A multi-horizon comparison of density forecasts for the S&P 500 using index returns and option prices

Mark B. Shackleton, Stephen J. Taylor and Peng Yu

Department of Accounting and Finance, Lancaster University, England

E-Mail: m.shackleton@lancaster.ac.uk, s.taylor@lancaster.ac.uk, p.yu1@lancaster.ac.uk

May 2006

Abstract

We compare density forecasts of the S&P 500 index from 1991 to 2004, obtained from option prices and daily and five-minute index returns, over seven horizons ranging from one day to twelve weeks. Risk-neutral forecasts are derived both from lognormal densities and by estimating the Heston stochastic volatility process from option prices, which provides a closed-form density for all future times. Out-of-sample methods, both parametric and non-parametric, are applied to transform the risk-neutral densities into real-world densities. These option-based densities are compared with historical densities defined by ARCH models.

We find the best forecasts are produced by the risk-transformations of the risk-neutral densities, for horizons of one, two and four weeks, while the best historical forecasts are generally superior for the one-day horizon, when forecast methods are ranked by the out-of-sample likelihood of observed index levels. For all risk-transformations, a mixture of the real-world densities and the historical densities obtained from five-minute returns has a higher likelihood than both components of the mixture, for horizons of one day, one week and two weeks.

The Kolmogorov-Smirnov and Berkowitz diagnostic tests show that the risk-transformed, option-based densities nearly always pass these tests, and they do so more often than the other density forecasting methods.

JEL classifications: C52, C53, G13.

Keywords: Density forecasting, Options, High frequency, Heston, S&P 500.

A multi-horizon comparison of density forecasts for the S&P 500 using index returns and option prices[1]

1. Introduction

Option prices reflect competitive opinions about the risk-neutral density of the underlying asset when a set of option contracts expire. Several empirical methods are able to convert option prices into an estimated risk-neutral density for a single expiry date, as has been illustrated by Jackwerth and Rubinstein (1996), Melick and Thomas (1997), Ait-Sahalia and Lo (1998), Bliss and Panigirtzoglou (2002) and Taylor (2005). The more difficult problem of estimating the risk-neutral dynamics of the underlying asset, at a specific moment in time from option prices for several expiry dates, has received much less attention. The most comprehensive study is the pricing and hedging paper by Bakshi, Cao and Chen (1997), that summarizes daily estimates for the S&P 500 index for the four-year period from 1988 to 1991. Our first contribution is to estimate the risk-neutral dynamics of Heston (1993) for S&P 500 futures prices, on each day in the fifteen years from 1990 to 2004 inclusive. We can then derive the risk-neutral densities for *all* time horizons, using our estimates of the Heston parameters.

Transformations from risk-neutral (Q) to real-world (P)[2] densities have been proposed and estimated by Bakshi, Kapadia and Madan (2003), Bliss and Panigirtzoglou (2004), Anagnou-Basioudis et al (2005) and Liu et al (2005). These real-world densities have,

[1] We thank Christoph Schleicher for several helpful comments and also seminar participants at the Bank of England, the MathFinance workshop in Frankfurt, the Juan Carlos III University of Madrid, the University of Manchester, the University of Warwick and the University of Zurich.

[2] Like Liu et al (2005), we prefer 'real-world' to alternative adjectives, such as 'subjective', 'objective', 'statistical', 'physical', 'true', 'risk-adjusted' and 'historical'. We use 'historical' to refer to densities that are obtained from time series of prices for the underlying asset.

however, been obtained for option expiry dates alone. Furthermore, only *ex post* information has been used to estimate transformation parameters. Our second contribution is to obtain and study *ex ante,* real-world densities for seven forecast horizons that range from one day to twelve weeks.

There is a vast literature that compares volatility forecasts obtained from historical asset prices and current option prices, surveyed by Poon and Granger (2003) and Taylor (2005). In contrast, we are only aware of two prior studies that make similar comparisons for density forecasts, namely Anagnou-Basioudis et al (2005) and Liu et al (2005) for small samples of forecasts for option expiry dates. Our third contribution is to compare ARCH and option-based forecasts for multiple horizons.

As option forecasts of volatility are often more accurate than historical forecasts, even when these are based upon intraday returns (Martens and Zein (2004), Jiang and Tian (2005)), we anticipate that a similar conclusion may apply to density forecasts. We provide the first results for density forecasts obtained from intraday returns. We find that real-world density forecasts obtained from option prices are superior to historical forecasts obtained from daily and intraday returns for horizons of one, two and four weeks, but historical forecasts rank very highly for the one-day horizon. Weighted combinations of historical and option densities outperform densities obtained from only one of the two sources of price information, however, for the shortest horizons of one day, one week and two weeks.

Our methodology requires us to specify a continuous-time, stochastic process for the underlying asset price, whose parameters can be estimated rapidly from daily panels of option prices. An appropriate process for a stock index must incorporate a stochastic volatility component, whose increments can have a general level of correlation with price increments. The price dynamics of Heston (1993) satisfy all of our requirements: these dynamics assume that the variance of asset prices follows a square-root process, and they provide closed-form

formulae for densities and option prices, based upon the numerical inversion of characteristic functions. Several previous studies have estimated the Heston parameters from S&P option prices, including Bakshi, Cao and Chen (1997), Nandi (1998) and Bates (2000).

More complicated price dynamics, that include jumps in prices and/or volatility, have been studied by Bates (1996, 2000), Bakshi, Cao and Chen (1997), Bollerslev and Zhou (2002), Duffie, Pan and Singleton (2000), Pan (2002), Eraker, Johannes and Polson (2003), Carr and Wu (2004), Eraker (2004) and Huang and Wu (2004). We do not consider jump components, firstly because it is difficult to estimate the additional parameters from daily panels of option prices and secondly because our transformations from risk-neutral to real-world densities are able to remove systematic mis-specifications of the risk-neutral densities.

It is possible that the mathematical sophistication of the Heston process and the extensions above may be counterproductive when the final goal is to produce real-world densities. Consequently, we also investigate transformations of risk-neutral, lognormal densities.

The positive risk premium for the aggregate equity market shows that some transformation must be applied to risk-neutral densities before appropriate, real-world, density forecasts can be made. Bliss and Panigirtzoglou (2004) evaluate single-parameter, utility transformations that can be motivated by a representative-agent model. They estimate the risk parameter by minimizing the diagnostic test statistic of Berkowitz (2001); this test assesses the uniformity and independence of observed, cumulative probabilities. Bliss and Panigirtzoglou find that *ex post*, real-world densities for the S&P 500 and FTSE 100 indices are a significant improvement upon their risk-neutral densities. Anagnou-Basioudis et al (2005) for the S&P 500, and Liu et al (2005) and Kang and Kim (2006) for the FTSE 100, also provide empirical results for utility transformations.

Liu et al (2005) estimate the two parameters of a more flexible transformation, by maximizing the *ex post*, likelihood of the observed index levels on the monthly, option expiry dates. We apply the same calibration transformation, but instead employ *ex ante* transformation parameters. These are obtained separately for seven forecast horizons, that are not restricted by the timing of option expiry dates. We also provide the first analysis of two further transformations from risk-neutral to real-world densities: one assumes the Heston dynamics apply in the real world by incorporating appropriate risk-premium functions, and the other applies a non-parametric calibration function.

All our density forecasting methods are described in Section 2. We consider historical densities obtained from ARCH models that are estimated from daily and intraday returns, risk-neutral (*Q*) densities that are either lognormal or provided by Heston's price dynamics, real-world (*P*) densities given by the three transformations of the *Q*-densities, and mixture densities that use all of the information derived from historical and option prices. The econometric methodology used to provide *ex ante* parameters and forecasts is presented in Section 3. We also present our criteria for making out-of-sample comparisons between the various sets of density forecasts.

The S&P 500 futures and options price data are described in Section 4. The empirical results are all contained in Section 5. We find that the best forecasts are given by the risk-transformations of the risk-neutral densities, for horizons of one, two and four weeks, when the ranking criterion is the out-of-sample likelihood of observed index levels. In contrast, the one-day ahead historical densities are superior to all but one set of *P*-densities. For horizons between six and twelve weeks, all the methods have similar out-of-sample likelihoods. A mixture of the best option-based and the best historical densities outperforms both components of the mixture, for the three shortest horizons. Standard diagnostic tests show

that the risk-transformed, option-based densities nearly always pass these tests, while the other density forecasting methods have more test failures. Finally, Section 6 concludes.

2. Density forecasts

2.1 Historical densities

By estimating ARCH models, the prices of the underlying asset up to and including time *t* can be used to produce historical density forecasts for the asset price at time $t+1$. One period of time defines a constant forecast horizon in this section, that may be one day, one week or several weeks. Supposing the underlying asset is a futures contract, we define one-period returns by $r_t = \log(F_t / F_{t-1})$; here F_{t-1} and F_t are end-of-period prices for the same contract.

A specific ARCH model uses price information I_t, known at the end of period *t*, to produce a parametric density function, $f(r_{t+1}|I_t)$, for the next return, r_{t+1}. The historical density for the next end-of-period price, F_{t+1}, is then:

$$g(F_{t+1}|I_t) = \frac{f(r_{t+1}|I_t)}{F_{t+1}}. \qquad (1)$$

We describe four specifications for the historical density $f(r_{t+1}|I_t)$.

The most elementary and credible ARCH model for a stock market index is the GJR(1, 1) model of Glosten, Jagannathan and Runkle (1993). The conditional variance h_t is then an asymmetric function of returns. We define the model as follows, with a constant conditional mean μ:

$$r_t = \mu + \varepsilon_t$$

$$\varepsilon_t = h_t^{1/2} z_t, \qquad z_t \sim \text{i.i.d.}(0,1),$$

$$\frac{h_t}{N_t} = \omega + \frac{(\alpha_1 + \alpha_2 d_{t-1})\varepsilon_{t-1}^2 + \beta_{GJR} h_{t-1}}{N_{t-1}}, \tag{2}$$

$$d_{t-1} = 1 \quad \text{if } \varepsilon_{t-1} < 0,$$
$$= 0 \quad \text{otherwise.}$$

The term N_t represents the number of trading days during period *t*, so that the conditional variance is proportional to the amount of trading time. Normal distributions for the i.i.d., standardized residuals z_t define a specification that we refer to as the GJR model. As it is well-known that fat-tailed, conditional distributions are preferable for daily horizons, we also define the GJR-t model by supposing the z_t have the standardized t-distribution, with degrees-of-freedom ν, as first evaluated in Bollerslev (1987).

Sums of squared intraday returns are superior to squared daily returns as measures of realized volatility (Andersen and Bollerslev, 1998, Andersen et al, 2001) and these sums can be used to improve volatility forecasts (Blair, Poon and Taylor, 2001, Martens and Zein, 2004). Let $Intra_t$ represent the total of some set of squared intra-period returns for period *t*. Then the Intra and Intra-t models are here defined by the conditional variance equation:

$$\frac{h_t}{N_t} = \omega + \frac{\gamma Intra_{t-1} + \beta_{Intra} h_{t-1}}{N_{t-1}} \tag{3}$$

with, respectively, conditional normal distributions and conditional t-distributions.

2.2 Risk-neutral densities

Almost all of the methods that estimate risk-neutral densities (RNDs) from option prices only provide densities for the underlying asset at option expiry times. To obtain densities for all future times, however, it is necessary to specify the risk-neutral dynamics of the underlying asset price. We consider two specifications. The first ignores the stochastic

property of volatility and simply assumes that prices follow geometric Brownian motion (GBM). All the RNDs are then lognormal. The second specifies a risk-neutral volatility process that ideally provides theoretical option prices that are compatible with observed prices.

The stochastic volatility process of Heston (1993) is a natural candidate, because it provides satisfactory, closed-form, theoretical option prices. The continuous-time, risk-neutral dynamics for a futures price, F_t, are given by supposing that the stochastic variance, V_t, follows the square-root process of Cox, Ingersoll and Ross (1985):

$$dF/F = \sqrt{V}\,dW_1,$$

and

$$dV = \kappa(\theta - V)dt + \xi\sqrt{V}\,dW_2, \tag{4}$$

with correlation ρ between the increments of the two Wiener processes, $W_{1,t}$ and $W_{2,t}$. The time *t* is now measured in years. The special case of GBM, with constant volatility θ, occurs when $V_0 = \theta$ and $\xi = 0$.

Several futures contracts are traded at the same time, with different expiry dates. We suppose that their prices satisfy standard, no-arbitrage, equations that imply the same variance process and the same parameters ($\kappa, \theta, \xi, \rho$) are applicable to all contracts.

Heston (1993) provides an analytic formula for the characteristic function of $\log(F_T)$, for given initial values F_0 and V_0 and any positive time *T*. Our notation for this characteristic function is $\tilde{g}(\psi) = E^Q[\exp(i\psi \log(F_T))]$, with ψ a real number and *Q* the risk-neutral measure. The following inversion formula then gives the risk-neutral density of F_T, denoted by $g_{Q,T}(x)$, for positive values of *x*:

$$g_{Q,T}(x) = \frac{1}{\pi x}\int_0^\infty \mathrm{Re}[\exp(-i\psi \log(x))\tilde{g}(\psi)]\,d\psi. \tag{5}$$

A straightforward numerical integration is required for each value of *x*. Heston (1993) also proves that the fair price of a European call option, whose strike is *K*, has the form:

$$c(F_0, K) = e^{-rT}\left(F_0 P_1(F_0, K) - K P_2(F_0, K)\right) \tag{6}$$

where *r* is the risk-free rate, $P_2(F_0, K)$ is the risk-neutral probability that the option expires in-the-money and $P_1(F_0, K)$ is a different probability for the same event when a different measure is applied. Both $P_1(F_0, K)$ and $P_2(F_0, K)$ are obtained by inversion formulae that require numerical integration.

