Normal mixture diffusion with uncertain volatility: Modelling short- and long-term smile effects

Carol Alexander [*]

ISMA Centre, School of Business, University of Reading, UK

Abstract

This paper introduces a parameterization of the normal mixture diffusion (NMD) local volatility model that captures only a short-term smile effect, and then extends the model so that it also captures a long-term smile effect. We focus on the 'binomial' NMD parameterization, so-called because it is based on simple and intuitive assumptions that imply the mixing law for the normal mixture log price density is binomial. With more than two possible states for volatility, the general parameterization is related to the multinomial mixing law. In this parsimonious class of complete market models, option pricing and hedging is straightforward since model prices and deltas are simple weighted averages of Black–Scholes prices and deltas. But they only capture a short-term smile effect, where leptokurtosis in the log price density decreases with term, in accordance with the 'stylised facts' of econometric analysis on ex-post returns of different frequencies and the central limit theorem. However, the last part of the paper shows that longer term smile effects that arise from uncertainty in the local volatility surface can be modeled by a natural extension of the binomial NMD parameterization. Results are illustrated by calibrating the model to several Euro–US dollar currency option smile surfaces.
© 2003 Elsevier B.V. All rights reserved.

[*] Tel.: +44-1189-316431.
  E-mail address: c.alexander@ismacentre.rdg.ac.uk (C. Alexander).

1. Introduction

The Black–Scholes (BS) implied volatility smile effect arises from an apparent under-pricing of out-of-the-money (OTM) puts and calls by the Black and Scholes (1973) formula for simple European options under constant volatility. This is because the market does not believe the BS assumption that the price process is a geometric Brownian motion (or equivalently, that the log price is an arithmetic Brownian motion) with constant volatility. In particular, if leptokurtosis were present in the price process, the likelihood of large price changes would be higher than that assumed by a Brownian motion with the same volatility. Thus, if option traders believe in a leptokurtic price process, they will place a greater value on OTM puts and calls, giving market prices that are higher than BS model prices. Consequently the BS implied volatility of these options will be higher than the BS implied volatility of at-the-money (ATM) options. [1]

Of course, leptokurtic or skew price densities are not the only reason for smile or skew effects in the data. An 'apparent' smile can arise when there are market imperfections, such as an exchange that quotes stale prices of OTM options that are too high, being left over from a time when the market was at that level. Smile effects in long-term options could be accentuated by traders making prices that include hedging costs, which are greater for OTM than ATM options; models with stochastic interest rates (respectively, dividends) are likely to explain a good fraction of the longer term smile in currency (respectively, equity) options; liquidity premia also play a role because – although the market for very in-the-money (ITM) options is usually quite thin, many investors preferring to trade in the underlying – market prices of these options can differ from BS prices because of the large cash amounts involved with these transactions; and finally, credit effects can accentuate the skew in equity option smiles because there is a positive probability of the price being zero.

Often the smile effect is greatest for near term options but decreases with maturity. In this case the market data corroborate the stylised facts that emerge from econometric research. Examination of historical returns sampled at different frequencies has shown that (a) the excess kurtosis estimated from unconditional historical returns densities increases with sampling frequency [2] and (b) there is strong evidence to support a non-constant conditional volatility model, and this implies leptokurto-

---

[1] Similarly, if skewness is present in the price process, the likelihood of large price falls (in equities) – or rises (in commodities) – would be higher than that assumed by a Brownian motion with the same volatility. Thus traders will place a greater value on OTM puts – or calls, for commodities – and the BS implied volatility of these options will be increased.

