20-23% on 12 September. The middle part of the smile rose relatively more than the tails. The associated RNDs are shown in figure 11. The larger uncertainty as indicated by the lift of the smile cause a widening RND. At the same time, the RND became considerable more skewed to the left, so that more probability mass was assigned to lower three-month interest rates. The attacks made the market much more uncertain as to how far the interest rate would drop before the options' expiry date in December.

## 6 Concluding remarks

We derived a simple and easy-to-implement variant of Shimko's (1993) implied Black-Scholes volatility smile approach to extracting RNDs. The suggested technique ensures smooth and flexible RNDs in a yet very parsimonious formulation. We also suggested using a slightly more complicated measure of curvature when penalizing curvature in the estimation process.

The constant horizon RND approach seems more promising from an operational point of view than the more common fan-charts, which do not deal with the fixed expiry date problem. The RNDs give a far more detailed picture of the underlying market expectation than the simple forward rate average.

An important topic for future research is comparison of the different techniques to extract RNDs, and not least what criterion should be used. We chose the Shimko (1993) approach due to its flexibility, however, this also means that it is extremely data dependent. And the data from the options market is often not of the quality one could hope for. This might call for relying more on parametric assumptions about e.g. underlying price dynamics or (other) restrictions derived from economic theory.

## References


[1] Aït-Sahalia, Yacine and Andrew W. Lo (1998), "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices", The Journal of Finance vol LIII, no. 2, pp. 49 – 547.

[2] Aït-Sahalia, Yacine and Andrew W. Lo (2000), "Non-parametric Risk Man-agement and Implied Risk Aversion", Journal of Econometrics vol 94, pp. 9 – 51.

[3] Aït-Sahalia, Yacine, Yubo Wang and Francis Yared (2001), "Do Option Markets Correctly Price the Probabilities of Movements of the Underlying Asset?", Journal of Econometrics 102, pp. 67 – 110.

[4] Arrow, Kenneth (1964), "The Role of Securities in the Optimal Allocation of Risk Bearing", Review of Economic Studies 31, pp. 91 – 96.

[5] Bahra, Bhupinder (1997), "Implied Risk-neutral Probability Density Functions From Option Prices: Theory and Applications", Bank of England Working Paper Series no 66.

[6] Breeden, Douglas and Robert H. Litzenberger (1978), "Prices of State-contingent Claims Implicit in Option Prices", Journal of Business 51, pp. 621 – 651.

[7] Campa, José M., P.H. Kevin Chang and Robert L. Reider (1997), "Implied Exchange Rate Distributions: Evidence From the OTC Option Markets", NBER Working Paper no. 6179.

[8] Campell, John Y., Andrew W. Lo and A. Craig MacKinlay (1997), "The Econometrics of Financial Markets", Princeton University Press, Princeton, New Jersey.

[9] Clews, Roger, Nikolaos Panigirtzoglou and James Proudman (2000), "Recent Developments in Extracting Information From Option Markets", Bank of England Quarterly Bulletin, February, pp. 50 – 60.

[10] Cochrane, John H. (2001), "Asset Pricing", Princeton University Press, Princeton, New Jersey.

[11] Cooper, Neil (1999), "Testing Techniques for Estimating Implied RNDs From the Prices of European-style Options", Bank of England Working Paper.

[12] Cox, John, and Stephen Ross (1976), "The Valuation of Options for Alter-native Stochastic Processes, Journal of Financial Economics 3, pp. 145 – 166.

[13] Debreu, Gerard (1959), "Theory of Value", John Wiley and Sons, New York.

[14] Harrison, J. Michael and David Kreps (1979), "Martingales and Arbitrage in Multiperiod Securities", Journal of Economic Theory 20, pp. 381 – 408.

[15] Huang, Chi-fu and Robert H. Litzenberger (1988), "Foundations for Financial Economics", Prentice Hall, Englewood Cliffs, New Jersey.

[16] Hull, John C. (1989), "Options, Futures, and Other Derivatives", Prentice Hall, Upper Saddle River, New Jersey.

[17] Jackwerth, Jens Carsten (2000), "Recovering Risk Aversion From Option Prices and Realized Returns", The Review of Financial Studies 13, pp. 433 – 451.

[18] Jondeau, Eric and Michael Rockinger (2000), "Reading the Smile: The Message Conveyed by Methods Which Infer Risk Neutral Densities", Journal of International Money and Finance 19, pp. 885 – 915.

