FINANCIAL OPTIONS RESEARCH CENTRE

University of Warwick

The Relation Between Implied and Realised Probability Density Functions

Iliana Anagnou, Mascia Bedendo,
Stewart Hodges and Robert Tompkins

February 2003

Financial Options Research Centre
Warwick Business School
University of Warwick
Coventry
CV4 7AL
Phone: (0)24 76 524118

FORC Preprint: 2003/124

The Relation Between Implied and Realised Probability Density Functions

Iliana Anagnou[*], Mascia Bedendo[*], Stewart Hodges[*] and Robert Tompkins[#]

[*] Financial Options Research Centre, University of Warwick, Coventry, CV4 7AL, Telephone: +44-2476-524118, Facsimile: +44-2476-524167, email: forcmb@wbs.warwick.ac.uk.

[#] Department of Finance, University of Technology, Vienna, Doeltergasse 33, A-1220, Vienna, Austria, Telephone: +43-1-726-0919, Facsimile: +43-1-729-6753, email: rtompkins@ins.at.

The Relation Between Implied and Realised Probability Density Functions

ABSTRACT

In this paper, we test the ability of risk-neutral densities (RNDs) extracted from option data to produce correct forecasts of the true densities of the underlying assets at expiry. Implied RNDs are estimated with both parametric and non parametric methods. A number of new testing procedures to assess the efficiency and unbiasedness of these density forecasts are presented. The forecasting performance of RNDs is also compared to that of return distributions simulated from GARCH-type models. Our findings suggest that implied RNDs represent poor forecasts of the actual densities. However their forecasting performance improves substantially after adjusting for the risk premium.

Risk management requires measuring, monitoring and controlling risk. In the financial markets, these activities could be enhanced by the possession of reliable forecasts of future possible states of the world. Given that such states are probabilistic, of particular interest is the market consensus for the probability density function of future returns.

The most logical source of such probability density functions is the options market because options prices reflect forward-looking distributions of asset prices. Given option prices corresponding to a sufficient range of strike prices, it is possible to infer the risk-neutral probability density function (RND) associated with the underlying asset. There is now a significant literature on the extraction of this density function. Even though the density function is the risk neutral one (reflecting prices, which may include risk premia, and not just expectations) it is an important open question as to whether the risk neutral distribution provides useful information concerning the probabilities of future outcomes. If it does contain such predictive information, this would allow investors to refine their investment strategies and allow regulators to pre-emptively intervene in financial markets to avert market turbulence.

To date, most of the work in this area has concentrated on the estimation of RNDs, hypothesised how such forecasts could be used and provided limited anecdotal evidence of their predictive ability. Much less attention has been devoted to the rigorous *ex-post* assessment of the quality of the RNDs as forecasts of the actual density of the underlying asset at expiry. This research considers this problem: whether or not such RNDs provide an unbiased and good forecast of the distribution of future prices for the underlying.

We will estimate the RNDs for options on the US Dollar/ British Pound and the S&P 500 for a period from 1986 to 2001 and assess how well these forecast the actual density of the underlying assets. The first step is the choice of an approach for the estimation of the RND and it is not obvious which of the alternatives should be used. Bahra (1997), Jondeau and Rockinger (1998) and Coutant, Jondeau and Rockinger (2001), among many others, have previously compared a variety of alternative methods for RND and the general conclusion is that both non-parametric and parametric approaches will yield sensible densities. We will not directly examine this question, but will choose a variety of methods for RND determination, which include approaches suggested by financial regulators. We will also examine for the first time a parametric approach based upon a special case of the generalised hyperbolic (GH) family: the Normal Inverse Gaussian (NIG) distribution, originally introduced by Barndorff-Nielsen (1997, 1998). By estimating RNDs in a variety of ways, we aim to assess if our ultimate hypothesis to be tested (the efficiency of RND forecasts) is dependent upon the choice of RND estimation, and not to directly compare the alternative approaches. To evaluate the RND predictions through time, we will introduce a more robust and rigorous means of testing the hypotheses of efficiency and unbiasedness, through a variety of distributional tests that have shown more statistical power than the standard Komolgorov-Smirnov statistic. These techniques will also be adapted to the case where the distribution is truncated at the extremes of the range of available strikes.

We find that for both the S&P 500 and the US$/BP, the RND is neither an efficient nor an unbiased estimator of the realised probability density function (PDF). Our results are consistent with a number of previously published papers. Using a data set similar to ours, Weinberg (2001) found that the RND for the S&P 500 is a biased forecast with the greatest source of error in the mean. As with his study, we adjust the mean of the distribution to account for the risk premium. In contrast to his results, for one method of mean adjustment (a naïve substitution of the average risk premium), the S&P 500 RND remains a biased forecast of the realised density. We also adjust the RNDs on the basis of some specifications for the investors' utility function. In contrast to our previous approach to mean adjustment, we are unable to reject the assumption that implied densities constitute unbiased and efficient forecasts. This result is consistent with Bliss and Panigirtzoglou (2001).

To provide a comparison between implied RNDs and alternative ways of producing density forecasts for future returns of the underlying asset, quarterly distributional forecasts for returns have been generated according to GARCH-type assumptions for the volatility process. The higher rejection of the null hypothesis of "goodness" of the forecast, recorded for the implied densities, compared to the historical-based distributions, reinforces our conclusion that unadjusted RNDs do not constitute an efficient forecasting tool.

The paper is organised as follows. The next section provides a brief review of the related literature. This is followed by an examination of alternative methods of determining RNDs and the presentation of methodologies chosen for testing their forecasting performance. We then present the data sets used, results of the alternative methods of RND estimation and the results of the forecast tests. A section is then dedicated to the comparison with the historical-based density forecasts. The paper finishes with conclusions and suggestions for further research.

I. Related Literature

Initially, much of the research on the informational content of option prices concentrated on the ability of implied volatility to predict the future realised volatility (for the time horizon of the option's life). The initial research examined common stock options (Latané and Rendleman

(1976), Trippi (1977), Chiras and Manaster (1978) and Beckers (1981)). Later, more emphasis was placed on stock index options and options on futures. Examples of such research include Canina and Figlewski (1993), Day and Lewis (1988,1992), Harvey and Whaley (1992) and Christensen and Prabhala (1998). Most of these studies examined the S&P 100 stock index. Options on the S&P 500 index futures were considered by Park and Sears (1985). The conclusions of these studies are contradictory. Canina and Figlewski (1993) finds that for the S&P 100 market, the implied volatility contained little information for predicting future realised volatility. On the other hand, Christensen and Prabhala (1998) concluded exactly the opposite.

For options on other financial assets, the findings have also been contradictory. For the German Government Bond market, Neuhaus (1995) reached similar conclusions as Canina and Figlewski (1993). While for options on currency futures, Jorion (1995) concluded that implied volatility does have substantial predictive power. Nevertheless, most studies have found that the implied volatility is an upwardly biased estimate of the realised volatility.

However, the existence of implied volatility smiles complicates these tests. Given that different implied volatilities are observed at different strikes, it is not obvious which one should be selected as the forecast. In most of these studies, the at-the-money implied volatility is selected and the smile structure is ignored. Many authors have found evidence for convex patterns of implied volatilities as a function of strike prices (in a wide variety of markets). These include Shastri and Tandon (1986), Kemna (1989), Xu and Taylor (1995) and Heynen, Kemna and Vorst (1994). Furthermore, such patterns have changed after periods of unusual market turbulence. Rubinstein (1994) examined options on the S&P 500 index and found that prior to the 1987 stock market crash the implied volatilities displayed a symmetrical pattern relative to the price of the underlying asset. After the 1987 crash, the pattern changed with the prices of lower strike price options increasing substantially relative to higher strike price options. This result has been interpreted as an expectation of more mass in the lower half of the expected probability density.

Such a pattern has been coined the "volatility skew". Dumas, Fleming and Whaley (1998) confirm this finding and find other time dependent divergences in the implied volatility patterns.

Given that such patterns exist, a number of approaches have been proposed to correct for this effect. One such approach is parametric: to derive an analytic option pricing model which extends the approach of Black and Scholes (1973). A number of authors have proposed that option prices can be seen as the sum of a Black-Scholes price plus adjustment terms which depend on the higher moments of the underlying security stochastic process. Specifically, Jarrow and Rudd (1982) used a generalised Edgeworth expansion, Corrado and Su (1996, 1997) applied a Gram-Charlier expansion, while Madan and Milne (1994), and Abken, Madan and Ramamurtie (1996) employed a Hermite polynomial expansion. Recently, Abadir and Rockinger (1997) proposed a Kummer's function adjustment to the normal density function. These approaches require additional unobservable parameters (for the higher moments) to price the options.

Related parametric approaches propose alternative functional forms for option prices. One of the first papers that examined the (RND) distributions implied by option prices was Bates (1991), who assumed a jump model parameterised using a Least Squares Method. In a similar vein, Rubinstein (1994), Derman and Kani (1994) and Dupire (1994) take as given the market prices of options and find the density function that is consistent with those prices. These methods propose alternative forms: Dupire (1993) and Derman and Kani (1994) assume diffusion processes, while Rubinstein (1994) suggests the use of a binomial framework. All three papers assume that the volatility is a deterministic function of time and the strike price. However, tests by Dumas, Fleming and Whaley (1998) reject the existence of such a deterministic relationship and this line of research has now been called into question.

Since that time, a number of alternative functional forms for the RND have been proposed. Ritchey (1990), Bahra (1997), Melick and Thomas (1997), Söderlind and Svensson (1997), Gemmill and Saflekos (2000) have suggested a mixture of lognormals. Sherrick, Garcia and Tirupattur (1996) have modelled RNDs with a three parameters Burr distribution. Aparicio and Hodges (1998) have used a Generalised Beta.

The third method, which is non-parametric and is attributed to Shimko (1991, 1993), relies on the fact that the risk-neutral density function is equal to the second derivative of the option price relative to the strike price. Breeden and Litzenberger (1978) were the first to demonstrate this. Assuming some functional form φ_T for call option prices, *C,* differentiating this form twice with respect to the strike prices, *K,* will provide the relationship between the risk neutral probability density and the smile surface:

$$\varphi_T(K) = \frac{\partial^2 C(K,T)}{\partial K^2} \tag{1}$$

This approach requires a continuum of observable option prices. Since these do not exist in practice, Shimko suggests interpolating the prices of market-traded options and then deriving the risk-neutral density function from these interpolated values. Shimko (1993) suggests the use of a least squares method to fit a quadratic function to the volatility smile. This approach can be expressed as:

$$\hat{\sigma}(K,\tau) = \alpha(\tau) + \beta_1(\tau)K + \beta_2(\tau)K^2 + \varepsilon \tag{2}$$

where K is the strike price, τ is point in time at which the implied volatility smile is estimated, and the coefficients of the regression measure the intercept of the regression, the first and second order effects can be interpreted as capturing higher moments of the distributional function (β_1 for the skewness and β_2 for the kurtosis).

Examples of non-parametric estimations of the RNDs can be found in Shimko (1993), Ait-Sahalia and Lo (1998), Aparicio and Hodges (1998), Malz (1997), Campa, Chang and Reider (1998). Jackwerth (1999) provides a review and a comparative analysis of the different approaches, concluding that the results are not too dissimilar.

To date most of the published work on RNDs has concentrated on estimation issues whereas much less emphasis has been placed on the formal assessment of the goodness of such estimates. The most common instruments used to test for the accuracy of fit are the pricing errors, computed as difference between the observed option prices and the option prices obtained from the estimated implied RNDs, or alternatively, as difference between the observed and the estimated implied volatility. These residuals are then averaged to compute aggregate indicators which place more weight on larger errors than on smaller ones, such as the mean squared error (MSE), the mean squared percentage error (MSPE) and the root mean squared error (RMSE). Orthogonality tests have also been performed to check for predictability or/and existence of patterns in the pricing errors.

The analysis of the forecast errors is usually accompanied by the computation of summary statistics for the implied risk-neutral densities. The summary statistics conventionally adopted are both traditional and robust measures of 1) location; 2) dispersion; 3) asymmetry; 4) fat-tailness; 5) various tail percentiles. The average values and the distributions for these statistics are then contrasted for alternative models. Most of the authors agree in finding similar values for the first two moments across various models and large discrepancies in the higher moments of the distributions, which seems to suggest that measures of skewness and kurtosis are highly model-dependent and, therefore, quite unstable.

Some work on the robustness and stability of the estimates of the implied probability density functions and their summary statistics has been done by Bliss and Panigirtzoglou (2002) and Cooper (1999). Aparicio and Hodges (1998) have also investigated the time series properties of RNDs. However, very little attention has been devoted by the existing literature to the *ex-post* assessment of the quality of the entire RNDs as forecasts of the actual density of the underlying asset at expiry.

Most research in this area has concentrated on the ability of RNDs to predict single events. The most popular events to be examined are stock market crashes (see Bates (1991), Gemmill (1996), Malz (1996), Jackwerth and Rubinstein (1996), Melick and Thomas (1997), Gemmill and Saflekos (2000). Other studies have considered more general economic news (Bahra (1997)), British general elections (Gemmill and Saflekos (2000)), announcements of interest rate changes by the Federal Reserve (McManus (1999)).

As with the literature on the information content of implied volatility, the findings are contradictory. Most papers, which have examined stock market turbulences, reject the hypotheses that they were predicted by options prices. Bates (1991) concluded that options on the S&P 500 did not predict the 1987 stock market crash. Gemmill (1996) reached similar conclusions for options on the FTSE 100. Bates (2000) further reports that since the 1987 crash, the RND of the S&P 500 has consistently over-estimated left tail events. He concludes that investors retain fears of a future stock market crash even though this has failed to occur. For options prices to be consistent with crash risk, events similar in order and magnitude to the 1987 stock market crash would have to occur every two to three years. Gemmill and Saflekos (2000) concludes that RNDs are not able to predict future market turbulence but react to it. Essentially RNDs are backward and not forward looking. Using a different methodology, Shiratsuka (2001) also found that the implied PDFs for the Japanese stock market did not "extract useful information automatically from the shape of an implied probability distribution" (page 18). He concludes that any application to regulators (for the conduct of monetary policy) of the implied PDF approach is extremely limited.

On the other hand, Bahra (1997) and McManus (1999) seem to provide some anecdotal evidence of the forecast ability of RNDs for changes in Government interest rate policy. Campa, Chang and Reider (1998) claim that RNDs correctly model the market expectations of currency realignments.

For all these studies, an extremely limited number of events have been examined. A fully systematic and formal analysis of the forecasting power of the RNDs over a long period, that exploits the tools introduced by recent developments in the theory of density forecast evaluation, has not been attempted, to our knowledge. However, this would be desirable, since we believe that the investigation of the entire density would lead to a more accurate and complete evaluation of the goodness of the distributional forecasts provided by the RNDs, in view of both a comparative study of different specifications, and a more general assessment of the usefulness of implied RNDs for forecasting.

While this paper was being written, we became aware of similar work in this area by Weinberg (2001). As in our research, Weinberg (2001) examined the predictive ability of RNDs (on US Dollar/Yen, US Dollar/Deutsche Mark and the S&P 500 options) for long time periods and not solely for individual events, through formal goodness-of-fit tests based on the empirical distribution function (EDF). He concludes that the RNDs for these markets are not efficient / unbiased forecasts of realised densities. More specifically, he compared the forecasting performance of a volatility smoothing method with that of a Black and Scholes lognormal distribution through the investigation of Anderson-Darling, Cramér-von Mises and Watson statistics. Given that EDF tests require independent realisations, Weinberg chose to enlarge the sample size by computing the EDF statistics with a Monte Carlo technique. Our study differs from this work under two aspects. First we produce a much wider battery of tests, more informative on the nature of the misspecification in the RNDs when rejection of the hypothesis of predictive accuracy occurs. Secondly, for constructing the test, we only rely on the actual total number of quarterly contracts available in the sample for a given underlying. By working only with non-overlapping observations, we restrict our data set severely, but the requirement of independence is met and no further adjustments for interdependence are needed.

Other related research is by Bliss and Panigirtzoglou (2001), who also examined options on the S&P 500 (and options on the Financial Times 100 stock index, FTSE 100) for a similar time period to ours. In contrast to our study, which is restricted to quarterly non-overlapping periods, they examine a variety of option expiration horizons from one week to six weeks.

It is well known that the representative agent's subjective density function can be obtained by dividing the risk neutral density by the (normalised) marginal utility of the agent. It is only in the unlikely case of a risk neutral representative agent that no adjustment is necessary. Bliss and Panigirtzoglou (2001) assume that an assumed utility function is stationary with the result that with time varying RNDs will yield time varying subjective density functions. They estimate the utility function that provides the best possible fit between the subjective and realised densities. As with our study, Bliss and Panigirtzoglou (2001) reject RNDs as unbiased efficient forecasts of realised densities, but find that the utility adjusted densities perform better. In this study, we also consider a power utility adjustment for the RND and also find that it performs better. As opposed to Bliss and Panigirtzoglou (2001), we are unwilling to conclude that the subjective density forecast are "good" forecasts but merely that there is insufficient evidence to prove that they are "bad" forecasts.

II. Alternative Methods For Estimating Risk-Neutral Densities

We follow four alternative approaches to the estimation of the risk-neutral distribution function: three parametric and one non-parametric approach. Each will be examined separately.

A. Estimating RND using a Generalised Beta Approach

The parametric approach is based on the assumption that the risk-neutral distribution belongs to a general family distribution and its unknown parameters are estimated from the option or asset data. For this estimation method, we first employ the Generalised Beta of the second kind (*GB2*). *GB2* is defined (McDonald (1984)) as follows:

$$GB2(y;a,b,p,q) = \frac{|a|y^{ap-1}}{b^{ap}\mathrm{B}(p,q)\left[1+(y/b)^a\right]^{p+q}}, \tag{3}$$

for $y > 0$, and $GB2(y;a,b,p,q) = 0$ otherwise.

The expressions for its distribution function and h^{th} moment respectively are,

$$\frac{\left(\frac{(y/b)^a}{1+(y/b)^a}\right)^p}{pB(p,q)}\,{}_2F_1\left[\begin{matrix} p,1-q; \dfrac{(y/b)^a}{1+(y/b)^a} \\ p+1; \end{matrix}\right] \tag{4}$$

$$\mathrm{E}(y^h) = \frac{b^h\mathrm{B}(p+h/a,q-h/a)}{\mathrm{B}(p,q)} \tag{5}$$

where $\mathrm{B}(p,q)$ denotes the beta function, as given by $\mathrm{B}(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}dt$ and ${}_2F_1$ represents the generalised hypergeometric series defined by:

$${}_pF_q\left[\begin{matrix} a_1,\ldots,a_p;x \\ b_1,\ldots,b_q; \end{matrix}\right] = \sum_{i=0}^{\infty}\frac{(a_1)_i\ldots(a_p)_i}{(b_1)_i\ldots(b_p)_i}\frac{x^i}{i!} \tag{6}$$

The mean as well as the shape of the *GB2* is determined by the four parameters of the distribution. The parameter b affects directly the mean of the *GB2*,

$$\mathrm{E}(y) = \frac{b\mathrm{B}(p+1/a,q-1/a)}{\mathrm{B}(p,q)} \tag{7}$$

For large values of the parameter a, the value of b will be close to that of the mean of the distribution. The other three parameters of the *GB2* have a direct effect on the shape of the distribution. The parameter a is associated to the speed with which the tails of the density function approach the x-axis. Large values of a entail faster approach to the axis. The term aq determines the fatness of the distribution. No moments of order equal to or higher than aq will exist[1]. The parameters p,q must be strictly positive for the beta function to be defined and work interactively in determining the skewness of the distribution. A decrease in the value of p leads to an effect that is the opposite of that of q.