[2] Intra-day and daily returns commonly exhibit highly significant excess kurtosis, but this usually disappears when returns are taken over a month or more (Goodhart and O'Hara, 1997; Gencay et al., 2001).

sis in the unconditional density, even if the conditional densities are normal. [3] Both (a) and (b) predict that leptokurtosis in the ex-post returns density increases with sampling frequency, but for a different reason. [4]

If the market held beliefs about the future evolution of the underlying price based only upon observations of past returns, a dynamic process with time-varying and state dependent volatility and possibly also heavy tails in the conditional densities of returns may be used: either or both assumptions would lead to leptokurtosis in the price density that is decreasing as the term increases. Many non-normal GARCH models will capture this, but there are alternatives. Other types of price processes that are more general than geometric Brownian motion can capture short-term smile effects. For example in the Merton (1976) jump diffusion model, leptokurtosis in price densities will decrease with term.

In many markets a persistence of smile effects into three months or longer maturity implied volatilities is commonly observed, even though the underlying ex-post returns densities are approximately normal at the three-month sampling frequency. However, there is no inconsistency here. Price densities will be leptokurtic even when ex-post returns are unconditionally normally distributed if there is uncertainty about the future volatility in the price process. There is no contradiction with the stylised facts that emerge from econometric analysis. By 'uncertain volatility' we do not mean that there is uncertainty in the past, present and future volatility of the price process – the uncertainty is only in the minds of traders because they do not know the average volatility of the price process in the future – and this uncertainty cannot be observed in ex-post returns. If, as is often the case, the uncertainty about volatility increases with term, so that traders are less certain about average volatility over the next year than they are about average volatility over the next month, then ex-ante price densities could become more leptokurtic as the term increases.

Most of the research on uncertainty in volatility has focused on the stochastic volatility models of Hull and White (1987), Heston (1993) and many others; however, Avellaneda et al. (1995) introduced an alternative approach where forward volatilities are not represented by a stochastic process, but instead each forward volatility is restricted to lie within a given range, and the uncertainty about its value within this range is represented by an unconditional volatility distribution. Neither class of models will allow for arbitrage-free pricing because the new uncertainty introduces market incompleteness. In this case option prices will include a risk premium so there is no unique model price if traders have differing attitudes to risk. For liquid options,

---

[3] The vast literature on estimating conditional densities of returns using models in the generalized autoregressive conditional heteroscedasticity (GARCH) family provides overwhelming evidence for non-constant conditional volatility (Bollerslev, 1986, 1987; Baillie and Bollerslev, 1989; Bollerslev et al., 1992, 1994). Also, conditional heteroscedastic effects become more pronounced as the frequency of returns increases (Baillie and Bollerslev, 1990; Dacorogna et al., 1998).

[4] These findings also concur with the central limit theorem: i.e. that the sum of non-normal variables tends towards a normal variable: if X_i have independent identical distributions with mean 0, variance σ^2 and excess kurtosis κ, then $Y = (X_1 + \cdots + X_n)$ has a distribution with mean 0, variance $n\sigma^2$ and kurtosis $3 + \kappa/n$, so the kurtosis approaches 3 and the excess kurtosis approaches zero as n increases.

and short-term ATM options in particular, risk premia on the buy and sell side are likely to be small and similar; buyers or sellers can close their portfolios quickly if they wish to. But for longer term OTM options where trading is sparse, the bank that writes the option may include a substantial risk premium in the price, depending on their risk attitude, and so the market incompleteness introduced by uncertain volatility can cause large bid-offer spreads at longer maturities.

This paper focuses on two distinct smile effects, both associated with leptokurtosis in the price density: one effect that is largest for very short dated options, where the associated leptokurtosis in the price density decreases with term, and another effect that is due to volatility uncertainty, where leptokurtosis in the price density and the associated smile effect in implied volatilities can increase with term. The prices observed in the market will reflect a mixture of these two effects and the model that is described in this paper captures both of them, whilst also distinguishing between them.

First, an elegant (multinomial) parameterization of the normal mixture diffusion (NMD) local volatility model of Brigo and Mercurio (2000, 2001a,b, 2002) is introduced and used to describe a leptokurtic price process in a complete market setting. The resulting specification of the local volatility is intuitive, parsimonious and straightforward to calibrate to the market smile. In this model, arbitrage-free option prices are obtained because no market incompleteness is introduced. Our first example calibrates the simple 'binomial' NMD, which has only three parameters, using an analytic relationship between the local volatility model prices and the BS prices for standard European options. Option sensitivities for this model are easily calculated as weighted averages of BS sensitivities.