2.3 Real-world densities

The risk-neutral density will always be incorrectly specified if it is used to make statements about real-world probabilities determined by a real-world measure *P*. Although a risk premium that compensates for price risk ensures this conclusion, it is possible that there is also a volatility risk premium. Furthermore, the assumption of a square-root process for volatility is at best a convenient approximation.

Transformations from risk-neutral to real-world densities rely on assumptions. These can be provided by a representative agent model, by specifying risk-premia functions or by calibration theory. We prefer the additional flexibility provided by either two risk-premium functions or the two-parameter calibration transformation of Fackler and King (1990) and Liu et al (2005) to the one-parameter utility transformation of Bliss and Panigirtzoglou (2004) and Anagnou-Basioudis et al (2005). We also investigate a non-parametric calibration transformation.

a) Risk-premium transformations

The Heston dynamics in (4) describe a risk-neutral, affine diffusion process. An affine real-world process is defined by including linear drift terms in both the price and the variance equations, thus:

$$dF/F = \lambda_1 V dt + \sqrt{V}\, d\tilde{W}_1,$$

and

$$dV = [\lambda_2 V + \kappa(\theta - V)]dt + \xi\sqrt{V}\, d\tilde{W}_2 . \tag{7}$$

The assumption of linear functions for the risk premia ensures the availability of analytic formulae for the real-world, characteristic functions of future prices. The inversion formula (5) then provides real-world densities $g_{P,T}(x)$ that depend on the premium coefficients, λ_1 and λ_2.

b) Calibration theory

At time 0, suppose $g_{Q,T}(x)$ and $G_{Q,T}(x)$ respectively denote the risk-neutral density and cumulative distribution function (c.d.f.) of the random variable F_T, and then define $U_T = G_{Q,T}(F_T)$. Following Bunn (1984), let the calibration function $C_T(u)$ be the real-world c.d.f. of the random variable U_T; our notation emphasizes that the calibration function depends on the forecast horizon *T*. Now consider the real-world c.d.f. of F_T. With "Pr" referring to real-world probabilities, this c.d.f. is

$$\Pr(F_T \le x) = \Pr(G_{Q,T}(F_T) \le G(x)) = \Pr(U_T \le G_{Q,T}(x)) = C_T(G_{Q,T}(x)) . \tag{8}$$

Consequently, the real-world c.d.f. of F_T is:

$$G_{P,T}(x) = C_T(G_{Q,T}(x)) . \tag{9}$$

Also, if $u = G_{Q,T}(x)$, the real-world density of F_T is given by

$$g_{P,T}(x) = \frac{d}{dx} C_T(G_{Q,T}(x)) = \frac{du}{dx}\frac{dC_T(u)}{du} \\ = g_{Q,T}(x) c_T(u), \tag{10}$$

with $c_T(u)$ representing the density of U_T.

These formulae for the real-world density and c.d.f. can be implemented providing we make the assumption that the calibration function $C_T(u)$ is invariant through time and hence can be estimated.

c) A parametric calibration function

Our preferred parametric calibration function is the c.d.f. of the Beta distribution, recommended by Fackler and King (1990) in their innovative study of densities obtained from commodity option prices. The c.d.f. is defined by the incomplete beta function:

$$C_T(u) = \int_0^u v^{j-1}(1-v)^{k-1} dv / B(j,k). \tag{11}$$

Here the constant $B(j,k)$ is defined by $B(j,k) = \Gamma(j)\Gamma(k)/\Gamma(j+k)$.

There are two calibration parameters, *j* and *k*, that are expected to depend on the horizon *T*. The special case $j = k = 1$ defines a uniform distribution and then the risk-neutral and real-world densities are identical. Liu et al (2005) show that the calibration transformation may define the same risk transformation as a utility function within a representative agent model. The necessary and sufficient conditions for an implicit risk-averse utility function are $k \le 1 \le j$ (with $j \ne k$).

From (10), the real-world density is

$$g_{P,T}(x) = \frac{G_{Q,T}(x)^{j-1}(1-G_{Q,T}(x))^{k-1}}{B(j,k)} g_{Q,T}(x). \tag{12}$$

This equation can be evaluated rapidly using numerical methods.

d) A non-parametric calibration function

An empirical calibration function can be estimated using a sample of *n* observations, $\{u_1, u_2, ..., u_n\}$, from the distribution of $U_T = G_{Q,T}(F_T)$, when we can assume these observations are i.i.d. with c.d.f. given by $C_T(u)$.

Let $\phi(.)$ and $\Phi(.)$ respectively denote the density and the c.d.f. of the standard normal distribution; also let $\Phi^{-1}(.)$ denote the inverse function defined by $\Phi^{-1}(\Phi(y)) = y$. We transform the observations u_i, whose domain is from 0 to 1, to new variables $y_i = \Phi^{-1}(u_i)$ and then fit a nonparametric, kernel c.d.f. to the set $\{y_1, y_2, ..., y_n\}$. We use a normal kernel, with bandwidth *B*, and so obtain the kernel density and c.d.f. respectively as:

$$\hat{h}_T(y) = \frac{1}{nB}\sum_{i=1}^{n}\phi\left(\frac{y - y_i}{B}\right), \qquad \text{and}$$

$$\hat{H}_T(y) = \frac{1}{n}\sum_{i=1}^{n}\Phi\left(\frac{y - y_i}{B}\right). \tag{13}$$

The empirical calibration function is then

$$\hat{C}_T(u) = \hat{H}_T(\Phi^{-1}(u)) \tag{14}$$

and, from (9), the real-world c.d.f. becomes

$$G_{P,T}(x) = \hat{C}_T(G_{Q,T}(x)). \tag{15}$$

With $u = G_{Q,T}(x)$ and $y = \Phi^{-1}(u)$, the real-world density is then:

$$\begin{aligned} g_{P,T}(x) &= \frac{d}{dx}\hat{H}_T(\Phi^{-1}(G_{Q,T}(x))) = \frac{dy}{dx}\frac{d\hat{H}_T(y)}{dy} = \frac{du}{dx}\frac{dy}{du}\hat{h}_T(y) \\ &= \frac{g_{Q,T}(x)\hat{h}_T(y)}{\phi(y)}. \end{aligned} \tag{16}$$

Once more, it is easy to evaluate this density using numerical methods.

2.4 Mixture densities

At some general time *t*, both ARCH densities and option-based densities may contain incremental information about the asset price at a later time $t+T$. Consequently, we also evaluate the mixture density:

$$g_{mix,T}(x) = \alpha g_{P,T}(x) + (1-\alpha) g_{ARCH,T}(x), \quad 0 \le \alpha \le 1. \tag{17}$$

As option traders know the historical price information, it is possible that $\alpha = 1$ if the transformations are able to translate an "efficient" risk-neutral density into the best possible real-world density. At the other extreme, $\alpha = 0$ might occur if option prices contain no real-world information that is incremental to the historical record of asset prices.

3. Empirical methods

3.1 Estimation of parameters

The historical, the risk-neutral and the real-world densities are all parametric. We always use *ex ante* estimates of parameters, to ensure that all density forecasts are evaluated out-of-sample. Consequently, all parameters required at time *t* are estimated by only using information available at time *t*. While the parameters of the Heston process are estimated daily from option prices by minimizing a least-squares function, all the other parameters are estimated by maximizing the log-likelihood function of selected, observed asset prices.

The parameters of the continuous-time, risk-neutral processes for asset prices are estimated at the end of each trading day. The estimated volatility of the GBM process is provided by the simplest credible estimate, namely the end-of-day, nearest-the-money implied volatility for the nearest-to-expiry options. For the Heston process defined by (4), at the end of day *n* we estimate the initial variance V_n, the three volatility parameters, κ_n, θ_n

and ξ_n, and the correlation ρ_n between the price and volatility differentials. Suppose N_n European, call[3] option contracts are traded on day *n*, labeled by $i = 1,..., N_n$, with strikes $K_{n,i}$, expiry times $T_{n,i}$ and market prices $c_{n,i}$; also, suppose $F_{n,i}$ is the futures price for the asset after $T_{n,i}$ years. Then the five *Q*-parameters are estimated by minimizing

$$\sum_{i=1}^{N_n} (c_{n,i} - c(F_{n,i}, K_{n,i}, T_{n,i}, V_n, \kappa_n, \theta_n, \xi_n, \rho_n))^2 . \tag{18}$$

with *c*(.) the Heston pricing formula, given by (6). We denote these risk-neutral parameters estimates obtained at time *n* by Θ_n.

Maximum likelihood estimates

The parameters of the risk-premium functions that are assumed in (7), namely λ_1 and λ_2, are estimated separately for each of seven forecast horizons. Each horizon defines a set of time periods. For one of these sets, at the end of period *s* corresponding to day n_s, we can numerically evaluate the real-world density $g_{P,s,T}(x|\lambda_1, \lambda_2, \Theta_{n_s})$ for the asset price *T* years later, at the end of period $s+1$. The *ex ante* maximum likelihood estimates (MLEs) of λ_1 and λ_2 at time *t* are given by maximizing the log-likelihood function of the observed asset prices F_s at the ends of periods $s = 1,2,...,t.$ Thus we maximize:

$$\log L(F_1,...,F_t|\lambda_1, \lambda_2) = \sum_{s=0}^{t-1} \log(g_{P,s,T}(F_{s+1}|\lambda_1, \lambda_2, \Theta_{n_s})) . \tag{19}$$

Likewise, the parameters of the parametric calibration function (11), namely *j* and *k*, are also estimated separately for each of several forecast horizons. The risk-neutral density

[3] We only include call prices in the estimation function, (18). We explain in Section 4.3 that the put prices in our database are converted to equivalent European call prices, using the put-call parity relationship, and are then included in (18).

$g_{Q,s,T}(x|\Theta_{n_s})$ and its c.d.f are used to evaluate the real-world density $g_{P,s,T}(x|j,k,\Theta_{n_s})$ given by (12). The *ex ante* MLEs of *j* and *k* at time *t* are also given by maximizing the log-likelihood function of the available, observed asset prices, i.e.:

$$\log L(F_1,...,F_t|j,k) = \sum_{s=0}^{t-1} \log(g_{P,s,T}(F_{s+1}|j,k,\Theta_{n_s})). \tag{20}$$

The non-parametric calibration function is re-estimated at the end of each period *t*. The observed prices define the cumulative, risk-neutral probabilities, $u_{s+1} = G_{Q,s,T}(F_{s+1}|\Theta_{n_s})$, from which we can obtain the real-world density by using (13) and (16). The bandwidth *B* in (13) should decrease as *t* increases. We have used the standard formula of Silverman (1986), $B = 0.9\sigma_y / t^{0.2}$, with σ_y the standard deviation of $\{y_s = \Phi^{-1}(u_s), 1 \le s \le t\}$.

The ARCH densities for one-period returns, specified by (2) and (3), have the general form $f(r_s|I_{s-1},\vartheta)$, that depends on a parameter vector ϑ and a set I_{s-1} of historical returns. The MLE at time *t* is the vector $\hat{\vartheta}_t$ that maximizes the log-likelihood function of all the returns since some earlier time τ (assumed to precede the first available option prices):

$$\log L(r_\tau,...,r_1,...,r_t|\vartheta) = \sum_{s=\tau}^{t} \log(f(r_s|I_{s-1},\vartheta)). \tag{21}$$

From (1), the *ex ante* density of the next end-of-period price, F_{t+1}, is then given by

$$g_{ARCH,t}(x|I_t,\hat{\vartheta}_t) = f(r|I_t,\hat{\vartheta}_t)/x, \quad \text{with } r = \log(x/F_t). \tag{22}$$

Similarly, the MLE of the mixture parameter α , that determines the weights given to option-based and historical densities in (17), can be obtained by maximizing an appropriate log-likelihood function. We use a two-step method. The first-step provides estimates of all the parameters except α . Then, at time *t* we will know the observed values of the

components of the mixture, for example we know $\tilde{g}_{P,s} = g_{P,s,T}(F_{s+1}|\hat{j}_s, \hat{k}_s, \Theta_{n_s})$ and $\tilde{g}_{A,s} = g_{ARCH,s}(F_{s+1}|I_s, \hat{\vartheta}_s)$ for times $0 \le s < t$. The MLE of α at time *t* is given by the number $\hat{\alpha}_t$ that maximizes

$$\log L(F_1, F_2, ..., F_t|\alpha) = \sum_{s=0}^{t-1} \log(\alpha\, \tilde{g}_{P,s} + (1-\alpha)\tilde{g}_{A,s}). \tag{23}$$

The *ex ante* mixture density for F_{t+1} is then

$$\hat{\alpha}_t g_{P,t,T}(x|\hat{j}_t, \hat{k}_t, \Theta_{n_t}) + (1-\hat{\alpha}_t) g_{ARCH,t}(x|I_t, \hat{\vartheta}_t). \tag{24}$$

3.2 Evaluation of the density forecasts

The forecasts are assessed using several numerical criteria, that include the out-of-sample likelihood and the values of diagnostic test statistics. For a fixed forecast horizon, suppose method *m* provides a series of density forecasts $g_{m,t}(x)$, made at times $v \le t \le w$, that are to be evaluated at times from $v+1$ to $w+1$ inclusive.