[19] Lancaster, Peter and Kestutis Salkauskas (1986), "Curve and Surface Fitting", Academic Press, London.

[20] Malz, Allan M. (1996), "Using Option Prices to Estimate Realignment Probabilities in the European Monetary System: The Case of Sterling-Mark", Journal of International Money and Finance 15, pp. 717 – 748.

[21] Melick, W.R. and C.P. Thomas (1997), "Recovering an Asset's Implied PDF From Option Prices: An Application to Crude Oil During the Gulf Crises", Journal of Financial and Quantitative Analysis 32(1), pp. 91 – 115.

[22] Ross, Stephen (1976), "Options and Efficiency", Quarterly Journal of Economics 90, pp. 75 – 89.

[23] Shimko, David (1993), "Bounds of Probability", RISK 6, pp. 33 – 37.

[24] Wand, M.P. and M.C. Jones (1995), "Kernel Smoothing", Chapman & Hall, London.

## Appendix A.

Show (8), which is reproduced here for convenience,

$$e^{-rT}[F\phi(d_1) - X\phi(d_2)] = e^{-rT}[F\phi(d_1) - F\phi(d_1)] = 0.$$

We do this by showing that $F\phi(d_1) = X\phi(d_2)$:

$$\begin{aligned}
X\phi(d_2) &= X\phi(d_1 - \sigma\sqrt{T}) && (14)\\
&= X\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(d_1-\sigma\sqrt{T})^2} && (15)\\
&= X\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(d_1^2-\sigma^2 T-2d_1\sigma\sqrt{T})} && (16)\\
&= X\phi(d_1)e^{(-\frac{1}{2}\sigma^2 T+d_1\sigma\sqrt{T})} && (17)
\end{aligned}$$

Now, recall that $d_1$ is given by

$$d_1 = \frac{\ln(\frac{F}{X}) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \tag{18}$$

from which

$$\frac{F}{X} = e^{(-\frac{1}{2}\sigma^2 T+d_1\sigma\sqrt{T})}. \tag{19}$$

Insert this in (17):

$$\begin{aligned}
X\phi(d_2) &= X\phi(d_1)\frac{F}{X} && (20)\\
&= F\phi(d_1), && (21)
\end{aligned}$$

which is the desired result.