Because there are so few model parameters, the model will not fit market prices exactly, and its main practical application may be for hedging liquid options that will be marked-to-market. [5] Of course, local volatility models are capable of fitting market prices exactly. This paper, however, does not seek to demonstrate the power of such models for practical problems. Rather, it seeks to provide a conceptual insight to volatility analysis – by distilling the essence of the local volatility approach, we find that it is intuitive to simplify such a model so that it captures only a short-term smile effect. We first show that the multinomial NMD parameterization explains only the short-term component of the smile – thus it is a deterministic volatility model that supports the 'stylised facts' of econometrics. We then distinguish this short-term smile component from the longer term smile effect that arises from uncertainty in future volatility. These longer term smile effects are introduced into the multinomial NMD parameterization by adding an uncertainty to the mixing law. Finally, several examples of calibration to market data on Euro–US dollar option prices reveal how uncertainty about the volatility of this exchange rate has changed over the different dates in our sample.

---

[5] In this context, the precision of the calibration should still be very important, as Prof. Emanuel Derman has so eloquently explained in "The Future of Modelling" (Goldman Sachs).

The structure of this paper is as follows: Section 2 reviews the literature on local volatility models; Section 3 explains the normal mixture diffusion (NMD) local volatility framework introduced by Brigo and Mercurio (2000, 2001a,b, 2002) and Brigo et al. (2002) where price densities are finite lognormal variance mixtures; Section 4 describes how a multinomial parameterization of the NMD will capture a term structure of kurtosis, so that it only explains the short-term smile effects described above. Section 5 presents the results of calibrating the model to a currency option smile surface, and discusses how much of the observed behaviour of option prices can be attributed to this short-term smile model. Section 6 extends the binomial NMD parameterization to include uncertainty in the mixing law and we demonstrate that this model can explain longer maturity smile effects in currency option market data. In Section 6, we investigate the robustness of the parameterization and the goodness of fit of the calibration. Section 7 summarizes the results and concludes.

2. Local volatility

The deterministic approach to non-constant volatility preserves market completeness by assuming the instantaneous or 'local' volatility of the price process is a deterministic function of time and the underlying asset values. Non-parametric local volatilities may be calibrated to current market prices of options using finite difference schemes. The 'implied tree' approach was pioneered by Dupire (1994, 1997) and subsequently extended to trinomial trees by Derman et al. (1996). [6] One first interpolates and extrapolates the implied volatility surface and then uses a finite difference solution of the BS equation to extract local volatilities for each node in the tree. Several refinements of the finite difference schemes employed for the model resolution have been proposed, the most stable of which appear to be the Crank–Nicholson scheme for trinomial lattices used by Andersen and Brotherton-Radcliffe (1997).

Direct calibration with non-parametric local volatilities will provide an exact fit to the current market data, but the local volatilities (and therefore also the model hedge ratios) are typically very sensitive to the interpolation and extrapolation methods used, particularly for the 'wings' of the implied volatility smile and for longer maturity options. If the smile is very pronounced, local squared volatilities may become negative, necessitating the use of some ad hoc procedures. Moreover, with incomplete and/or stale option price data, the calibrated local volatility surfaces can be excessively 'spikey' and consequently can give large variations in delta from day to day. For this reason, regularization methods have been used by Avellaneda et al. (1997), and Bouchouev and Isakov (1997, 1999) amongst others, to obtain the smoothest possible fit to the interpolated implied volatility surface. But still, with such an exact fit, the local volatility surface may jump considerably over time and the approach provides no foresight of future movements in the local volatility

---

[6] See also Breeden and Litzenberger (1978), Rubinstein (1994).

surface. These local volatilities are unlikely to perform well in out-of-sample tests and the direct calibration approach may be of limited practical use for hedging purposes.