We use the Generalised Beta of the second kind because it provides greater flexibility. Compared to the lognormal, it allows for either positive or negative skewness and also for the existence of even infinite moments. Furthermore, *GB2* can be replaced by any distribution from a wide range because it includes a wide range of well-known distributions. The generalised gamma as well as the lognormal density are limiting cases of the *GB2* as the value of q approaches infinity. Moreover, distributions such as the chi square, exponential, gamma, Burr type 12, and Burr type III can be expressed as both limiting and special cases. Finally, using the parametric estimation allows having estimates beyond the corresponding strike range.

Under the assumption that the risk neutral distribution belongs to the distribution family of the *GB2*, the values of the call and put options can be expressed as follows:

$$C_i(K_i) = e^{-r\tau} \int_{K_i}^{+\infty} (S_T - K_i) GB2(S_T; a, b, p, q) dS_T \,, \quad i = 1..n \tag{8.1}$$

$$P_j(K_j) = e^{-r\tau} \int_{0}^{K_j} (K_j - S_T) GB2(S_T; a, b, p, q) dS_T \,, \quad j = 1..m \tag{8.2}$$

$$S_t = e^{-r\tau} \int_{0}^{+\infty} S_T GB2(S_T; a, b, p, q) dS_T \tag{8.3}$$

where n and m denote the number of call and put option market prices with the same maturity, for different strikes and for a given contract per day.

We estimate the four parameters of the *GB2* distribution by minimising the sum of the squared errors between observed market option prices and the theoretical option prices calculated accordingly to the above equations. Because calls and puts are evaluated from the same distribution, they are both involved in the estimation of the parameters of the *GB2* distribution. Moreover, the forward price of the underlying is treated as an additional observation so that we ensure that the mean of the distribution is the forward price. The estimation of the four parameters of the *GB2* distribution involves minimising the following objective function:

$$\underset{a,b,p,q}{Min} \sum_{i=1}^{n} \left[C(K_i, \tau) - \hat{C}_i \right]^2 + \sum_{j=1}^{m} \left[P(K_j, \tau) - \hat{P}_j \right]^2 + \left[E(S_T) - F \right]^2 \tag{9}$$

The estimation of the parameters involves a non-linear least squares minimisation problem. Using the recovered implied volatilities and the Black-Scholes option pricing formula, we calculate the pseudo-European option prices. These are used as the observed market prices in the minimisation problem.

B. Estimating RND using a Normal Inverse Gaussian Approach

Many recent studies (Barndorff-Nielsen (1997, 1998), Rydberg (1997)) have suggested that the distribution of logarithmic asset returns[2] can be well fitted by a particular distribution belonging to the class of generalised hyperbolic (GH) densities, the Normal Inverse Gaussian. The NIG density function is given by:

$$NIG(x; \alpha, \beta, \mu, \delta) = \pi^{-1} \alpha e^{\delta\sqrt{(\alpha^2 - \beta^2)} - \beta\mu} q\left(\frac{x-\mu}{\delta}\right)^{-1} K_1\left\{\delta\alpha q\left(\frac{x-\mu}{\delta}\right)\right\} e^{\beta x} \tag{10}$$

where $q(x) = \sqrt{(1 + x^2)}$, $\mu \in \mathbf{R}$, $\delta > 0$, $0 \leq |\beta| \leq \alpha$ and K_1 is the modified Bessel function of third order and index 1. The moment generating function of a NIG possesses a nice and neat expression:

$$M(u; \alpha, \beta, \mu, \delta) = \exp\left[\delta\left\{\sqrt{\alpha^2 - \beta^2} - \sqrt{(\alpha^2 - (\beta + u)^2)}\right\} + \mu u\right] \tag{11}$$

which allows a convenient way of mapping between the parameters of the distribution and the first four moments. As an intuitive explanation of the parameters, α represents a steepness parameter, β is an asymmetry parameter, δ describes the scale and μ the location of the distribution. A further advantage of adopting a NIG specification consists in its property of being closed under convolution, not shared by alternative GH distributions.

Despite its ability to account for crucial features of the distributions of financial asset returns such as semi-heavy tails and asymmetry, and its parsimony in the number of parameters to be estimated, the NIG has never been proposed in the past as a possible parametric specification for the RNDs implied from option prices. We believe that the flexibility and tractability of such distribution make it appropriate for our purposes. Therefore we estimate the parameters of the implied NIG RNDs by solving a non-linear least squares minimisation problem of the kind in (9), where:

$$C_i(K_i) = e^{-r\tau} \int_{K_i}^{+\infty} (S_0 e^{x_T} - K_i) NIG(x_T;\alpha,\beta,\mu,\delta)\, dx_T\,, \quad i = 1..n \tag{12.1}$$

$$P_j(K_j) = e^{-r\tau} \int_{0}^{K_j} (K_j - S_0 e^{x_T}) NIG(x_T;\alpha,\beta,\mu,\delta)\, dx_T\,, \quad j = 1..m \tag{12.2}$$

under the constraint that the mean of the implied distribution for the underlying equals the forward price, exactly as before.

C. Estimating RND using a Two-lognormal Mixture

Many studies indicate the mixture of lognormal distributions as a good candidate to represent the RND function, given its flexible specification that allows approximating quite a wide range of shapes. In fact, this functional form seems to be the preferred one by the policy-makers in quite a few industrialised countries. Even though we think that such a distribution is not the most appropriate parametric form in this context, given the large number of parameters to be estimated, we included it in our analysis for completeness and to provide an alternative parametric specification to be compared with the generalised beta and the NIG.

We recall that the prices of European call and put options at time t can be written as follows:

$$C(K_i,\tau) = e^{-r\tau} \int_{K_i}^{\infty} q(S_T)(S_T - K_i)\, dS_T \tag{13.1}$$

$$P(K_j,\tau)=e^{-r\tau}\int_0^{K_j} q(S_T)(K_j-S_T)dS_T \tag{13.2}$$

Here we assume that the density function $q(S_T)$ is given by a mixture of two lognormal density functions, that is

$$q(S_T)=\sum_{i=1}^{2}\left[\theta_i L(\alpha_i,\beta_i;S_T)\right] \tag{14}$$

where $L(\alpha_i,\beta_i; S_T)$ is the i^{th} lognormal density with parameters α_i and β_i:

$$\alpha_i=\ln S+\left(\mu_i-\frac{1}{2}\sigma_i^2\right)\tau \text{ and } \beta_i=\sigma_i\sqrt{\tau} \quad \text{for } i=1,2. \tag{15}$$

Given the assumption made on $q(S_T)$, equations (13.1) and (13.2) can be expressed as follows:

$$C_i(K_i,\tau)=e^{-r\tau}\int_{K_i}^{\infty}\left[\theta L(\alpha_1,\beta_1;S_T)+(1-\theta)L(\alpha_2,\beta_2;S_T)\right](S_T-K_i)dS_T \tag{16.1}$$

$$P_j(K_j,\tau)=e^{-r\tau}\int_0^{K_j}\left[\theta L(\alpha_1,\beta_1;S_T)+(1-\theta)L(\alpha_2,\beta_2;S_T)\right](K_j-S_T)dS_T \tag{16.2}$$

The estimates for the five parameters α_1, β_1, α_2, β_2, θ, are obtained by minimising the deviation of the theoretical prices for both calls and puts given by (16.1) and (16.2) from the market prices, across the available range of strikes. The total sum of squared errors for call and put options is minimised by means of a non-linear least squares optimisation routine. The minimisation problem then becomes:

$$\underset{\alpha_1,\alpha_2,\beta_1,\beta_2,\theta}{Min}\sum_i\left[C(K_i,\tau)-\hat{c}_i\right]^2+\sum_j\left[P(K_i,\tau)-\hat{p}_i\right]^2+\left[\theta e^{\alpha_1+0.5\beta_1^2}+(1-\theta)e^{\alpha_2+0.5\beta_2^2}-e^{r\tau}S\right]^2 \tag{17}$$

subject to β_1, $\beta_2>0$ and $0\leq\theta\leq1$.

Bahra (1997) derives closed-form solutions to equations (16.1) and (16.2):

$$C(K_i,\tau)=e^{-r\tau}\left\{\theta\left[e^{\alpha_1+0.5\beta_1^2}N(d_1)-K_iN(d_2)\right]+(1-\theta)\left[e^{\alpha_2+0.5\beta_2^2}N(d_3)-K_iN(d_4)\right]\right\} \tag{18.1}$$

$$P(K_j,\tau)=e^{-r\tau}\left\{\theta\left[-e^{\alpha_1+0.5\beta_1^2}N(-d_1)+K_jN(-d_2)\right]+(1-\theta)\left[-e^{\alpha_2+0.5\beta_2^2}N(-d_3)+K_jN(-d_4)\right]\right\} \tag{18.2}$$

where:

$$d_1=\frac{-\ln K+\alpha_1+\beta_1^2}{\beta_1}, \qquad d_2=d_1-\beta_1 \tag{18.3}$$

$$d_3 = \frac{-\ln K + \alpha_2 + \beta_2^2}{\beta_2}, \qquad d_4 = d_3 - \beta_2 \tag{18.4}$$

The existence of a closed-form solution represents a nice feature of the mixture of lognormal densities, obviating the need for numerical integration in these equations and therefore, ensuring higher computational tractability and greater accuracy in the minimisation process that leads to the estimation of the implied RNDs.

D. Estimating RND using B-Splines

In contrast to the parametric approach, the non-parametric makes no distributional assumptions about the risk-neutral distribution. The implied volatility curve is modelled and the risk-neutral density function is recovered from the second derivative of the pricing formula with respect to the strike price, according to Breeden and Litzenberger (1978).

To estimate non-parametrically the risk-neutral distribution, we use Shimko's approach (1993) but in a more flexible way. To calculate derivatives, we need smooth option-pricing functions. Therefore, we want to obtain a volatility curve that is as smooth as possible and also fits the given data set as closely as possible.

For this purpose, we use a linear combination of cubic B-spline functions. We want the smoothest function that lies within the given tolerance *tol* of the data. We estimate a cubic smoothing spline f (De Boor (1978)) such that:

$$D^m f(t) = \int_{x(1)}^{x(n)} \left[D^m f(t)\right]^2 dt \qquad x(1) < t < x(n) \tag{19}$$

is smallest, for which:

$$E(f) = \sum_{i=1}^{n} (O_i - O(f(x_i))^2 \leq tol \tag{20}$$

with $m = 2$ leading to the cubic smoothing spline. The variables O_i and $O(f(x_i))$ reflect options prices (both calls and puts will be subsequently examined in this research). $D^m f(t)$ corresponds to a roughness measure and *tol* to the penalty imposed on the roughness of the approximation. As the level of the tolerance varies, the spline changes within the two extreme cases of the least-squares straight-line approximation and the natural cubic interpolating spline. Given that the available strike prices used to fit the cubic spline occur over a limited range (i.e. do not extend from zero to positive infinity), we are lacking information regarding the extreme tails of the distribution. While a number of solutions to this have been proposed to fill in missing data in the tails (for example Jackwerth and Rubinstein (1996)), we have chosen not to model the tails. Therefore, for the non-parametric specification, our analysis is limited to the estimation and testing only of the truncated density. The B-spline estimation method provides us with the truncated RND.

III. Testing Forecast Ability of Risk-Neutral Densities

As stated previously, we propose to apply a battery of goodness-of-fit techniques designed to assess density forecasts for the purpose of investigating whether a sequence of observed values of future prices comes from the estimated RND.

A. Density forecast evaluation: the Probability Integral Transform (PIT) approach

The key device in the field of density forecast evaluation is the probability integral transform (PIT) approach, which dates back to Fisher (1930) and Rosenblatt (1952), and was subsequently adopted by Dawid (1984).

Given a sequence of one-step ahead density forecasts $p_t(y_t|\Omega_t)$ of the conditional density $f_t(y_t|\Omega_t)$, the probability integral transform of the realisation of the process y_t taken with respect to the density forecast is:

$$\begin{aligned} z_t &= \int_{-\infty}^{y_t} p_t(u)du \\ &= P_t(y_t), \qquad t = 1,\ldots,n \end{aligned} \tag{21}$$

If the forecasts and the true densities coincide, assuming that the Jacobian of the transformations is non-zero over the support of the realisations, with continuous partial derivatives, then the sequence of the PITs z_t is distributed as i.i.d. U(0,1) (see Diebold, Gunther and Tay (1998)).

In the PIT approach the evaluation of density forecasts translates into assessing whether the series of the probability integral transforms is i.i.d. U(0,1). For this purpose, Diebold, Gunther and Tay (1998) suggest a visual approach, claiming that formal joint goodness-of-fit tests of i.i.d. U(0,1), as well as related separate tests of i.i.d. and U(0,1) are not constructive in revealing the nature of the misspecification when rejection of the null hypothesis occurs. Therefore, they use histograms to evaluate unconditional uniformity and correlograms of both the levels and the powers of the series z_t to detect inaccuracies in modelling the linear and non-linear dynamics of the true process.

Formal goodness-of-fit testing procedures based on the PIT, have been employed more often in the relevant literature. The most popular amongst these techniques is certainly the Kolmogorov-Smirnov test. Cnkovic and Drachman (1996) use the Kuiper statistic. Noceti, Smith and Hodges (2000) present a comparative study on the power of several alternative techniques, mainly based on the PIT, to detect misspecifications of different type in the forecasted distribution. As a result, they point out the lack of power displayed by the Kolmogorov -Smirnov test against alternative methods.

Stressing the inadequacy of the tests usually associated with the PIT approach for realistic sample sizes, Berkowitz (2001) suggests a modification of Rosenblatt's techniques. A further transformation to normality is applied to the series z_t of probability integral transforms:

$$x_t = \Phi^{-1}(P_t(y_t)) \tag{22}$$

If the sequence of z_t is i.i.d. U(0,1), that of x_t must be i.i.d. N(0,1). At this point, conventional testing techniques for normally distributed data, whose statistical properties are well documented also for small sample sizes, can be applied. Berkowitz proposes to use log-likelihood ratio tests, which, however, implicitly rest on the assumption of normality of the transformed series. Instead, De Raaij and Raunig (2002) choose to test a simple autoregressive model, which embeds the i.i.d. standard normal as a special case.

B. The PIT Approach and the Evaluation of Risk-Neutral Densities

The PIT result holds even if the forecasting model changes through time, and regardless of the specific distributional form of the realisations y_t, and of the way the forecasts have been obtained. This latter aspect makes this approach particularly suitable to our problem of evaluating the forecasting performance of RNDs, since RND functions are extracted from option prices and may not be generated from a specific model. In the context of interest, p_t denotes the sequence of RNDs[3] and y_t the sequence of observed values for the underlying at expiry of the contract.

Although the PIT approach seems to be both a flexible and a rigorous method to assess the adequacy of RNDs as a forecasting tool for the actual distribution of the underlying asset, an important issue arises here. The set of option prices and strikes across which the derivation of the RNDs is made is discrete and, for some contracts, quite narrow. When we choose a parametric approach to estimate the RNDs, the tails of the resulting distribution outside the range of available strikes are implicitly obtained and, in general, no additional extrapolation needs to be performed.[4] Therefore, in presence of an entirely specified density function, series of PIT z_t can be computed to assess the forecasting performance of RNDs, following exactly the procedure described above.

On the contrary, the implementation of non-parametric techniques only consents to recover the implied RNDs within the range of available exercise prices. The resulting density is then truncated on both tails, and we only know the probability mass and conditional mean of each tail.

In this case, the PIT approach is applicable only to the truncated density, and the assessment of the forecasting power is confined to the body of the distribution. Since the probability integral transforms are obtained from a truncated, rather than an entire density, a modification of (21) becomes necessary. The "truncated" version of the PIT can be expressed as follows:[5]

$$z_t^* = \frac{P_t(y_t) - P_t(X_{mn,t})}{P_t(X_{mx,t}) - P_t(X_{mn,t})}, \qquad X_{mn,t} \leq y_t \leq X_{mx,t} \tag{23}$$

where $X_{mn,t}$ and $X_{mx,t}$ denote, respectively, the minimum and maximum strike available at time t on a certain contract and $P_t(.)$ is the value of the cumulative distribution associated with the density forecast. As before, if the estimate of the truncated RND function coincides with the actual process followed by the underlying at expiry within the range of available strikes, the series of z_t^* is i.i.d. U(0,1).

However, in this context, the PIT technique provides a useful instrument for judging the forecasting performance of RNDs only in the body of the distribution. To guarantee a complete evaluation of the forecasting power of the informational content that can be extracted from traded option prices, we need to integrate the PIT analysis with testing procedures applicable to the tails of the distribution. Since we only possess information about the probability mass below the minimum strike and above the maximum strike, the natural testing techniques seem to be the ones used for evaluating probability forecasts.[6]

C. Testing the Tails of the RND Functions

The most common measures of accuracy for probability forecasts are the so called "scoring rules", based on the distance between the probability forecast P_t formulated at time $t\text{-}1$ for an event at time t (in our case, the forecasted probability mass in a given tail) and a binary variable R_t which assumes value of 1 if the event occurs (specifically, if the actual realisation of the underlying falls in the tail), and value of 0 otherwise.

Amongst the various "scoring rules", we chose to work with the Brier's (1950) quadratic probability score, also simply called the Brier score:

$$B_n = \frac{1}{n}\sum_t 2(P_t - R_t)^2 \tag{24}$$

Clearly, B_n assumes values between 0 and 2 and more accurate probability forecasts are reflected in smaller values for the score. To assess whether B_n departs significantly from its expected value $\sum_t P_t(1-P_t)$, Seillier-Moiseiwitsch and Dawid (1993) suggest the test statistic:

$$Y_n^B = \sum_t (1-2P_t)(R_t - P_t) \Big/ \left[\sum_t (1-2P_t)^2 P_t(1-P_t)\right]^{1/2} \tag{25}$$

which is asymptotically distributed as a standard normal.[7]

In what follows the combination of PIT approach and techniques for probability forecasts evaluation is adopted to assess the quality of not only RNDs extracted via non-parametric methods (B-splines), but also those modelled according to a parametric functional form (generalised beta, NIG and mixture of lognormals). For the latter, this "combined" analysis is additional to the application of the PIT approach on the entire RND, and consents a more immediate comparison with the non-parametric specification, in terms of forecasting performance.

D. Testing the Body of the RND Functions

In the present work, various alternative techniques suggested by the relevant literature have been implemented to test the series of both truncated and entire probability integral transforms. Since we restricted our attention to non-overlapping data, the sample size of the PIT series is very small, equal to 60 observations. Many studies (Berkowitz (2001) and Noceti, Smith and Hodges (2000)) have shown that for such a small sample size, most test statistics generally display very little statistical power to reject the density forecast under the null when it is false. Therefore, we think that the comparison of findings obtained from several alternative testing procedures might be more informative on the robustness of the results themselves.