Our preferred method is the one that maximizes the out-of-sample, log-likelihood of observed asset prices. This equals

$$L_m = \sum_{t=v}^{w} \log(g_{m,t}(F_{t+1})). \tag{25}$$

The same criterion is used by Bao, Lee and Saltoglu (2004) to compare ARCH density forecasting methods applied to equity indices. The out-of-sample, log-likelihood is a special case of the weighted log-likelihood criterion used by Amisano and Giacomini (2005) to test for differences between the accuracy of competing forecasts.

Note that if one of the methods, say method *M*, correctly specifies the densities then it will have the highest expected log-likelihood. This follows from a property of the information criterion of Kullback and Liebler (1951), also called relative entropy; it is defined by

$$E[\log(g_{M,t}(F_{t+1})/g_{m,t}(F_{t+1}))] = \int_0^\infty g_{M,t}(x)\log(g_{M,t}(x)/g_{m,t}(x))\,dx\,. \qquad (26)$$

The relative entropy is positive whenever the two densities are continuous and distinct. Consequently, $E[L_M] > E[L_m]$ for $m \neq M$ and we may expect the sample value of L_M to exceed that of L_m when the number of forecasts made is sufficiently large. When none of the methods correctly specifies the densities, maximizing the likelihood criterion will select the densities that are nearest to the true densities according to the information criterion (Bao et al, 2004).

Several diagnostic tests are available to assess the adequacy of a set of forecasts. We focus on tests that use the time series of observed cumulative probabilities, $\{u_{v+1},...,u_{w+1}\}$, as recommended by Diebold, Gunther and Tay (1998). For a general method *m* these probabilities are defined by

$$u_{t+1} = \int_0^{F_{t+1}} g_{m,t}(x)dx\,. \qquad (27)$$

We check first whether or not the values of *u* are consistent with i.i.d. observations from the uniform distribution between zero and one. The Kolmogorov-Smirnov (KS) test is used, that relies on the maximum difference between the sample and theoretical cumulative functions. The sample c.d.f. of $\{u_{v+1},...,u_{w+1}\}$, evaluated at *u*, is the proportion of outcomes less than or equal to *u*, i.e.:

$$\tilde{C}(u) = \frac{1}{w-v+1}\sum_{t=v+1}^{w+1} S(u-u_t)\,, \qquad (28)$$

with $S(x) = 1$ if $x \geq 0$, and $S(x) = 0$ if $x < 0$. The test statistic is then

$$\mathrm{KS} = \max_{0\leq u\leq 1}\left|\tilde{C}(u) - u\right|. \qquad (29)$$

Secondly, we use the test of Berkowitz (2001) to check whether the values of $y = \Phi^{-1}(u)$ are consistent with the null hypothesis of i.i.d. observations from a standard normal distribution. The alternative hypothesis for the test is a stationary, Gaussian, AR(1) process with no restrictions on the mean, variance and autoregressive parameters. The test is decided by comparing a likelihood-ratio statistic (LR3) with χ_3^2; LR3 is defined by the maximum log-likelihood of $\{y_{v+1},..., y_{w+1}\}$ for the alternative hypothesis minus the log-likelihood for the null[4].

4. Data

The underlying assets for the density forecasts are futures contracts written on the S&P 500 index, traded at the CME. We investigate density forecasts for the futures price, rather than the spot level of the index, because of the availability of contemporaneous, settlement prices for futures and options contracts. A second advantage of working with futures data is that it is not necessary to consider dividend payments on the stocks that define the index. Option prices for S&P 500 futures have also been studied by Bates (2000), Bliss and Panigirtzoglou (2004) and Anagnou-Basioudis et al (2005).

4.1 Futures prices

End-of-day settlement prices and intraday prices for S&P 500 futures contracts are studied from 28 April 1982 until 31 December 2004. The settlement and intraday prices are

[4] Bliss and Panigirtzoglou (2004) estimate their risk transformation parameters by minimizing the Berkowitz test statistic. We only use the Berkowitz test as a diagnostic test. Extensive Monte Carlo simulations, that we will report elsewhere, compare density parameters estimated either by maximizing the likelihood function or by minimizing the Berkowitz statistic. The likelihood estimator is never less accurate than the Berkowitz alternative and it is much more accurate for estimating the mixture parameter α. Also see Bao et al (2004) who document theoretical advantages from applying likelihood methods to prices and not to the values of $y = \Phi^{-1}(u)$.

respectively provided by the CME and Price-Data.com. S&P 500 futures are contracts written on the index level on the third Fridays of March, June, September and December. All returns are calculated from the nearest contract, except on the expiry days and on the Thursdays that precede them when the next contract is used.

The high-frequency, realized variances are calculated from five-minute returns. This frequency provides a satisfactory trade-off between maximizing the accuracy of volatility estimates and minimizing the bias attributable to microstructure effects (Bandi and Russell, 2006). As the S&P 500 futures contracts are traded from 08:30 to 15:15 local time at the CME, we use 81 intraday returns for each day. The realized variance for day *t* is defined as the sum of the squares of the five-minute returns $r_{t,i}$:

$$Intra_t = \sum_{i=1}^{81} r_{t,i}^2 . \tag{30}$$

4.2 Interest rates

Three-month, six-month and one-year U.S. Treasury bill rates have been converted to continuously compounded rates. The risk-free rate *r* employed in an option pricing formula is then the three-month rate for option lives up to three months, otherwise the rate is given by linear interpolation.

4.3 Option prices

We study the prices of options on S&P 500 futures contracts for fifteen years, from 2 January 1990 to 31 December 2004. These prices post-date the crash of October 1987, after which the skewness of risk-neutral densities became significantly negative (Jackwerth and Rubinstein (1996)).

We consider prices for each option-on-futures contract that expires on the same Friday as its underlying futures contract. We do not use the prices of the remaining option contracts that expire one or two months earlier than the futures, because they are less actively traded. Option contracts with seven or less calendar days until expiry are excluded.

Call and put settlement prices for the same strike and expiry date should theoretically contain the same information. Either the call or the put will be out-of-the-money (OTM), except for the rare occasions when both are at-the-money (ATM). We choose to only use the information provided by the prices of OTM and ATM options, since the in-the-money option contracts are usually less actively traded.

The option contracts are American. We obtain equivalent European option prices from the American prices, that have the same implied volatility when the pricing formulae are those of Black (1976) and Barone-Adesi and Whaley (1987). The early exercise premia are small for OTM options and hence only very small errors can be created by approximating the premia by using the formulae of Barone-Adesi and Whaley. Finally, the put-call parity equation for options on futures is used to obtain equivalent European call prices from the European OTM put prices.

We analyze 435,100 option prices for the 3,777 trading days from 1990 to 2004. The average number of option prices studied per day is 115, made up of 45 OTM calls and 70 OTM puts. The number of different expiry dates available on any day is 2, 3 or 4 and the average number is 3.1. Table 1 summarizes the quantity, the moneyness (defined as F/K), and the time-to-expiry of the contracts that provide the observed prices. There are far more prices for deep-OTM put options than for deep-OTM call options, reflecting the greater demand for put options.

4.4 Volatility comparisons

Figure 1 shows, as dark dots, the time series of the implied volatilities obtained from the nearest-to-the-money options that are nearest to expiry. Figure 1 also shows, as light dots, the time series of the annualized, intraday, realized standard deviation, defined by $\sqrt{252 Intra_t}$.

The intraday and option measures of volatility move together, as expected, but the implied level is higher than the intraday, realized level. The first reason for this is that the intraday measure excludes the price variation from the market's close until the market reopens. From 1990 until 2004, the average of the daily realized variances (0.93×10^{-4}) equals 80% of the variance of the daily returns (1.17×10^{-4}). The second reason is a systematic difference between historical variance and risk-neutral variance: the average of the squared implied volatilities on Figure 1, converted to daily units (1.32×10^{-4}) is 113% of the variance of the daily returns. Stated as annualized, standard deviations, these average measures of volatility are 15.3% for Intra, 17.2% for daily returns and 18.3% for near-the-money options. The higher level for risk-neutral volatility is to be expected, because of the evidence supporting a negative, volatility risk premium in U.S. equity indices; this has been documented in several papers, including Jackwerth and Rubinstein (1996), Bakshi and Kapadia (2003) and Bollerslev, Gibson and Zhou (2005).

5. Empirical results

Density forecasts are evaluated for the fourteen years from January 1991 until December 2004 inclusive. The option prices during 1990 are only used to contribute to the *ex ante* information that is required to estimate the transformations from risk-neutral to real-world densities.

The density forecasts are made for seven horizons: one trading day and one, two, four, six, eight and twelve weeks. The first forecast for each horizon is made on Wednesday, 3 January 1990. The forecasts for the multi-day horizons do not overlap and they are all made on Wednesdays. The forecast density at time *t* for time $t+T$ always refers to the first futures contract that matures at least one calendar day after time $t+T$.

5.1 Illustrative density plots

The final one-day ahead densities, calculated from the information available on 30 December 2004, are shown on Figures 2a, 2b and 2c. Figure 2a shows the four historical densities; the conditional t-densities have higher peaks and fatter tails than the conditional normal densities. Figures 2b and 2c respectively show how the risk-transformations change the shapes of the lognormal and the Heston risk-neutral densities. The labels P1, P2 and P3 for the real-world densities respectively refer to the parametric calibration, the non-parametric calibration and the risk-premium transformations. It can be seen that each transformation increases the peak of the density and decreases the probability of a large price change, consistent with the real-world density having a lower standard deviation than the risk-neutral density.

Illustrative densities for the longer horizon of four weeks, calculated on 17 November 2004, are shown on Figures 3a, 3b and 3c. The real-world standard deviations are again less than the risk-neutral levels. From Figures 2c and 3c it is seen that the illustrative P1-densities are very similar to the corresponding P3-densities.

5.2 Historical densities

The parameters of the one-day ahead, ARCH densities have been estimated from daily and intraday return data that commences on 4 January 1988 and thus post-dates the

crash of October 1987. Referring to equation (2), the averages of the *ex ante*, GJR parameter estimates, used in the densities from 1991 onwards, include $\alpha_1 = 0.032$, $\alpha_2 = 0.043$ and $\beta = 0.913$. For the more credible GJR-t specification, with degrees-of-freedom ν, the averages include $\nu = 4.70$, $\alpha_1 = 0.009$, $\alpha_2 = 0.046$ and $\beta = 0.960$, with the persistence parameter equal to $\alpha_1 + 0.5\alpha_2 + \beta = 0.992$. The corresponding averages for the Intra-t specification, given by (3), are $\nu = 4.76$, $\gamma = 0.145$ and $\beta = 0.864$.

The ARCH densities for the one-week and longer periods are estimated from data that commences on 28 April 1982. All the averages for the degrees-of-freedom parameter indicate a high level of excess kurtosis in the conditional distributions. The averages of ν are 6.20 and 7.22 for the one-week returns, respectively for the GJR-t and Intra-t models, and they are 4.95 and 4.18 for the longest return period of twelve weeks.

Comparisons of the log-likelihoods of the four ARCH specifications favor the Intra-t model for all return periods, as will be shown in Section 5.7.

5.3 Risk-neutral Heston parameters

Several previous studies have estimated the parameters of Heston's continuous-time process from S&P 500 index levels and/or option prices. The risk-neutral, price dynamics are, in our notation:

$$dF/F = \sqrt{V}\,dW_1 \quad \text{and} \quad dV = \kappa(\theta - V)dt + \xi\sqrt{V}\,dW_2,$$

with correlation ρ between the two Wiener processes.

As researchers use a variety of markets (underlying and/or options), derivatives prices (none, spot options or futures options), different sample periods and different estimation methodologies, it is not surprising that their parameter estimates are rather diverse. We compare the risk-neutral estimates of κ, θ, ξ and ρ in Bakshi, Cao and Chen (1997),

Nandi (1998) and Bates (2000), provided in Table 2, with our risk-neutral estimates that are summarized in Table 3. We also tabulate historical and real-world parameters for five studies in Table 2. The estimates in Table 2 are for complete sample periods, while our estimates are obtained separately for each trading day. Our only non-trivial constraint is that the "half-life" parameter of the variance process, equal to $\log(2)/\kappa$ years, is at least one week, i.e. $\kappa \leq 36$.

The stochastic variance reverts towards the level θ. Our median estimate is 0.0452, which is equivalent to a volatility of 21.3%, compared with the risk-neutral estimates of 0.040 in Bakshi et al (1997), 0.028 in Nandi (1998) and 0.067 in Bates (2000); the real-world estimates in Table 2 are generally lower, which is consistent with the negative risk premium for S&P 500 volatility mentioned in Section 4.4.

The rate of reversion towards θ is determined by κ. Our median estimate of κ is 4.15; the "half-life" parameter of the variance process is then two months. Previous estimates of κ are generally lower, ranging from 0.93 in Chernov and Ghysels (2000) to 5.81 in Eraker, Johannes and Polson (2003).

