MPS-RR 1999-19

ISSN 1398-599X

Risk neutral densities of the 'Christmas tree' type

Niels Væver Hartvig, Jens Ledet Jensen* and Jan Pedersen
Department of Theoretical Statistics,
Departments of Mathematical Sciences and MaPhySto*,
Ny Munkegade, DK-8000 Aarhus C, Denmark

Abstract

From observed bid and ask prices of European call and put options we estimate the risk neutral density of a stock at some future time $t > 0$. We restrict attention to a class of densities with heavy tails and use a Bayesian formulation in order to study the variation in the distributions fitting the data. From the fitted risk neutral density we also consider the inverse problem of finding the volatility in a diffusion model for the price process. Finally, we apply our methods to data on the S&P 500 index.

Key words: Christmas tree densities; risk neutral density; Markov chain Monte Carlo; inverse problems; diffusion model;

JEL: C13, C15, C63, G12; *AMS*: 60J60, 62G05, 65C05, 65U05.

1 Introduction

In the celebrated standard Black-Scholes model the evolution of a stock price is modelled as a geometric Brownian motion. This in particular implies that it is possible to price derivative assets uniquely; moreover, the log stock price process is normal and has independent increments. However, starting with Fama (1965) several statistical investigations have revealed that the latter two properties are unrealistic. In particular, the tails of the increments are heavier than the normal ones. Alternative models have been proposed over the last decades, and let us just mention a few proposals, all of which result in heavy tails. For instance, Eberlein and Keller (1995) consider Levy processes where the increments follow a hyperbolic distribution; a similar model is considered in Rydberg (1997) where the increments follow a Normal Inverse Gaussian distribution, while Bibby and

*Centre for Mathematical Physics and Stochastics, funded by a grant from the Danish National Research Foundation.

Sørensen (1997) propose a diffusion process with a hyperbolic distribution as the invariant measure. Alternatively, one may introduce stochastic volatility or consider 'GARCH-type' models, see the overview by Bollerslev, Engle and Nelson (1994) or Barndorff-Nielsen (1997,1998), and the references therein.

The papers mentioned above all model the 'true' distribution of stock prices; i.e. the evolution under the so-called true probability measure. (This measure is also referred to as the 'physical' or 'objective' probability measure). However, it is well known that when it comes to pricing of derivative assets, the true probability measure is in a sense irrelevant. Specifically, standard financial theory tells us that a derivative asset should be priced as the expected payoff (properly discounted), not under the true measure, but instead under the so-called *risk neutral probability measure*.

In the literature one encounters basically two different approaches for determining the risk neutral measure. Using the first approach one has to model the evolution of the stock under the true measure. If the market is complete this will produce a unique risk neutral measure. In contrast, when incompleteness is present, there are many candidates for the risk neutral measure, and one has to pick one particular measure. Several risk neutral measures have been proposed in the latter case, including the minimal risk neutral measure (Schweizer 1991), the variance optimal measure (Schweizer 1996, Delbaen and Schachermayer 1996), a measure determined by the Esscher transform (Gerber and Shiu 1994) and a measure determined by utility-based arguments (Davis 1994). Using the second approach one may leave the distribution of the stock price (more or less) unspecified under the true measure and instead try to infer the risk-neutral measure directly from observed derivative prices. In the present paper we use this approach, and take as input observed prices (or, rather, observed bid and ask prices) of European calls and puts with the same time to expiration, t say. We then propose a new methodology for estimating the risk neutral density of the stock at time t. A typical data set is shown in the first five columns of Table 1. As a second part of our study we consider the problem of determining the volatility in a diffusion model for the stock price from observed bid and ask prices. The idea here is to use the data to determine the risk neutral density at time t and, from the latter, to determine the volatility. Before describing our approach in detail and giving an outline of our paper we provide a brief overview of related work.

1.1 Related work

From Breeden and Litzenberger (1990) it is well known how to infer the risk neutral density $p(\cdot)$ of the stock price at some future time $t > 0$, provided that one knows the prices $\tilde{\pi}(K, p; C)$, for all positive strikes K, of European calls that expire at time t. Actually, since $\tilde{\pi}(K, p; C) = \pi(K, p; C)\exp(-rt)$ with r the riskless interest rate and $\pi(K, p; C) = \int_K^\infty (s - K)p(s)ds$ we have simply that

	call		put			
Strike	Bid	Ask	Bid	Ask	Price	Volatility
0	(349.94)	(350.85)			350.31	
100					253.97	0.435
250	109.00	110.00			109.98	0.270
275	85.88	86.88			86.60	0.246
300	63.50	64.50			63.98	0.223
310	54.88	55.88			55.29	0.213
315	50.63	51.63			51.05	0.209
320	46.50	47.50			46.89	0.204
325	42.50	43.50			42.83	0.199
330	38.50	39.50	6.38	6.88	38.87	0.194
335	34.75	35.75	7.38	7.88	35.02	0.190
340	31.00	32.00	8.50	9.00	31.31	0.185
345			9.63	10.13	27.75	0.180
350	24.13	25.13	11.00	11.38	24.35	0.175
355	20.88	21.88	12.63	13.38	21.14	0.171
360	17.88	18.63	14.63	15.13	18.12	0.166
365	15.13	15.88	16.25	17.00	15.32	0.161
370	12.50	13.25	18.63	19.38	12.76	0.157
375	10.13	10.88	21.00	22.00	10.46	0.153
380			23.63	24.63	8.42	0.149
385			26.50	27.50	6.64	0.145
390					5.15	0.141
395					3.92	0.138
400			37.00	38.00	2.95	0.135
410					1.65	0.133
430					0.53	0.133

Table 1: Data from 10 pm on June 25 1990 for options with half a year to expiration. The data are in the first five columns with the values at strike zero obtained as explained in Section 5. The last two columns give the call price $\tilde{\pi}(K, p; C)$ and the implied volatility based on a estimate of the risk neutral probability density from the posterior distribution.

$p(s) = \frac{d^2\pi(s,p;C)}{ds^2}$. Determining p is therefore a linear inverse problem. The main obstacles in solving this inverse problem are that only a finite number of calls are traded and that we do not necessarily know the price itself but rather the bid and ask prices. A number of different approaches have been suggested for solving the problem. Shimko (1993) proposes to interpolate with a smooth function between the known prices at different strikes before calculating the second derivative and then augments p with lognormal tails outside the interval of strike values. Longstaff (1990) assumes that a finite number of calls are available and that $p(\cdot)$ is constant between the strike values used. Instead of finding the second derivative of $\pi(K,p;C)$ a set of linear equations are solved. This method is very unstable, see the review in Rubinstein (1994).

An approach somewhat more related to the one we shall use is described in the paper Rubinstein (1994) and implemented in Jackwerth and Rubinstein (1995). The data available are bid and ask prices for finitely many European call options with the same time to expiration $t > 0$. Jackwerth and Rubinstein make a fine discretization of the risk neutral density p and then find an estimate by minimizing the distance (in some sense) between p and a prior (typically log normal) density, subject to the demand that call prices calculated under the risk neutral distribution must fall within the bid-ask spreads. They also consider a method where minimizing the distance to a prior distribution is replaced by the requirement of maximizing the smoothness, that is, the integral of the squared value of the second derivative of $p(\cdot)$ should be small. One can formulate the approach of Jackwerth and Rubinstein in the way that a regularized solution fitting the data is looked for, where the regularization is either the closeness to a prior distribution or is the smoothness of the solution itself. The latter method is the standard method in linear inverse problems and is often called Tikhonov regularization.