Testing the uniformity of z_t series

We start by computing empirical distribution function (EDF) test statistics, based on the vertical difference between the EDF and the theoretical DF of the null. Let $z_{(1)} < z_{(2)} < \ldots < z_{(n)}$ be the values of the z_t series (both truncated and non truncated) arranged in ascending order. Some measures are computed as the supremum of that vertical difference: the most famous amongst these measures is the Kolmogorov-Smirnov statistic, introduced by Kolmogorov (1933):

$$D = \max\left(D^{+}, D^{-}\right) \tag{26}$$

where: $D^{+} = \max_i \left\{i/n - z_{(i)}\right\}$; $\quad D^{-} = \max_i \left\{z_{(i)} - (i-1)/n\right\}$

A similar measure is the Kuiper statistic (Kuiper (1962)), defined as:

$$V = D^{+} + D^{-} \tag{27}$$

Many other tests are based on a quadratic measure of the difference between the EDF and the distribution function under the null. Two of the most common statistics belonging to this class are the Watson U^2 statistic, introduced by Watson (1961) and the Anderson-Darling (1952) A^2 test, which are respectively defined as follows:

$$U^2 = \sum_i \left\{z_{(i)} - (2i-1)/(2n)\right\}^2 + 1/(12n) - n(\bar{z} - 0.5)^2 \tag{28}$$

(where $\bar{z} = \sum_i z_i / n$)

$$A^2 = -n - (1/n)\sum_i (2i-1)\left[\log z_{(i)} + \log\left\{1 - z_{(n+1-i)}\right\}\right] \tag{29}$$

Comparative studies of the statistical performance of various EDF tests (D'Agostino and Stephens (1986) and Noceti, Smith and Hodges (2000)) suggest that D and A^2 are particularly appropriate for the identification of errors in the mean, U^2 and V are more appropriate for misspecifications in the variance, whereas A^2 shows more power when the forecasted distribution departs from the true frequency distribution in the tails. Since our aim is to detect generic deviations from the null hypothesis of simple uniformity of the z_t series, we need to focus on test statistics that possess good power against a wide range of alternatives. D'Agostino and Stephens (1986) identify Watson's test as the best amongst the EDF statistics for this purpose. They also recommend the Neyman-Barton smooth test of order 2 (N_2) as a general omnibus test for uniformity. This statistic, based on likelihood ratio methods, is defined as:

$$v_j = 1/\sqrt{n}\sum_i \pi_j(z_i), \tag{30}$$

$$\pi_1(z) = 2\sqrt{3}y;$$

$$\pi_2(z) = \sqrt{5}(6y^2 - 0.5),$$

$$y = z - 0.5$$

$$N_2 = \sum_{j=1}^{2} v_j^2$$

and asymptotically distributed as $\chi^2(2)$.

Testing the normality of x_t series

Berkowitz (2001) shows that most of the testing procedures described above are not powerful enough for small samples. Therefore, we apply the inverse probability transformation in (22) to the z_t series. The resulting x_t series should be i.i.d. N(0,1) if the RNDs coincide with the actual distribution of the underlying asset. Both graphical methods and formal goodness-of-fit tests can be implemented to test this hypothesis.

Berkowitz suggests to test the null hypothesis against a first-order autoregressive alternative given by:

$$x_t - \mu = \rho(x_{t-1} - \mu) + \varepsilon_t \tag{31}$$

such that the null of i.i.d. (0,1) translates into $\mu = 0$, $\rho = 0$, and var(ε_t) = $\sigma^2 = 1$. He proposes the use of log-likelihood ratio tests. The log-likelihood function associated with the model in (31) presents the following formulation:

$$L(\mu,\sigma^2,\rho)=-\frac{1}{2}\log(2\pi)-\frac{1}{2}\log[\sigma^2/(1-\rho^2)]-\frac{(x_1-\mu/(1-\rho))^2}{2\sigma^2/(1-\rho^2)}+$$

$$-\frac{n-1}{2}\log(2\pi)-\frac{n-1}{2}\log(\sigma^2)-\sum_{t=2}^{n}\left(\frac{(x_t-\mu(1-\rho)-\rho x_{t-1})^2}{2\sigma^2}\right) \quad (32)$$

In order to test the hypothesis of independence across the observations, the likelihood ratio test can be expressed as:

$$LR1=-2\left(L(\hat{\mu},\hat{\sigma}^2,0)-L(\hat{\mu},\hat{\sigma}^2,\hat{\rho})\right)\sim\chi^2(1) \quad (33)$$

Similarly, the likelihood ratio (LR) test for the joint hypothesis of independent observations with zero mean and unit variance, can be formulated as:

$$LR2=-2\left(L(0,1,0)-L(\hat{\mu},\hat{\sigma}^2,\hat{\rho})\right)\sim\chi^2(3) \quad (34)$$

Although the LR tests seem adequate for our problem since they should possess good statistical power against general alternatives also for small sample sizes, they implicitly maintain the assumption of normality of the x_t series, instead of verifying it explicitly.

The normality of the transformed probability series is then assessed via the Jarque-Bera and the Doornik-Hansen tests. The Jarque-Bera statistic (1980), probably the most common test for normality, is formulated as:

$$JB=n\left[\beta_1/6+(\beta_2-3)^2/24\right]\sim\chi^2(2) \quad (35)$$

where $\sqrt{\beta_1}$ is the sample skewness and β_2 the sample kurtosis.

Arguing that the statistics $\sqrt{\beta_1}$ and β_2 are not independent, except for very large sample sizes, Doornik and Hansen (1994) propose an alternative test for normality based on transformed measures of skewness (z_1) and kurtosis (z_2), expressed as:

$$DH=z_1^2+z_2^2\sim\chi^2(2) \quad (36)$$

At completion of our testing experiments, we have also carried out diagnostic tests on the single parameters of the autoregressive model in (31), which should be more informative on the nature of the violations of the assumptions made, when rejection of the joint null hypothesis occurs. First we estimate the parameters of the model using OLS, and then we apply *t*-statistics to test the individual hypothesis $\mu = 0$ and $\rho = 0$, as well as chi-square statistics to test $\sigma^2 = 1$.[8]

To absence of both linear and non-linear dependence in the series of modified probability transforms (both from the truncated and the entire RNDs) has been ascertained by plotting the correlograms of x and x^2 series.[9] Even though such analysis can provide some intuition on potential misspecification of the dynamics of the forecasts, in general there is no one-to-one correspondence between the dependence pattern shown by the x series and that of the underlying. However, the study of the dependence in x is important also because the distribution theory for many of the conventional goodness-of-fit tests rests on an assumption of independence of the variable, which should be tested *a priori*.

IV. Data Sources

The estimates of the RNDs are obtained from quarterly prices of call and put options on:

- Standard & Poor's 500 index future, for the period March 86 – September 2001;
- US Dollar/British Pound, for the period March 86 – September 2001;

We chose to work with quarterly data to ensure non-overlapping observations and to have as many strike prices as possible. Specifically, we identified all expiration dates for the quarterly expiration cycle during this period (March, June, September, December). On each expiration date, we recorded the settlement levels of the futures contract expiring on that day, the futures contract with exactly three months to expiration and all available option prices on this three-month futures contract[10]. Since the options are American style, the Barone-Adesi and Whaley (1987) approximation has been used to recover the implied volatilities, which have then been plugged back into the Black-Scholes formula to calculate the pseudo-European option prices needed to compute the RNDs.

As is standard, all options prices which traded at the minimum level at the relevant market or which allowed a butterfly arbitrage were excluded (see Jackwerth and Rubinstein

(1996)). Furthermore, to reduce the potential problem of nonsynchronous prices for the options and underlying futures, only those implied volatilities from the available out-of-the-money (OTM) option contracts (not admitting arbitrage) were utilised. Bates (1991) and Gemmill (1996) have shown that much greater deviations occur in the implied volatilities for in-the-money (ITM) options relative to the OTM options. They suggest that this is due to futures and ITM option prices not being recorded simultaneously. Thus, if the strike price was equal to or below the underlying futures price, put options were examined; otherwise call options were examined.[11] The interest rate inputs were obtained from the Federal Reserve Bank in New York (US Dollar Treasury Bill rate).

With this available data, we then estimated the RND from the options prices and then compared the probability density forecast to the actual realised underlying futures price that occurred in three month's time. Given we examined fifteen years of data on a quarterly basis, we had 60 observations in our analysis.

V. Results of the RND Estimation And Forecastability

A. Discussion of RND Estimates

Employing four alternative methods, three parametric and one non-parametric, we have estimated the RNDs for options in the S&P 500 and the US Dollar/British Pound from 1986 to 2001.

Table I contains summary statistics for the time series of realised quarterly log-returns and for the forecasted densities of the log-returns both extracted from option prices and simulated from GARCH-type models for the volatility (as discussed in section VI). The first part of the table refers to the currency contract and the second half to the index contract. For this latter contract, the statistics have been computed for both unadjusted RNDs and RNDs adjusted through Girsanov's transformation to account for the risk premium in the equity market, as described in subsection B.2. To provide a consistent comparison between moments of the observed point realisations and those of the density forecasts, we proceeded as follows. Sample conventional measures of location, dispersion, skewness and kurtosis have been computed on realised log-returns. Equivalent statistics have then been calculated for the distribution obtained as an equally weighted mixture of the single density forecasts estimated for each quarter. This method seems more appropriate than simply averaging the summary statistics of the individual densities across time. However, simple averages of time series of the statistics have been computed as well, and the average values for skewness and kurtosis are reported in the last two columns of the table.[12] The higher absolute figures obtained for skewness and fat-tailedness for the mixture can be explained with Jensen's inequality: since higher statistics are computed as ratios of moments, the average of those ratios (method of time series) is smaller than the ratio of those average moments (mixture method).

The RNDs of the currency options display very similar moments, whatever specific functional assumption is made. The distributions exhibit slightly negative skewness and moderate excess kurtosis. As expected, these moments underestimate the sample moments of the observations, when the September 92 (EMU ejection of the British Pound) is included in the sample. When this event is excluded, the kurtosis is estimated quite closely, whereas the forecasted and realised skewnesses exhibit opposite sign.

For the S&P 500 options, implied NIG densities turn out to be more negatively skewed and much more leptokurtic than the alternative specifications. The Generalised Beta distribution shows instead the smaller values for the higher moments, however larger than the corresponding moments recorded for the currency RNDs, as expected. Not surprisingly, the mean of unadjusted RNDs is smaller than the sample mean since it does not account for the risk premium. The values for dispersion and higher moments instead exceed the sample counterparts, supporting the generally accepted conclusion that implied densities overestimate the skewness and fat-tailedness observed in equity markets.

For illustrative purposes, Figures 1 and 2 show plots of some pre- and post- October 87 crash implied RNDs for both the US$/BP and the S&P 500 options, under the NIG specification.[13] The densities implied from the currency options look very similar for both periods, almost symmetric and slightly fat-tailed. As expected, the shape of the S&P 500 implied distributions change substantially after the crash, becoming remarkably negatively skewed and more fat-tailed.

Figures 3 to 6 report time series of the skewness and excess kurtosis for the index and currency RNDs. For the implied RNDs on the index, higher moments become much more volatile after the 87 crash. The NIG densities display larger values of negative skewness and excess kurtosis, which are also the most volatile. Less pronounced and volatile higher moments characterise the other models, in particular the mixture of lognormals. For the currency options, the three parametric models display very similar values for the higher moments, which are much less pronounced and less volatile through time than those recorded for the index options. Table II displays some summary statistics for the parameters of the RNDs estimated according to the three parametric specifications.

As mentioned before, we have only estimated the truncated density for the non-parametric specification. To compare the non-parametric to the three parametric approaches, we have truncated the mixture of the lognormals, the NIG and the Generalised Beta at the lowest and highest strikes. As an illustrative example, Figures 7 and 9 show the implied volatility smiles, as produced by the B-spline approximation, for the S&P 500 and the US$/BP options on the 18th of June 1992 and 4th of June 1992 respectively. Figures 8 and 10 provide the corresponding truncated RNDs for all the four estimation methods. For the currency, the four truncated RNDs almost coincide. They differ for the S&P 500 options. The implied NIG and GB2 densities are almost identical to the B-spline approximation, whilst the mixture of lognormals substantially differs.

In comparing the four alternative specifications, we remind the reader that the obvious disadvantage of the B-spline approximation approach is that it does not recover the tails of the RND outside the range of the available strikes. On the other hand, it provides greater flexibility. No assumptions need be made for the underlying asset distribution and there is a better fit to the implied volatility structure; however, with a danger of overfitting. In contrast to the non-parametric approach, the three parametric specifications assume that the RND belongs to a distribution family. Amongst the parametric distributions, we express preference for the NIG and, to a slightly less extent, to the generalised beta, for their flexibility, the small number of parameters to be estimated and, especially for the NIG, the excellent fit to the implied volatility structure.[14] Despite its computational tractability and greater accuracy in the estimation of its parameters, the mixture of the lognormals is not as flexible as the alternative specifications.

B. Are Implied RNDs Good Predictors? A Discussion of the Results.

Tables IIIa and IIIb present the results from the tests conducted on the tails outside the range of available strikes. For the right tail, the left tail and the combination of both tails we have compared the frequency with which actual observations fall in those areas with the probability mass assigned by the B-splines and the truncated versions of the parametric distributions to the tails. We have also computed the test statistic based on the Brier's score Y_n^B at 5% and 1% confidence level.

Tables IVa and IVb display the results on the goodness-of-fit tests for uniformity of the probability integral transforms z_t, computed for each contract type and both parametric and non-parametric specifications for the RNDs. The critical values (at both 5% and 1% confidence level) for D, A^2, U^2 and V were taken from D'Agostino and Stephens (1986).

The results of the tests on the modified PIT (after a transformation to normality) for all the estimation methods are shown in Tables Va and Vb. The first two columns report the *t*-tests on the parameters of the model in (31), with p-values in brackets. The third column displays the chi-test on the variance. The two normality statistics and the two log-likelihood ratio tests follow.

Some Q-Q plots of the empirical cumulative distribution of the probability integral transforms z_t against their theoretical cumulative density (45° line) are also plotted in Figures 11-14 for both contracts, under the NIG specification. The purpose is to provide a visual and more intuitive tool to facilitate understanding of how the estimated implied RNDs differ from the actual distribution of the underlying, when rejection of the null hypothesis occurs.

B.1. RNDs from Currency Options

When the implied RNDs for options on the US dollar/British Pound contract are defined only within the range of strikes (B-spline and truncated version of the parametric forms), the probability forecast for the right tail is rejected. More specifically, the figures in Table IIIa indicate that the forecasted probability mass in the right tail is a downward biased estimate of the frequency with which the observed values for the underlying exceed the maximum strike. The hypothesis that the body of the distribution well represents the corresponding portion of the actual distribution cannot instead be rejected for the truncated versions of the parametric models. In Table IVa, a rejection occurs for the B-spline, as suggested by the U^2, V, and N_2 statistics, which could be caused by a bias in the variance, given the low value for the estimated variance (0.6776) [see Table Va].[15] The good forecasting performance of implied RNDs modelled with parametric, non-truncated, specifications can never be rejected according to distributional tests for uniformity. However, when the transformation in (31) is applied, the normality of the resulting series is rejected by Jarque-Bera and Doornik-Hansen tests, as a consequence of the misspecification in the tails discussed before. The non-normality of the transforms for the (entire) parametric distributions is also observable from the visual inspection of the Q-Q plot in Figure 11. This suggests that the previous failure to reject was probably due to the low statistical power of uniformity tests rather than to the good fit of the implied RND to the actual density.

B.2. RNDs from Index Options

The unadjusted battery of tests conducted on the S&P 500 options seem, at first sight, to lead to quite different conclusions. In Table IIIa, the tests on the tails for both the B-spline and the truncated parametric distributions seem to indicate that the probability mass left in both tails represents a good forecast of the actual frequency with which realisations of the underlying at expiry fall into those tails. On the contrary, the values assumed by D, A^2, N_2 (in Table IVa) lead to a rejection of the predictive performance of the body of the implied RNDs for all specifications, parametric and non-parametric.[16] A deeper investigation via diagnostic tests confirms the suspicion that the principal reason for rejecting is a bias in the mean,[17] with the addition of a bias in the variance for the B-spline. This result is not surprising, given the historical evidence of the presence of a risk premium in equity indices which implies that the mean of the risk-neutral distribution always understates the mean of the actual distribution. We propose two measures of adjustment, in an attempt to remove the risk-premium effect from the actual realisations for the underlying.

A simple mean-adjustment for the risk premium

The first adjustment consists of a simple shift in the mean of the distribution. Let:

$$AR = \frac{1}{n}\sum_{t=1}^{n}\frac{O_t}{F_t} \qquad (37)$$

be the average value of the ratios of the actual observation (in three months time) to the forward price observed at the point in time when the RND was estimated (at the beginning of the period).

The "adjusted" realisations, corrected for the risk premium, are computed as: $O_t^* = O_t \cdot \frac{1}{AR}$.[18]

This type of adjustment possesses the nice property of keeping the actual mean of the distribution equal to the forward price, without affecting the volatility, which remains unchanged. This feature turns out to be particularly desirable when, as in our case, we only want to correct for a bias in the mean. All the tests have then been recomputed on the adjusted observations for the underlying and the same implied RNDs as before. The findings are now in line with those for the currency options. The bias in the mean has been removed. Again the right tail is misspecified, as indicated by the tests of the tails for the truncated distributions,[19] by the rejection of normality assumption for the entire parametric specifications, displayed by the normality tests, and by the Q-Q plots (Figure 13). In particular, the probability assigned to the tails overestimates the actual frequency of observations in these areas. Most of the relevant test statistics also agree in detecting a bias in the variance, which turns out to be significantly smaller than one.

A more rigorous adjustment

Following Bliss and Panigirtzoglou (2001), the second adjustment is an application of Girsanov's theorem and requires specific assumptions on the representative investor's utility function $U(S_T)$. It is well known that when markets are complete and frictionless and a single risky asset is traded, the subjective density function $q(S_T)$ can be related to the risk-neutral density function $p(S_T)$ through:[20]

$$q(S_T) = \frac{\dfrac{p(S_T)}{U'(S_T)}}{\int \dfrac{p(x)}{U'(x)} dx} \tag{38}$$

Common choices for $U(S_T)$ are the:

power utility function:
$$U(S_T) = \frac{S_T^{1-\gamma} - 1}{1-\gamma} \tag{39}$$

exponential utility function:
$$U(S_T) = -\frac{e^{-\gamma S_T}}{\gamma} \tag{40}$$

Both utility functions depend on a single parameter γ. The power specification constitutes a sensible choice to make given its analytical tractability and its constant relative risk aversion (RRA), measured by the parameter γ. A utility function of the exponential type is the specification assumed in the application of the popular Esscher transform. However, it has constant absolute risk aversion, whilst the relative risk aversion is time varying, depending on both γ and the realisations for S_T. Given that the specific parametric forms hypothesised for the RNDs do not support an exponential transformation, and that we give preference to constant RRA utility functions, we have chosen to work with the power utility function.