The kurtosis of returns is primarily controlled by the "volatility of volatility" parameter ξ. Estimates obtained solely from option prices, such as our median value of 0.79, the 0.74 of Bates (2000) and the 1.28 of Nandi (1998), are much higher than those that are obtained from asset returns and option prices; a typical lower estimate is the 0.22 of Eraker (2004), that is obtained from returns data alone.

Our median estimate of the correlation ρ is -0.66. This is similar to the average, risk-neutral estimate of -0.64 in Bakshi et al (1997), which is far more negative than their estimate of -0.28 obtained from time series of asset returns and changes in implied volatilities.

5.4 Risk-neutral densities

The one-day ahead, risk-neutral densities, now denoted by $g_{Q,t}(x)$, provide cumulative distribution functions $G_{Q,t}(x)$ that are evaluated at the next, observed futures prices, F_{t+1}, to define observed probabilities defined by $u_{t+1} = G_{Q,t}(F_{t+1})$. As expected, the observed probabilities are incompatible with a uniform distribution. They are calculated separately for the lognormal and the Heston risk-neutral densities.

The sample c.d.f. calculated from a time series $\{u_{t+1}\}$ is denoted by $\tilde{C}(u)$ and defined by (28). We show the differences between sample and uniform probabilities, $\tilde{C}(u) - u$, as the dark curve on Figure 4b for the Heston densities. It is obvious that there are too few outcomes for *u* near to either zero or one; only 5.7% of the observed *u*-values are below 0.1 and only 6.6% of them are above 0.9. The maximum value of $\left|\tilde{C}(u) - u\right|$ equals 6.6%; this value of the Kolmogorov-Smirnov test statistic rejects the null hypothesis of a uniform distribution at the 0.01% level. The deviations $\tilde{C}(u) - u$ for the lognormal densities are similar, as can be seen from Figure 4a.

The shape of the deviation curves can be explained primarily by the fact that the risk-neutral, standard deviations are, on average, significantly higher than the historical standard deviations, as we observed in Section 4.4. Consequently, the risk-neutral probabilities of extreme price changes exceed the real-world probabilities.

5.5 Calibration transformations

A non-parametric estimate of the density of the probabilities u_{t+1} (from 1991 to 2004) is provided by differentiating the empirical, non-parametric, calibration function shown in (13). This estimated density, $\hat{c}(u)$, is shown by the light curves on Figures 5a and 5b, respectively for the lognormal and Heston cases.

The time series averages of the *ex ante* estimates for the parametric calibration transformation applied to the risk-neutral Heston densities are $j = 1.434$ and $k = 1.412$. The corresponding calibration density can be found from (11) and it is plotted as the dark curve on Figure 5b. It can be seen that the parametric and non-parametric calibration densities are similar, except near the end points of the distribution. The *ex ante* estimates of *j* and *k* vary between 1.3 and 1.6 and nearly always have $j > k$. The corresponding *ex ante* estimates for the risk-neutral lognormal densities are between 1.15 and 1.4 and their time series averages are 1.280 and 1.247, which define the calibration density shown by the dark curve on Figure 5a.

The calibration method is intended to produce real-world densities whose observed probabilities $\{u_{t+1}\}$ are uniformly distributed. After applying the parametric and the non-parametric calibration transformations, the deviations $\tilde{C}(u) - u$ estimated *ex ante* from all the data are shown as light curves on Figures 4a and 4b. It can be seen that these deviations are much nearer to zero than those for the risk-neutral densities, particularly for the non-parametric transformation.

Similar results and conclusions are obtained for the one-week ahead densities. The time series averages of the Heston-estimates of *j* and *k* are 1.424 and 1.409 respectively. For horizons of two or more weeks, the average Heston-estimate of *j* is between 1.45 and 1.58, and it is always more than the average estimate of *k* which ranges from 1.30 to 1.43.

Figure 6a shows the four-week ahead deviations $\tilde{C}(u) - u$ for the risk-neutral lognormal densities and their derived real-world densities, while Figure 6b shows the corresponding deviations that are based upon the risk-neutral Heston densities.

5.6 Risk-premium transformations

The third transformation of the risk-neutral Heston densities into real-world densities adjusts the drift rates of the price and the volatility and thereby incorporates both price and volatility risk premia. The premia coefficients λ_1 and λ_2 in the bivariate diffusion (7) have been estimated separately for each of the seven horizons. These estimates should be similar across horizons if the assumed risk-neutral and real-world dynamics are correct. We find that the seven estimates of λ_1 (the return risk premium per unit variance) are indeed similar, including 2.41, 2.25 and 2.86 for the one-day, one-week and two-weeks horizons estimated from the entire sample from 1991 to 2004. All the full-period estimates of λ_2 (the variance risk premium per unit of variance) are negative, which is compatible with the evidence elsewhere for a negative volatility premium. The estimates, however, are approximately proportional to the reciprocal of the forecast horizon, varying from -197 for the one-day horizon to -4.2 for the 12-week horizon. This empirical effect is consistent with the real-world variance at time *t* being systematically lower than the estimated initial level V_t of the stochastic process for the risk-neutral variance.

With $\lambda_1 > 0$ and $\lambda_2 < 0$, the risk-premium transformation ensures that the means and the standard deviations of the real-world densities are respectively above and below their risk-neutral counterparts. For the one-day horizon, Figure 4b shows that the risk-premium transformation reduces the magnitudes of the deviations $\tilde{C}(u) - u$ as expected. It can also be seen that the deviation curves for the risk-premium and the parametric calibration transformations are very similar.

A risk transformation of the risk-neutral lognormal densities has also been investigated. Only the single risk parameter λ_1 is then available, which improves the means but not the standard deviations of the densities. Consequently, the transformation only changes the log-likelihoods by minor amounts and so the results are not reported.

5.7 Likelihood comparisons

Table 4 shows the log-likelihoods of the observed futures prices from January 1991 until December 2004, for eleven *ex ante* density forecasting methods. These log-likelihoods are given for non-overlapping forecasts, made for seven horizons that range from one day to twelve weeks. We define the benchmark log-likelihoods as the values for the simplest historical method, namely the GJR densities. Table 4 shows the log-likelihood values in excess of the benchmark levels, for all other methods.

Initially, we consider the log-likelihoods of the four historical methods. These values are always higher for conditional t-densities than for the matched conditional normal densities. They are also always higher for densities obtained from high-frequency returns than for the matched densities obtained from one-period returns. Consequently, the best of the four methods is the Intra-t method for all seven horizons. At the shortest horizon of one day, incorporating non-normality adds more to the log-likelihood than incorporating intraday price information. The relative contributions of non-normality and intraday prices are equal for the one-week horizon, while intraday prices contribute more for horizons that are two weeks or longer.

Likelihood comparisons are now made between eight methods, which define Intra-t densities and seven option-based sets of densities. We refer to the risk-neutral densities as the lognormal-Q and the Heston-Q densities. The real world densities are labeled lognormal-P1, lognormal-P2, Heston-P1, Heston-P2 and Heston-P3; P1 refers to the parametric calibration transformation, P2 to the non-parametric calibration transformation and P3 to the risk-premium transformation.

One-day horizon

The log-likelihoods, in excess of the GJR benchmark value, are as follows in descending order: Heston-P2 127, Intra-t 123, Heston-P1 104, lognormal-P2 101, Heston-P3 94, lognormal-P1 73, lognormal-Q 27 and Heston-Q -2. These numbers are summarized in five remarks. First, the Intra-t densities obtained from high-frequency returns have high log-likelihoods compared with the option-based densities. Second, the Heston P-densities have higher log-likelihoods than the lognormal P-densities, the differences being 30.4 for P1 and 27.5 for P2. Third, the non-parametric risk transformation P2 is superior to the parametric transformation P1, with the differences respectively equal to 23.5 and 27.4 for the Heston and lognormal cases. Fourth, the risk-premium transformation P3 ranks behind the statistical transformations P1 and P2. Finally, as expected the Q-densities are seen to be far inferior to the P-densities, reflecting the fact that the transformations are able to diminish the impact of systematic mis-specification in the risk-neutral densities.

Horizons from one to four weeks

The *ex ante* best method depends on the horizon, being Heston-P2 for one day, Heston-P1 for one week, lognormal-P2 for two weeks and Heston-P3 for four weeks. The absence of a uniformly best method reflects the similarity of the log-likelihoods for the five option-based, P-densities, for all but the shortest horizon; the five one-week numbers in Table 4 range from 32.8 to 41.5, for two weeks from 22.4 to 27.8 and for four weeks from 15.8 to 23.2.

Next we note that the P-densities are superior to the Intra-t densities for 14 of the 15 possible comparisons when the horizon is between one and four weeks inclusive; for the single exception, the one-week Intra-t densities are slightly ahead of the lognormal-P1 densities. For the one-week horizon, the excess log-likelihood for Intra-t is 34.4 and the

average for the five sets of P-densities is 37.3. The corresponding comparisons are 19.2 versus 25.9 for two weeks and 12.3 versus 18.6 for four weeks.

Each set of P-densities always outperforms the corresponding set of Q-densities. The differences vary between 15.8 and 33.0 for one week, 7.6 and 14.2 for two weeks and from 0.9 to 7.2 for four weeks.

Horizons from six to twelve weeks

The differences between the log-likelihoods of the various methods decrease as the horizon increases, primarily because the numbers of non-overlapping forecasts decrease. The best methods for the longer horizons are Heston-P3 for six weeks, lognormal-P1 for eight weeks and Heston-Q for twelve weeks, while the worst is always lognormal-Q. The differences between the best and the worst methods are 3.6, 4.5 and 3.8 for these horizons.

Mixtures

Mixture densities are defined by (17). We now consider the log-likelihoods for mixtures defined by a fraction α of an option-based density added to a fraction $1-\alpha$ of the best historical density specification, which is given by the Intra-t model. Table 5 shows the log-likelihoods of *ex ante* mixture densities and the time series averages of the *ex ante* estimates of α. The mixtures always have a higher log-likelihood than their component densities for horizons of one day, one week and two weeks. The mixtures are, however, often inferior for horizons of four or more weeks.

The one-day mixtures of the Heston-P2 and the Intra-t densities give a 51% weight, on average, to the options-based densities and the remaining 49% weight to the historical densities. The log-likelihood of these mixtures is 143 above the benchmark level, which is far above the 127 of the Heston-P2 densities and the 123 of the Intra-t densities. The usual

likelihood-ratio test values for the null hypotheses $\alpha = 0$ and $\alpha = 1$ are then 41 and 32 respectively, which clearly reject these hypotheses when comparisons are made with the χ_1^2 distribution[5]. The same test conclusions are obtained for all of the other one-day mixture densities.

The average weight given to the five sets of P-densities increases as the horizon increases. These averages are 42% for the one-day horizon, 50% for one week, 61% for two weeks and 66% for four weeks. The null hypothesis $\alpha = 0$ is rejected at the 5% level for all three Heston P-mixtures for all horizons up to four weeks, but $\alpha = 1$ is only rejected at this level by all three mixtures for the one-week horizon.

5.8 Diagnostic tests

The Kolmogorov-Smirnov (KS) test statistic, defined by (29) as the maximum value of $\left|\widetilde{C}(u) - u\right|$, can be used to test the null hypothesis that a set of densities are correctly specified. This test makes the assumption that the observed cumulative probabilities are observations of independent random variables. Figures 4a and 4b show that there are high values of $\left|\widetilde{C}(u) - u\right|$ for the risk-neutral densities (RNDs) for short horizons, so that the KS test establishes that these densities are indeed mis-specified.

Table 6 lists the percentage p-values for the KS test for the eleven *ex ante* density forecasting methods, for each of the seven horizons. As the null hypothesis is rejected at the α% level whenever $p < \alpha$, it can be checked that 24 of the 77 test values reject the null hypothesis at the 5% level. Two-thirds of the 24 rejections occur for densities that might be expected to be mis-specified, namely the RNDs and the ARCH densities that are conditionally normal. It is noteworthy that the P2-densities obtained by applying the non-

[5] The standard, asymptotic theory needs to be interpreted cautiously, because we re-estimate α for each time period.

parametric transformation to the RNDs have the most satisfactory p-values: 9 of the 14 p-values exceed 50% and their minimum is 18%, while the values for the one-day horizon are 47% for Heston-P2 and 88% for lognormal-P2 compared with 6% for the highest value given by the other nine sets of one-day densities. We also note that the Heston-P1 and the Heston-P3 densities provide satisfactory p-values, except for the one-day horizon when the p-values are less than 0.5%.

The likelihood-ratio test statistic of Berkowitz (2001), denoted by LR3, tests the null hypothesis that the numbers $y_t = \Phi^{-1}(u_t)$ are i.i.d. observations from a standard normal distribution against the alternative that they are from a stationary, Gaussian, AR(1) process with no restrictions on the mean, variance and autoregressive parameters. Table 7 contains the values of LR3 and the MLEs of the variance and autoregressive parameters, once more for the eleven *ex ante* density forecasting methods and for each of the seven horizons.