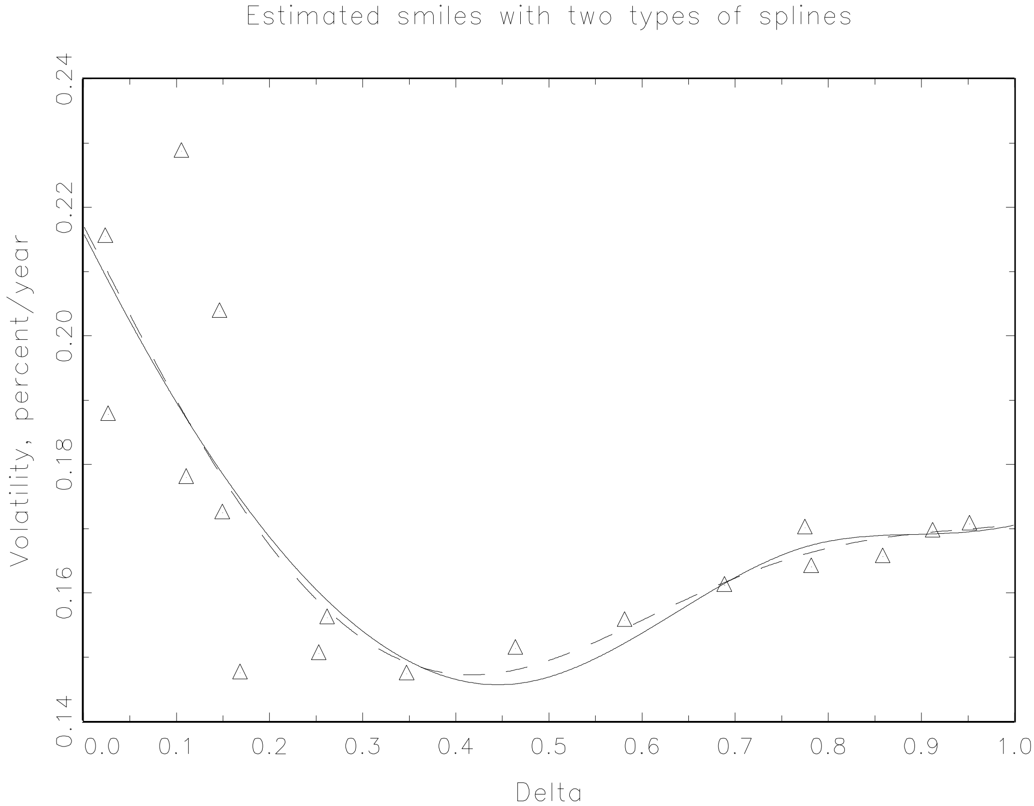


Figure 1: Two types of impled volatility smiles

Note: The solid curve is a smoothed cubic spline with four knot points. The dashed line is a spline of fourth order polynomials with one knot point. Data are prices of 3M Euribor futures options with expiry in September 2001, quoted on LIFFE on 4 April 2001.

Source: Bloomberg.

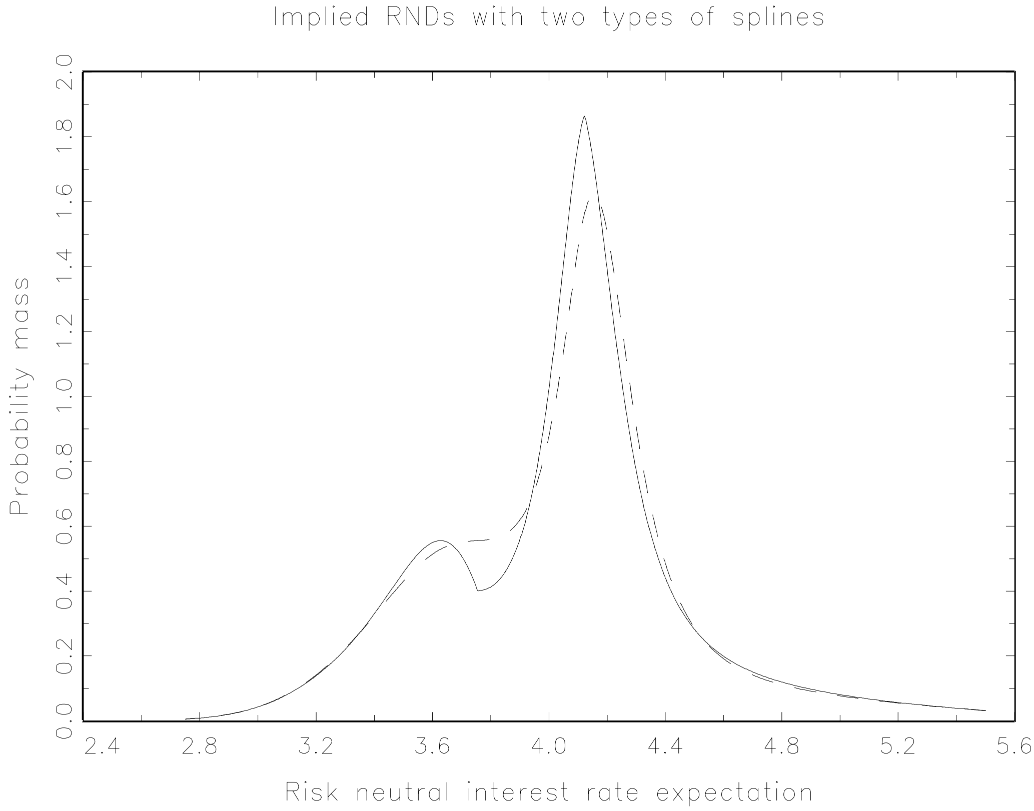


Figure 2: Implied RNDs

Note: The solid and dashed RND correspond to the, respectively, solid and dashed smile in figure 1.