In this paper we propose a different kind of regularization than the one used by Jackwerth and Rubinstein. We use a Bayesian formulation where we choose a prior on a suitable set of densities and then consider the posterior distribution given the data. The main point of our method is the choice of a class of densities. We define the class of interest in terms of the density ϕ for the log price rather than the density p for the price itself. Thus $\phi(x) = \log(p(\exp(x))\exp(x))$ and we require that $\log(\phi)$ has heavy tails of the form discussed above for the increments. We will use the name 'Christmas tree' densities for the considered class and the log-concave densities will be a subclass. When modelling the log-density instead of the density itself we no longer have a linear inverse problem and the price to pay for this is that it is not easy to find an estimate in the class considered. We solve this by simulating from the posterior distribution using a Markov chain Monte Carlo technique. In this way we are also able to quantify the uncertainty in our estimate.

Above we have been concerned with the risk neutral density at time t. More generally it is also of interest to price e.g. American options, and to this end we need the risk neutral distribution of the entire stock price process. Building on the work described above Rubinstein (1994) proposes to first estimate a discretized version of the density p at time t and afterwards specify a unique so-called 'Implied Binomial Tree' which reproduces the risk neutral density of the stock at time t. A very different approach is based on the diffusion model

$$dX_u = \mu(u)du + \sigma(X_u)dW_u \tag{1}$$

for the log price process X_u. The parameter $\sigma(\cdot)$ is the volatility as a function of log price and $\mu(u)$ is a drift. This typically results in a complete market, and the unique risk neutral measure is fully determined once σ is specified. More precisely, under the risk-neutral measure the log stock price evolves according to

$$dX_u = [(r-d) - \frac{1}{2}\sigma(X_u)^2]du + \sigma(X_u)dW_u^Q, \tag{2}$$

where d is the fixed dividend rate and W_u^Q is a standard Wiener process under the risk neutral measure. Let $V(x, u; K)$ denote the price at time u, of a strike K call expiring at time t, when the log stock price is x. If we let x_0 be the observed log stock price it then follows that $\tilde{\pi}(K, p; C) = V(x_0, 0; K)$. It is seen from (2) that for fixed K the function V solves the following partial differential equation in (x, u)

$$-\frac{\partial V}{\partial u} + rV = \frac{1}{2}\sigma^2(x)\frac{\partial^2 V}{\partial^2 x} + [(r-d) - \frac{1}{2}\sigma(x)^2]\frac{\partial V}{\partial x}. \tag{3}$$

The problem of determining $\sigma(\cdot)$ from knowledge of $\tilde{\pi}(\cdot, p; C)$ using (3) is known in mathematics as an inverse problem for parabolic partial differential equations. When $\tilde{\pi}(K, p; C)$ is known for K in an interval ω, and we assume that σ is known outside ω as well as on a subinterval of ω, Bouchouev and Isakov (1997) show that (3) determines σ uniquely. Also they show a stability result for the mapping from $\tilde{\pi}(\cdot, p; C)$ to σ and propose a numerical procedure for finding σ in the case where t is small. (In fact, they first transform (2) into a PDE in (s, K), but this is not important here). Dupire (1994) uses (3) to determine σ when European call options of all strike prices and all expiration times are available. Similar ideas are described in Derman and Kani (1994). Actually, the two last mentioned papers let the volatility be time dependent as well as price dependent. In the more challenging case where one has data only for a few expiration times and strikes Lagnado and Osher (1997) regularize the inverse problem for finding σ by including the L^2-norm of the Laplacian of σ in a functional to be minimized. They then use a gradient descent procedure to find σ. Judging from their figures this method seems to work well in the region of observed strike values and less so outside this region. Bouchouev and Isakov (1998) consider the same problem and further references can be found in their paper.

In this paper we use an alternative approach for inferring the volatility. We will find $\sigma(\cdot)$ such that the transition density of the diffusion (2) fits the estimate of the density ϕ described above for the log price. To this end we use the well known fact that the transition density, $\phi(x, u; \sigma)$, $u \geq 0$, of the diffusion (2) can be found by solving the (Kolmogorov forward) partial differential equation

$$\frac{\partial \phi}{\partial u} = \frac{1}{2} \frac{\partial^2}{\partial^2 x}[\sigma^2(x)\phi] - \frac{\partial}{\partial x}[(r - d) - \frac{1}{2}\sigma(x)^2)\phi]. \tag{4}$$

The initial condition at time zero is that all mass is concentrated in x_0. Within the diffusion setting we have that the transition density $\phi(\cdot, t; \sigma)$ in fact equals $\phi(\cdot)$. Thus, if we have an estimate of the latter then this can be used as input in (4) to determine an estimate of the volatility. We regularize this inverse problem by letting σ be a spline, that is, a function which has a continuous piecewise linear second derivative and where we choose the points (the knots of the spline), at which there is a change in the third derivative of σ, from the shape of ϕ. For a given spline σ we solve the forward problem in (4) numerically to get the corresponding density $\phi(\cdot, t; \sigma)$. The spline for which the distance between $\phi(\cdot, t; \sigma)$ and the estimate of $\phi(\cdot)$ is minimized is our estimate of σ. For the data that we consider it turns out that we can get an almost perfect fit to the estimate of ϕ.

The paper is organized as follows. In Section 2 we define the class of densities we will use for the risk neutral density and set up a Bayesian framework for studying the density. In Section 3 we describe a Markov Chain Monte Carlo method for producing samples from the posterior distribution for the risk neutral density. In Section 4 we turn to the method used for inferring the volatility from the risk neutral density. Section 5 is the application section where we demonstrate the use of the methods suggested, and we conclude with a discussion section.

2 The model

We consider a financial market in which a stock is continuously traded on a finite horizon $[0, t]$. Let S_u denote the stock price at time u, and for simplicity let $S = S_t$. In the market we assume that the stock pays out dividends at a fixed rate $d \geq 0$; i.e. the holder of the stock at time u receives the dividend $dS_u du$ on the interval $[u, u + du]$. Moreover, in the economy there is a bank account with fixed interest rate $r \geq 0$.

Now, as mentioned in the introduction, financial theory shows that the price, $\tilde{\pi}(f)$, of some European asset, f, expiring at time t is $\tilde{\pi}(f) = E^Q[f]\exp(-rt)$, where Q denotes the risk neutral probability measure. We assume that S under Q has density p with respect to the Lebesgue measure. Recall that Q, which is often also referred to as the Equivalent Martingale Measure, must be defined such that S_u, properly discounted, is a Q-martingale. This, in the presence of dividends, implies that

$$S_0 = \exp(-(r - d)t)E^Q[S] = \exp(-(r - d)t)\int_0^\infty sp(s)ds. \tag{5}$$

Define

$$\pi(K,p;C) = \int_K^\infty (s-K)p(s)ds$$

and

$$\pi(K,p;P) = \int_0^K (K-s)p(s)ds = K - \pi(0,p;C) + \pi(K,p;C). \tag{6}$$

The prices of a call option and a put option with strike K are

$$\tilde{\pi}(K,p;C) = \pi(K,p;C)\exp(-rt), \quad \tilde{\pi}(K,p;P) = \pi(K,p;P)\exp(-rt),$$

respectively, for $K > 0$. For the theory leading to these price formulas we refer to e.g. Duffie (1996) or Bensoussan (1984).

The data we have available are the current value S_0 of the stock and the bid and ask prices for a number of call and put options at different strike values. Our aim is to 'invert' the relations above to obtain the risk neutral density p. This problem is ill-posed since we only have finitely many data points and we have only the bid and ask prices not the price itself. This means that there will be many choices of p fitting the data. We therefore need to regularize the problem, and do this by requiring the density p to belong to a suitable chosen class which we now describe.