Therefore, the transformation in (38) has been applied to the time series of all the RNDs obtained under our three parametric models, for different choices of the parameter γ. Summary statistics for the resulting risk-adjusted density functions for values of γ equal to 1 and 2 are displayed[21] in Table I. As expected, the adjustment had the effect of pushing the mean of the implied densities towards the observed one as well as reducing the gap between implied and sample higher moments, which remain, however, overestimated.

All the tails and distributional tests have then been re-run on the risk-adjusted densities, and the outcomes are reported in Tables IIIb, IVb, and Vb. Our findings indicate that we can no longer reject, at 5% confidence level, the hypothesis that the risk-adjusted density functions represent unbiased/efficient forecasts for the actual distribution of the underlying at expiry. The only exception is given by the truncated version of the adjusted mixture of lognormal density, with $\gamma = 1$, which is rejected by LR1 and LR2 tests. These results clearly denote an improvement in the quality of the density forecasts consequent upon an appropriate adjustment of the original RNDs to account for the risk premium. However, we suspect that the inability to reject is largely due to the low statistical power of our testing procedures for the small sample size considered.

We conclude this because Bliss and Panigirtzolou (2001) found that the parameters for the power (and exponential) utility function varies as a function of the levels of the at-the-money (ATM) volatility and also varies depending upon the term to expiration of the option.[22] This would suggest that splitting the available data to test these results would be warranted. However, given that we have already utilised all available non-overlapping data, it would probably be infeasible to split the data to test these additional hypotheses: we would most probably not have enough observations to prove or disprove the additional hypotheses.

Our main results can therefore be summarised as follows. For the contracts under examination, unadjusted implied RNDs extracted from option prices do not represent an appropriate forecast of the true distribution of the underlying at expiry, whatever functional form (parametric or non-parametric) is chosen to model them. Not surprisingly, the main reason for rejecting the predictive ability of RND is given by misspecification of the tails outside the range of strike, on which very little information is available. The second reason (however, not common to all specifications) seems to be a bias in the variance, i.e. an underestimation of the real variance of the actual process. The worst specification turns out to be the B-spline, which tends to overfit the actual option prices. The mixture of lognormals, the NIG and the generalised beta exhibit very similar forecasting performances, only slightly better for the latter under the majority of the tests implemented. However, parametric methods are clearly superior to the B-spline, perhaps suggesting that parametric specifications are more appropriate for the purpose under investigation.

The rejection of the predictive ability of the implied RNDs is stronger for the S&P 500 options. This is not altogether surprising, given the higher risk premium in this market. However, things change when an adjustment for the risk premium is introduced. If a simple mean-adjustment has the main effect of shifting the bias from the mean to the tails, a more rigorous adjustment in the shape of the original RND translates into a sensible improvement of the predictive ability of the resulting adjusted distribution. This indicates that not only unadjusted RNDs from index options should not be used to forecast the actual underlying densities, but also that particular attention should be paid to the specific methods employed to adjust to account for the risk-premium.

As far as the testing techniques are concerned, our findings provide evidence of the usefulness of a complete battery of distributional tests to assess the predictive ability of implied RNDs. Given the small size of the sample of observations and estimated RNDs, some tests might lack statistical power. Therefore, a more robust and informative appraisal can be carried out by implementing a whole set of complementary techniques.

VI. Assessing The Relative Performance Of RNDs: Comparison To Historical-Based Density Forecasts

Even though RNDs fail to provide unbiased and efficient forecasts in absolute terms, it could well be that they represent the best we can do in comparison with alternative ways of producing the density forecasts. It is therefore of interest to assess the performance of forecasts made on an entirely different principle.

In order to investigate this, we have used a GARCH-type model to produce distributional forecasts for quarterly returns and then assessed their performance using the same methods as before. Working with both the S&P500 and the US$/BP data, we estimated various GARCH specifications. The estimates were made using only data available at the time the forecast was made.

As customary, we assume that the evolution of the (demeaned) log-returns h_t follows the conditional process:

$$h_t = \sigma_t \varepsilon_t \tag{41}$$

where ε_t is an i.i.d. process with zero mean and unit variance and the conditional volatility σ_t is a time varying, positive and measurable function of the information set at time $t-1$. Regarding the particular process that describes the conditional variance σ^2_t, we first investigate a simple GARCH(1,1) specification as suggested by Bollerslev (1986):

$$\sigma_t^2 = \alpha_0 + \alpha_1 h_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{42}$$

where $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ and $\alpha_1 + \beta_1 < 1$ to ensure stationarity. A vast literature (see, e.g., Hsieh (1989), Baillie and Bollerslev (1989)) indicates that the normal specification, originally employed to model the conditional distribution, does not account for the significant fat-tailness exhibited by the empirical data. Therefore, following Baillie and Bollerslev (1989), we have also chosen to use a Student's t distribution for the error.

When modelling the standardised returns on the index, the leverage effect observed in the data (for which positive and negative shocks have a different impact on the volatility) must also be taken into account. Several variants of GARCH models that introduce asymmetry have been proposed by the relevant literature (TARCH, EGARCH, GJR, etc.) and many studies provide evidence of their superior performance at describing equity time series (see Brailsford and Faff (1996), Loudon, Watt and Yadav (2000)). After attempting different models,[23] we have opted for a very simple asymmetric GARCH that seems to guarantee robust estimates and reliable forecasts. The conditional variance is modelled according to the following process:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \left(h_{t-1} - \kappa_1\right)^2 + \beta_1 \sigma_{t-1}^2 \tag{43}$$

where κ_1 represents the asymmetry parameter, which models the leverage effect for positive values.

At the beginning of each quarter for the time period of interest we have estimated the model in (42) with both normal and Student's t errors and the model in (43) with Student's t errors[24] over the past series of daily returns. The series of past returns employed for the estimation include daily returns from April 82 to the beginning of the quarter under exam. Therefore the length of the estimation window increases as we consider more recent quarters. We chose to fit our models over long return series in order to minimise the impact of extraordinary high or low volatility periods on the estimation process and, therefore, to obtain robust and stable estimates for the coefficients which constitute an essential basis for reliable forecasts.

Average values for the time series of the coefficient estimates, together with their dispersion figures, are displayed in table VI. In line with the general findings for these models, the sum of the AR and MA parameters is close to one, and the GARCH lag coefficients are large, indicating that shocks to conditional variance die out very slowly. The average figures for the degrees of freedom suggest that conditional fat-tailed distributions are very close to the boundary between finite and infinite kurtosis (which characterises Student's t with four or less degrees of freedom). As expected, the sign of the asymmetry coefficient is positive and, therefore, consistent with the leverage effect observed in equity markets. Perhaps with the exception of the number of degrees of freedom, there is little variation amongst the estimates.

At the beginning of every quarter, the estimated coefficients are then used to produce recursive daily forecasts of the returns over the following three months according to:

$$\hat{r}_{t+1} = \mu + \hat{\sigma}_{t+1}\varepsilon_{t+1} \tag{44}$$

with

$$\hat{\sigma}_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 h_t^2 + \hat{\beta}_1\sigma_t^2 \tag{45}$$

or

$$\hat{\sigma}_{t+1}^2 = \hat{\alpha}_0 + \hat{\alpha}_1 (h_t - \hat{\kappa}_1)^2 + \hat{\beta}_1\sigma_t^2 \tag{46}$$

where ε_{t+1} are random numbers generated from a standard normal (for the GARCH model only) and from a standardised Student's t (for both the GARCH and the asymmetric GARCH models) with the estimated degrees of freedom.[25] The daily forecasts are then added up together to compute the three-month return forecast and the process is repeated 10,000 times to obtain a distribution for the simulated returns.

Summary statistics for mixtures of these simulated quarterly distributions are shown in Table I, and contrasted to the corresponding statistics for implied RNDs and observed log-returns. For the currency contract, the GARCH specification with fat-tailed errors produces moments similar to the implied RNDs' ones. However, statistics closer to the empirical moments are obtained from the GARCH model with normal errors. For the S&P 500 contract, the most striking feature is the gap between the kurtosis of the mixture distribution and the average kurtosis from the time series of the distributional forecasts. In fact, the mixtures built on any of the GARCH specifications exhibit higher fat-tailness than both the empirical sample and the mixtures obtained from the implied RNDs.[26] In contrast, only the asymmetric GARCH specification with t errors enables us to produce a mixture density with a significant negative skewness. The degree of asymmetry however, not only is far smaller than the asymmetry of the implied density counterpart, but also underestimates the asymmetry observed in the data.

In order to test how closely our simulated distributions for the log-returns fit the actual distribution, we have run the entire battery of tail and distributional tests on the density forecasts from the various models. The results are reported in tables VII-IX.

The most appropriate specification for the currency contract seems to be the GARCH(1,1) with fat-tailed errors.[27] As for the implied RNDs, we record a misspecification of the right tail, due to an underestimate of the probability of the observation falling beyond the maximum strike. However, in contrast to the implied forecasts, the good forecasting performance of both entire and truncated versions of the simulated distributions can never be rejected according to distributional tests for uniformity and normality. These results suggest that the forecasts obtained by simulating from a simple GARCH(1,1) process with Student's *t* errors do a better job in predicting the actual distribution of returns than the distributional forecasts extrapolated from option prices.

The best choice for the log-returns on the index contract amongst our alternatives is the asymmetric GARCH with fat-tailed conditional distribution. As for the currency contract, the findings from our test statistics suggest that the distributional forecasts produced according to this model outperform the (unadjusted) implied risk-neutral density forecasts. When the distribution is split into tails and body, and each component tested separately, the goodness-of-fit is never rejected, even though the results of the tail tests might indicate a misspecification in the right tail. Rejection occurs only for the entire distribution from the Watson's statistic and from the normality tests on the transformed PIT series.

In this section we have assessed a variety of GARCH models estimated from historical daily returns. We found that although some methods of implementing GARCH forecasts did not produce robust results, some quite simple specifications could provide better forecasts that risk-neutral densities implied from option prices. Our finding that the historically estimated forecasts usually dominate the forecasting performance of the RNDs supports the conclusion that these RND forecasts seem not to be efficient.

VII. Conclusions And Suggestions For Further Research

A significant recent literature considers the risk neutral density function (RND) associated with options prices. An informationally efficient options market, with sufficiently low risk premia, could provide more precise information than other financial indicators about future levels of underlying market prices and the amount of probability attributed to any given realisation. Bahra (1996, 1997), McManus (1999) among others present the theoretical argument for this relationship and suggest that RNDs could be used by financial regulators to avert future market turbulence. Whether or not RNDs actually provide such information is the empirical issue considered here. Previous research on this question has tended to rely upon single extreme events. The general conclusion is that for stock markets, RNDs do not provide useful information about future market distributions. Evidence for foreign exchange and interest rate markets is mixed. Some studies find antecodal evidence for useful information, while others do not. In this research we have asked whether or not such RNDs provide an unbiased forecast of the distribution of future returns for options on the S&P 500 and British Pound/US Dollar.

Using a variety of methods to determine RNDs, we find that pure RND forecasts are biased estimators of realised density functions. As with recent work by Weinberg (2001) and Bliss and Panigirtzolou (2001), all the moments of the RNDs are biased estimates of the realised moments of the probability density function. The major source of error for the S&P 500 was initially found to be the mean value. As we may have an inappropriate drift adjustment (under the risk neutral measure), the substitution of the actual mean of the density, corrects this error and shifts the error to the higher moments. In both cases, we reject the hypothesis that RNDs for either the S&P 500 or the currency are unbiased forecasts of the actual probability density function. Only after adjusting the shape of the original RNDs for the S&P 500, through a Girsanov's transformation, are we unable to reject the hypothesis that the resulting subjective, risk-adjusted densities are unbiased. This is also similar to findings of Bliss and Panigirtzolou (2001). However, it would be dangerous to conclude that an inability to reject the hypothesis of efficient forecast ability of the subjective distribution should be interpreted as acceptance of the subjective distribution as a "good" forecast. In our tests, the statistical tests are close to significance levels that reject the efficient forecast hypothesis. In addition, the results of Bliss and Panigirtzolou (2001) reject their own assumption of a stationary utility function. Therefore, we conclude that instead of accepting the hypothesis that the utility adjusted RND is an efficient forecast of realised density functions, we are merely unable to reject it. At the very least, there is little disagreement that the RND density function is a poor estimate *prima facia* of the realised density functions.

In support of our conclusion, we have also found that the forecasting performance of distributions simulated from GARCH-type specifications for the volatility process is usually superior to that of implied RNDs. Therefore also in relative terms, RNDs do not appear to be efficient predictors of the true densities.

The results of this study, that simple RNDs provide only biased information about future market dynamics, should be interpreted in the context of a number of other studies that have reached similar conclusions in slightly different contexts. Gemmill and Saflekos (2000) show that RNDs are reactive and not predictive, and studies of the predictive content of implied volatilities (such as Canina and Figlewski (1993), Christensen and Prabhala (1998) and more recently Hansen, Christensen and Prabhala (2001)) universally estimate them to be biased predictors of future volatility, even where some information content is found. In general it seems that implied volatilities look backwards more than forwards, and typically respond to past shocks, providing exaggerated forecasts of the future.

These considerations lead us to concur with Shiratsuka (2001) that the utility of RNDs for financial regulators is extremely limited. The use of pure, unadjusted RNDs as indicators of future market turbulence could be counterproductive. Such information would not indicate to regulators when market turbulence was imminent. If recent market price dynamics were subdued, it appears that the RNDs would be benign, provided no indication of future market turbulence. If such market turbulence subsequently occurred, options markets would react to this with the RNDs assuming density functions consistent with this shock. Therefore, if regulators reacted to RNDs as though they were predictive, it could be that after the occurrence of a market crisis, they would assume further crises would occur and act accordingly. Such reaction may be counter productive if actions taken by regulators send signals to the market that further market turbulence were expected. Our research suggests that RNDs do not provide such unbiased forecasts of future market turbulence and are unsatisfactory indicators of such events.

REFERENCES

Abadir, Karim, and Michael Rockinger, 1997, Density-Embedding Function, University of York, Working Paper.

Abken, Peter A., Dilip B. Madan, and Sailesh Ramamurtie, 1996, Estimation of Risk Neutral Densities by Hermite Polynomial Approximation: With an Application to Eurodollar Future Options, Federal Reserve Bank of Atlanta, Working Paper 96-5.

Aït-Sahalia, Yacine, and Andrew W. Lo, 1998, Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, *Journal of Finance* 52, 499-548.

Aït-Sahalia, Yacine, and Andrew W. Lo, 2000, Nonparametric Risk Management and Implied Risk Aversion, *Journal of Econometrics* 94, 9-51.

Anderson, Thomas W., and David A. Darling, 1952, Asymptotic Theory of Certain Goodness-of-Fit Criteria Based on Stochastic Processes, *Annals of Mathematical Statistics* 23, 193-212.

Aparicio, Silio D., and Stewart D.Hodges, 1998, Implied Risk-Neutral Distribution: A Comparison of Estimation Methods, FORC Preprint 1998/95, University of Warwick.

Bahra, Bhupinder, 1996, Probability Distributions of Future Asset Prices Implied by Option Prices, *Bank of England Quarterly Bulletin*, 299-311.

Bahra, Bhupinder, 1997, Implied Risk-Neutral Probability Density Functions from Options Prices: Theory and Application, Bank of England, Working Paper.

Baillie, Richard, and Tim Bollerslev, 1989, The Message in Daily Exchange Rates: a Conditional-Variance Tale, *Journal of Business and Economic Statistics* 7, 297-305.

Barndorff-Nielsen, Ole E., 1997, Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling, *Scandinavian Journal of Statistics* 24, 1-13.

Barndorff-Nielsen, Ole E., 1998, Processes of Normal Inverse Gaussian Type, *Finance and*

Stochastic 2, 41-68.

Barone-Adesi, Giovanni, and Robert E. Whaley, 1987, Efficient Analytical Approximation of American Option Values, *Journal of Finance* 42, 301-320.

Bates, David S., 1991, The Crash of '87: Was it Expected? The Evidence from Options Market, *Journal of Finance* 46, 1009-1044.

Bates, David S., 2000, Post –'87 crash fears in the S&P 500 futures options markets, *Journal of Econometrics* 94, 181-238.

Beckers, Stan, 1981, Standard Deviations Implied in Option Prices as Predictors of Future Stock Price Variability, *Journal of Banking and Finance* 5, 363-382.

Bera, Anil K., and Carlos M. Jarque, 1980, Efficient Tests for Normality, Homoskedasticity and Serial Independence of Regression Residuals, *Economic Letters* 6, 255-259.

Berkowitz, Jeremy, 2001, Testing Density Forecasts with Applications to Risk Management, *Journal of Business and Economic Statistics* 19, 465-474.

Black, Fischer, and Myron Scholes, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637-654.

Bliss, Robert, and Nikolaos Panigirtzoglou, 2002, Testing the Stability of Implied Probability Density Functions, *Journal of Banking and Finance* 26, 381-422.

Bliss, Robert, and Nikolaos Panigirtzoglou, 2001, Option-Implied Risk Aversion Estimates: Robustness and Patterns, Federal Reserve Bank of Chicago, Working Paper 2001-15.

Bollerslev, Tim, 1986, Generalised Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics* 31, 307-327.

Bookstaber, Richard M., and James B. McDonald, 1987, A general distribution for describing security price returns, *Journal of Business* 60, 401-424.

Brailsford, Tim, and Robert Faff, 1996, An evaluation of volatility forecasting techniques, *Journal of Banking and Finance* 20, 419-438.

Breeden, Douglas T. and Robert H. Litzenberger, 1978, Prices of State Contingent Claims Implicit in Option Prices, *Journal of Business* 51, 621-651.

Brier, Glenn W., 1950, Verification of Forecasts expressed in terms of probability, *Monthly Weather Review* 78, 1-3.

Campa, José M., Kevin Chang, and Robert L. Reider, 1998, Implied Exchange Rate Distributions: Evidence from OTC Option Markets, *Journal of International Money and Finance* 17, 117-160.

Canina, Linda, and Stephen Figlewski, 1993, The Informational Content of Implied Volatilities, *The Review of Financial Studies* 6, 659-681.

Chang, Kevin, and William R. Melick, 1999, Background Note, in *Estimating and Interpreting Probability Density Functions: Proceeding of the workshop held at the BIS on 14 June 1999* BIS: Basle 1999.

Chiras, Daniel P., and Steven Manaster ,1978, The Information Content of Option Prices and a Test of Market Efficiency, *Journal of Financial Economics* 6, 213-234.

Christensen, Bent J., and Nagpurnanand R. Prabhala, 1998, The relation between implied and realized volatility, *Journal of Financial Economics* 50, 125-150.

Cnkovic, Cedomir, and Jordan Drachman, 1996, Quality Control, *RISK* 9, 138-143.

Cooper, Neil, 1999, Testing techniques for estimating implied RNDs from the prices of European and American options, *mimeo*, Bank of England.

Corrado, Charles J., and Tie Su, 1996, Skewness and kurtosis in S&P 500 Index returns implied by option prices, *Journal of Financial Research* 19, 175-192.

Corrado, Charles J., and Tie Su, 1997, Implied volatility skews and stock return skewness and kurtosis implied by stock option prices, *The European Journal of Finance* 3, 73-85.