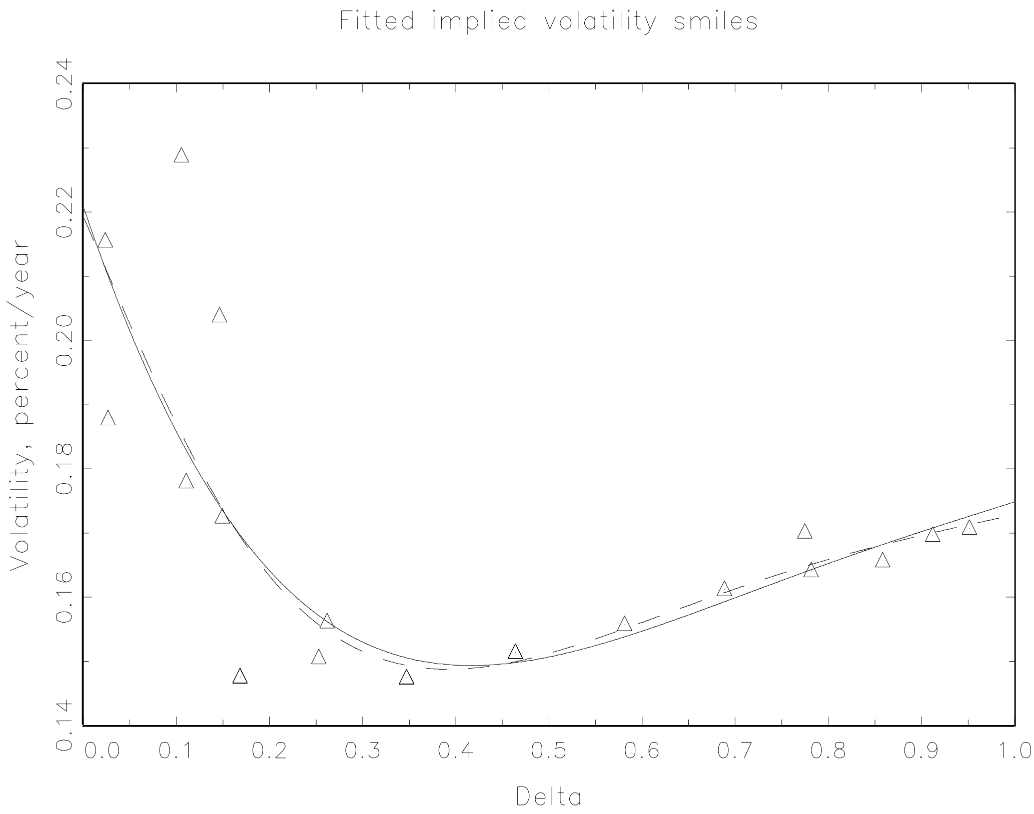


Figure 3: Implied volatility smiles

Note: The solid line is the estimated smile using the second order derivative as the penalty measure for curvature. The dashed line is estimated using the measure in (13) Data are prices of 3M Euribor futures options with expiry in September 2001, quoted on LIFFE on 4 April 2001.

Source: Bloomberg.