Let $X = \log(S)$ and let $\phi(x) = p(\exp(x))\exp(x)$ be the corresponding density. We choose a division $z_1 < z_2 < \cdots < z_l$ and set $z_0 = -\infty$ and $z_{l+1} = \infty$, and model $\log(\phi(x))$ as a continuous function which is linear on any interval (z_i, z_{i+1}), $i = 0, \ldots, l$. We can then write

$$\log(\phi(x)) = \begin{cases} a_1 + \beta_1(x - z_1) & x < z_1 \\ a_i + \beta_{i+1}(x - z_i) & z_i \le x < z_{i+1}, i \ge 1, \end{cases} \tag{7}$$

where $a_i = a_{i-1} + \beta_i(z_i - z_{i-1})$ for $i = 2, \ldots, l$, and the value of a_1 is determined from the requirement that ϕ is a density, that is

$$1 = \exp(a_1)\left\{\frac{1}{\beta_1} + \sum_{j=2}^{l+1}\exp\left(\sum_{i=2}^{j-1}\beta_i(z_i - z_{i-1})\right)\frac{\exp(\beta_j(z_j - z_{j-1})) - 1}{\beta_j}\right\}. \tag{8}$$

We must here require that

$$\beta_1 > 0 \text{ and } \beta_{l+1} < -1,$$

where the last requirement is needed for the existence of $E^Q[S]$. The number of points l in the division must be larger than the number of strike values in the data and is chosen sufficiently large to make the piecewise linear structure in $\log(\phi)$ 'look' smooth.

The 'Christmas tree' class of densities is defined as follows. Let

$$w_i = \beta_i - \beta_{i+1}, \ i = 1, \ldots, l, \quad w_0 = -\beta_1.$$

If there exists $1 \leq k_1 < k_2 \leq l$ such that

$$\begin{array}{ll} w_i < 0 & 1 \leq i < k_1 \\ w_i >= 0 & k_1 \leq i \leq k_2 \\ w_i < 0 & k_2 < i \leq l, \end{array} \tag{9}$$

then we say that ϕ is a Christmas tree density. The densities given by (9) will be our most general class of densities. The subclass with $k_1 = 1$ and $k_2 = l$ is the class of log-concave densities with a continuous piecewise linear log density. Note that there is a one-to-one correspondence between p, ϕ, $\beta = (\beta_1, \ldots, \beta_{l+1})$ and $w = (w_0, \ldots, w_l)$ within the class of Christmas tree densities; in particular, the signs of the w_i's determine k_1 and k_2. We will therefore use the different parametrizations freely in the formulas below.

To get from a Christmas tree density to the call prices we define

$$p_j = \begin{cases} \exp(a_1)/\beta_1 & j = 1 \\ \exp(a_{j-1})\{\exp(\beta_j(z_j - z_{j-1})) - 1\}/\beta_j & 2 \leq j \leq l+1, \end{cases}$$

and

$$q_j = \begin{cases} \exp(a_1 + z_1)/(\beta_1 + 1) & j = 1 \\ \exp(a_{j-1} + z_{j-1})\{\exp((\beta_j + 1)(z_j - z_{j-1})) - 1\}/(\beta_j + 1) & 2 \leq j \leq l+1. \end{cases}$$

Then the call price at strike $K = \exp(z_m)$ is

$$\pi(\exp(z_m), p; C) = \sum_{j=m}^{l} q_j - \exp(z_m) \sum_{j=m}^{l} p_j. \tag{10}$$

Having defined the Christmas tree class of functions the next problem is to find good candidates in this class to describe our data. We could here proceed in a similar way as in Jackwerth and Rubinstein (1995) and find the one with certain optimality properties. However, we find it more natural to adopt a formulation where we can study the variation in those functions ϕ that provide a reasonable fit to the data. We therefore adopt a Bayesian formulation in which ϕ becomes a random quantity through a prior distribution on the Christmas tree class of densities. This allows us to study the distribution of ϕ given the information provided by the data, that is, the posterior distribution of ϕ. As an example we have in Figure 1 plotted two realizations of $\log(\phi)$ from the posterior distribution for one of the data sets that we consider. It is not obvious how to sample directly from the posterior distribution. Instead we describe in the next section a Markov chain Monte Carlo algorithm that will produce samples from this distribution.

What is a good prior distribution on the Christmas tree class of densities? We want a weak prior that simply favours 'smooth' densities. To this end we design a very simple prior through a distribution on $(w_0, \ldots, w_l)$. If we consider the class of log-concave densities $(k_1, k_2) = (1, l)$ will be fixed. Similarly, we can also

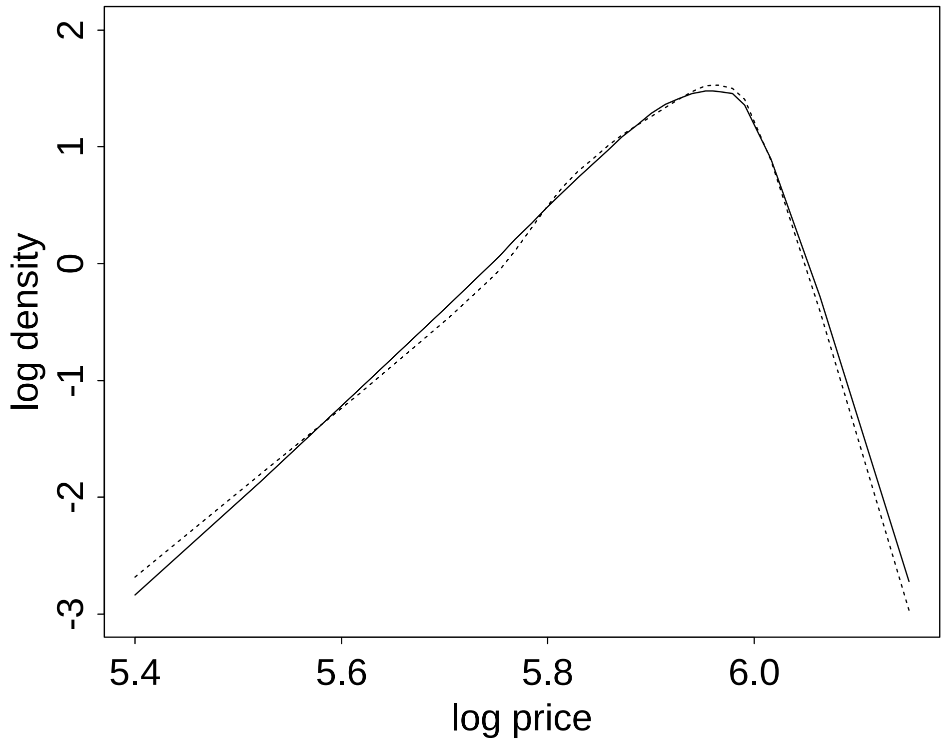

Figure 1: Two realizations from the posterior distribution of the log density $\phi(\cdot)$ plotted as function of log price.

consider the subclasses where either $k_1 = 1$ or $k_2 = l$ are fixed. In the general case we take (k_1, k_2) to be uniformly distributed on $1 \leq k_1 \leq k_2 \leq l$. Conditionally on the values of (k_1, k_2) we take $|w_0|$ as an exponential distribution with mean γ and $|w_i|$, $i = 1, \ldots, l$, as an exponential distribution with mean λ. All the $|w_i|$'s are independent. The restriction of this measure to the space of possible values of w, see (15) below, will be our prior measure with the corresponding density denoted by $\psi(w)$. Note that the w_i's, loosely speaking, are the second derivative of $\log(\phi)$, and that the exponential distribution favours small values.

In a Bayesian formulation we finally need to describe the distribution of the data given the value of p or, equivalently, given the value of ϕ. Basically what we want here is to say that the option prices calculated under ϕ should fall within the bid-ask spreads with no preference as to where in the spreads they fall. Let B_i, A_i be the observed bid and ask prices at strike $K_i > 0$, $i = 1, \ldots, n$, where $i = 1, \ldots, n^C$ are calls and $i = n^C + 1, \ldots, n = n^C + n^P$ are puts. Note that the strike zero call is not a traded asset. However, we describe in Section 5 an approach to establish bounds (B_0, A_0) for $\tilde{\pi}(0, p; C)$. We treat these bounds as the bid-ask spread for the strike zero call, and by letting $K_0 = 0$ this leaves us with data for $n+1$ options.