Coutant, Sophie, Eric Jondeau, and Michael Rockinger, 2001, Reading PIBOR Futures Options Smiles: the 1997 French Snap Election, *Journal of Banking and Finance* 25, 1957-1987.

D'Agostino, Ralph B., and Michael A. Stephens, 1986, *Goodness-of-Fit Techniques* (Marcel Dekker Inc., New York).

Dawid, Philip A., 1984, Statistical Theory: The prequential approach, *Journal of the Royal Statistical Society, ser. A.* 147, 278-292.

Day, Theodore E. and Craig M. Lewis, 1988, The Behavior of the Volatility Implicit in the Prices of Stock Index Options, *Journal of Financial Economics* 22, 103-122.

Day, Theodore E. and Craig M. Lewis, 1992, Stock Market Volatility and the Information Content of Stock Index Options, *Journal of Econometrics* 52, 289-311.

De Boor, Carl, 1978, *A Practical Guide to Splines* (Springer-Verlag, New York).

De Raaij, Gabriela, and Burkhard Raunig, 2002, Evaluating Density Forecasts of Stock Market Returns, Bank of Austria, working paper n. 59.

Derman, Emanuel, and Iraj Kani, 1994, Riding on the Smile, *RISK* 7, 32-39.

Diebold, Francis X., and José A. Lopez, 1996, Forecast Evaluation and Combination, in G.S. Maddala and C.R. Rao, eds.: *Handbook of Statistics* (North-Holland, Amsterdam).

Diebold, Francis X., Todd A. Gunther, and Anthony S. Tay, 1998, Evaluating Density Forecasts, with applications to Financial Risk Management, *International Economic Review* 39, 863-883.

Doornik, Jurgen, and Henrik Hansen, 1994, An Omnibus Test for Univariate and Multivariate Normality, Nuffield College, Oxford University, working paper.

Dumas, Bernard, Jeff Fleming, and Robert E. Whaley, 1998, Implied Volatility Functions: Empirical Tests, *Journal of Finance* 53, 2059-2106.

Dupire, Bruno, 1993, Arbitrage Pricing with Stochastic Volatility, SORT Research Paper, Paribas Capital Markets.

Dupire, Bruno, 1994, Pricing with a Smile, *RISK* 7, 18-20.

Fisher, Ronald A., 1930, Inverse Probability, *Proceedings of the Cambridge Philosophical Society* 36, 528-535.

Gemmill, Gordon, 1996, Did Option Traders Anticipate the Crash? Evidence from Volatility Smiles in the UK with US Comparisons, *Journal of Futures Markets* 16, 881-897.

Gemmill, Gordon, and Apostolos Saflekos, 2000, How Useful are Implied Distributions? Evidence from Stock Index Options, *The Journal of Derivatives* 7, 83-98.

Hansen, Charlotte S., Bent J. Christensen, and Nagpurnanand R. Prabhala, 2001, Accounting for the Overlapping Data Problem in the Implied-Realized Volatility Regression, Working Paper, University of Aarhus, Denmark.

Harvey, Campbell R., and Robert E. Whaley, 1992, Market volatility prediction and the efficiency of the S&P 100 index option market, *Journal of Financial Economics* 31, 43-73.

Heynen, Ronald, Angelica Kemna and Tom Vorst, 1994, Analysis of the Term Structure of Implied Volatilities, *Journal of Financial and Quantitative Analysis* 29, 31-56.

Hsieh, David A., 1989, Modelling heteroskedasticity in daily exchange rates, *Journal of Business and Economic Statistics* 7, 307-317.

Jackwerth, Jens C., 1999, Option-Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review, *The Journal of Derivatives* 7, 66-82.

Jackwerth, Jens C., and Mark Rubinstein, 1996, Recovering probability distributions from option prices, *Journal of Finance* 51, 1611-1631.

Jarrow, Robert A. and Andrew Rudd, 1982, Approximate Option Valuation for Arbitrary Stochastic Processes, *Journal of Financial Economics* 10, 347-369.

Jondeau, Eric, and Michael Rockinger, 1998, Estimating Gram-Charlier Expansions under Positivity Constraints, Working Paper n. 653, HEC.

Jorion, Philippe, 1995, Predicting Volatility in the Foreign Exchange Market, *Journal of Finance* 50, 507-528.

Kemna, Angelica, 1989, An Empirical Test of the Option Pricing Model Based on Transactions Data of the European Options Exchange, in Robert Guimares, *et.al.*, eds.: *A Re-Appraisal of Market Efficienc,* (Springer, Berlin).

Kolmogorov, Andrey N., 1933, Sulla determinazione empirica di una legge di distribuzione, *Giornale Ist.Ital. Attuari* 4, 83-91.

Kuiper, Nicolaos H., 1962, Tests concerning random points on a circle, *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen* 63, 38-47.

Latané, Henry A., and Richard J. Rendleman, Jr., 1976, Standard Deviations of Stock Price Ratios Implied in Option Prices, *Journal of Finance* 31, 369-381.

Loudon, Geoffrey, Wing Watt, and Pradeep K. Yadav, 2000, An empirical analysis of alternative parametric ARCH models, *Journal of Applied Econometrics* 15, 117-136.

Madan, Dilip B., and Frank Milne, 1994, Contingent Claims Valued and Hedged by Pricing and Investing in a Basis, *Mathematical Finance* 4, 223-245.

Malz, Allan M., 1996, Using Option Prices to Estimate Realignment Probabilities in the European Monetary System, in Christian Dunis, ed.: *Forecasting Financial Markets*, (John Wiley & Sons, Chichester).

Malz, Allan M., 1997, Estimating the Probability Distribution of Future Exchange Rates from Options Prices, *The Journal of Derivatives* 5, 20-36.

McDonald, James B., 1984, Some generalized functions for the size distribution of income, *Econometrica* 52, 647-663.

McManus, Des J., 1999, The Information Content of Interest Rate Futures Options, Working Paper Bank of Canada, presented at the BIS workshop on Estimating and Interpreting Probability Density Functions, 14 June 1999.

Melick, William R. and Charles P. Thomas, 1997, Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis, *Journal of Financial and Quantitative Analysis* 32, 91-115.

Neuhaus, Holger, 1995, The Information Content of Derivatives for Monetary Policy, Discussion Paper 3/95, Economic Research Group of the Deutsche Bundesbank.

Noceti, Pablo, Jeremy Smith, and Stewart D. Hodges, 2000, An Evaluation of Tests of Distributional Forecasts, FORC Preprint 2000/102, University of Warwick.

Park, Hun Y., and Stephen R. Sears, 1985, Estimating Stock Index Futures Volatility Through the Prices of Their Options, *Journal of Futures Markets* 5, 223-238.

Powell, Mike J.D., 1981, *Approximation Theory and Methods* (Cambridge University Press).

Reinsch, Christian H., 1967, Smoothing by spline functions, *Numerische Mathematik* 10, 177-183.

Ritchey, Robert J., 1990, Call Option Valuation for Discrete Normal Mixtures, *Journal of Financial Research* 13, 285-286.

Rosenblatt, Murrey, 1952, Remarks on a multivariate Transformation, *Annals of Mathematical Statistics* 23, 470-472.

Rubinstein, Mark, 1994, Implied Binomial Trees, *Journal of Finance* 49, 771-818.

Rydberg, Tina H., 1997, The Normal Inverse Gaussian Lévy Process: Simulation and Approximation, *Communications in Statistics: Stochastic Models* 13, 887-910.

Seillier-Moiseiwitsch, Françoise, and Philip A. Dawid, 1993, On Testing the Validity of Sequential Probability Forecasts, *Journal of the American Statistical Association* 88, 355-359.

Shastri, Kuldeep, and Kishore Tandon, 1986, An Empirical Test of a Valuation Model for American Options on Futures Contracts, *Journal of Financial and Quantitative Analysis* 12, 377-392.

Sherrick, Bruce J., Philip Garcia, and Viswanath Tirupattur, 1996, Recovering Probabilistic Information from Options Markets: Tests of Distributional Assumptions, *Journal of Futures Markets* 16, 545-560.

Shimko, David, 1991, Beyond Implied Volatility: Probability Distributions and Hedge Ratios Implied by Option Prices, Working Paper, University of Southern California.

Shimko, David, 1993, Bounds of probability, *RISK* 6, 34-37.

Shiratsuka, Shigenori, 2001, Information Content of Implied Probability Distributions: Empirical Studies on Japanese Stock Price Index Options, IMES Discussion Paper 2001-E-1, Institute for Monetary and Economic Studies, Bank of Japan.

Söderlind, Paul, and Lars E.O. Svensson, 1997, New Techniques to Extract Market Expectations From Financial Instruments, *Journal of Monetary Economics* 40, 383-429.

Trippi, Robert R., 1977, A Test of Option Market Efficiency Using a Random-Walk Valuation Model, *Journal of Economics and Business* 29, 93-98.

Watson, George S., 1961, Goodness-of-fit tests on a circle, *Biometrika* 48, 109-114.

Weinberg, Steven A., 2001, Interpreting The Volatility Smile: An Examination of the Informational Content of Options Prices, International Finance Discussion Papers, No. 706, Federal Reserve Board, Washington, D.C.

Xu, Xinzhong, and Stephen Taylor, 1995, Conditional Volatility and the Informational Efficiency of the PHLX Currency Options Market, *Journal of Banking and Finance* 19, 803-821.

[1] This provides a tool for constructing leptokurtic distributions.

[2] Here we assume $S_T = S_0 \exp(X_T)$, therefore $X_T = \ln(S_T) - \ln(S_0)$.

[3] The estimated RNDs are non-overlapping (one at the beginning of every quarter, for the contract expiring at the end of the same quarter), in order to ensure consistency with the requirement of independent observations in the PIT analysis.

[4] In fact, the tails are approximated in this case as well. However, if the family of the functional form chosen for modelling the RNDs is not too rich, there will normally be a unique fit for the tails (see Aparicio and Hodges (1998)).

[5] The analogy with the formulation for the truncated version of a density function is evident.

[6] For a good review, see Diebold and Lopez (1996).

[7] Since an accurate forecast produces a small value for B_n, a one-sided test is more appropriate here.

[8] The use of t-tests in the context under examination can still be justified asymptotically, even if the t-statistic is not exact for autoregressive specifications, or for non-Gaussian error terms.

[9] Since the correlograms do not exhibit statistically significant autocorrelation, they have been omitted here.

[10] Clearly, for the first date in the analysis, the current expiring futures contract was not used.

[11] In the instance that put and call options with the same time to expiration and same striking prices have different implied volatilities, this indicates that Put-Call Parity is violated and that an arbitrage opportunity may exist. In reality, it would most probably suggest that one of the option prices might be "old". From the previously quoted references, this would most probably be the in-the-money option. Given that liquidity problems should not exist when dealing in the underlying futures, it would be a simple matter to combine the out-of-the-money options with a position in the futures contract to create an in-the-money option with exactly the same implied volatility. It might be possible for markets where selling the underlying asset is prohibited, one would have to examine put and call smiles separately. However, the restriction of this research to options on actively traded futures contracts precludes this case and thus, the smiles we have estimated are not two branches glued together at the at-the-money level, but (by Put-Call parity) seamless.

[12] Mean and dispersion have been omitted since their figures were not significantly different from those of the mixture density.

[13] Analogue plots have been drawn for the two other parametric specifications, but we decided not to report them here as they closely resemble the plots in figures 1 and 2.

[14] As for the B-spline approximation, we run the risk of overfitting.

[15] However, the null hypothesis of unit variance cannot be rejected at 5% confidence level.

[16] Under some testing procedures, the rejection occurs more often for the truncated version than for the full parametric specification. This could be explained by the fact that these tests place more weight on the tails of the distribution, which are the original ones for the full distribution and which are somehow "redetermined" for the truncated one. If the outcomes from the tail tests suggested that the original tails are not misspecified, it is less likely that the full density specification is rejected by the distributional tests than the truncated one.

[17] The bias in the mean is also evident from the Q-Q plots (Figure 12), as the empirical distribution of the transforms lies below the theoretical 45° line.

[18] Note that that with lognormal distributions and power utility, the adjustment would take this simple form.

[19] Since the observed values have been adjusted, whilst the truncation points (corresponding to the minimum and the maximum strikes) have been left unchanged, only the tail frequencies have been affected, and the probability forecasts in the tails stay the same.

[20] The denominator represents a normalisation for the subjective density function to integrate to one.

[21] The risk adjustments and the corresponding tests have been done for the integer range of values for γ between 1 and 6. However, since the test outcomes do not differ substantially, we only report the results for $\gamma = 1, 2$, which are the most sensible values to assume for the constant RRA parameter.

[22] The fundamental assumption made by Bliss and Panigirtzolou (2001) is that the utility function is stationary. In their research, they find that this is not the case. Their paper shows that the when the ATM implied volatility is low, the tails of the distribution are underestimated and when the ATM implied volatility is high, the tails are overestimated.

[23] In particular we abandoned the EGARCH model because, even though we could obtain stable estimates, the resulting forecasted distributions were far too volatile.

[24] The asymmetric model was estimated only for the index.

[25] Somewhat arbitrarily we assigned values for μ of zero for the currency and 7% on an annual basis for the index, rather than allow our estimates to be distorted by sampling error.

[26] With the exception of the NIG specification.

[27] These findings are consistent with Hsieh's (1989) results on modelling foreign exchange rates.

Figure Captions

Figure 1. Implied RNDs from currency options

A sample of implied Risk-Neutral densities of Normal Inverse Gaussian type extracted from currency options before and after the '87 crash.

Figure 2. Implied RNDs from S&P 500 futures options

A sample of implied Risk-Neutral densities of Normal Inverse Gaussian type extracted from options on the index future before and after the '87 crash.

Figure 3. Skewness Time Series for Implied RNDs - S&P 500

Time series of skewness of the implied Risk-Neutral densities extracted from options on the index future. Comparison amongst the three alternative parametric specifications: mixture of lognormals, generalised beta, normal inverse gaussian.

Figure 4. Excess Kurtosis Time Series for Implied RNDs - S&P 500

Time series of excess kurtosis of the implied Risk-Neutral densities extracted from options on the index future. The excess kurtosis is computed with respect to a normal distribution (coefficient of kurtosis – 3). Comparison amongst the three alternative parametric specifications: mixture of lognormals, generalised beta, normal inverse gaussian.

Figure 5. Skewness Time Series for Implied RNDs – US$ / BP

Time series of skewness of the implied Risk-Neutral densities extracted from currency options. Comparison amongst the three alternative parametric specifications: mixture of lognormals, generalised beta, normal inverse gaussian.

Figure 6. Excess Kurtosis Time Series for Implied RNDs – US$ / BP

Time series of excess kurtosis of the implied Risk-Neutral densities extracted from currency options. The excess kurtosis is computed with respect to a normal distribution (coefficient of kurtosis – 3). Comparison amongst the three alternative parametric specifications: mixture of lognormals, generalised beta, normal inverse gaussian.

Figure 7.

Plot of implied volatilities from option market prices and approximated with a B-spline function. Data on S&P 500 future options at 18/6/1992.

Figure 8.

Plot of implied risk-neutral densities, modelled as B-spline, mixture of lognormals, generalised beta and normal inverse gaussian, truncated at the bounds of the range of available strikes. Data on S&P 500 future options at 18/6/1992.

Figure 9.

Plot of implied volatilities from option market prices and approximated with a B-spline function. Data on currency options at 4/6/1992.

Figure 10.

Plot of implied risk-neutral densities, modelled as B-spline, mixture of lognormals, generalised beta and normal inverse gaussian, truncated at the bounds of the range of available strikes. Data on currency options at 4/6/1992.

Figure 11. Q-Q plots: currency options

Q-Q plots of the empirical distribution function for the probability integral transforms (calculated with respect to the probability forecast) vs the theoretical distribution function of a uniform distribution (45° line). The density forecast is modelled as a Normal Inverse Gaussian density. Data on currency options. Sub a) is for the entire density forecast, sub b) is for the truncated density forecast (within the range of available strike prices).

Figure 12. Q-Q plots: S&P 500 future options (unadjusted for risk premium)

Q-Q plots of the empirical distribution function for the probability integral transforms (calculated with respect to the probability forecast) vs the theoretical distribution function of a uniform distribution (45° line). The density forecast is modelled as a Normal Inverse Gaussian density. Data on S&P 500 future options. Sub a) is for the entire density forecast, sub b) is for the truncated density forecast (within the range of available strike prices).

Figure 13. Q-Q plots: S&P 500 future options (mean-adjusted for risk premium)

Q-Q plots of the empirical distribution function for the probability integral transforms (calculated with respect to the probability forecast) vs the theoretical distribution function of a uniform distribution (45° line). The density forecast is modelled as a Normal Inverse Gaussian density. Data on S&P 500 future options. Observations mean-adjusted for the risk premium. Sub a) is for the entire density forecast, sub b) is for the truncated density forecast (within the range of available strike prices).

Figure 14. Q-Q plots: S&P 500 future options (adjusted for risk premium - Girsanov)

Q-Q plots of the empirical distribution function for the probability integral transforms (calculated with respect to the probability forecast) vs the theoretical distribution function of a uniform distribution (45° line). The density forecast is modelled as a Normal Inverse Gaussian density. Data on S&P 500 future options. Probability forecasts are adjusted for the risk premium through a Girsanov's transformation assuming a power utility function with $\gamma = 1$. Sub a) is for the entire density forecast, sub b) is for the truncated density forecast (within the range of available strike prices).

Table I. Summary Statistics for Actual and Forecasted Distributions

Summary statistics for time series of realised quarterly log-returns (*sample*) and for the forecasted densities both extracted from option prices and simulated from GARCH-type models. For the index contract, the statistics are computed for both unadjusted RNDs and RNDs adjusted through Girsanov's transformation with power utility function. Sample conventional measures of location, dispersion, skewness and kurtosis are computed on realised log-returns. Equivalent statistics are calculated for the distribution obtained as an equally weighted mixture of the single density forecasts estimated for each quarter. Simple averages of time series of the skewness and kurtosis are also reported in the last two columns of the table.