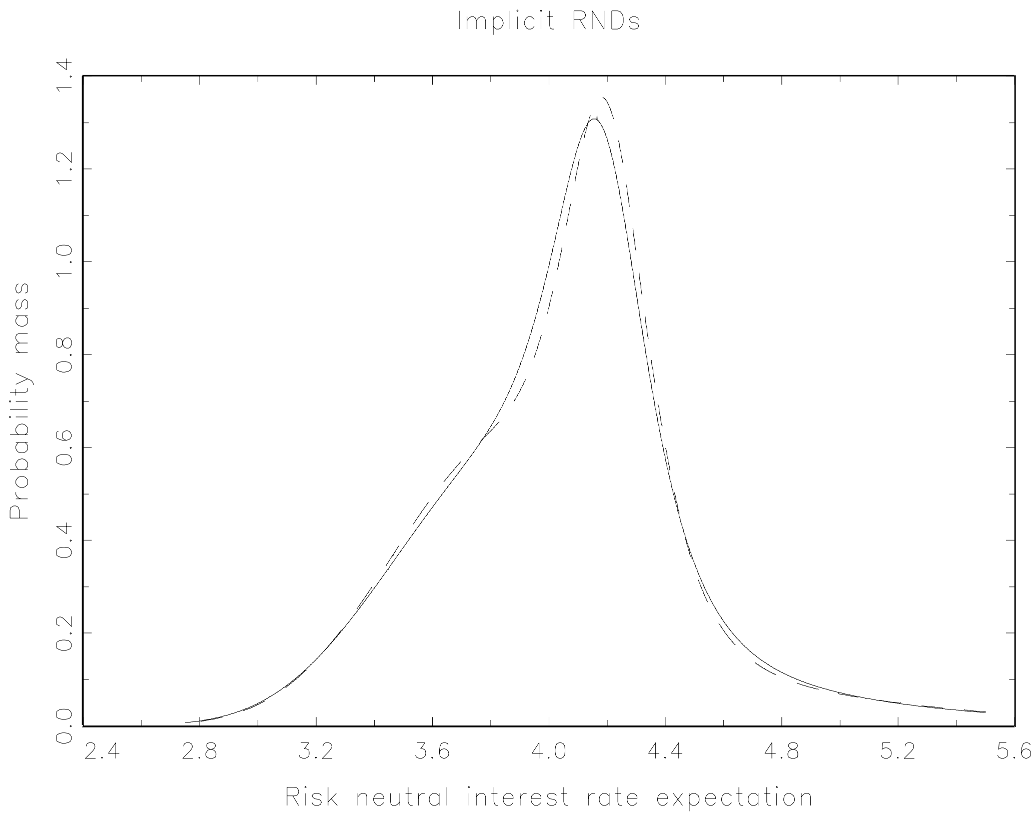


Figure 4: Implicit RNDs

Note: The solid and dashed RNDs correspond to the, respectively, solid and dashed smile in figure 3.

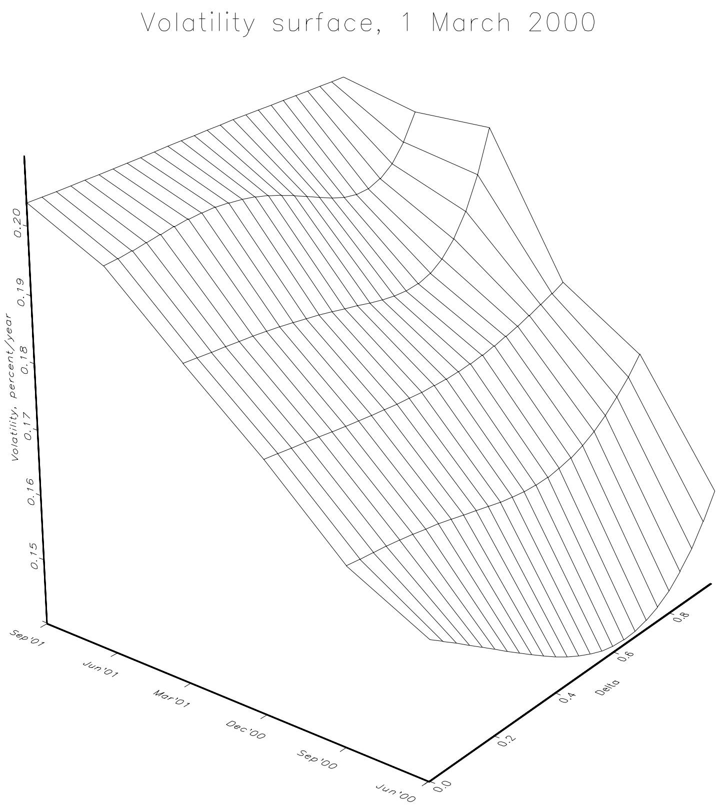


Figure 5: Volatility surface

Note: The volatility surface is computed from prices of 3M Euribor futures options with expiry in June, September and December 2000, and March, June and September 2001, quoted on LIFFE on 1 March 2000.

Source: Bloomberg.

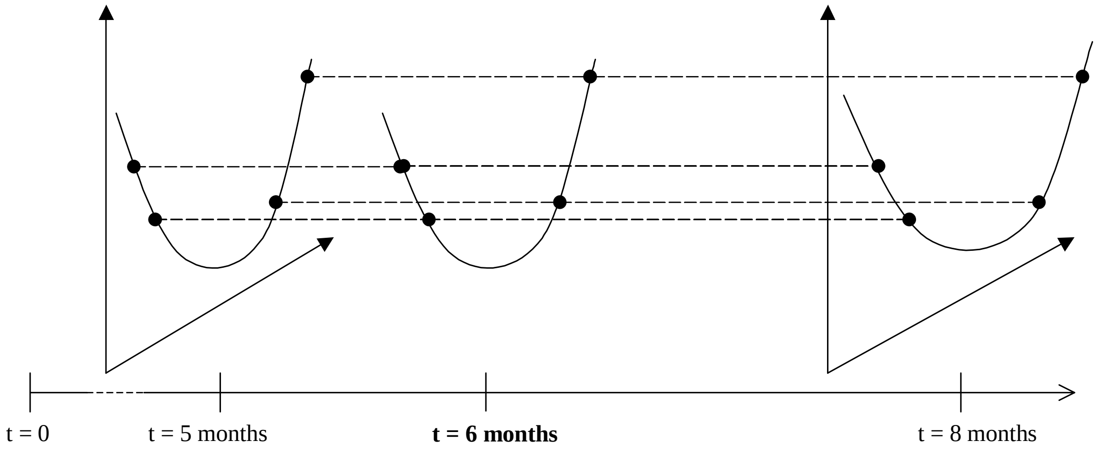


Figure 6: Construction of constant horizon smiles

Note: The four interpolation points are only for illustration. The application uses 2500 interpolation points.