Now, instead of modelling $B_i, A_i,\ i = 0, \ldots, n$, directly we consider the following transformation of the data. Define $\bar{B}_i = \exp(rt)B_i$, $\bar{A}_i = \exp(rt)A_i$ and

$$Y_i = (\bar{A}_i + \bar{B}_i)/2, \quad \delta_i = (\bar{A}_i - \bar{B}_i).$$

Clearly, except for the $\exp(rt)$ term, the variables Y_i and δ_i denote the midpoint and the size of the i'th bid-ask spread, respectively. The distribution of these variables are modelled as follows. First, we choose for simplicity to let all the δ_i's be constants. This leaves us with the distribution of the midpoints Y_i given ϕ and δ_i. Letting Y_i be uniformly distributed on the interval $(\pi(K_i, p; \nu_i) - \delta_i/2, \pi(K_i, p; \nu_i) + \delta_i/2)$, where ν_i is C for $i \leq n^C$ and P otherwise, the scenario described above appears; that is, the i'th option price falls within the i'th spread with no preference as to where it falls. More generally, though, for the use in the MCMC algorithm, we do as follows. Let $\alpha > 0, \xi > 0$ be two parameters and consider the density $f(y_i; K_i, \nu_i|\delta_i, \phi)$ for Y_i defined by

$$f(y_i; K_i, \nu_i|\delta_i, \phi) = \exp\{-\alpha \cdot 1(g_i > 0) - \xi g_i\}/Z_i(\alpha, \xi), \tag{11}$$

where

$$\begin{aligned} g_i &= \max\{0, \pi(K_i, p; \nu_i) - (y_i + \delta_i/2), y_i - \delta_i/2 - \pi(K_i, p; \nu_i)\} \\ &= \max\{0, \pi(K_i, p; \nu_i) - \bar{A}, \bar{B} - \pi(K_i, p; \nu_i)\} \end{aligned} \tag{12}$$

and $Z_i(\alpha, \beta)$ is a normalizing constant for the density, $Z_i(\alpha, \xi) = 2\exp(-\alpha)/\xi + \delta_i$. This density corresponds to the i'th option price being 'almost' uniformly distributed within the i'th spread. However, with a small probability the price falls outside the spread. By choosing $\alpha = \xi = \infty$ the uniform distribution is in fact recovered, but, in order to enhance mixing for the MCMC algorithm we describe in the next section, it is fruitful to consider the more general specification in (11). Note that if we sample from (11) but only use the subsample where all option prices fall within the spreads ($g_i = 0$ for all i), then this corresponds to sampling from the uniform distribution $\alpha = \xi = \infty$.

Assuming that the Y_i's are independent the likelihood function of the data given the density ϕ is then from the description above

$$L(\phi) = \prod_{i=0}^{n} f(y_i; K_i, \nu_i|\delta_i, \phi).$$

Finally, the posterior density for ϕ is $\omega(\phi) = L(\phi)\psi(w)/Z(y)$ for some normalizing constant $Z(y)$. In the next section we describe an MCMC algorithm to simulate observations from the density ω.

3 An MCMC algorithm

For a general description of the MCMC methodology used in this section we refer to Gilks, Richardson and Spiegelhalter (1996). The idea is to construct a Markov chain for which the invariant distribution is the posterior density ω. This is done by proposing a change of the current state and then accept this change by an appropriate probability.

We parametrize our model through w, where k_1 and k_2 are then defined from the signs of the w_i's, and propose a change by adding a small random amount to w. Define

$$\xi_i = \begin{cases} w_i - \epsilon & i \leq k_1, i \geq k_2 \\ \max\{w_i - \epsilon, 0\} & k_1 < i < k_2 \end{cases}$$

and

$$\eta_i = \begin{cases} \min\{w_i + \epsilon, 0\} & i < k_1 - 1, i > k_2 + 1 \\ w_i + \epsilon & k_1 - 1 \leq i \leq k_2 + 1. \end{cases}$$

Let $u_0, \ldots, u_l$ be independent and uniformly distributed on $(0, 1)$. We then propose a move to

$$\tilde{w}_i = \xi_i + u_i(\eta_i - \xi_i), \quad i = 0, \ldots, l. \tag{13}$$

If, as an example, $w_{k_1} < \epsilon$ the transition from w to $\tilde{w}$ gives the possibility of changing the value of k_1 to $k_1 + 1$. Similarly, k_1 can be changed to $k_1 - 1$ and k_2 can be changed to either $k_2 + 1$ or $k_2 - 1$. If $\tilde{w}_{k_1 - 1} > 0$ and $\tilde{w}_{k_1} < 0$ the transition to $\tilde{w}$ will be rejected because the new value does not satisfy the Christmas tree condition.

The Metropolis-Hastings ratio is for the case (13)

$$\frac{L(\tilde{\phi})\psi(\tilde{w})/J(\tilde{w})}{L(\phi)\psi(w)/J(w)}, \tag{14}$$

with $\tilde{\phi}$ the value of ϕ corresponding to w replaced by $\tilde{w}$, and where

$$J(w) = \prod_{i=0}^{l} (\eta_i - \xi_i).$$

We denote the transition mechanism defined through (13) and (14) by P_1.

To discuss the ergodicity properties of the above defined MCMC procedure we first make a restriction on the possible values of β corresponding to changing the support of the prior ψ. The restrictions we impose on β are as follows,

$$c_1 < \beta_1, \quad \beta_{k_1} < c_2, \quad \beta_{k_2+1} > -c_2, \quad \beta_{l+1} < -1 - c_1, \tag{15}$$

where $0 < c_1 < c_2 < \infty$ are constants. Under the restrictions (15) the space of possible values of β or w is bounded. With $L(\phi)$ defined through (11) and with the prior defined in terms of exponential distributions for $|w_i|$ we find that $L(\phi)$

and $\psi(w)$ are greater than $\delta_1 > 0$, say, on the space of possible values of β. We can therefore find $\delta_2 > 0$ so that the Metropolis-Hastings ratio (14) is always greater than δ_2. Therefore, we can also find $\delta_3 > 0$ so that the density of P_1 with respect to Lebesgue measure is greater than δ_3 in a ϵ-neighbourhood of the present position. We can next find k and $\delta_4 > 0$ so that the density of the k-fold iterate P_1^k is greater that δ_4 in all of the state space for w. This shows that P_1 is uniformly ergodic.

To enhance mixing we have also employed a simulated tempering algorithm (Geyer and Thompson, 1995) based on the above Metropolis-Hastings algorithm. The motivation for doing this is as follows. Our main interest is the case of (11) with $\alpha = \xi = \infty$. The set of allowable log densities then become a very 'thin' region of $\mathbf{R}^{l+1}$ and it is difficult to move around in this region. Instead we could try with finite values of α and ξ and then subsample those cases with $m_1 = 0$, where the variable m_1 counts the number of option prices outside the spreads,

$$m_1 = \sum_{i=0}^{n} 1(g_i > 0).$$

This will allow us to get outside the allowable set, corresponding to $m_1 = 0$, and thereby we will be able to move more freely. However, experience shows that when first $m_1 > 0$ it is very difficult to get back to $m_1 = 0$ unless α and ξ are large. We therefore suggest a tempering algorithm with two temperatures, $T = 1$ and $T = 2$. The temperature $T = 1$ will be given through (11) with α and ξ large, whereas the temperature $T = 2$ shall give us the possibility of having $m_1 > 0$ and at the same time to be able to get back to $m_1 = 0$. For each temperature we use (13) and (14) with $L(\phi)$ replaced by $\exp(l(T,\phi))$ to propose a change of ϕ. To define $l(T,\phi)$ we first introduce

$$G_1 = \sum_{i=0}^{n} g_i \quad \text{and} \quad G_2 = \sum_{i=0}^{n} h_i,$$

with h_i defined as

$$h_i = \max\{0, \pi(K_i,p;\nu_i) - (\bar{A}_i - \delta_i/8), (\bar{B}_i + \delta_i/8)) - \pi(K_i,p;\nu_i)\}.$$

We then take

$$l(2,\phi) = \kappa - \alpha_2(G_1 + G_2) - \xi_2 \min\{m_1, 2\}$$

and

$$l(1,\phi) = -\alpha_1 G_1 - \xi_1 m_1,$$

where $l(1,\phi)$ corresponds to using (11).