			Mean	St. Dev.	Skew.	Kurt.	Skew.*	Kurt.*
US$ / British Pound								
Sample (full)			0.0101	0.0602	-0.4020	6.3930	-	-
Sample (ex sept. 92)			0.0123	0.0521	0.3494	4.2397	-	-
Implied RNDs		Mixture	-0.0017	0.0586	-0.1396	5.5477	-0.0909	4.2137
		GB2	-0.0017	0.0581	-0.0638	4.9472	-0.0515	3.9291
		NIG	-0.0017	0.0584	-0.0953	4.8392	-0.0769	3.9793
Simulated		GARCH *n*	3.02E-05	0.0557	-0.0005	3.7678	0.0022	3.3581
		GARCH *t*	1.13E-05	0.0574	-0.0178	5.2279	-0.0104	4.8566
S&P 500								
Sample (full)			0.0180	0.0793	-1.2059	6.2415	-	-
Implied RNDs	Unadjusted	Mixture	-0.0053	0.1054	-1.5161	9.1559	-1.1248	5.3709
		GB2	-0.0046	0.1031	-1.3865	8.5881	-1.1969	6.3864
		NIG	-0.0056	0.1101	-2.5713	27.0292	-1.5612	9.4931
	Adjusted	Mixture	0.0001	0.1003	-1.3342	8.7682	-1.0881	5.1365
	$\gamma = 1$	GB2	0.0003	0.0996	-1.2391	8.2547	-1.0428	5.2622
		NIG	-0.0002	0.1052	-2.0411	19.5538	-1.4932	8.5213
	$\gamma = 2$	Mixture	0.0052	0.0980	-1.2032	8.1518	-1.0698	5.4313
		GB2	0.0055	0.0965	-1.0753	8.0381	-0.9905	5.4368
		NIG	0.0052	0.0990	-1.5528	12.9225	-1.3385	8.0068
Simulated		GARCH *n*	0.0160	0.1071	0.0569	11.9356	-0.0055	5.1237
		GARCH *t*	0.0159	0.0947	0.0693	13.8484	-0.0122	6.9809
		AGARCH	0.0159	0.0924	-0.3171	12.0915	-0.3404	5.7866

Table II. Summary Statistics for RNDs Parameters

Average and dispersion values for the time series of parameters of implied RNDs for the three parametric specifications, for both the currency and the index contract. The statistics have been computed as sample averages and sample standard deviations of the time series of the parameters.

		Mixture of lognormals			Generalised Beta			NIG	
		Mean	St. Dev		Mean	St. Dev		Mean	St. Dev
US$ / British Pound									
Overall	μ_1	7.3631	0.1043	a	32.6826	18.0630	α	46.0802	40.0090
	σ_1	0.0496	0.0314	b	1588.2151	156.0428	β	-4.7392	15.2142
	μ_2	7.3591	0.0951	p	1.6016	1.5059	μ	0.1599	0.2473
	σ_2	0.0798	0.0272	q	1.6478	1.4258	δ	0.0142	0.0679
	θ	0.7104	0.1828						
S&P 500									
Pre-Crash Oct. 87	μ_1	5.5941	0.1098	a	18.0844	13.4017	α	20.8044	9.6114
	σ_1	0.0818	0.0183	b	282.2138	24.9645	β	-3.3220	2.6088
	μ_2	5.5264	0.1317	p	1.8253	1.3471	μ	0.1992	0.0831
	σ_2	0.1434	0.0506	q	2.9418	2.2263	δ	0.0326	0.0331
	θ	0.6847	0.0789						
Post-Crash Oct. 87	μ_1	6.4558	0.5372	a	101.2404	20.7728	α	21.6629	7.6972
	σ_1	0.0567	0.0164	b	756.3794	423.6070	β	-14.1291	6.7670
	μ_2	6.3105	0.5118	p	0.1291	0.0477	μ	0.0854	0.0345
	σ_2	0.1242	0.0385	q	0.3734	0.2107	δ	0.0688	0.0372
	θ	0.7860	0.0587						
Overall	μ_1	6.3029	0.5905	a	86.5218	37.4625	α	21.5105	7.9885
	σ_1	0.0612	0.0192	b	671.9585	425.1849	β	-12.2117	7.4817
	μ_2	6.1714	0.5556	p	0.4288	0.8489	μ	0.1056	0.0635
	σ_2	0.1276	0.0411	q	0.7895	1.2368	δ	0.0624	0.0388
	θ	0.7680	0.0732						

Table IIIa. Tail Tests

Tests of misspecification of the tails of the distributions outside the range of available strikes. For the right tail, the left tail and the combination of both tails, the frequency with which actual observations fall in those areas and the probability mass assigned by the B-splines and the truncated versions of the parametric specifications to the tails, have been reported. The test statistic based on the Brier's score Y_n^B (suggested by Seillier-Moiseiwitsch and Dawid (1993)) at 5% and 1% confidence level has also been computed. The critical values for the test are those of a standard normal distribution. Actual realisations and density forecasts are from the currency option contract and the S&P 500 contract, unadjusted to account for the impact of the risk premium.

Test statistics			**Right tail**			**Left tail**			**Both tails**		
			Frequency	Probability forecast	Y_n^B	Frequency	Probability forecast	Y_n^B	Frequency	Probability forecast	Y_n^B
5% conf. level					1.6449			1.6449			1.6449
1% conf. level					2.3263			2.3263			2.3263
US$ / BP	B-spline		0.1194	0.0566	**2.1281*	0.0299	0.0431	*-0.3084*	0.1493	0.0997	*1.1111*
	Mixture lognormals	(truncated)	0.1194	0.0549	**2.2102*	0.0299	0.0430	*-0.3302*	0.1493	0.0980	*1.1484*
	Generalised Beta	(truncated)	0.1194	0.0541	**2.2410*	0.0299	0.0424	*-0.3099*	0.1493	0.0965	*1.1997*
	NIG	(truncated)	0.1194	0.0548	**2.2222*	0.0299	0.0433	*-0.3437*	0.1493	0.0982	*1.1419*
S&P500 *(unadjusted)*	B-spline		0.0645	0.0777	*-0.0583*	0.0161	0.0343	*-0.6143*	0.0806	0.1121	*-0.4217*
	Mixture lognormals	(truncated)	0.0645	0.0780	*-0.2070*	0.0161	0.0364	*-0.6840*	0.0806	0.1144	*-0.5862*
	Generalised Beta	(truncated)	0.0645	0.0742	*-0.0660*	0.0161	0.0330	*-0.5761*	0.0806	0.1072	*-0.3949*
	NIG	(truncated)	0.0645	0.0757	*-0.1234*	0.0161	0.0337	*-0.6075*	0.0806	0.1094	*-0.4800*

* rejected at 5% confidence level.

Table IIIb. Tail Tests

Tests of misspecification of the tails of the distributions outside the range of available strikes. For the right tail, the left tail and the combination of both tails, the frequency with which actual observations fall in those areas and the probability mass assigned by the B-splines and the truncated versions of the parametric specifications to the tails, have been reported. The test statistic based on the Brier's score Y_n^B (suggested by Seillier-Moiseiwitsch and Dawid (1993)) at 5% and 1% confidence level has also been computed. The critical values for the test are those of a standard normal distribution. Data is from the S&P 500 future contract. In the top half of the table, actual observations are mean-adjusted to account for the risk-premium. In the bottom half of the table, actual observations are unadjusted whilst density forecasts are adjusted for the risk-premium, through a Girsanov's transformation, by assuming a power utility function with coefficient of risk aversion $\gamma = 1, 2$.

Test statistics			**Right tail**			**Left tail**			**Both tails**		
			Frequency	Probability forecast	Y_n^B	Frequency	Probability forecast	Y_n^B	Frequency	Probability forecast	Y_n^B
5% conf. level					1.6449			1.6449			1.6449
1% conf. level					2.3263			2.3263			2.3263
S&P500 mean adjusted	B-spline		0.0168	0.0777	**-1.7875*	0.0161	0.0343	*-0.6143*	0.0287	0.1121	**-1.9654*
	Mixture lognormals	truncated	0.0168	0.0780	**-1.7916*	0.0161	0.0364	*-0.6840*	0.0287	0.1144	**-1.9775*
	Generalised Beta	truncated	0.0168	0.0742	**-1.7317*	0.0161	0.0330	*-0.5761*	0.0287	0.1072	**-1.9181*
	NIG	truncated	0.0168	0.0757	**1.7438*	0.0161	0.0337	*-0.6075*	0.0287	0.1094	**1.9215*
S&P500 fully adjusted	Mixture lognormals	$\gamma = 1$	0.0645	0.0883	*-0.4540*	0.0161	0.0281	*-0.4091*	0.0806	0.1164	*-0.6212*
		$\gamma = 2$	0.0645	0.0995	*-0.6825*	0.0161	0.0218	*-0.1474*	0.0806	0.1213	*-0.6965*
	Generalised Beta	$\gamma = 1$	0.0645	0.0884	*-0.3176*	0.0161	0.0252	*-0.2915*	0.0806	0.1096	*-0.4314*
		$\gamma = 2$	0.0645	0.0954	*-0.5525*	0.0161	0.0196	*-0.0248*	0.0806	0.1149	*-0.5181*
	NIG	$\gamma = 1$	0.0645	0.0858	*-0.3681*	0.0161	0.0254	*-0.3045*	0.0806	0.1112	*-0.4961*
		$\gamma = 2$	0.0645	0.0966	*-0.5938*	0.0161	0.0195	*-0.0268*	0.0806	0.1161	*-0.5676*

* rejected at 5% confidence level.

Table IVa. Uniform Distributional Tests H_0: $z \sim$ i.i.d. U(0,1)

Goodness-of-fit tests for the density forecasts modelled as implied risk neutral densities of both parametric and non-parametric type. For parametric RNDs, GOF tests were computed on both the entire implied distribution and the one truncated at the bounds of the range of available strikes. GOF tests are based on uniformity tests of the probability integral transforms z_i: Kolmogorov-Smirnov (D), Anderson-Darling (A^2), Watson (U^2), Kuiper (V), Neyman-Barton of order 2 (N_2). Critical values (at both 5% and 1% confidence level) for D, A^2, U^2 and V were taken from D'Agostino and Stephens (1986). Critical values for N_2 are those of a $\chi^2(2)$. Observations and density forecasts are from the currency option contract and the S&P 500 contract, unadjusted to account for the impact of the risk premium.

Test statistics			D	A^2	U^2	V	N_2
5% conf. level			1.3580	2.4920	0.1870	1.7470	5.9915
1% conf. level			1.6280	3.8570	0.2680	2.0010	9.2103
US$ / BP							
	B-spline		1.1704	1.3305	*0.2101	*1.8189	5.2446
	Mixture lognormals						
		truncated	1.0711	0.9918	0.1537	1.5762	3.4711
		full	1.2019	1.5066	0.0899	1.3311	2.8267
	Generalised Beta						
		truncated	0.9929	0.9344	0.1378	1.4917	3.1033
		full	1.1875	1.5160	0.0798	1.3158	2.7967
	NIG						
		truncated	1.0421	0.9780	0.1489	1.5492	3.3900
		full	1.1865	1.5177	0.0849	1.3159	2.8274
S&P500							
(unadjusted)	B-spline		*1.4018	*2.9740	*0.1997	*1.8423	*6.9183
	Mixture lognormals						
		truncated	*1.4222	*3.4280	0.1684	1.7401	*6.6247
		full	*1.3364	*2.5544	*0.1923	1.6642	5.7182
	Generalised Beta						
		truncated	1.2515	*3.1139	0.1284	1.4637	*6.1668
		full	1.2617	2.3233	0.1414	1.4659	5.2380
	NIG						
		truncated	1.3413	*3.1801	0.1474	1.6025	*6.2891
		full	1.3317	2.3951	0.1705	1.6101	5.4749

* rejected at 5% confidence level.

Table IVb. Uniform Distributional Tests H_0: $z \sim$ i.i.d. U(0,1)

Goodness-of-fit tests for the density forecasts modelled as implied risk neutral densities of both parametric and non-parametric type. For parametric RNDs, GOF tests were computed on both the entire implied distribution and the one truncated at the bounds of the range of available strikes. GOF tests are based on uniformity tests of the probability integral transforms z_t: Kolmogorov-Smirnov (D), Anderson-Darling (A^2), Watson (U^2), Kuiper (V), Neyman-Barton of order 2 (N_2). Critical values (at both 5% and 1% confidence level) for D, A^2, U^2 and V were taken from D'Agostino and Stephens (1986). Critical values for N_2 are those of a $\chi^2(2)$. Data is from the S&P 500 future contract. In the top half of the table, actual realisations are mean-adjusted to account for the risk-premium. In the bottom half of the table, implied RNDs are adjusted for the risk-premium, through a Girsanov's transformation, by assuming a power utility function with coefficient of risk aversion $\gamma = 1, 2$.

Test statistics			D	A^2	U^2	V	N_2
5% conf. level			1.3580	2.4920	0.1870	1.7470	5.9915
1% conf. level			1.6280	3.8570	0.2680	2.0010	9.2103
S&P500	Mixture lognormals						
mean adjusted		truncated	0.7816	0.9685	0.1205	1.5038	3.8592
		full	1.0112	1.0023	*0.2023	1.6845	*6.7412
	Generalised Beta						
		truncated	0.8347	1.2673	0.1697	1.5496	5.1874
		full	0.0995	1.8277	*0.2462	*1.8317	*7.2213
	NIG						
		truncated	0.8439	1.1201	0.1528	1.5872	4.8237
		full	1.0195	1.7389	*0.2279	*1.7473	*7.3278
S&P500							
fully adjusted	Mixture lognormals						
		truncated	1.2257	2.2117	0.1224	1.5940	4.3194
$\gamma = 1$		full	1.1319	1.4176	0.1444	1.5019	3.8273
	Generalised Beta						
		truncated	1.0661	1.9670	0.0888	1.3205	3.9606
		full	0.9959	1.2182	0.1003	1.2442	3.2466
	NIG						
		truncated	1.1028	2.0225	0.1048	1.4079	4.0920
		full	1.0089	1.2966	0.1259	1.3257	3.5912
	Mixture lognormals						
$\gamma = 2$		truncated	1.0328	1.3215	0.0875	1.4547	2.6036
		full	0.8366	0.7662	0.1079	1.2507	2.7586
	Generalised Beta						
		truncated	0.9099	1.1388	0.0606	1.2067	2.3155
		full	0.7939	0.5993	0.0733	1.1102	2.0760
	NIG						
		truncated	0.9011	1.1786	0.0725	1.2531	2.4574
		full	0.7867	0.6865	0.0945	1.1553	2.5204

* rejected at 5% confidence level.

Table Va. Normal Distributional Tests H_0: $x \sim$ i.i.d. N(0,1)

Goodness-of-fit tests for the density forecasts modelled as implied RNDs of both parametric and non-parametric type. GOF tests are based on normality tests of the modified probability integral transforms x_t (after a transformation to normality). The first two columns report the t-tests for $\mu = 0$, $\rho = 0$ of the model $x_t - \mu = \rho(x_{t-1} - \mu) + \varepsilon_t$, with p-values in brackets. The third column displays the chi-test on the variance ($Var(\varepsilon_t) = \sigma^2 = 1$). The Jarque-Bera and Doornik-Hansen tests for normality of the residuals follow. Their critical values are those of a $\chi^2(2)$. LR1 and LR2 denote the two log-likelihood ratio tests of serial independence and of joint serial independence, zero mean, and unit variance. Critical values are, respectively, from $\chi^2(1)$ and $\chi^2(3)$. Observations and density forecasts are from the currency option contract and the S&P 500 contract, unadjusted to account for the impact of the risk premium.

Test statistics			μ	ρ	σ^2	J-B	D-H	LR1	LR2
5% conf. level			(>0.05)	(>0.05)		(>0.05)	5.9915	3.8415	7.8147
1% conf. level			(>0.01)	(>0.01)		(>0.01)	9.2103	6.6349	11.3449
US $ / BP									
	B-spline		0.0678 (0.4667)	-0.1946 (0.1517)	0.6776	0.3098	2.9851	2.0746	6.4250
	Mixture lognormals	truncated	0.0642 (0.5213)	-0.1732 (0.2027)	0.7572	0.4140	2.5221	1.6315	4.2031
		full	0.1450 (0.2328)	-0.0316 (0.7959)	0.9859	*0.0364	*7.1201	0.0528	2.0036
	Generalised Beta	truncated	0.0702 (0.4885)	-0.1717 (0.2067)	0.7723	0.4591	2.3032	1.6030	3.9385
		full	0.1520 (0.2151)	-0.0280 (0.8186)	0.9929	0.0785	*5.9985	0.0391	2.1515
	NIG	truncated	0.0659 (0.5119)	-0.1732 (0.2028)	0.7635	0.4309	2.5294	1.6305	4.0930
		full	0.1487 (0.2235)	-0.0324 (0.7911)	0.9958	*0.0336	*7.2972	0.0555	2.0916
S&P500 *(unadjusted)*	B-spline		0.2361 (0.0684)	0.1094 (0.4198)	*0.6567	0.6796	1.2221	0.7357	*9.1551
	Mixture lognormals	truncated	*0.2881 (0.0386)	0.1014 (0.4549)	0.7470	0.6741	0.7546	0.6375	*8.4560
		full	0.2153 (0.0733)	0.0189 (0.8858)	0.7738	0.1659	3.5481	0.0218	5.4918
	Generalised Beta	truncated	0.2855 (0.0517)	0.1355 (0.3177)	0.7771	0.8716	0.3378	1.0994	*8.1476
		full	0.2234 (0.0820)	0.0567 (0.6655)	0.8186	0.3089	2.9475	0.1862	4.9505
	NIG	truncated	*0.2808 (0.0483)	0.1130 (0.4049)	0.7681	0.7830	0.6162	0.7801	*7.9044
		full	0.2134 (0.0833)	0.0310 (0.8134)	0.7969	0.1960	3.4374	0.0573	5.0050

* rejected at 5% confidence level.

Table Vb. Normal Distributional Tests H_0: $x \sim$ i.i.d. N(0,1)

GOF tests for the density forecasts modelled as implied RNDs of both parametric and non-parametric type. GOF tests are based on normality tests of the modified probability integral transforms x_t (after a transformation to normality). The first two columns report the t-tests for $\mu = 0$, $\rho = 0$ of the model $x_t - \mu = \rho(x_{t-1} - \mu) + \varepsilon_t$, with p-values in brackets. The third column displays the chi-test on the variance ($Var(\varepsilon_t) = \sigma^2 = 1$). The Jarque-Bera and Doornik-Hansen tests for normality of the residuals follow. Their critical values are those of a $\chi^2(2)$. LR1 and LR2 denote the two log-likelihood ratio tests of serial independence and of joint serial independence, zero mean, and unit variance. Critical values are, respectively, from $\chi^2(1)$ and $\chi^2(3)$. Data is from the S&P 500 future contract. In the top half of the table, actual realisations are mean-adjusted to account for the risk-premium. In the bottom half of the table, implied RNDs are adjusted for the risk-premium, through a Girsanov's transformation, by assuming a power utility function with coefficient of risk aversion $\gamma = 1, 2$.