The MLEs of the autoregressive parameter are all between -0.006 and 0.005 for the sets of one-day forecasts, so they provide no evidence to doubt that the time series $\{u_t\}$ are composed of independent observations. The corresponding MLEs for the one-week forecasts are, however, all between -0.12 and -0.08 inclusive; they reject the null hypothesis that the autoregressive parameter is zero at the 5% level for all the historical and all the option-based sets of densities. There is no significant evidence of time-series dependence for horizons of two weeks or longer.

The MLEs of the variance parameter are usually near one, as required for correctly specified densities. The low estimates for the RNDs, such as the 0.68 and the 0.79 for the one-day ahead Heston-Q and lognormal-Q forecasts, are a direct consequence of historical volatility being lower (on average) than risk-neutral volatility.

The null distribution of LR3 is χ_3^2 and thus a test value is significant at the 5% level if it exceeds 7.81. Table 7 shows that the null is always rejected for the lognormal-Q densities

at the 5% level. For the other methods, rejections of the null at the 5% level occur for various sets of one-day, one-week and two-week forecasts; there are no rejections for the longer horizons, which may reflect low power when few forecasts are evaluated. The only density forecasting methods whose densities always pass the LR3-test at the 5% level are the Heston-P1 and Heston-P3 methods. The highest significant values of LR3 are for the Q-densities, which can be explained by the substantial difference between the MLEs of their variances and the null value of one. Almost all of the other significant values of LR3 may be explained either by an incorrect normal assumption about the conditional shape of a historical density or by the negative estimates of the autoregressive parameter for the one-week horizon.

6. Conclusions

Option-based density estimation methods only provide results within a risk-neutral context and most methods require the forecast horizon to coincide with an option expiry date. In contrast, we have provided the first evidence that it is possible to construct informative, real-world densities for many forecast horizons by using currently available price information.

Jiang and Tian (2005) have shown that the information content of option prices is higher than that of daily and intraday index values when forecasting the volatility of the S&P 500 index. Our study shows that the same conclusion does not apply to *ex ante* density forecasts of the S&P 500 index when the forecast horizon is only one day, but it does apply for intermediate horizons between one and four weeks inclusive.

For the intermediate horizons we find that three transformations of the risk-neutral density, estimated from index levels, option prices and Heston's pricing formula, all provide real-world densities that are more informative than the historical densities estimated from ARCH models and high-frequency returns. We say the real-world densities are more

informative because they rank higher according to the out-of-sample likelihood criterion. The real-world densities also pass almost all of the standard diagnostic tests that we have evaluated.

Option prices are extrapolated, as a function of time-to-expiry, to obtain the one-day-ahead risk-neutral densities. This may explain why the best historical densities are relatively more successful than the real-world densities for the shortest horizon. A mixture of real-world and historical densities, however, outperforms the components of the mixture both for the one-day horizon and also for horizons equal to one or two weeks.

References

Ait-Sahalia, Y. and A.W. Lo, 1998, Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance*, 53, 499-547.

Amisano, G. and R. Giacomini, 2005, Comparing density forecasts via weighted likelihood ratio tests, *Journal of Business Economics and Statistics*, forthcoming.

Anagnou-Basioudis, I., M. Bedendo, S.D. Hodges and R. Tompkins, 2005, Forecasting accuracy of implied and GARCH-based probability density functions, *Review of Futures Markets*, 11, 41-66.

Andersen, T.G. and T. Bollerslev, 1998, Answering the skeptics: Yes standard volatility models do provide accurate forecasts, *International Economic Review*, 39, 885-905.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens, 2001, The distribution of stock return volatility, *Journal of Financial Economics*, 61, 43-76.

Bakshi G., C.G. Cao, and Z. Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance*, 52, 2003-2049.

Bakshi G. and N. Kapadia, 2003, Delta-hedged gains and the negative market volatility risk premium, *Review of Financial Studies*, 16, 527-566.

Bakshi, G., N. Kapadia and D.B. Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies*, 16, 101-143.

Bandi, F.M. and J.R. Russell, 2006, Separating microstructure noise from volatility, *Journal of Financial Economics*, 79, 655-692.

Bao, Y., T.-H. Lee and B. Saltoglu, 2004, Comparing density forecast models, *Journal of Forecasting*, forthcoming.

Barone-Adesi, G. and R.E. Whaley, 1987, Efficient analytic approximation of American option values, *Journal of Finance*, 42, 301-320.

Bates, D., 1996, Jumps and stochastic volatility: exchange rate processes implicit in Deutsche mark options, *Review of Financial Studies*, 9, 69-107.

Bates, D., 2000, Post-'87 Crash fears in S&P 500 futures options, *Journal of Econometrics*, 94, 181-238.

Benzoni. L., 2002, Pricing options under stochastic volatility: an empirical investigation, working paper, University of Minnesota.

Berkowitz, J. 2001, Testing density forecasts with applications to risk management, *Journal of Business and Economic Statistics*, 19, 465-474.

Black, F., 1976, The pricing of commodity contracts, *Journal of Financial Economics*, 3, 167-179.

Blair, B.J., S. Poon and S.J. Taylor, 2001, Forecasting S & P 100 volatility: the incremental information content of implied volatilities and high frequency index returns, *Journal of Econometrics*, 105, 5-26.

Bliss, R. and N. Panigirtzoglou, 2002, Testing the stability of implied probability density functions, *Journal of Banking and Finance*, 26, 381-422.

Bliss, R. and N. Panigirtzoglou, 2004, Recovering risk aversion from options, *Journal of Finance*, 59, 381-422.

Bollerslev, T., 1987, A conditionally heteroskedastic time series model for security prices and rates of return data, *Review of Economics and Statistics*, 59, 542-547.

Bollerslev, T., M.S. Gibson and H. Zhou, 2005, Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities, working paper, Duke University and Federal Reserve Board.

Bunn, D.W., 1984, *Applied Decision Analysis*, McGraw-Hill.

Carr, P. and L. Wu, 2004, Time-changed Lévy processes and option pricing, *Journal of Financial Economics*, 71, 113-141.

Chernov, M., and E. Ghysels, 2000, A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation, *Journal of Financial Economics*, 56, 407-458.

Cox, J.C., J.E. Ingersoll and S.A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica*, 53, 385-407.

Diebold, F.X., T.A. Gunther and A.S. Tay, 1998, Evaluating density forecasts with applications to financial risk management, *International Economic Review*, 39, 863-883.

Duffie, D., J. Pan, and K.J. Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica*, 68, 1343–1376.

Eraker, B., 2004, Do equity prices and volatility jump? reconciling evidence from spot and option prices, *Journal of Finance*, 59, 1367-1403

Eraker, B., M.S. Johannes, and N.G. Polson 2003, The impact of jumps in returns and volatility, *Journal of Finance*, 53, 1269–1300.

Fackler, P. L. and R. King, 1990, Calibration of option-based probability assessments in agricultural commodity markets, *American Journal of Agricultural Economics*, 72, 73-83.

Glosten, L.R., R. Jagannathan and D. Runkle, 1993, Relationship between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, 48, 1779-1801.

Heston, S., 1993, Closed-form solution of options with stochastic volatility with application to bond and currency options, *Review of Financial Studies*, 6, 327–343.

Huang, J.-Z. and L. Wu, 2004, Specification analysis of option pricing models based on time-changed Lévy processes, *Journal of Finance*, 59, 1405-1439.

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

Jiang, G.J. and Y.S. Tian, 2005, The model-free implied volatility and its information content, *Review of Financial Studies*, 18, 1305-1342.

Kang, B.J. and T.S. Kim, 2006, Option-implied risk preferences: an extension to wider classes of utility functions, *Journal of Financial Markets*, 9, 180-198.

Kullback, L. and R.A. Leibler, 1951, On information and sufficiency, *Annals of Mathematical Statistics* 22, 79-86.

Liu, X., M. B. Shackleton, S. J. Taylor and X. Xu, 2005. Closed-form transformations from risk-neutral to real-world distributions, working paper, Lancaster University.

Martens, M. and J. Zein, 2004, Predicting financial volatility: high-frequency time-series forecasts vis-à-vis implied volatility, *Journal of Futures Markets*, 24, 1005-1028.

Melick, W.R. and C.P. Thomas, 1997, Recovering an asset's implied PDF from option prices: an application to crude oil during the Gulf crisis, *Journal of Financial and Quantitative Analysis*, 32, 91-115.

Nandi, S., 1998, How important is the correlation between returns and volatility in a stochastic volatility model? empirical evidence from pricing and hedging in the S&P 500 index options market, *Journal of Banking and Finance*, 22, 589-610.

Pan, J., 2002, The jump risk premia implicit in options: evidence from an integrated time-series study, *Journal of Financial Economics*, 63, 3-50.

Poon, S.-H. and C.W.J. Granger, 2003, Forecasting volatility in financial markets, *Journal of Economic Literature*, 41, 478-539.

Silverman, B.W., 1992, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall.

Taylor, S.J., 2005, *Asset Price Dynamics, Volatility, and Prediction*, Princeton University Press.

Table 1

Summary statistics for the S&P 500 futures option dataset

Information about the numbers of out-of-the-money (OTM) options on S&P 500 futures, from 1990 to 2004

		Total number	**Average options per day**	**Max number per day**	**Min number per day**
Calls		171,383	45	157	8
Puts		263,717	70	173	8
Overall		435,100	115	255	29
Number of cross-section			3.1	4	2
Moneyness\Maturity	*F/K*	1 month	Between 1 and 6 months	6 month	Subtotal
Deep OTM put	<0.90	10,800 (2.48%)	89,779 (20.63%)	39,879 (9.17%)	140,458 (32.28%)
OTM put	0.90–0.97	8,743 (2.01%)	52,427 (12.05%)	23,964 (5.51%)	85,134 (19.57%)
Near the money	0.97–1.03	7720 (1.77%)	47,325 (10.88%)	20,206 (4.64%)	75,251 (17.30%)
OTM call	1.03–1.10	6,881 (1.58%)	45,519 (10.46%)	19,178 (4.41%)	71,578 (16.45%)
Deep OTM call	>1.10	2,483 (0.57%)	42,253 (9.71%)	17,943 (4.12%)	62,679 (14.41%)
Subtotal		36,627 (8.42%)	277,303 (63.73%)	121,170 (27.85%)	435,100 (100.00%)

Table 2

Recent empirical studies of the Heston model using S&P 500 index data

The continuous-time dynamics for prices and volatility are

$$dS/S = \mu dt + \sqrt{V}\,dW_1 \quad \text{and}$$

$$dV = \kappa(\theta - V)dt + \xi\sqrt{V}\,dW_2\,,$$

for the spot index, with $dW_1 dW_2 = \rho dt$. The parameters have been estimated for the risk-neutral (Q) and real world (P) measures.

Paper	Measure	κ	ρ	ξ	θ	Data	Estimation method	Sample period	Paper Objective
Bakshi et al (1997)	*Q*	1.15	-0.64	0.390	0.040	Spot option	RMSE	1988-1991	Hedging and pricing
Nandi (1998)	*Q*	3.29	-0.79	1.280	0.028	Spot option	Likelihood on group-specific errors	1991-1992	Volatilities and returns correlation
Bates (2000)	*Q*	1.49	-0.57	0.742	0.067	Futures option	Likelihood on group-specific errors	1988-1993	Options and returns consistency
Chernov and Ghysels (2000)	*P*	0.93	-0.02	0.061	0.007	Spot option and returns	EMM	1985-1994	Hedging and pricing
Benzoni (2002)	*P*	3.93	-0.60	0.078	0.013	Spot option and returns	EMM	1996-1997	Pricing
Pan (2002)	*P*	5.3	-0.57	0.380	0.024	Spot option and returns	GMM	1989-1996	Pricing
Eraker (2004)	*P*	4.80	-0.57	0.22	0.049	Returns	MCMC	1987-1996	Pricing
Eraker et al (2003)	*P*	5.81	-0.40	0.143	0.023	Spot option and returns	MCMC	1987-1999	Pricing
This study	*Q*	4.93	-0.66	0.93	0.05	Futures option	RMSE	1990-2004	Density forecasts

Table 3

Summary statistics for the daily estimates of the Heston parameters

Estimates are summarised for the risk-neutral dynamics

$$dF/F = \sqrt{V}\,dW_1 \quad \text{and}$$

$$dV = \kappa(\theta - V)dt + \xi\sqrt{V}\,dW_2\,,$$

with $dW_1 dW_2 = \rho dt$. The parameters are estimated each day from 1990 to 2004, for the out-of-the-money options on S&P 500 futures, by minimizing the mean squared error (MSE) of the fitted option prices.

Parameters	Mean	Median	Max	Min	Standard Deviation
$\sqrt{V_t}$	0.1898	0.1787	0.6785	0.0558	0.0741
κ	4.9292	4.1528	36*	0.1940	3.6598
ρ	-0.6590	-0.6624	-0.3610	-0.9710	0.0875
ξ	0.9296	0.7925	7.3848	0.3243	0.5160
θ	0.0505	0.0452	0.2747	0.0169	0.0273
MSE	0.1621	0.0472	3.4433	0.0000	0.3126

*The constraint $\kappa \leq 36$ is applied.

Table 4

Log-likelihoods for sets of density forecasts

The numbers tabulated are the log-likelihoods of the GJR density forecasts and the log-likelihoods of the other sets of forecasts in excess of the GJR benchmark values. The risk transformation P1 refers to the parametric calibration transformation, P2 to the nonparametric calibration transformation, and P3 to the risk-premia transformation.