In the runs in Section 5 we have taken $\alpha_1 = \xi_1 = 20$, $\alpha_2 = 10$, $\xi_2 = 1$ and $\kappa = 2$. Furthermore, we have taken $\epsilon = 0.01$ in the updating (13) except when $T = 2$ and $m_1 = 0$ where we have used $\epsilon = 0.1$. Here α_2 and ξ_2 have been chosen to ensure that under the temperature $T = 2$ we spend time both in $m_1 = 0$ and in $m_1 > 0$. The higher value of ϵ when $m_1 = 0$ is used to make a fast transition from $m_1 = 0$ to $m_1 > 0$. The parameter κ is introduced to control the jumps between the two temperatures. A proposed jump from $T = 1$ to $T = 2$ is accepted with probability

$$a_{12} = \min\{1, \exp(\eta)\}$$

and a jump from $T = 2$ to $T = 1$ is accepted with probability

$$a_{21} = \min\{1, \exp(-\eta)\},$$

where

$$\eta = \kappa + \xi_1 m_1 - \xi_2 \min\{m_1, 2\} + (\alpha_1 - \alpha_2)G_1 - \alpha_2 G_2.$$

With the parameter values mentioned above we find that η is almost one if $m_1 > 0$ and $T = 1$. Thus we will immediately jump to $T = 2$ and stay there until $m_1 = 0$. When $m_1 = 0$ we have $\eta = \kappa - \alpha_2 G_2$ and κ controls how often we jump between the two temperatures. We now give a couple of numbers to illustrate the behaviour of the algorithm on the data in Table 1. In 93 percent of the cases we jump from $T = 1$ to $T = 2$ without having a change of ϕ in between and in 5 percent there will be one change of ϕ. In 97 percent of the jumps from $T = 1$ to $T = 2$ the waiting time is less than 5 steps in the algorithm, the mean is 1.6 and the maximum seen in a run of length 300,000 was 85. For jumps from $T = 2$ to $T = 1$ we find that 98.5 percent happens without a change of ϕ and in 1.1 percent m_1 becomes positive while staying at $T = 2$. The maximum value of m_1 seen in a run of length 300,000 was 6. For those cases with $m_1 > 0$ the waiting time to return to $T = 1$ has mean 182 with 5 percent being larger than the mean and the maximum seen was 56,000. With the parameter values described above the algorithm spends 63 percent of the time in $T = 2$. In Figure 2 we have illustrated the behaviour of the algorithm by running 2.8 million steps and collecting the value of the posterior density for ϕ when $T = 1$ and $m_1 = 0$. In the plot we have taken every 100'th value of the density.

For the actual applications of the MCMC algorithm described in Section 5 we have started by using (14) based on the likelihood from (11) with $\alpha = \xi = 20$. These high values of α and ξ will force an initial estimate of the density to move to the region where all prices, calculated from the estimated density, will fall in the bid-ask spreads. From here on we then use the simulated tempering algorithm described above.

4 Determination of the volatility

Above we have described how to obtain an estimate (rather: a random value from the posterior distribution) of the risk neutral density ϕ from the 'Christmas tree' class of densities. Abusing notation we denote this estimate by ϕ as well. In this

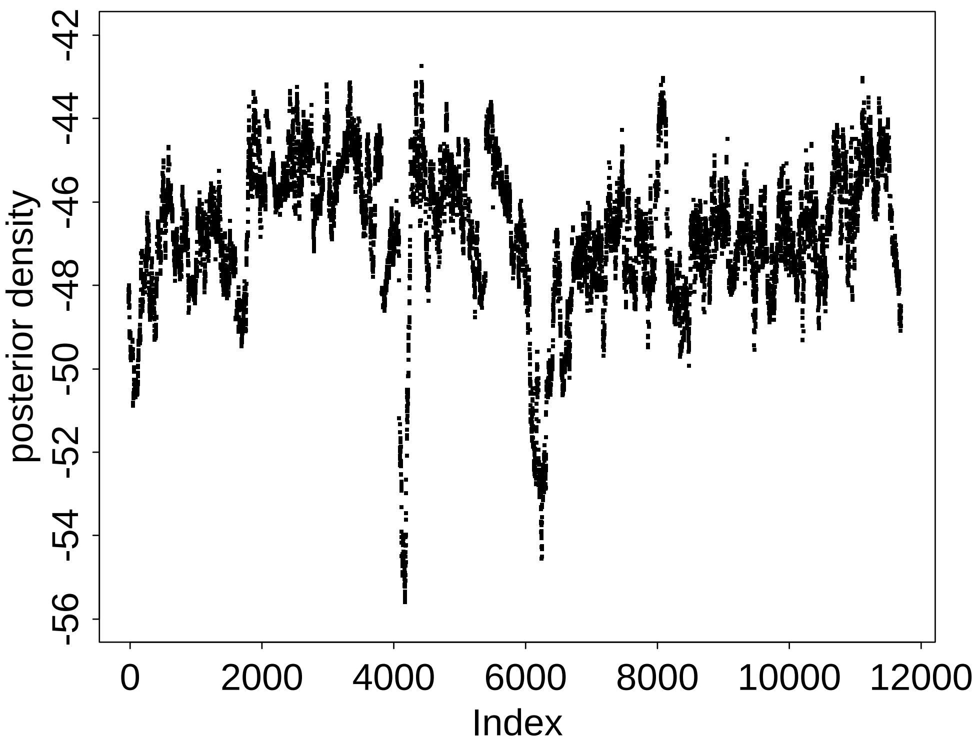

Figure 2: A sample of the posterior density for ϕ based on the simulated tempering algorithm.

section we assume that the stock price is a diffusion and we want to determine the diffusion coefficient such that the transition density matches ϕ.

More precisely, assume that the log stock price is a diffusion as specified in equation (2) with drift $r - d - \frac{1}{2}\sigma^2$ and with diffusion coefficient $\sigma(\cdot)$. The corresponding transition density $\phi(x, u; \sigma)$ is obtained by solving the Kolmogorov forward equation, see equation (4). We then search for a suitable value of σ such that the transition density at time t equals the estimated value of ϕ, i.e. such that

$$\phi(\cdot, t; \sigma) = \phi(\cdot). \tag{16}$$

However, two problems arise. First, as ϕ does not have smooth derivatives we cannot expect to obtain equality in (16). Instead we search for the value of σ for which the distance (in the sense of (17) below) between $\phi(\cdot, t; \sigma)$ and $\phi(\cdot)$ is minimized. Secondly, it is well known that the density obtained from (4) is seldom available explicitly. Instead we solve (4) numerically using the Crank-Nicholson scheme, see Duffie (1996), Poulsen (1999), and the references therein. Since the boundary condition at time zero is a one-point measure at x_0 we use instead the boundary condition that the distribution at time δ, with δ small, is a normal distribution with mean $(r - d - \frac{1}{2}\sigma(x_0)^2)\delta$ and variance $\sigma(x_0)^2\delta$, see (2). We take δ equal to the time step we use in the Crank-Nicholson procedure.