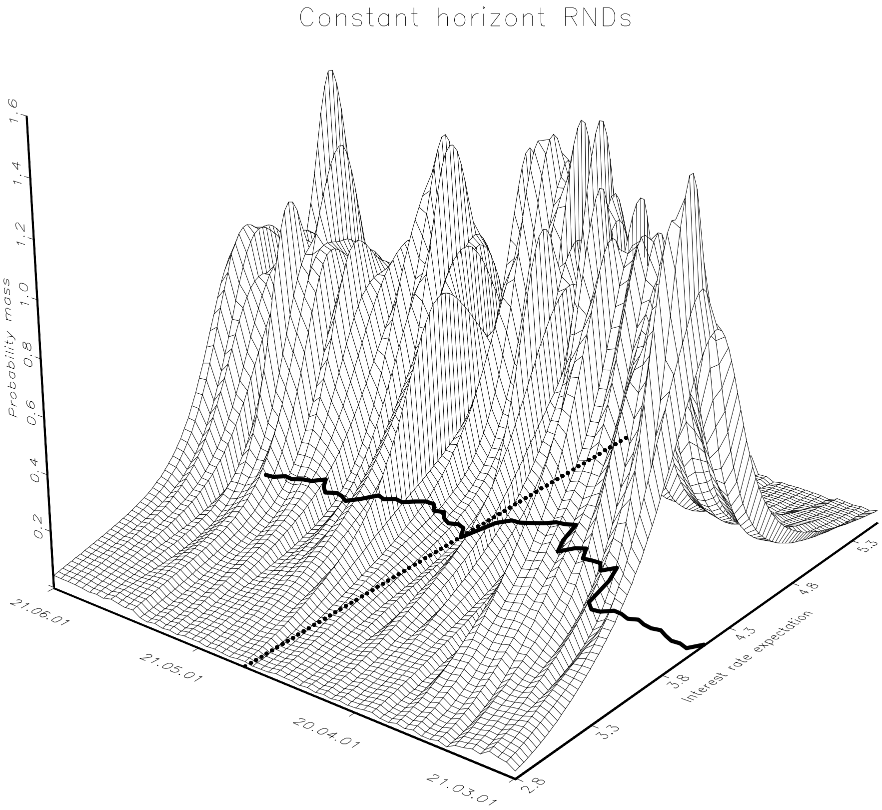


Figure 7: Constant horizont RNDs

Note: The solid line in the zero-plane is the forward interest rate. The dashed line indicates 10 May 2001, the date of the ECB's interest rate reduction of 25 bp.

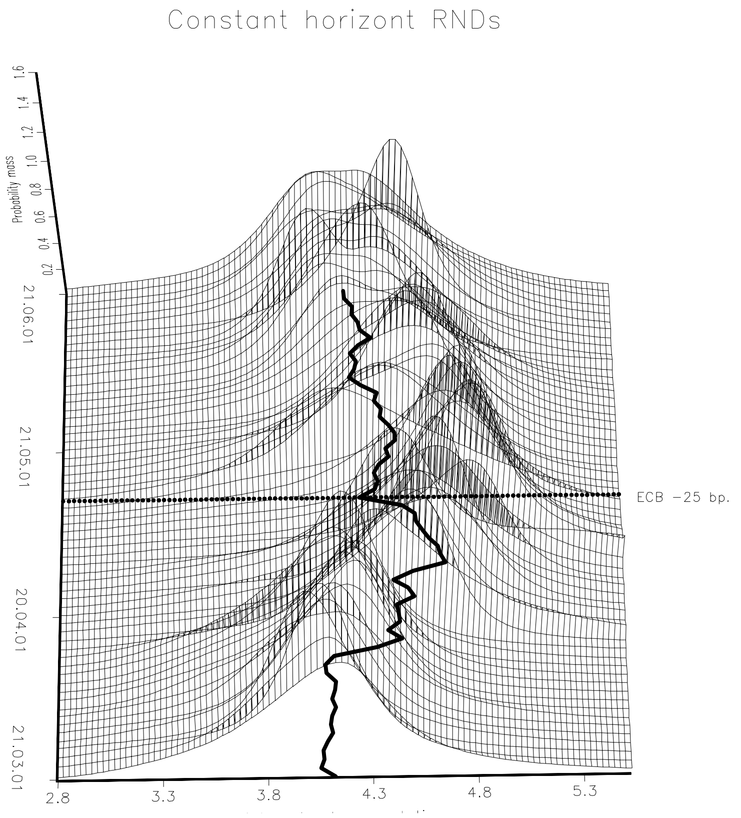


Figure 8: Constant horizont RNDs

Note: The solid line in the zero-plane is the forward interest rate. The dashed line indicates 10 May 2001, the date of the ECB's interest rate reduction of 25 bp.