Much research now focuses on the use of a parametric form for local volatilities. Pioneered by Cox and Ross (1976), in the parametric approach a functional form for the local volatility is chosen and the parameters are calibrated using only the available and reliable market prices. Typically, parameterized local volatilities will be smoother and more stable over time than those obtained by direct calibration. Although the model prices based on a parameterized local volatility surface will not exactly match the current market prices, these local volatilities could be more useful for hedging. If the calibration is sufficiently robust, we should gain some idea of the likely movements in the local volatility surface over time, and could modify hedge ratios accordingly. [7]

Recently a number of parametric and semi-parametric forms for local volatility have been proposed in the literature, including: simple polynomials (Dumas et al., 1998); cubic splines (Coleman et al., 1998); hyperbolic trigonometric functions (Brown and Randall, 1999); Hermite polynomials (McIntyre, 2001); and piecewise quadratic forms (Beaglehole and Chebanier, 2002) amongst many others. Whilst these all represent useful developments in the local volatility literature, the problem for practitioners is now how to choose the 'best' functional form for their purposes. It is not simply a question of choosing a functional form for local volatility that is flexible enough to provide a good fit to the observed smiles or skews in option market data, nor simply a question of ensuring that the parameterization is sufficiently parsimonious to be calibrated with accuracy. An important factor when choosing a functional form for local volatility is that it should reflect what we believe about the underlying price process: what is the risk neutral density of a price process with a given functional form for the local volatility, and does this density have appealing properties?

Whilst the log price process can always be specified as a Brownian with local volatility given by the calibrated chosen parametric form, it is difficult – if not impossible – to analyse the properties of the risk neutral price density. In general there will be no tractable functional form for the risk neutral density and, furthermore, model prices and hedge ratios may not be given in closed form.

Ritchey (1990) and Melick and Thomas (1997) introduced a finite normal variance mixture model for pricing options, and the finite normal mixture framework has since found many applications in finance – see Bingham and Kiesel (2002) for a survey. One particularly important application is that the local volatility can be linked to an analytically tractable price density, a result that was first stated by Brigo and Mercurio (2000). Subsequently Brigo and Mercurio (2001a) proved that if the risk neutral price density is a lognormal variance mixture, with the volatility in each lognormal density being a deterministic function of time, then the price process will

---

[7] In the presence of the smile the option delta is $\Delta_{BS} + \text{vega}\, \partial\sigma/\partial S$ where Δ_{BS} is the BS delta and σ is the implied volatility. Hence when hedging options it is important to have some idea of the movements in the smile as the underlying changes.

be a Brownian motion with a local volatility given by the volatility of a lognormal variance mixture. [8] Brigo et al. (2002) then extended this result to general lognormal mixtures, appropriate when there is skew in the log price density, and also proved the converse result, that if the local volatility is given as a weighted sum of average volatilities of deterministic volatility processes, the log price density will be a normal mixture. Brigo and Mercurio have named this model the 'Normal Mixture Diffusion' (NMD).

The NMD can be seen as an extension of the BS model where the volatility is not constant, but instead there are a finite number of continuous and bounded deterministic volatility processes; and price densities are not lognormal, but lognormal mixtures with a fixed mixing law. It is therefore not a stochastic volatility model, but a local volatility model with some very tractable properties. In particular, risk neutral option prices are just a weighted average of BS prices and option sensitivities will also be weighted averages of BS sensitivities. When several parameterized volatility processes are assumed the model can fit market prices almost exactly, so it has useful applications to pricing path dependent options. [9] On the other hand, a parsimonious parameterization of the NMD is likely to have more robust calibration, and so be more suitable for hedging purposes. By linking local volatility to the price process, the NMD has revived the literature on local volatility models and this paper now continues this line of research.