Forecast horizon	**Number of Obs.**	**GJR**	**GJR-*t***	**Intra**	**Intra-*t***	**Log normal**	**Risk-transformed Lognormal**		**Heston**	**Risk-transformed Heston**		
							P1	**P2**		**P1**	**P2**	**P3**
Data source		Daily returns	Daily returns	Intraday returns	Intraday returns	Options	Options	Options	Options	Options	Options	Options
1 day	3520	-11951.2	91.4	52.9	122.7	27.0	73.5	100.9	-2.4	103.9	127.4	93.7
1week	711	-2961.9	13.5	13.5	34.4	17.0	32.8	36.5	18.5	41.5	35.2	40.6
2 weeks	351	-1574.0	10.4	14.4	19.2	13.6	27.8	26.4	14.8	26.9	22.4	25.8
4 weeks	176	-853.6	4.1	7.6	12.3	12.5	13.4	15.8	16.0	20.3	20.2	23.2
6 weeks	115	-596.9	5.7	15.0	16.0	16.0	17.1	16.9	19.6	16.3	18.7	20.3
8 weeks	86	-446.1	1.5	2.8	5.3	4.9	9.4	9.3	6.7	7.8	7.5	7.5
12 weeks	58	-310.2	5.2	6.9	7.9	5.6	6.9	6.8	9.8	8.5	9.1	7.6

Table 5

Log likelihoods for mixtures of historical densities and option-based densities

Each log-likelihood is the value in excess of the GJR benchmark given in Table 4. The mixture densities are a fraction α of the option-based density plus a fraction $1-\alpha$ of the Intra-*t* density. α is estimated ex ante. The risk transformation P1 refers to the parametric calibration transformation, P2 to the nonparametric calibration transformation, and P3 to the risk-premia transformation.

Forecast Horizon	Intra-t only	Intra-*t* combined with													
		Lognormal						Heston							
		Q	Average α	P1	Average α	P2	Average α	*Q*	Average α	P1	Average α	P2	Average α	P3	Average α
1 day	122.7	134.4	23%	139.8	38%	132.1	39%	139.9	19%	151.5	42%	143.4	51%	151.5	38%
1 week	34.4	33.1	21%	38.4	58%	37.8	37%	34.7	15%	44.5	57%	38.0	37%	44.4	63%
2 weeks	19.2	19.4	38%	27.0	81%	25.6	44%	20.3	36%	27.5	58%	33.6	54%	27.0	69%
4 weeks	12.3	12.1	43%	13.5	80%	15.3	50%	16.3	78%	19.4	67%	18.3	57%	22.3	74%
6 weeks	16.0	16.3	59%	16.7	91%	16.1	73%	18.5	71%	18.9	58%	17.7	70%	19.9	98%
8 weeks	5.3	4.6	58%	9.4	100%	8.6	71%	6.2	69%	8.1	50%	7.1	53%	7.4	70%
12 weeks	7.9	6.8	34%	7.1	35%	7.4	36%	9.0	79%	9.2	54%	8.6	57%	8.2	69%

Table 6

Results from the Kolmogorov-Smirnov test

The tabulated numbers are the *p*-values for the Kolmogorov-Smirnov test of the null hypotheses that the variables u_t have a uniform distribution. The risk transformation P1 refers to the parametric calibration transformation, P2 to the nonparametric calibration transformation, and P3 to the risk-premia transformation.

Forecast horizon	Number of Obs.	GJR	GJR-*t*	Intra	Intra-*t*	Log normal	Risk-transformed Lognormal		Heston	Risk-transformed Heston		
							P1	P2		P1	P2	P3
1 day	3520	0.00%	0.75%	0.00%	6.31%	0.00%	0.06%	87.56%	0.00%	0.32%	46.72%	0.01%
1 week	711	0.36%	25.09%	0.10%	1.01%	0.04%	1.14%	74.35%	0.60%	96.08%	78.80%	67.58%
2 weeks	351	13.36%	76.69%	0.15%	88.00%	0.03%	21.60%	67.63%	3.52%	92.71%	66.42%	78.16%
4 weeks	176	13.09%	82.55%	0.60%	0.17%	0.04%	0.67%	24.86%	7.75%	82.81%	49.13%	69.38%
6 weeks	115	94.00%	83.30%	14.38%	5.65%	1.57%	14.28%	18.18%	54.23%	84.58%	80.00%	41.38%
8 weeks	86	59.42%	72.82%	5.77%	8.65%	1.40%	48.79%	98.95%	9.90%	84.75%	99.45%	51.56%
12 weeks	58	85.19%	89.06%	7.49%	8.86%	0.18%	21.62%	61.28%	31.29%	90.03%	91.03%	85.18%

Table 7

Berkowitz test values and parameters

The null hypothesis that the variables $y_t = \Phi^{-1}(u_t)$ are i.i.d with a standard normal distribution is tested against the alternative of an AR(1), Gaussian process. The tabulated numbers are the test statistic LR3 and the estimates of the AR and variance parameters. Stars indicate that the null is rejected at the 5% level, when LR3>7.81.

Forecast horizon		**GJR**	**GJR-*t***	**Intra**	**Intra-*t***	**Log normal**	**Risk-transformed Lognormal**		**Heston**	**Risk-transformed Heston**		
							P1	**P2**		**P1**	**P2**	**P3**
1 day	AR	0.00	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	0.00
	Variance	0.93	1.07	0.90	1.02	0.79	1.04	1.00	0.68	1.04	1.02	1.02
	LR3	9.78*****	8.44*****	19.06*****	0.91	96.82*****	2.58	0.09	241.46*****	3.72	1.02	0.64
1 week	AR	-0.11	-0.12	-0.10	-0.11	-0.10	-0.10	-0.09	-0.09	-0.09	-0.10	-0.08
	Variance	0.78	0.88	0.69	0.83	0.71	0.97	0.93	0.65	0.98	0.94	0.94
	LR3	26.57*****	13.93*****	50.37*****	20.38*****	46.13*****	8.51*****	8.27*****	61.63*****	6.14	8.50*****	5.52
2 weeks	AR	-0.03	-0.05	-0.05	-0.06	-0.02	-0.02	-0.01	-0.01	-0.01	0.00	0.00
	Variance	0.86	1.01	0.75	0.91	0.63	0.99	0.91	0.63	1.02	0.94	1.02
	LR3	3.90	0.89	17.11*****	2.72	35.95*****	0.94	2.11	32.32*****	0.34	1.39	0.33
4 weeks	AR	0.00	-0.01	0.06	0.04	0.04	0.04	0.04	0.06	0.05	0.05	0.08
	Variance	0.77	0.92	0.81	0.92	0.63	0.99	0.90	0.74	1.00	0.92	1.00
	LR3	6.04	0.98	6.07	4.01	19.61*****	0.93	1.43	7.39	0.71	1.41	1.29
6 weeks	AR	-0.22	-0.17	-0.09	-0.16	-0.14	-0.11	-0.10	-0.10	-0.09	-0.07	-0.03
	Variance	0.93	1.09	0.90	1.11	0.71	1.17	1.05	0.77	1.11	1.07	1.22
	LR3	6.03	4.19	4.39	7.45	11.63*****	3.81	1.89	6.48	1.89	1.74	2.50
8 weeks	AR	0.02	0.06	0.09	0.12	0.14	0.16	0.18	0.11	0.18	0.18	0.22
	Variance	0.85	1.24	0.83	0.85	0.54	0.94	0.79	0.74	0.93	0.80	0.91
	LR3	1.65	2.90	4.30	5.54	18.36*****	2.39	4.35	5.63	2.71	4.40	4.36
12 weeks	AR	-0.08	-0.04	0.03	0.02	0.06	0.08	0.11	0.09	0.10	0.15	0.19
	Variance	1.40	1.07	0.87	1.00	0.59	1.12	0.93	0.75	1.04	0.98	1.16
	LR3	4.41	0.46	4.46	3.98	10.95*****	1.06	0.82	4.92	0.90	1.62	3.18

Figure 1

Implied volatilities from the at-the-money, shortest-maturity options, and realized volatilities from intra-day returns

Dark dots are implied volatilities, light dots are realized volatilities

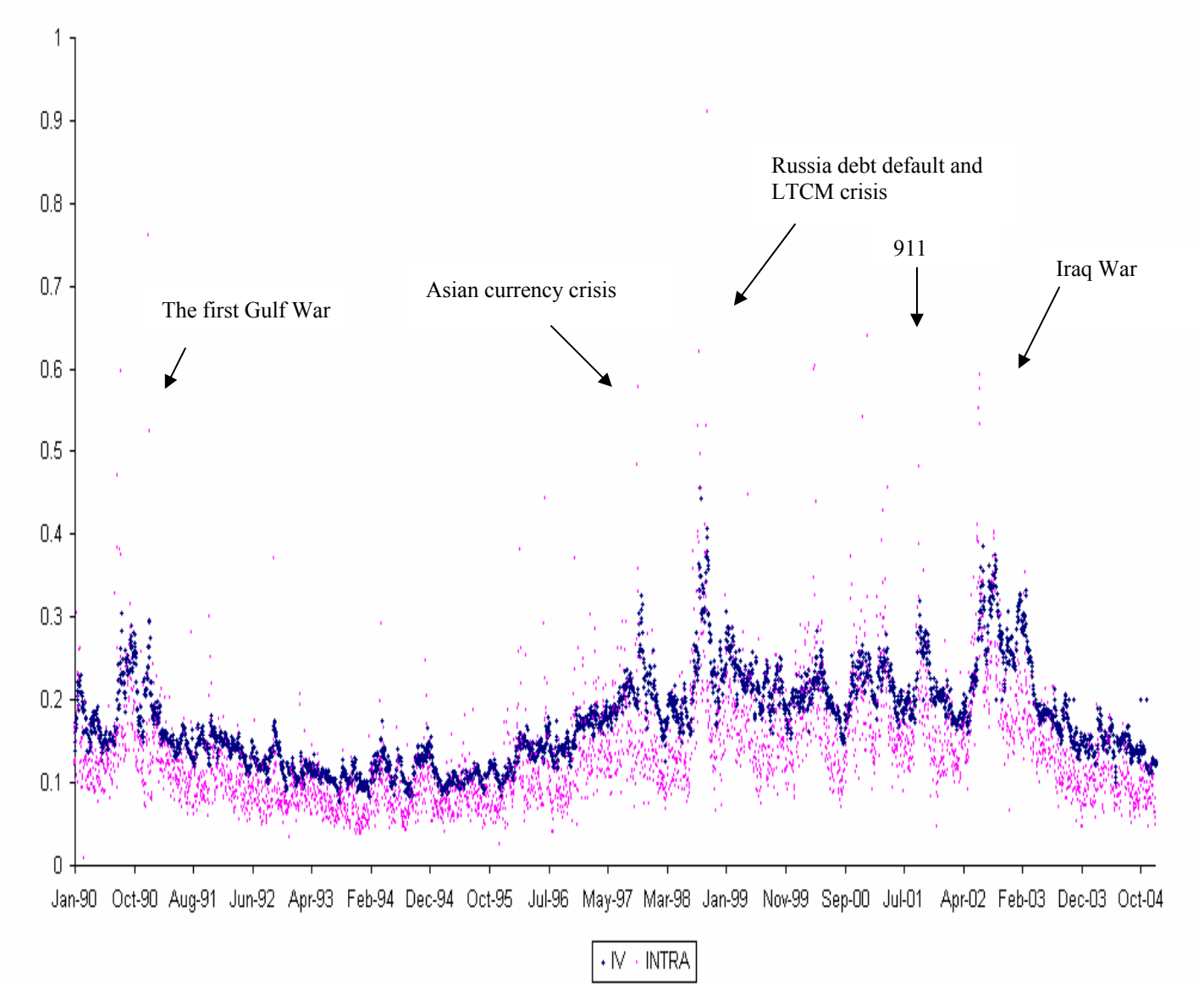

Figure 2a

One-day ahead density forecasts obtained from ARCH models on Dec 30th, 2004

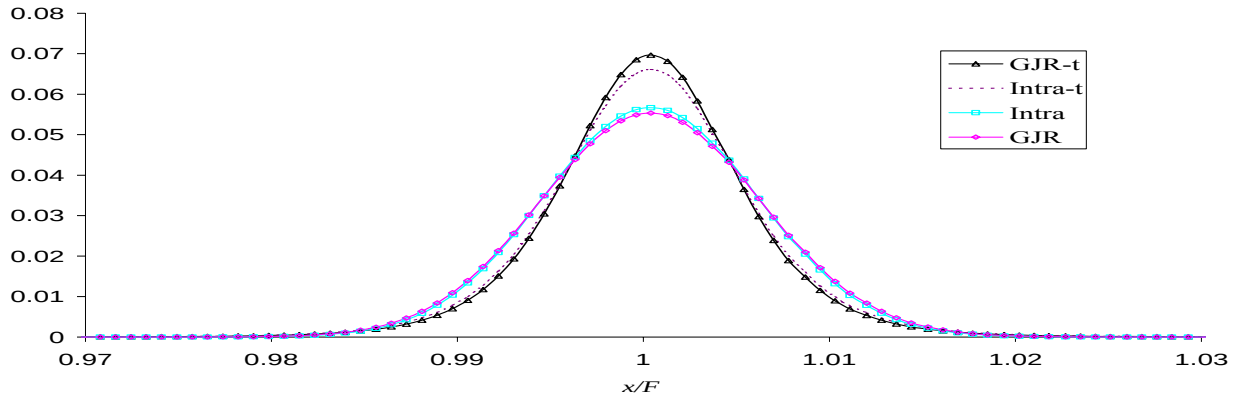

Figure 2b

One-day ahead density forecasts obtained from lognormal densities on Dec 30th, 2004