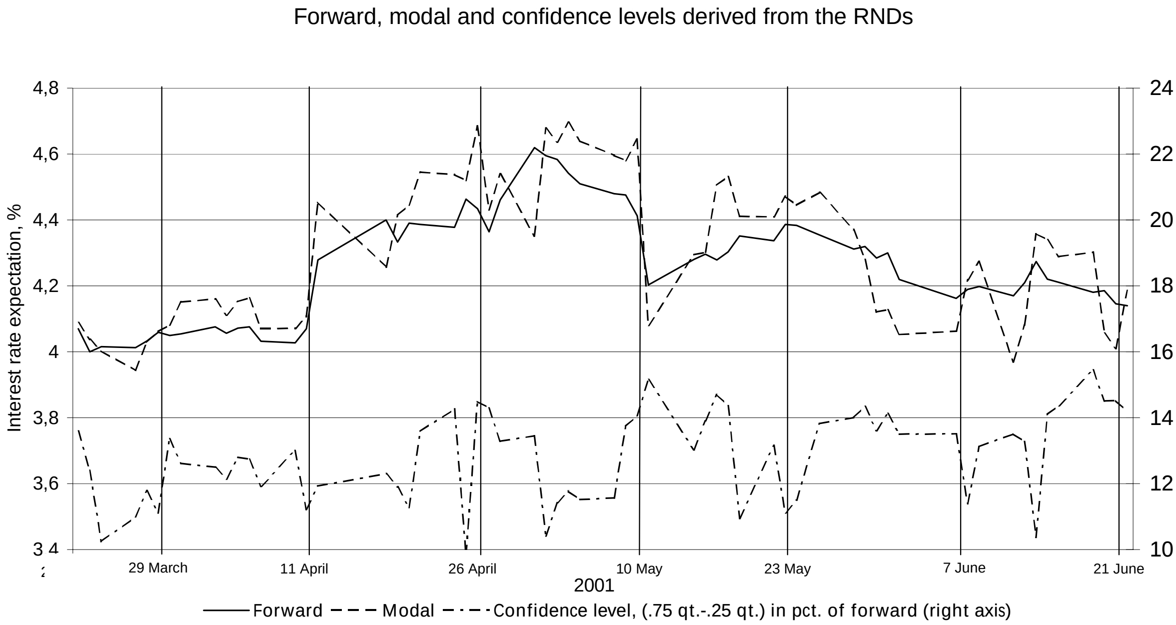


Figure 9: Statistics derived from the RNDs

Note: The lines corresponds to the meeting dates of ECB's Governing Council

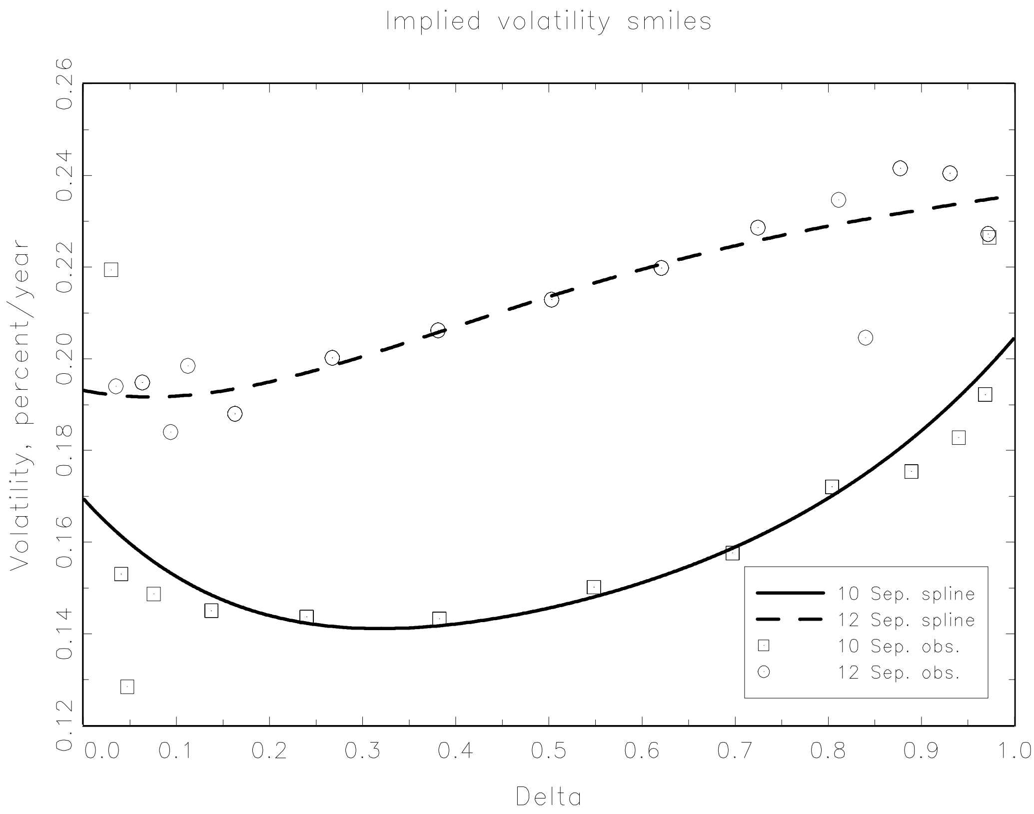


Figure 10: Implied volatility smiles

Note: Data are prices of 3M Euribor futures