We regularize the problem of finding σ by searching over a finite dimensional space. To be precise we have in the examples presented in Section 5 modelled σ by a spline function with 24 knots. Of these 8 are outside the interval we use to measure the closeness of $\phi(\cdot, t; \sigma)$ and the estimated density $\phi(\cdot)$, and more knots are used in the center of the distribution than in the tails. Using a spline the second derivative of σ is a continuous piecewise linear function. We measure the closeness of $\phi(\cdot, t; \sigma)$ and $\phi(\cdot)$ by the term

$$\sum_{i=1}^{k} c_i[\log(\phi(x_i, t; \sigma)) - \log(\phi(x_i))]^2, \tag{17}$$

where the c_i's are weights that emphasize the center of the distribution and, similarly, the points x_i are densely spaced in the center of the distribution. In the examples in Section 5 we have used $k = 31$ with x_1 and x_k chosen so that the log density has the value –6 at these points. This means that we cover 0.9993 of the probability mass. We have used a fairly primitive search routine. The idea is that we update the second derivative σ'' locally so that σ is not changed outside an interval. If the knots of the spline are denoted $\tilde{x}_i$ we propose a random change of $\sigma''(\tilde{x}_i)$. We then change $\sigma''(\tilde{x}_{i+j})$, $j = 1, 2, 3$ in such a way that σ is not changed outside the interval $[\tilde{x}_i, \tilde{x}_{i+3}]$.

The above outlined procedure has worked well in the examples. We have managed to get the discrepancy (17) very close to zero. One should remember here that the derivative of ϕ has discontinuities, which makes it impossible to have a perfect fit.

5 Applications

In this section we apply our methods on real data from the Chicago Board of options Exchange (CBOE). The data we consider are Bid and Ask quotes on European calls and puts on the S&P 500 Index. For simplicity we only consider data from June 25 1990, and quotes on options expiring in December 1990, i.e. quotes for options having approximately six months to expiration; whence we take $t = 0.5$. Specifically we consider only two data sets; first we have quotes as they appeared near 10 pm on June 25 1990 on 16 calls and 13 puts. The second data set consists of quotes on the same set of options but near 2 am, the same day. Besides we also know the Index level near 10 pm and 2 pm. The data for 10 am are given in Table 1, and we will below distinguish the two data sets as the morning data and the afternoon data.

In order to apply our methods we need to determine the value of the interest rate r and the dividend d. One idea is to use T-bill rates to determine r and use for d the observed dividends. However, we choose to determine these quantities essentially from the the put-call Parity. Let $(B_i^C, A_i^C, B_i^P, A_i^P)$ be the bid and ask for a call and a put, respectively, with the same strike K_i^c, $i = 1, \ldots, n_0$. Then the requirements that the prices should fall in the bid-ask spreads, i.e.

$$B_i^C < \pi(K_i^c, p; C)\exp(-rt) < A_i^C, \quad B_i^P < \pi(K_i^c, p; P)\exp(-rt) < A_i^P,$$

can from (6) be rewritten as

$$\begin{aligned} &\max\{B_i^C \exp(rt), B_i^P \exp(rt) + C_0 - K_i^c\} < \pi(K_i^c, p; C) \\ &< \min\{A_i^C \exp(rt), A_i^P \exp(rt) + C_0 - K_i^c\}, \end{aligned}$$

where $C_0 = \pi(0, p; C)$. Thus we must find r and C_0 such that

$$\begin{aligned} &\max\{B_i^C \exp(rt), B_i^P \exp(rt) + C_0 - K_i^c\} \\ &< \min\{A_i^C \exp(rt), A_i^P \exp(rt) + C_0 - K_i^c\} \quad \text{for } i = 1, \ldots, n_0. \end{aligned} \tag{18}$$

For the morning and afternoon data sets we get a fairly broad interval for the values of the riskless interest rate r satisfying (18). We pick a particular value by regressing $(B_i^C + A_i^C)/2 - (B_i^P + A_i^P)/2$ on K_i. For the morning data this gives $\exp(rt) = 1.038$ and for the afternoon data we find $\exp(rt) = 1.036$. We have in our investigations below used $\exp(rt) = 1.038$ for both data sets. Given the value of r we find the interval $(\bar{B}_0, \bar{A}_0)$ for C_0 for which (18) is satisfied. Since the strike zero call is not traded we use $B_0 = \bar{B}_0 \exp(-rt)$ and $A_0 = \bar{A}_0 \exp(-rt)$ as bid and ask prices for this call. For the morning data this gives $(B_0, A_0) = (349.94, 350.82)$ and for the afternoon data we get $(B_0, A_0) = (347.46, 348.52)$. Using the relation (5) for the current index S_0 we can transform the above two intervals into intervals for the dividend d. The morning and afternoon indexes are 355.48 and 352.88, respectively, and the intervals for d are $(0.0264, 0.0314)$ and $(0.0249, 0.0310)$, respectively.

The data we have give a fairly detailed knowledge of the left hand side of the distribution, covering 70-80 percent of the distribution. Still, though, it is possible to fit a Christmas tree density to the risk neutral density. In particular, we do not see the second mode at the lowest strike value as in Jackwerth and Rubinstein (1995) (by them denoted "crash-o-phobia"). This difference must be attributed to the different kind of regularization used. In Figure 1 we have plotted two realizations of $\log(\phi)$ for the morning data. The two realizations have been obtained by taking a long run of the MCMC algorithm and then selecting the two with the largest and smallest value of the posterior density, respectively. In our case the posterior density is proportional to the prior density (the likelihood only enters as an indicator function for the prices to be in the bid-ask spread), and the posterior density therefore reflects the smoothness of ϕ. We have only plotted two densities in Figure 1 since these represent the variation that we have seen, and it appears that the restriction to the class of Christmas tree densities

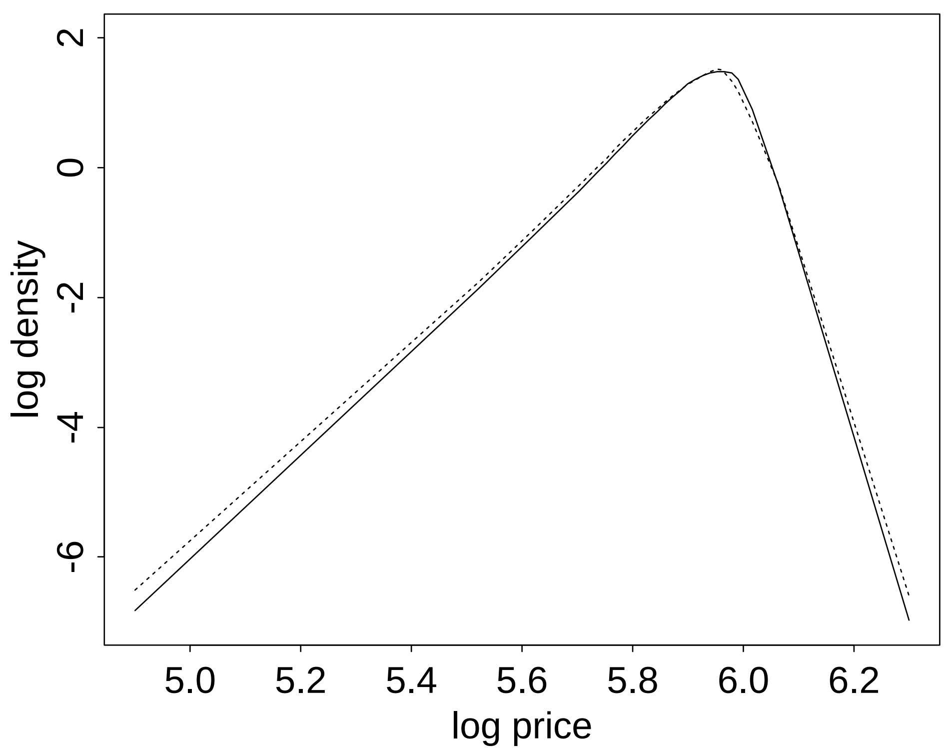

Figure 3: Log density $\phi(\cdot)$ as function of log price. The full drawn curve is a realization for the morning data and the dotted line is a realization for the afternoon data.