Test statistics			μ	ρ	σ^2	J-B	D-H	LR1	LR2
5% conf. level			(>0.05)	(>0.05)		(>0.05)	5.9915	3.8415	7.8147
1% conf. level			(>0.01)	(>0.01)		(>0.01)	9.2103	6.6349	11.3449
S&P500 *mean adjusted*	Mixture lognormals	truncated	-0.0167 (0.8821)	0.1741 (0.2164)	*0.6054	0.8065	0.0716	1.7302	*7.9393
		full	-0.0924 (0.3862)	0.0063 (0.9631)	*0.6437	0.0528	*5.9983	0.0173	5.3069
	Generalised Beta	truncated	-0.0241 (0.8183)	0.1894 (0.1743)	*0.5792	0.9215	0.6871	1.7879	*8.1876
		full	-0.0894 (0.3783)	0.0316 (0.8097)	*0.6397	*0.0392	*6.9745	0.0790	5.9825
	NIG	truncated	-0.0215 (0.8465)	0.1792 (0.1979)	*0.5983	0.8906	0.7503	1.7544	6.8452
		full	-0.0931 (0.3922)	0.0171 (0.8978)	*0.6428	*0.0412	*6.8825	0.0267	5.8162
S&P500 *Fully adjusted*	Mixture lognormals	truncated	0.2173 (0.1175)	0.1127 (0.4055)	0.7717	0.6829	0.7333	*3.8415	*7.8148
$\gamma = 1$		full	0.1274 (0.2894)	0.0339 (0.7962)	0.7940	0.1682	3.5017	0.0675	3.2305
	Generalised Beta	truncated	0.2151 (0.1420)	0.1461 (0.2806)	0.8015	0.8760	0.3458	1.2534	5.7529
		full	0.1362 (0.2899)	0.0705 (0.5906)	0.8400	0.3096	2.9095	0.2877	2.7514
	NIG	truncated	0.2103 (0.1383)	0.1245 (0.3581)	0.7914	0.7937	0.6074	0.9215	5.5187
		full	0.1256 (0.3088)	0.0454 (0.7294)	0.8177	0.1986	3.3902	0.1204	2.8174
$\gamma = 2$	Mixture lognormals	truncated	0.1485 (0.2876)	0.1238 (0.3605)	0.7968	0.6900	0.7199	0.8953	4.0339
		full	0.0426 (0.7273)	0.0480 (0.7142)	0.8146	0.1699	3.4655	0.1337	2.0007
	Generalised Beta	truncated	0.1467 (0.3195)	0.1562 (0.2482)	0.8264	0.8792	0.3561	1.4087	4.0092
		full	0.0519 (0.6905)	0.0832 (0.5252)	0.8616	0.3096	2.8824	0.4016	1.5770
	NIG	truncated	0.1417 (0.3210)	0.1357 (0.3160)	0.8152	0.8024	0.6039	1.0699	3.7808
		full	0.0412 (0.7422)	0.0587 (0.6540)	0.8387	0.2005	3.3537	0.2001	1.6650

* rejected at 5% confidence level.

Table VI. Average Parameter Values for Estimated Historical Distributions

For both the currency and the index options contracts, GARCH-type models have been estimated at the beginning of each quarter for the time period Mar 86-Sep 01. The series of past returns employed for the estimation include daily returns from April 82 to the beginning of the quarter under exam. Summary statistics for the time series of the coefficients of the estimated models are shown below. Sample mean values for the coefficients have been computed. Standard deviation values are in brackets.

	α_0	α_1	**asymm. coeff.**	β_1	**Degrees of freedom**
US $ / British Pound					
GARCH (1,1) normal errors	8.99E-07	0.0471		0.9376	
	(3.74E-07)	(0.0073)		(0.0122)	
GARCH (1,1) *t* errors	6.27E-07	0.0472		0.9449	5.4744
	(3.32E-07)	(0.0052)		(0.0098)	(1.1048)
S&P 500 future					
GARCH (1,1) normal errors	5.6E-06	0.1421	-	0.8311	-
	(3.6E-06)	(0.0467)	-	(0.0603)	-
GARCH (1,1) *t* errors	1.6E-06	0.0527	-	0.9370	4.8635
	(1.1E-06)	(0.0087)	-	(0.0129)	(1.0452)
Asymm. GARCH (1,1) *t* errors	1.1E-06	0.0560	0.0038	0.9287	5.0194
	(1.1E-06)	(0.0099)	(0.0010)	(0.0149)	(0.9931)

Table VII. Tail Tests for Simulated Distributional Forecasts

Tests of misspecification of the tails of the distributions outside the range of available strikes. For the right tail, the left tail and the combination of both tails, the frequency with which actual observations fall in those areas and the probability mass assigned by the density forecasts to the tails, have been reported. The test statistic based on the Brier's score Y_n^B (suggested by Seillier-Moiseiwitsch and Dawid (1993)) at 5% and 1% confidence level has also been computed. The critical values for the test are those of a standard normal distribution. Actual realisations and density forecasts are for the currency option contract and the S&P 500 contract. Density forecasts are obtained by Monte Carlo simulation of the distribution of log-returns based on previously estimated GARCH-type models for the volatility process.

Test statistics		**Right tail**			**Left tail**			**Both tails**		
		Frequency	Probability forecast	Y_n^B	Frequency	Probability forecast	Y_n^B	Frequency	Probability forecast	Y_n^B
5% conf. level				1.6449			1.6449			1.6449
1% conf. level				2.3263			2.3263			2.3263
US$ / BP										
	GARCH (1,1) normal errors	0.1194	0.0554	*2.1591	0.0299	0.0377	-0.1668	0.1493	0.0931	1.3230
	GARCH (1,1) *t* errors	0.1194	0.0563	*2.0590	0.0299	0.0391	-0.2373	0.1493	0.0954	1.2139
S&P 500										
	GARCH (1,1) normal errors	0.0645	0.1582	*-1.8923	0.0161	0.0199	0.0026	0.0806	0.1782	*-1.6454
	GARCH (1,1) t errors	0.0645	0.1266	-1.4186	0.0161	0.0130	0.5191	0.0806	0.1396	-1.1298
	Asymm. GARCH (1,1) *t* errors	0.0645	0.1229	-1.3156	0.0161	0.0141	0.4046	0.0806	0.1370	-1.0322

* rejected at 5% confidence level.

Table VIII. Uniform Distributional Tests for Simulated Distributional Forecasts H_0: $z \sim$ i.i.d. U(0,1)

Goodness-of-fit tests for the density forecasts obtained by Monte Carlo simulation of the distribution of log-returns based on previously estimated GARCH-type models for the volatility process. GOF tests are based on uniformity tests of the probability integral transforms z_i: Kolmogorov-Smirnov (D), Anderson-Darling (A^2), Watson (U^2), Kuiper (V), Neyman-Barton of order 2 (N_2). Critical values (at both 5% and 1% confidence level) for D, A^2, U^2 and V were taken from D'Agostino and Stephens (1986). Critical values for N_2 are those of a $\chi^2(2)$. Observations and density forecasts are from the currency option contract and the S&P 500 contract.

Test statistics			**D**	A^2	U^2	**V**	N_2
5% conf. level			1.3580	2.4920	0.1870	1.7470	5.9915
1% conf. level			1.6280	3.8570	0.2680	2.0010	9.2103
US $ / BP							
	GARCH (1,1) normal errors	truncated	1.0996	1.0934	0.1780	1.7395	4.2697
		full	1.0899	1.3629	0.0833	1.2173	2.5599
	GARCH (1,1) *t* errors	truncated	1.0201	0.9827	0.1586	1.6576	3.5691
		full	1.1384	1.3391	0.0733	1.2594	2.4308
S&P 500							
	GARCH (1,1) normal errors	truncated	**1.8050	**4.2127	**0.3170	**2.2270	*8.4397
		full	1.2752	*2.6595	**0.4430	**2.4478	**11.2613
	GARCH (1,1) *t* errors	truncated	*1.4613	*2.9512	*0.2087	*1.7678	5.6231
		full	1.0303	1.7241	*0.2568	*1.8188	5.7681
	Asymm. GARCH (1,1) *t* errors	truncated	1.2243	2.1189	0.1572	1.5821	4.3894
		full	0.9289	1.3345	*0.2034	1.6901	5.1812

* rejected at 5% confidence level.
** rejected at 1% confidence level.

Table IX. Normal Distributional Tests for simulated distributional forecasts H_0: x ~ i.i.d. N(0,1)

Goodness-of-fit tests for the density forecasts obtained by Monte Carlo simulation of the distribution of log-returns based on previously estimated GARCH-type models for the volatility process. GOF tests are based on normality tests of the modified probability integral transforms x_t (after a transformation to normality). The first two columns report the t-tests for $\mu = 0$, $\rho = 0$ of the model $x_t - \mu = \rho(x_{t-1} - \mu) + \varepsilon_t$, with p-values in brackets. The third column displays the chi-test on the variance ($Var(\varepsilon_t) = \sigma^2 = 1$). The Jarque-Bera and Doornik-Hansen tests for normality of the residuals follow. Their critical values are those of a $\chi^2(2)$. LR1 and LR2 denote the two log-likelihood ratio tests of serial independence and of joint serial independence, zero mean, and unit variance. Critical values are, respectively, from $\chi^2(1)$ and $\chi^2(3)$. Observations and density forecasts are from the currency option contract and the S&P 500 contract.

Test statistics			μ		ρ		σ^2	J-B	D-H	LR1	LR2
5% conf. level				(>0.05)		(>0.05)		(>0.05)	5.9915	3.8415	7.8147
1% conf. level				(>0.01)		(>0.01)		(>0.01)	9.2103	6.6349	11.3449
US $ / BP											
	GARCH (1,1) normal errors	truncated	0.0551	(0.5600)	-0.2068	(0.1269)	0.7177	0.3197	3.1117	2.3630	5.5006
		full	0.1341	(0.2778)	-0.0093	(0.9385)	0.9793	0.0866	*7.4266	0.0002	1.8084
	GARCH (1,1) *t* errors	truncated	0.0590	(0.5434)	-0.1993	(0.1417)	0.7463	0.4396	2.5794	2.1872	4.7786
		full	0.1422	(0.2427)	-0.0176	(0.8842)	0.9620	0.4197	3.7750	0.0113	1.9532
S&P 500											
	GARCH (1,1) normal errors	truncated	*0.2816	(0.0409)	0.1273	(0.3499)	0.6896	0.0857	4.3461	0.9462	*9.7887
		full	0.0164	(0.8713)	-0.0518	(0.6933)	0.6872	**0.0000	**13.8051	0.1542	4.9204
	GARCH (1,1) *t* errors	truncated	0.2331	(0.0864)	0.1807	(0.1807)	0.7842	0.1807	3.6286	0.3996	5.8067
		full	0.0309	(0.7812)	-0.0302	(0.8182)	0.7960	**0.0000	**10.5957	0.0521	2.2154
	Asymm. GARCH (1,1) *t* errors	truncated	0.1858	(0.1651)	0.0902	(0.5085)	0.7725	0.2803	2.7779	0.4698	4.8811
		full	0.0086	(0.9388)	-0.0151	(0.9084)	0.7900	**0.0007	*7.8341	0.0131	2.2365

* rejected at 5% confidence level.
** rejected at 1% confidence level.

Figure 1. Implied RNDs from currency options

A sample of implied Risk-Neutral densities of Normal Inverse Gaussian type extracted from currency options before and after the '87 crash.

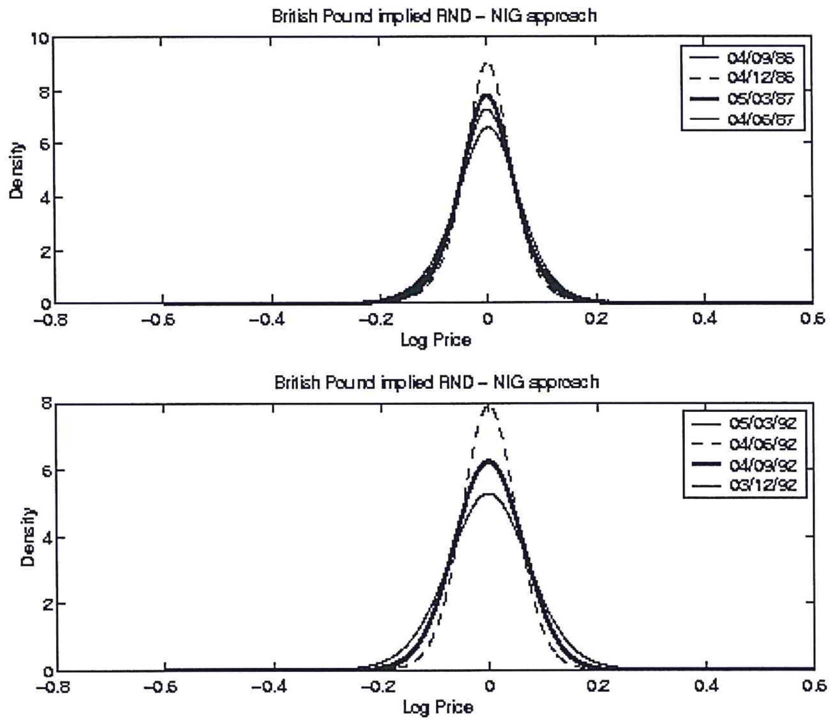


Figure 2. Implied RNDs from S&P 500 futures options

A sample of implied Risk-Neutral densities of Normal Inverse Gaussian type extracted from options on the index future before and after the '87 crash.

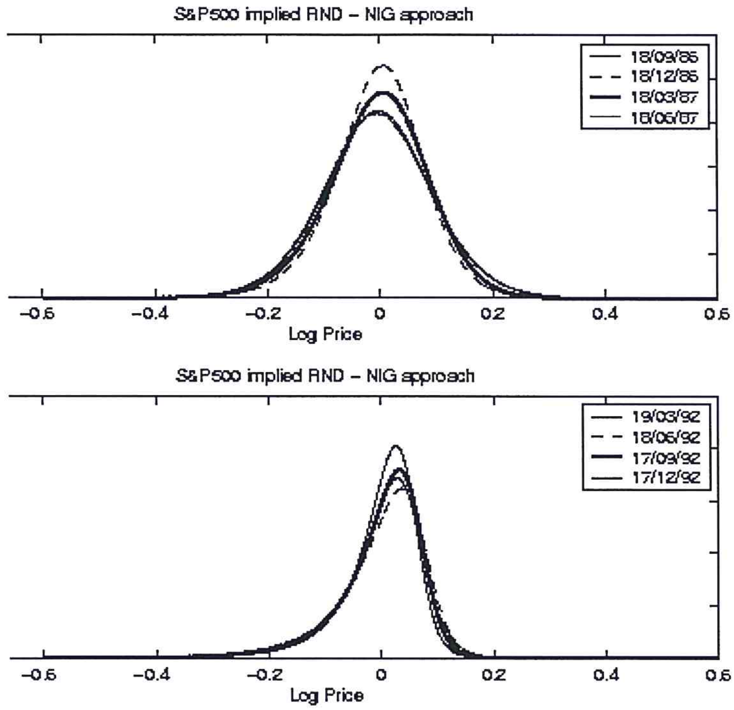


Figure 3. Skewness Time Series for Implied RNDs - S&P 500

Time series of skewness of the implied Risk-Neutral densities extracted from options on the index future. Comparison amongst the three alternative parametric specifications: mixture of lognormals, generalised beta, normal inverse gaussian.

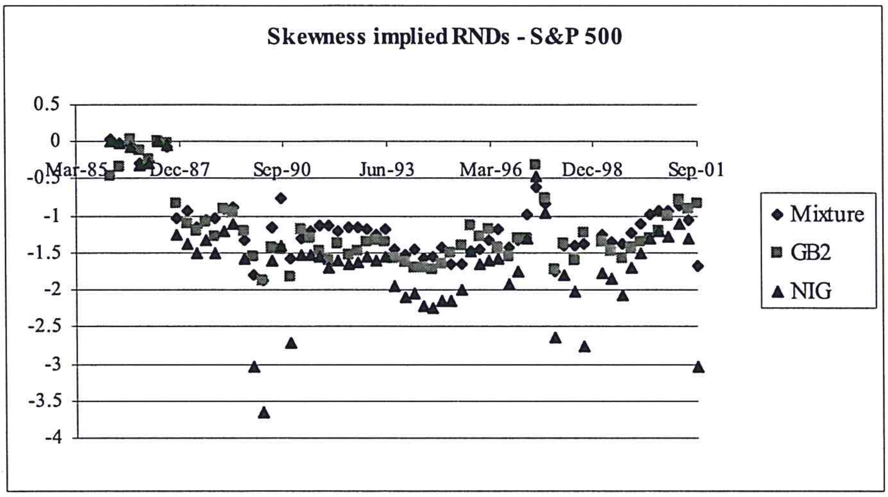


Figure 4. Excess Kurtosis Time Series for Implied RNDs - S&P 500

Time series of excess kurtosis of the implied Risk-Neutral densities extracted from options on the index future. The excess kurtosis is computed with respect to a normal distribution (coefficient of kurtosis – 3). Comparison amongst the three alternative parametric specifications: mixture of lognormals, generalised beta, normal inverse gaussian.

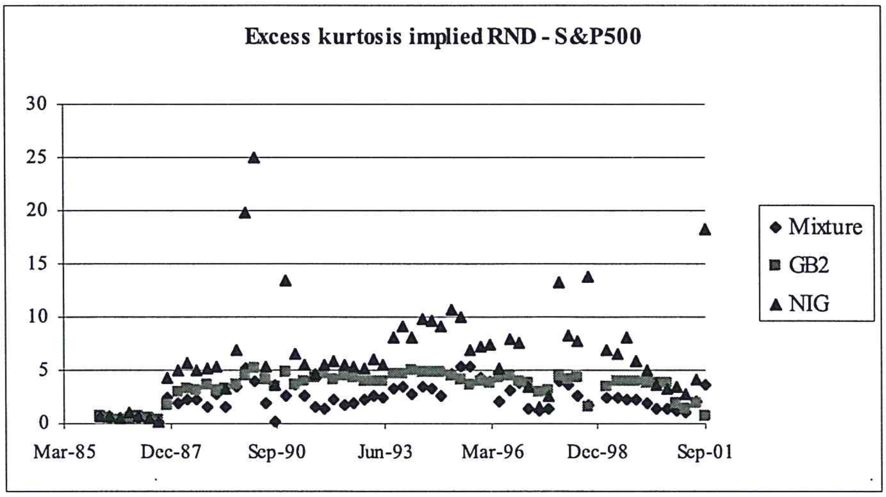


Figure 5. Skewness Time Series for Implied RNDs – US$ / BP

Time series of skewness of the implied Risk-Neutral densities extracted from currency options. Comparison amongst the three alternative parametric specifications: mixture of lognormals, generalised beta, normal inverse gaussian.

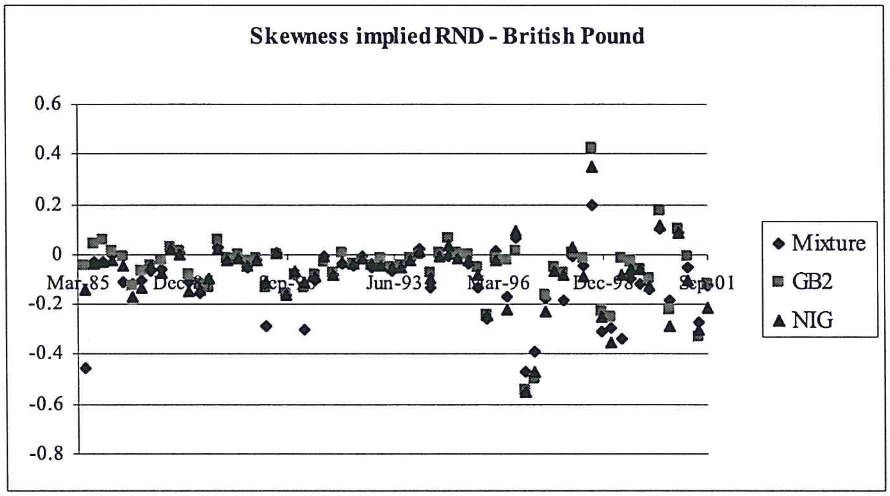


Figure 6. Excess Kurtosis Time Series for Implied RNDs – US$ / BP

Time series of excess kurtosis of the implied Risk-Neutral densities extracted from currency options. The excess kurtosis is computed with respect to a normal distribution (coefficient of kurtosis – 3). Comparison amongst the three alternative parametric specifications: mixture of lognormals, generalised beta, normal inverse gaussian.