options with expiry in December 2001, quoted on LIFFE on 10 and 12 September 2001.

Source: Bloomberg.

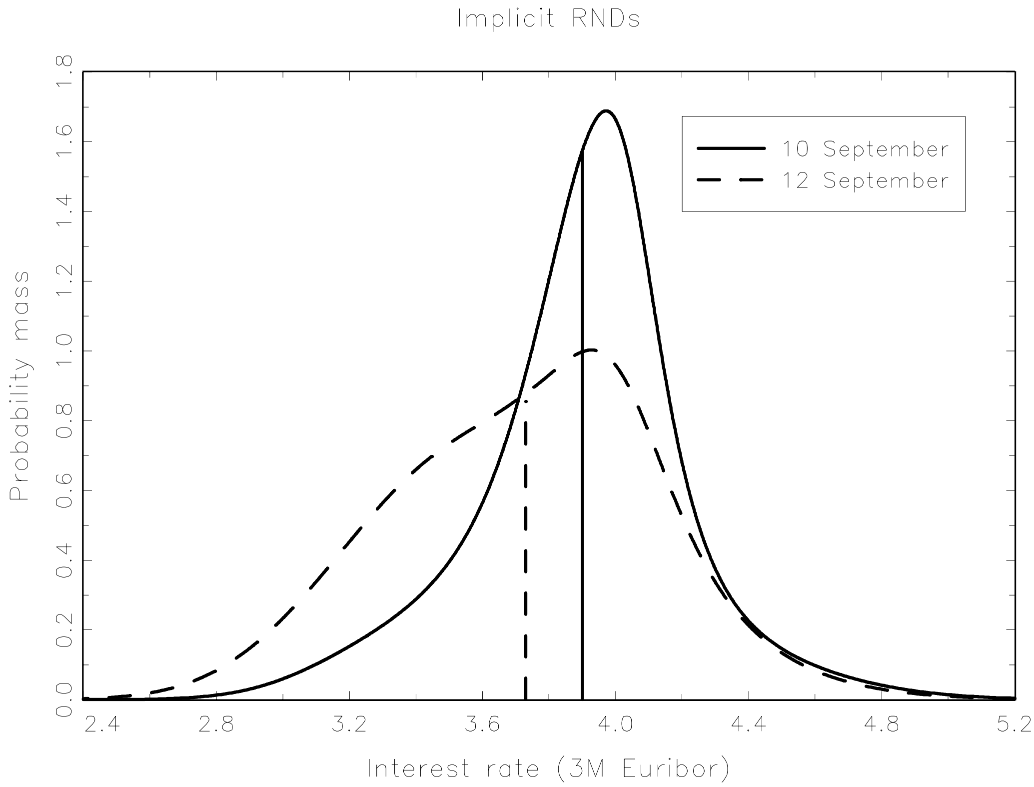


Figure 11: Implied RNDs

Note: Implied RNDs corresponding to the smiles in figure 10. The vertical lines indicate the forward interest rates.