3. The normal mixture diffusion (NMD)

Suppose that there are a fixed, finite number of continuous and bounded deterministic volatility processes $\sigma_1(t), \sigma_2(t), \ldots, \sigma_m(t)$. At any point in time t the t-period volatility associated with the ith instantaneous volatility is denoted $\sqrt{v_i(t)}$ where

$$v_i(t) = \int_0^t \sigma_i^2(s)\, \mathrm{d}s.$$

Denote by $X(t)$ the log price of an asset, such as an equity or index or exchange rate. So $X(t)$ is a random variable with some probability measure which, without loss of generality, can be assumed to be the risk neutral measure. [10] Assume that the dynamics of the log price process follow a diffusion with local volatility:

---

[8] One condition for their result is that the mixing law for the risk neutral density must be finite and discrete.

[9] Exact pricing of path dependent options (without analytic approximations) is not straightforward in the NMD: since transition densities are not known (only marginals are) one needs to resort to Monte Carlo simulation. But, because the calibration can be very accurate, this should result in good pricing of path-dependent options.

[10] Note that the use of the logarithm means that results here will be presented in terms of normal mixture log price densities and arithmetic diffusions rather than lognormal mixture price densities and geometric diffusions. However the mixing law is invariant, since when $X = \ln S$ the price density $g(s) = (1/s)f(\ln s)$ so if $f(\cdot)$ is a normal mixture with mixing law $[\lambda_1, \ldots, \lambda_m]$, then $g(\cdot)$ will be a lognormal mixture with the same mixing law.

$$dX = \mu(X,t)\,dt + \sigma(X,t)\,dB, \tag{1}$$

where B is a Brownian motion, and $\mu(X,t) = r - 1/2\sigma(X,t)^2$ and r is the risk neutral drift. Furthermore, assume that X has a normal mixture risk neutral density at every time t given by

$$f_t(x) = \sum_{i=1}^{m} \lambda_i \phi(x; rt - 1/2v_i(t), v_i(t)) = \sum_{i=1}^{m} \lambda_i \phi_{i,t}(x), \tag{2}$$

where ϕ denotes the normal density function and $\sum_{i=1}^{m} \lambda_i = 1$. Then Brigo and Mercurio (2001a) and Brigo (2002) prove that the local volatility in (1) that is consistent with the price densities (2) is given by [11]

$$\sigma(x,t) = \sqrt{\sum_{i=1}^{m} \lambda_{i,t}^* \sigma_i^2(t)},$$

where

$$\lambda_{i,t}^* = \lambda_i \phi_{i,t}(x)/f_t(x). \tag{3}$$

Since market completeness is preserved in this framework, arbitrage-free pricing of standard European options on X is straightforward. Absence of arbitrage implies that the option price is the discounted expectation of the pay-off under the risk neutral density (2). The simple form of this density allows this expectation to be expressed as a weighted average of expectations under normal densities. That is, the normal mixture option price will be a weighted average of BS option prices. For a European option with strike K and maturity T and underlying at x the normal mixture option price is given by

$$NM(x,K,T) = \sum_{i=1}^{m} \lambda_i BS(x,K,T,\eta_i(T)), \tag{4}$$

where $\eta_i(T) = \sqrt{v_i(T)/T}$. Note that hedging with European options is also straightforward in the NMD framework, since the expression (4) allows a simple representation of the option sensitivities as a weighted sum of the BS sensitivities.

The basis of the model calibration is to minimize some distance metric between model prices and current market prices of simple European options. But before this can be done it is first necessary to fix the number of instantaneous volatility processes m and parameterize the volatilities $\eta_i(t)$ for $i = 1, 2, \ldots, m$. One factor to take into account is the large number of model parameters: In addition to $m - 1$ weights $\lambda_1, \lambda_2, \ldots, \lambda_{m-1}$ [with $\lambda_m = 1 - (\lambda_1 + \lambda_2 + \ldots + \lambda_{m-1})$] each of the m volatility processes could have many parameters. Therefore, to reduce the number of parameters, Brigo and Mercurio (2000) have suggested setting $m = 2$ or 3, and assuming that $\eta_i(t) = c_i$ (a constant) for all t. In that case, the average variances and drifts in each