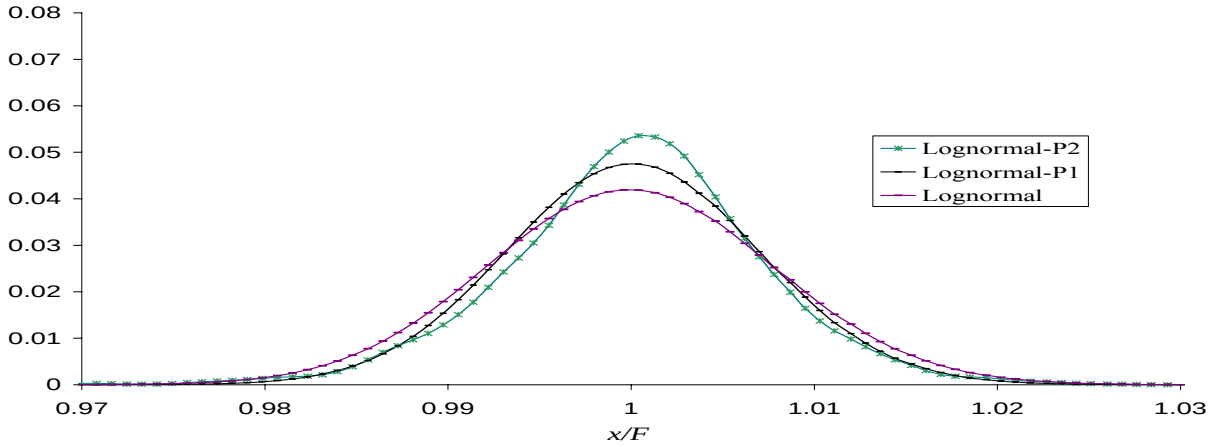

Figure 2c

One-day ahead density forecasts obtained from Heston's model on Dec 30th, 2004

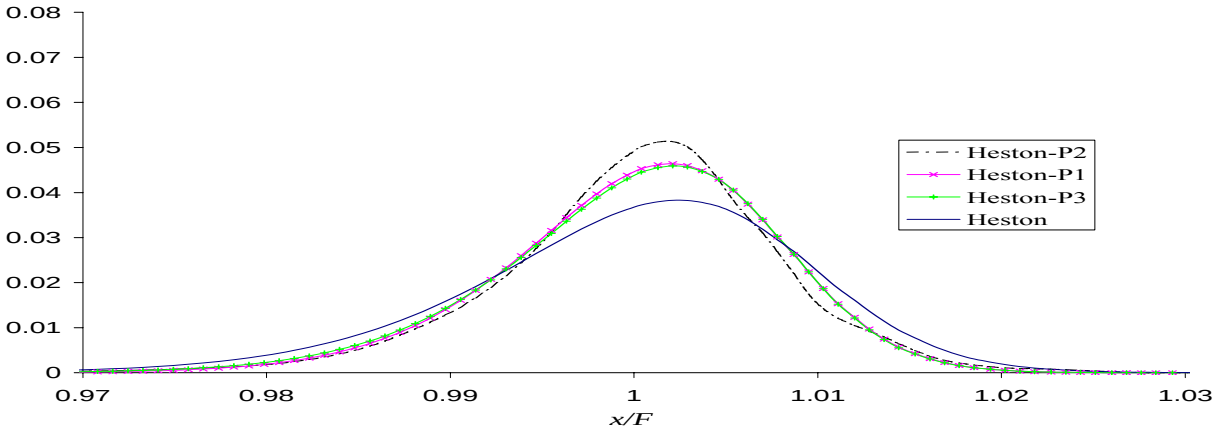

Figure 3a

Four-week ahead density forecasts obtained from ARCH models on Nov 17th, 2004

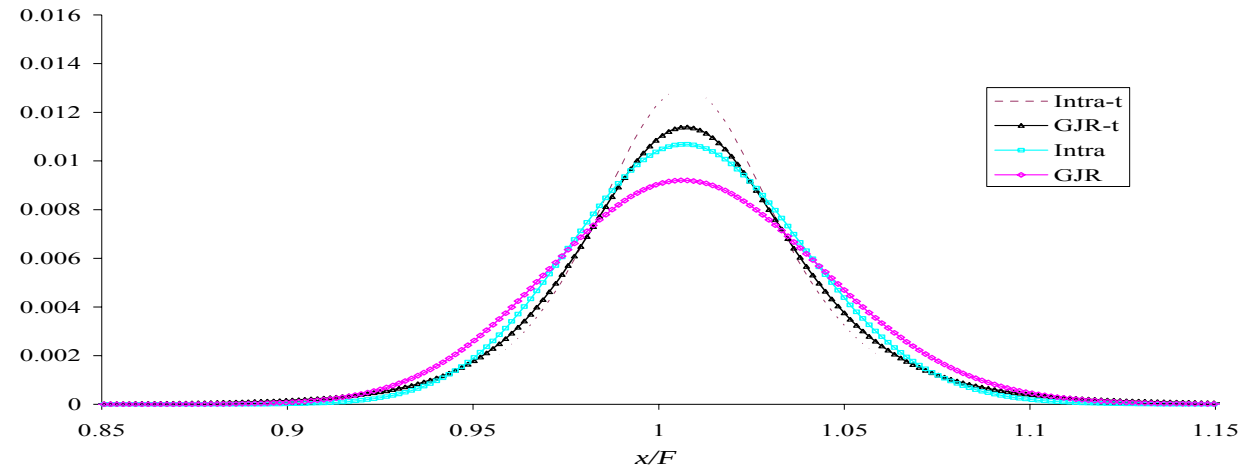

Figure 3b

Four-week ahead density forecasts obtained from lognormal densities on Nov 17th, 2004

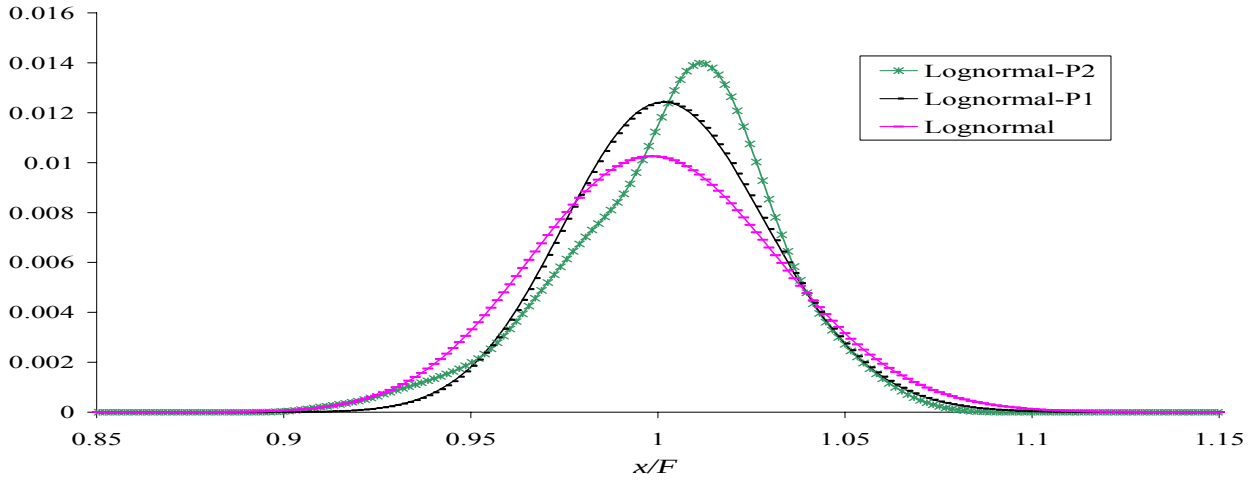

Figure 3c

Four-week ahead density forecasts obtained from Heston's model on Nov 17th, 2004

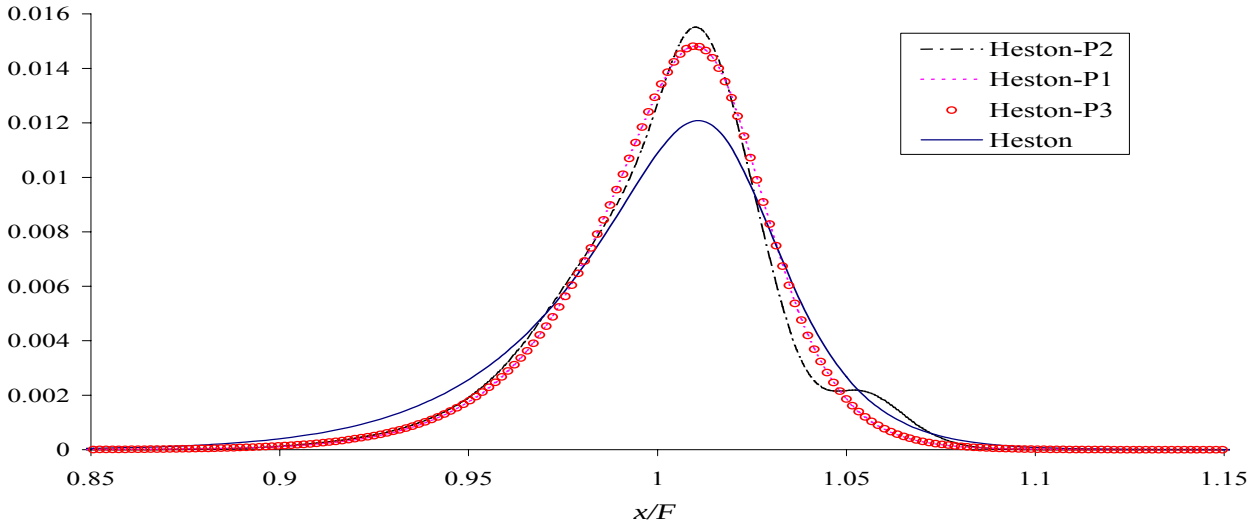

Figure 4a

The function $\widetilde{C}(u)-u$ for one-day forecasts obtained from lognormal densities and risk-transformations

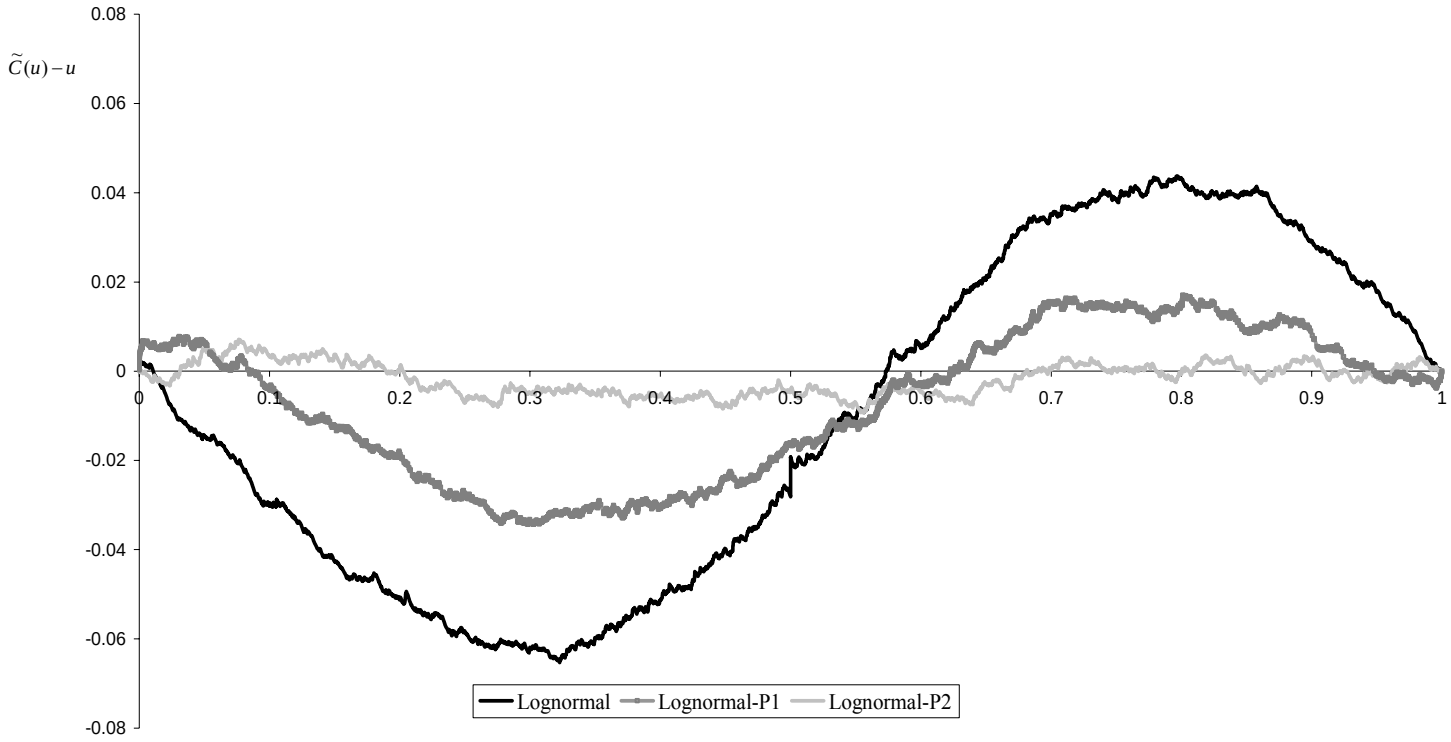

Figure 4b

The function $\widetilde{C}(u)-u$ for one-day forecasts obtained from Heston's model and risk-transformations