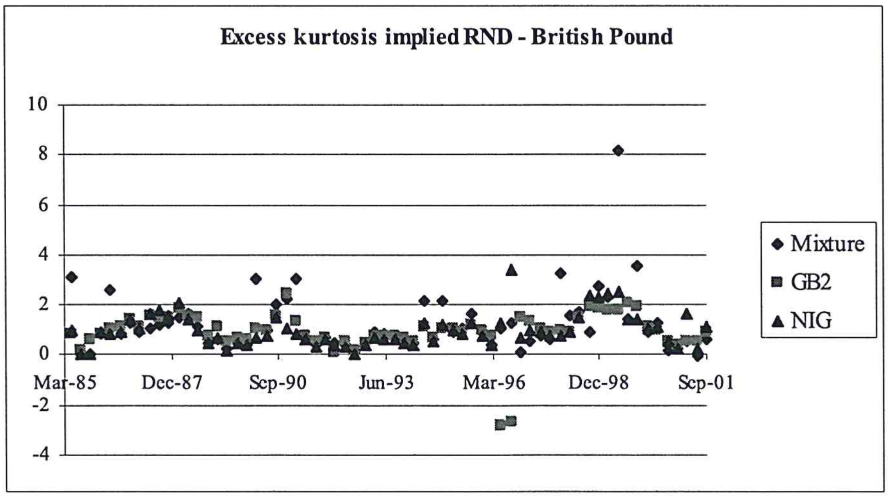


Figure 7.

Plot of implied volatilities from option market prices (small circles) and approximated with a B-spline function (line). Data on S&P 500 future options at 18/6/1992.

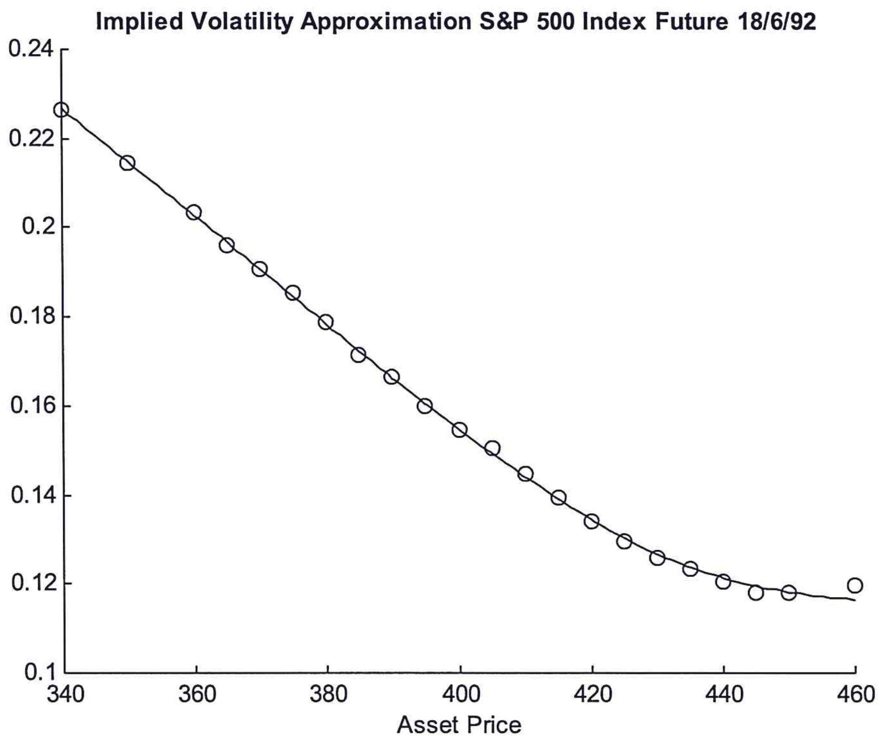


Figure 8.

Plot of implied risk-neutral densities truncated at the bounds of the range of available strikes. Data on S&P 500 future options at 18/6/1992.

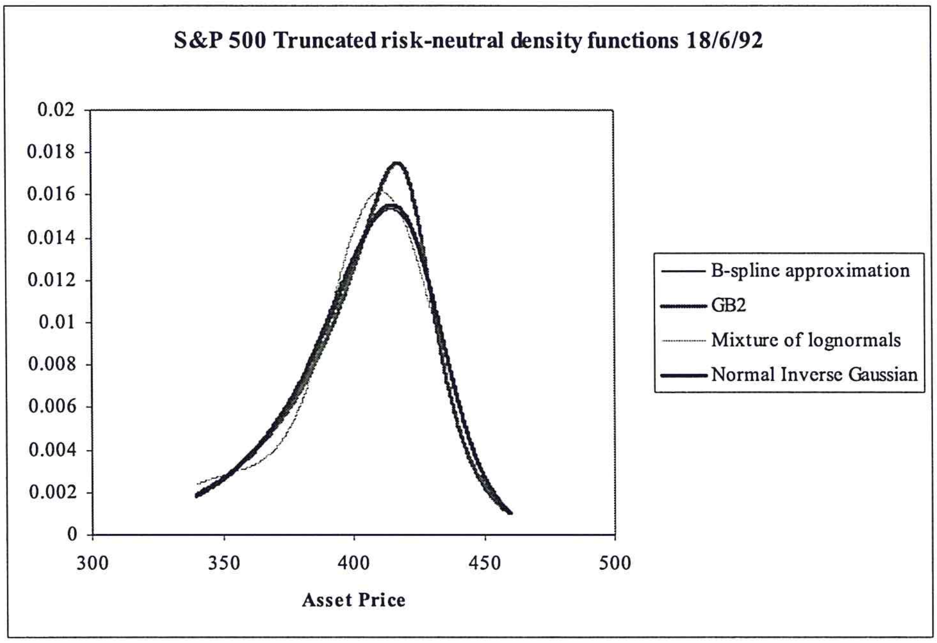


Figure 9.

Plot of implied volatilities from option market prices (small circles) and approximated with a B-spline function (line). Data on currency options at 4/6/1992.

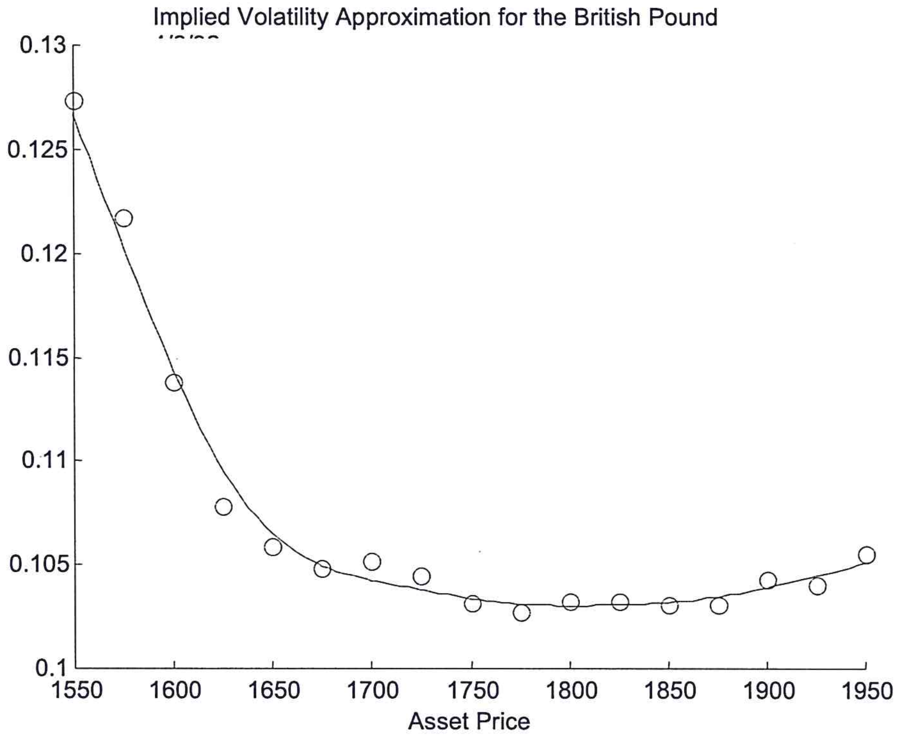


Figure 10.

Plot of implied risk-neutral densities truncated at the bounds of the range of available strikes. Data on currency options at 4/6/1992.

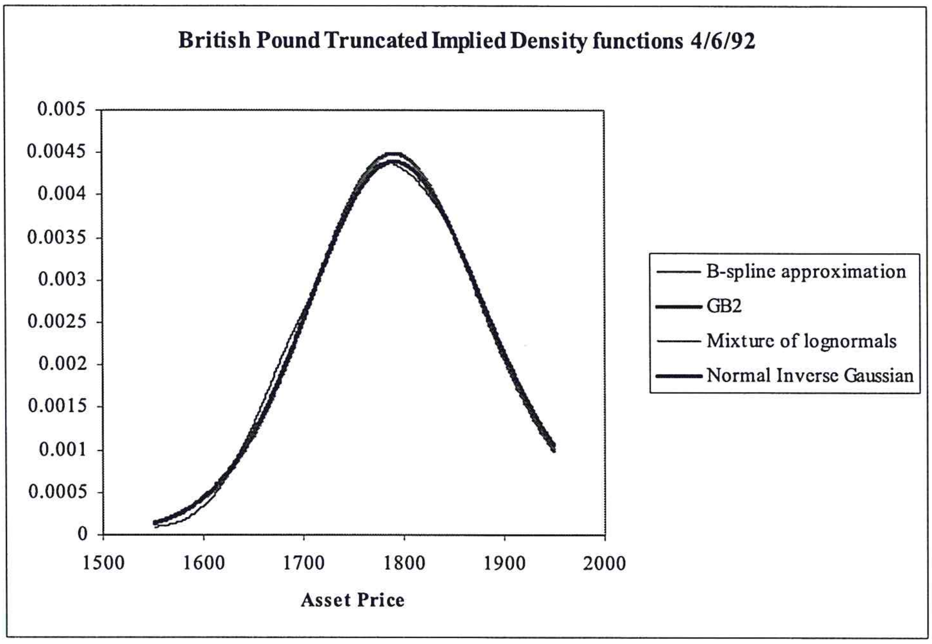
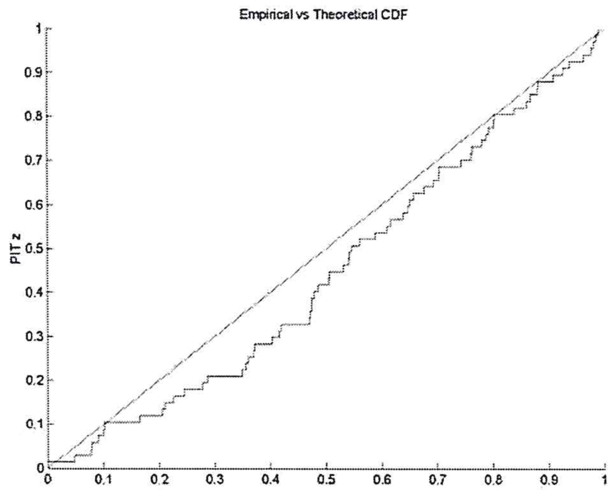


Figure 11. Q-Q plots: currency options

Q-Q plots of the empirical distribution function for the probability integral transforms (calculated with respect to the probability forecast) vs the theoretical distribution function of a uniform distribution (45° line). The density forecast is modelled as a Normal Inverse Gaussian density. Data on currency options. Sub a) is for the entire density forecast, sub b) is for the truncated density forecast (within the range of available strike prices).

a) Entire Distribution



b) Truncated Distribution

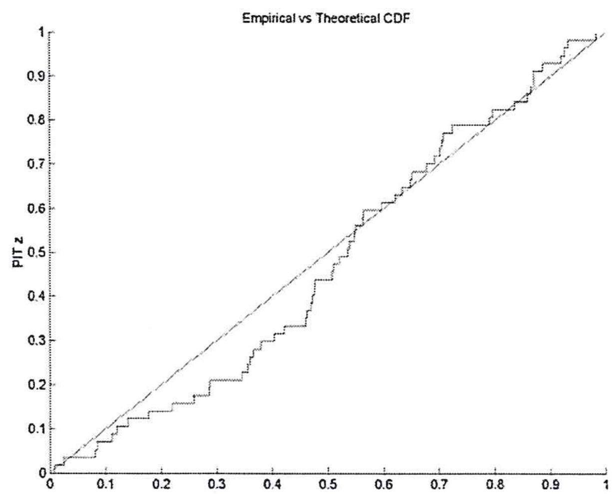
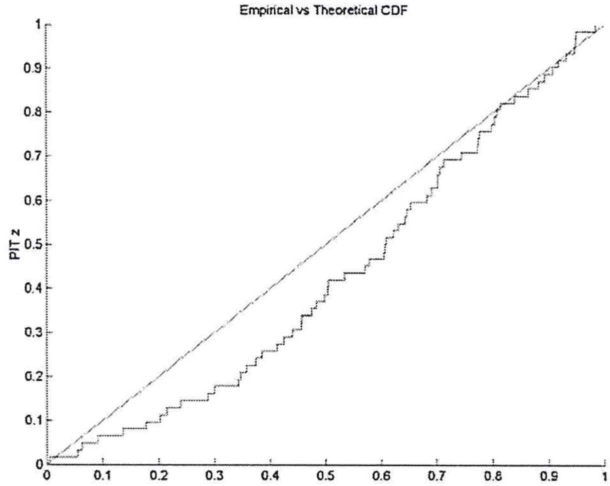


Figure 12. Q-Q plots: S&P 500 future options (unadjusted for risk premium)

Q-Q plots of the empirical distribution function for the probability integral transforms (calculated with respect to the probability forecast) vs the theoretical distribution function of a uniform distribution (45° line). The density forecast is modelled as a Normal Inverse Gaussian density. Data on S&P 500 future options. Sub a) is for the entire density forecast, sub b) is for the truncated density forecast (within the range of available strike prices).

a) Entire Distribution



b) Truncated Distribution

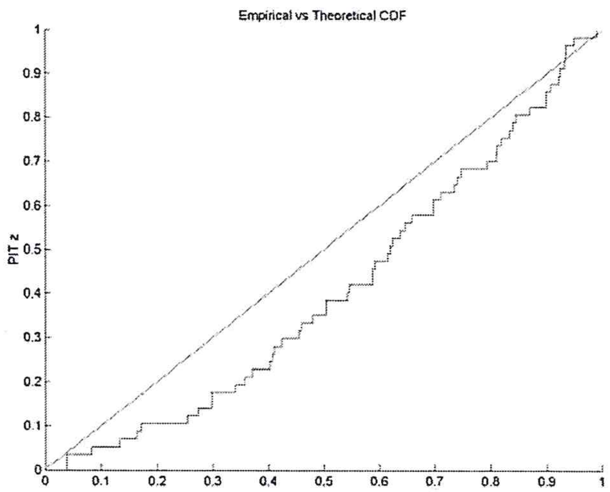
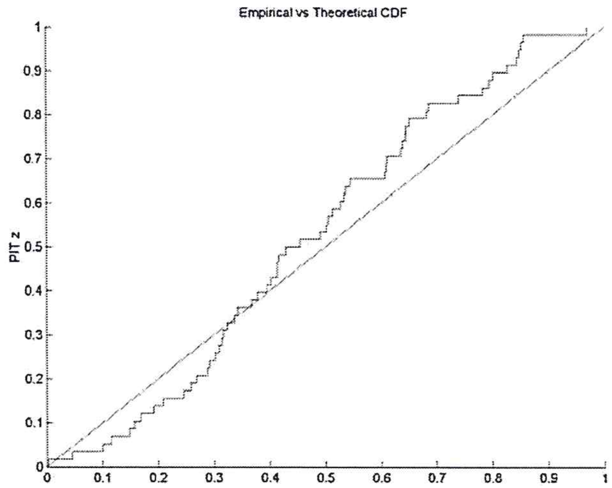


Figure 13. Q-Q plots: S&P 500 future options (mean-adjusted for risk premium)

Q-Q plots of the empirical distribution function for the probability integral transforms (calculated with respect to the probability forecast) vs the theoretical distribution function of a uniform distribution (45° line). The density forecast is modelled as a Normal Inverse Gaussian density. Data on S&P 500 future options. Observations mean-adjusted for the risk premium. Sub a) is for the entire density forecast, sub b) is for the truncated density forecast (within the range of available strike prices).

a) Entire Distribution



b) Truncated Distribution

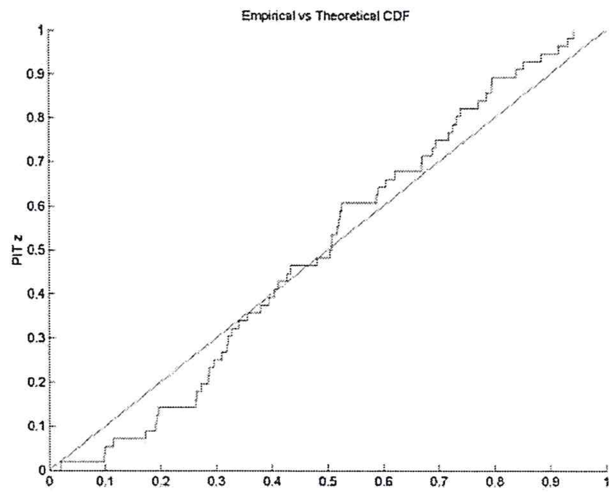
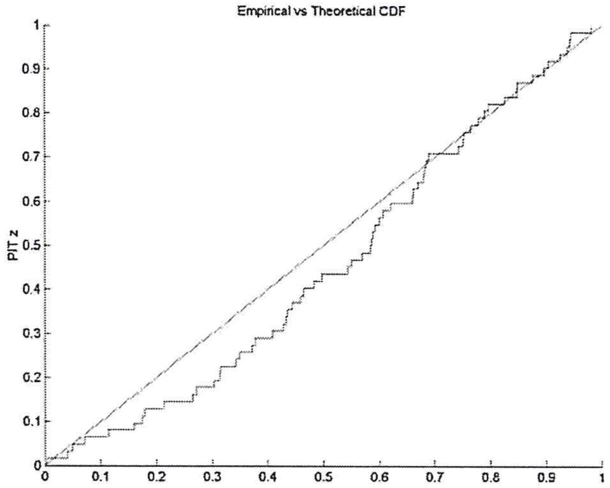


Figure 14. Q-Q plots: S&P 500 future options (adjusted for risk premium - Girsanov)

Q-Q plots of the empirical distribution function for the probability integral transforms (calculated with respect to the probability forecast) vs the theoretical distribution function of a uniform distribution (45° line). The density forecast is modelled as a Normal Inverse Gaussian density. Data on S&P 500 future options. Probability forecasts are adjusted for the risk premium through a Girsanov's transformation assuming a power utility function with $\gamma = 1$. Sub a) is for the entire density forecast, sub b) is for the truncated density forecast (within the range of available strike prices).

a) Entire Distribution



b) Truncated Distribution