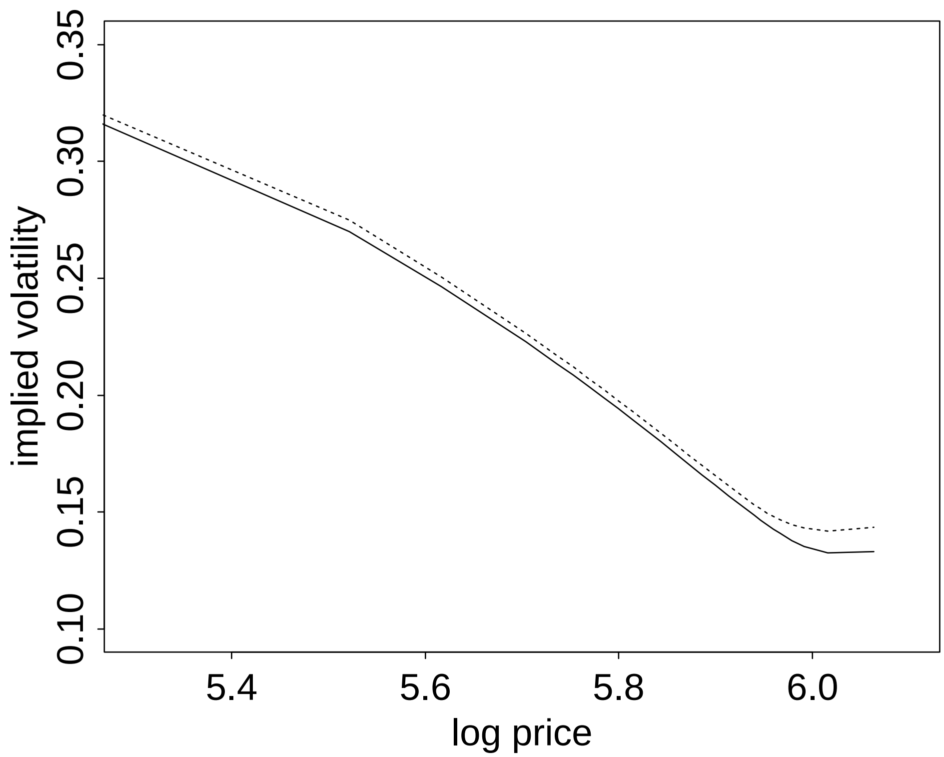

Figure 4: The implied volatilities based on the risk neutral densities shown in Figure 3. The full drawn curve is the implied volatility for the morning data and the dotted line is for the afternoon data.

leaves very few degrees of freedom for the density. As can be seen from Table 1 we have 21 strikes available for the morning data and we have modelled the log density (7) by 28 linear pieces adding one below the smallest strike and then adding some extras above strike 380. The parameters λ and γ in the prior density should be chosen in accordance with the number of linear pieces used for the log density. We have simply taken $\lambda = \gamma = 1$ and this is in fairly good agreement with the fact that we end up with $\sum |w_i|$ being of the order 30-35 and with the mean being $\gamma + 27\lambda$.

In Figure 3 we have plotted a realization of the log density for the log stock price for both the morning and the afternoon data sets. We have here taken the realizations with the largest value of the posterior density in a long run of the MCMC algorithm. As can be seen there is a shift in the densities in accordance with the change of the current index from 355.48 to 352.88.

In Figure 4 we have plotted the implied volatilities based on the realizations of the densities shown in Figure 3. Both of the implied volatility curves show a smooth smile.

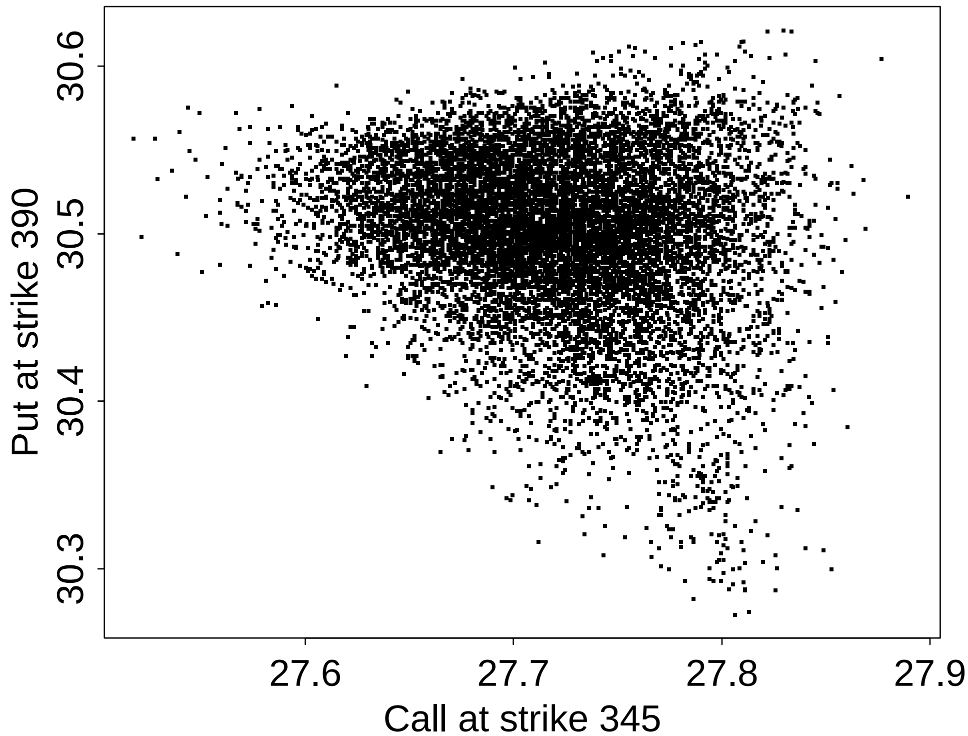

Figure 5: A sample from the MCMC simulation of joint values of the call price at strike 345 and the put value at strike 390.

To see the variation in the posterior distribution for the density p we have collected values of $\tilde{\pi}(345, p; C)$ and $\tilde{\pi}(390, p; P)$ for a long run of the MCMC algorithm. In Figure 5 the two values are plotted against one another.

We next turn to the estimation of the volatility $\sigma(\cdot)$ as described in Section 4 based on the equation (4). We have here used the two realizations of ϕ shown in Figure 3. Using these realizations of ϕ and the relation (5) we find that the value of the dividend d is 0.0293 and 0.0274, respectively for the morning and afternoon data. In Figure 6 we have plotted the estimates of $\sigma(\cdot)$ from the two data sets. The estimates have been obtained as described in Section 4. If we plot $\phi(\cdot, t; \sigma)$ in Figure 3 we will not be able to see the difference to the estimated functions plotted there, that is, the estimated σ reproduces the estimated density almost perfectly. It is noteworthy that the estimated volatility shown in Figure 6 is more regular and has a more appropriate tail behaviour than the estimates obtained in Lagnado and Osher (1997). Also we see that there is an increase in the volatility from the morning to the afternoon.