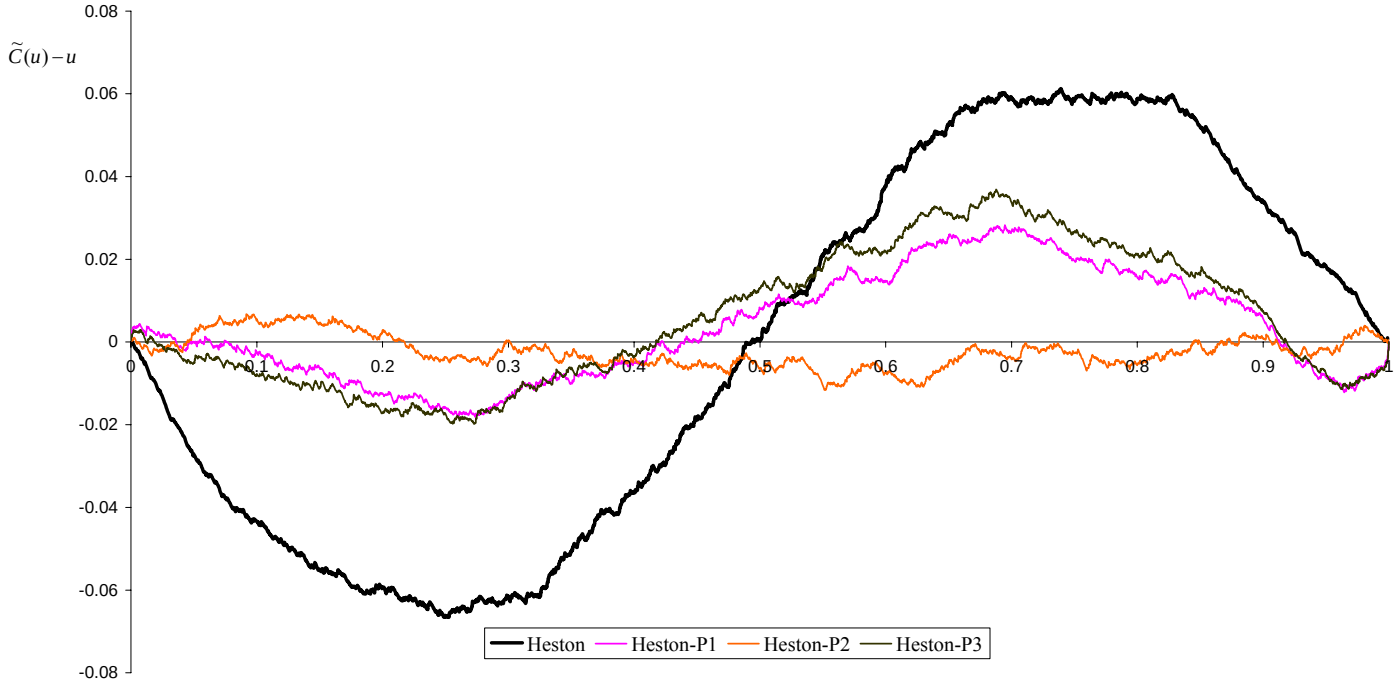

Figure 5a

The estimated calibration densities $\hat{c}(u)$ for cumulative probabilities u obtained from the one-day lognormal forecasts

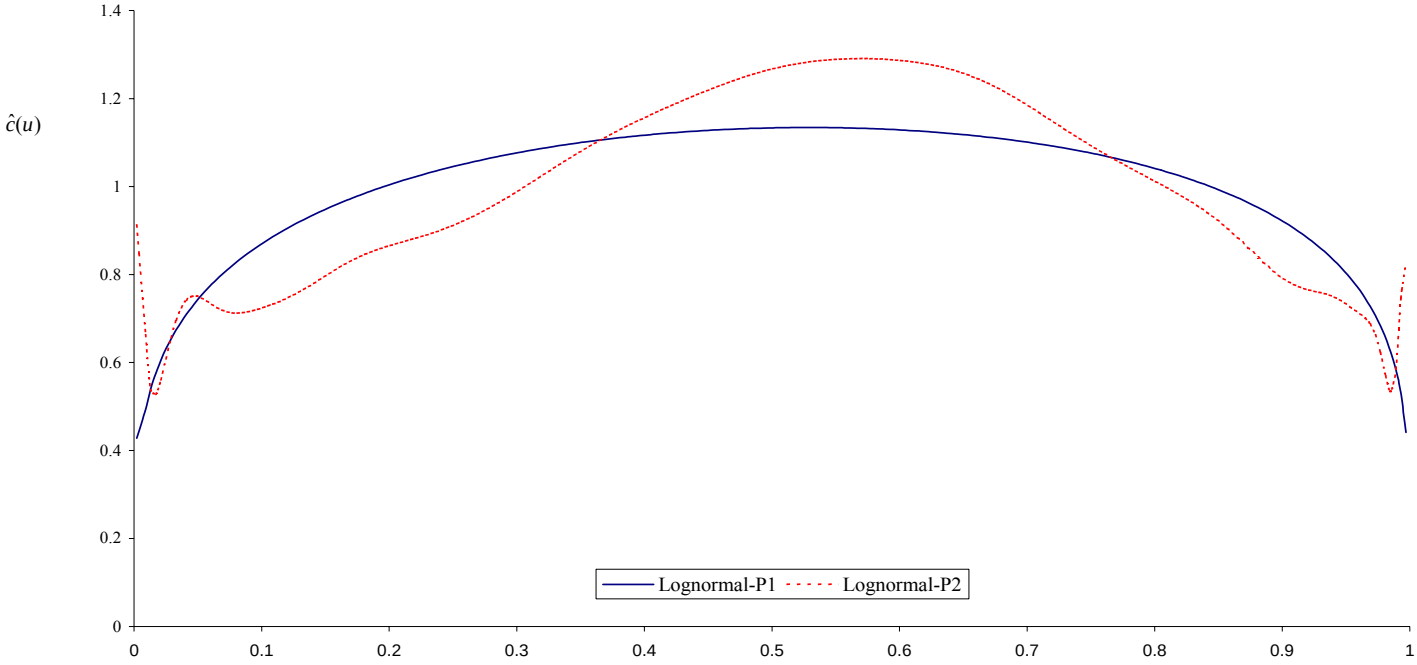

Figure 5b

The estimated calibration densities $\hat{c}(u)$ for cumulative probabilities u obtained from the one-day Heston forecasts

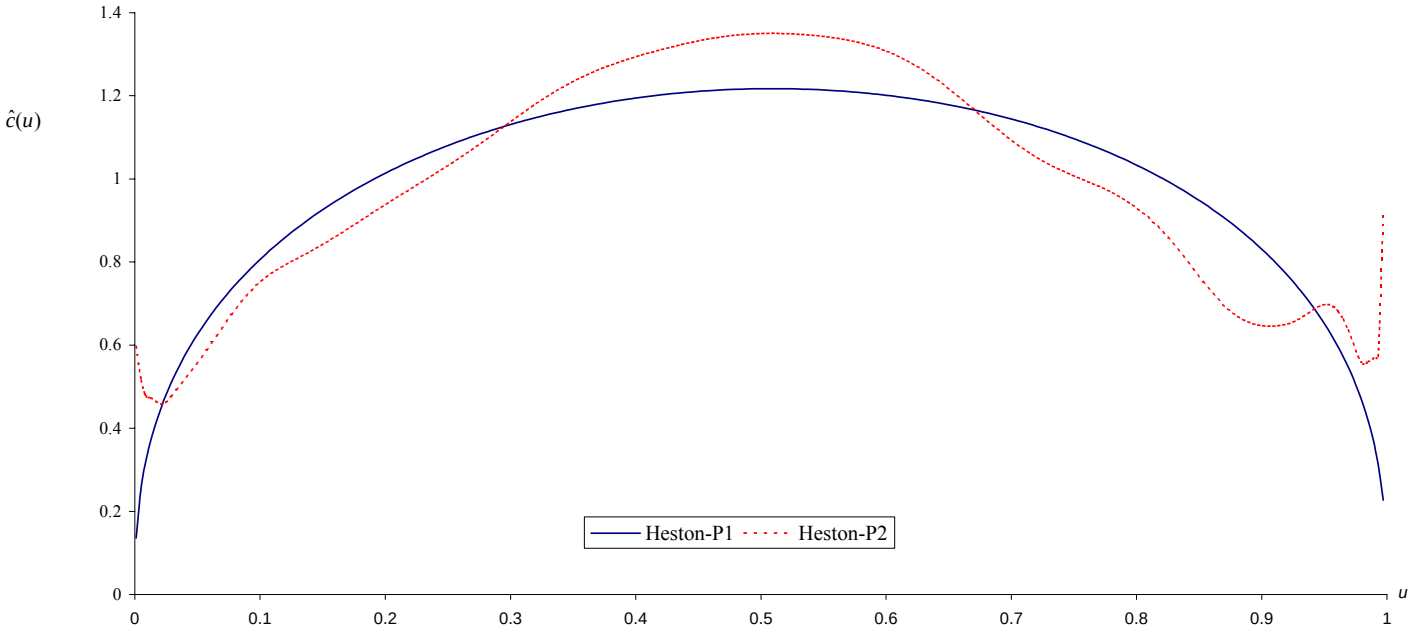

Figure 6a

The function $\widetilde{C}(u)-u$ for four-week forecasts obtained from lognormal densities and risk-transformations

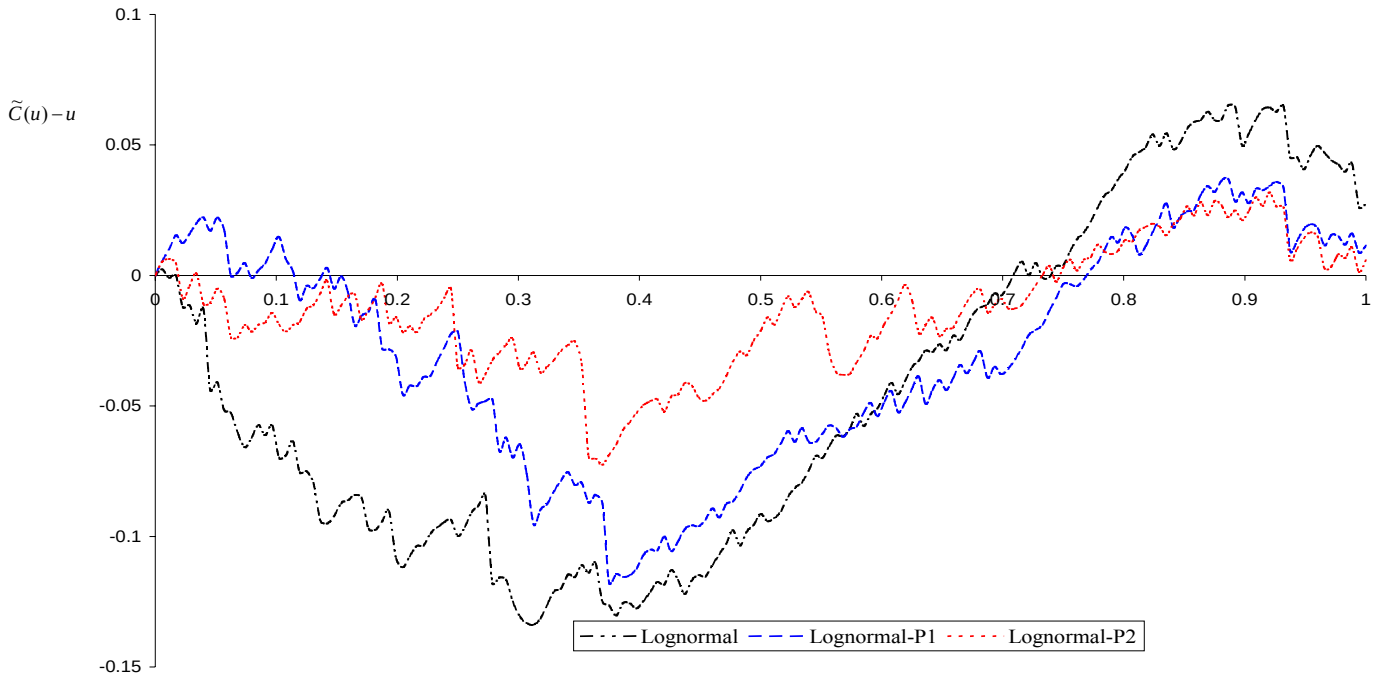

Figure 6b

The function $\widetilde{C}(u)-u$ for four-week forecasts obtained from Heston's model and risk-transformations