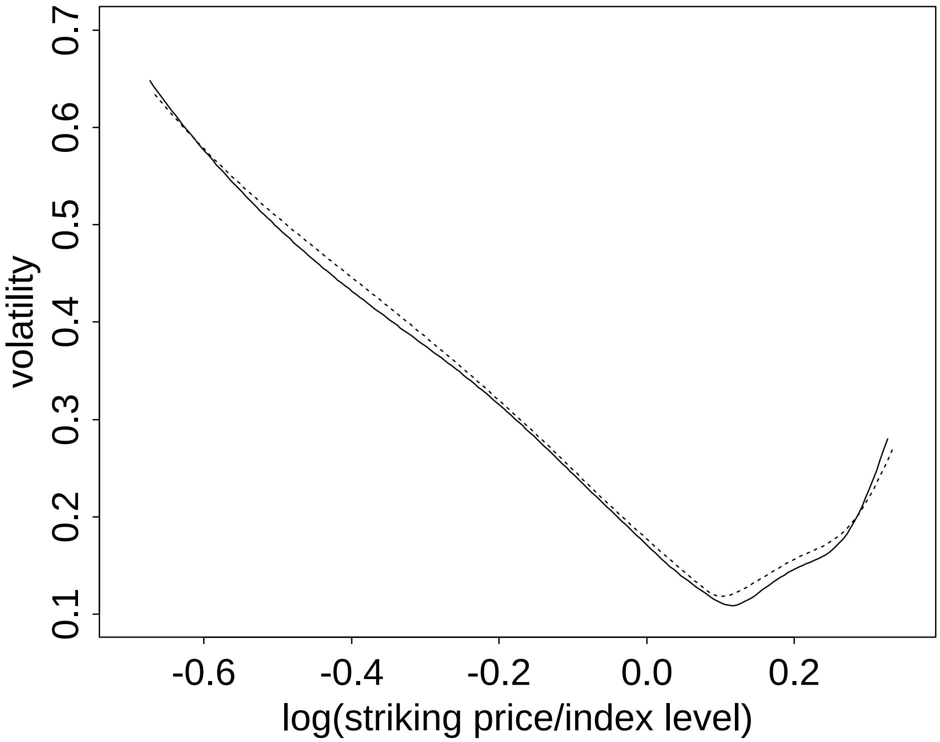

Figure 6: The volatilities based on the risk neutral densities shown in Figure 3 and on the relation (4) plotted against $\log(K/S_0)$, where K is the strike value and S_0 is the current index. The full drawn curve is the volatility for the morning data and the dotted line is for the afternoon data.

6 Discussion and open problems

Empirical investigations of stock prices have shown that the increments tend to have heavy tails under the true probability measure. In accordance with this we have chosen a class of densities for the risk neutral measure with the same tail behaviour and have managed to fit this to observed European call and put option prices. We introduced the class of Christmas tree densities and investigated the variability in the risk neutral density fitting the data through a Bayesian formulation. In order to sample from the posterior distribution of the risk neutral density we designed a Markov chain Monte Carlo algorithm. The risk neutral measures that we obtain do not show the 'crash-o-phobia' reported in Jackwerth and Rubinstein (1995).

Having a good estimate of the risk neutral distribution gives us the possibility of considering other questions. In particular we have in this paper focused on the diffusion model and considered the inverse problem of finding $\sigma(\cdot)$ from the risk neutral density. Using the entire estimated risk neutral density regularizes the solution for σ and in particular we obtain a more reasonable behaviour of σ in the tails than seen in Lagnado and Osher (1997).

Our investigation raises a number of questions. Here we list some of these and separate them into financial/statistical questions and mathematical questions.

Open financial and statistical problems:

1) It is a topic for future empirical research to study the shift in the risk neutral density over calendar time and over time to expiration. As a part of this research we need to improve the estimates of the interest rate r and the dividend d. For results in this direction see Jackwerth and Rubinstein (1995).

2) Moreover, describing the stock price process in a diffusion setting it will be interesting to study the change in volatility over time (as done very preliminary in Figure 6). It also seems relevant to investigate if it is reasonable to let the volatility be independent of time. If not, how should we then model time dependence ?

3) From a theoretical point of view it will be interesting to model stock prices under the true and risk neutral probability measure simultaneously. This must be done in such a way that the stock prices have heavy-tailed behaviour under both measures.

Open mathematical problems:

1) Can we characterize the class of functions $\pi(\cdot, p; C)$ that are solutions to equation (3) (we know that the function must be convex with certain limiting behaviour for $K \to 0$ and $K \to \infty$)?

2) Will any 'Christmas tree' density $p(\cdot)$ with continuous second derivative be the solution to equation (4) for some $\sigma(\cdot)$?

3) Is the inversion from $p(\cdot)$ to $\sigma(\cdot)$ ill-posed, and can we say how much regularization (say in terms of the number of knots in a spline) we need in terms of the uncertainty concerning $p(\cdot)$? Results in this direction can be found in Bouchouev and Isakov (1997)

Acknowledgement: The Data used in this study has been kindly supplied by CAF, Centre for Analytical Finance, University of Aarhus and Aarhus School of Business.

References

[1] O.E. Barndorff-Nielsen. Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24:1 – 13, 1997.

[2] O.E. Barndorff-Nielsen. Processes of normal inverse gaussian type. *Finance and Stochastics*, 2:41 – 68, 1998.

[3] A. Bensoussan. On the theory of option pricing. *Acta Applicandae Mathematicae*, 2:139 – 158, 1984.

[4] B.M. Bibby and M. Sørensen. A hyperbolic diffusion model for stock prices. *Finance and Stochastics*, 1:25 – 42, 1997.

[5] T. Bollerslev, R.F. Engle, and D.B. Nelson. Arch models. In *Handbook of Econometrics, Vol. IV*, pages 2959 – 3038. North-Holland, Amsterdam, 1994.

[6] I. Bouchouev and V. Isakov. The inverse problem of option pricing. *Inverse Problems*, 13:L11 – L17, 1997.

[7] I. Bouchouev and V. Isakov. Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. Preprint, 1998.

[8] D. Breeden and R. Litzenberger. Prices of state-contingent claims implicit in options prices. *Journal of Business*, 51:621 – 651, 1978.

[9] M.H.A. Davis. A general option pricing formula. Preprint, Imperial College, London, 1994.

[10] F. Delbaen and W. Schachermayer. The variance-optimal martingale measure for continuous processes. *Bernoulli*, 2:81 – 105, 1996.

[11] E. Derman and I. Kani. Riding on a smile. *RISK*, 7:32 – 39, 1994.

[12] D. Duffie. *Dynamic asset pricing theory*. Princeton University Press, 1996.

[13] B. Dupire. Pricing with a smile. *RISK*, 7:18 – 20, 1994.

[14] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1:281 – 299, 1995.

[15] E.F. Fama. The behaviour of stock market prices. *Journal of Business*, 38:34 – 105, 1965.

[16] H.U. Gerber and E.S.W. Shiu. Option pricing by esscher transforms. *Transactions of the Society of Actuaries*, 56:99 – 140, 1994.

[17] C. J. Geyer and E. A. Thompson. Annealing markov chain monte carlo with applications to ancestral inference. *Journal of the American Statistical Association*, 90:909 – 920, 1995.

[18] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman Hall, 1996.

[19] J.C. Jackwerth and M. Rubinstein. Recovering probability distributions from contemporaneous security prices. Preprint, 1995.

[20] R. Lagnado and S. Osher. Reconciling differences. *RISK*, 10:79 – 83, 1997.

[21] F. Longstaff. Martingale restriction tests of option pricing models. Preprint, University of California, Los Angeles, 1990.

[22] R. Poulsen. Approximate maximum likelihood estimation of discretely observed diffusion processes. CAF Working Paper no. 29, University of Aarhus, 1999.

[23] M. Rubinstein. Implied binomial trees. *The Journal of Finance*, 49:771 – 818, 1994.

[24] T.H. Rydberg. The normal inverse gaussian levy process: Simulation and approximation. *Communications in Statistics. Stochastic Models*, 13:887 – 910, 1997.

[25] M. Schweizer. Option hedging for semimartingales. *Stochastic Processes and their Applications*, 37:339 – 363, 1991.

[26] M. Schweizer. Approximation pricing and the variance-optimal martingale measure. *The Annals of Probability*, 24:206 – 236, 1996.

[27] D. Shimko. Bounds of probability. *RISK*, 6:33 – 37, 1993.