---

[11] Brigo (2002) explains how to mix normal densities with different drifts $r - 1/2\sigma_i^2$, showing that inserting the $1/2\sigma_i^2$ terms in the basic normal densities to be mixed amounts to adding $-1/2\sigma(X,t)^2$ into the drift term in the final log-price dynamics.

density of the mixture will not depend on t, but this is inconsistent with a term struc-
ture of kurtosis observed in the market. If the number of normal densities in the mix-
ture does not vary over time, and neither do their weights in the mixture, the only
way to model variation in kurtosis over time would be to build this into the param-
eterization of the volatilities.

4. Parameterization of the NMD to capture the short-term smile

To calibrate the NMD using (4) one has to assume a value for the number of vol-
atility processes m and to specify a form for each of the average volatilities $\eta_i(t)$ for
$i = 1, 2, \ldots m$. In the absence of further structure, these assumptions are quite arbi-
trary. We now propose to add further structure to the NMD so that both the number
of volatility processes and the behaviour of the average volatilities are fully deter-
mined. By restricting the possible values taken by each volatility process on any time
interval Δt, and linking the mixing law to these values, we derive a parsimonious
parameterization of the NMD that captures the term structure of kurtosis in price
densities, and the short-term smile effect, in a tractable manner. We shall make
two assumptions:

Assumption 1. The deterministic volatility processes $\sigma_1(t), \sigma_2(t), \ldots, \sigma_m(t)$ are
piecewise constant over a certain time interval Δt and each takes a value from the set
$[s_1, \ldots, s_d]$ in every time interval, with $d \leqslant m$.

This assumption implies a constant, bounded volatility over each interval Δt. It is
not a very restricting assumption, since the time interval can be made as short as you
wish. [12] The number m of volatility processes in the model and the average volatility
are determined by the maximum maturity of the options on X that are to be priced
and/or hedged and the length of the basic time interval Δt. For example, suppose that
there are only two possible volatility values, so the value of each of the volatility pro-
cesses can be either high ($s_1 = \sigma_H$) or low ($s_2 = \sigma_L$) in each time interval of length Δt.
If the maximum option maturity is $N\Delta t$ then $m = 2^N$. More generally, with d distinct
volatility values, $m = d^N$. [13]

The assumption that different volatility states can exist in a financial market is not
new. Epps and Epps (1976) introduced a behavioural model where different volatility
states arise from the different types of traders in the market, who have different

---

[12] In fact, its choice should depend on the kurtosis term structure of X: If daily or intra-day changes in
X have high excess kurtosis but this excess kurtosis disappears after a week or so, Δt might be taken to be 1
day. But if excess kurtosis is still strong in weekly data, Δt could be 1 week.

[13] If the maximum maturity of options on the underlying is 6-months (26 weeks) and the basic time
interval $\Delta t = 1$ week, there will be $m = 2^{26}$ different volatility processes in the two volatility state
(binomial) NMD. Note that, even in this most simple case, m is much greater than the two or three distinct
volatility processes that were previously used by Brigo and Mercurio (2000, 2001a,b, 2002) when
calibrating the NMD.

expectations regarding returns and volatilities according to which they form their own prices and trade. Later, in foreign exchange markets, McFarland et al. (1982) and in stock markets, Ball and Torous (1983) decompose the total change in price into ''normal'' and ''abnormal'' components, where the normal component is due to only a temporary imbalance between supply and demand, or the receipt of information causing only marginal price changes. The ''abnormal'' component, present in larger than normal price changes, is linked to trading volume, and depends on the arrival of new relevant information. More recently, the Markov switching GARCH model introduced by Hamilton and Susmel (1994) and Cai (1994) has been modified by Gray (1996) and Klaassen (1998) so that the process variance, instead of being equal to the variance of the existing regime (as it would be in a pure Markov Switching model), is a weighted average over the regime-specific variances.

Fig. 1 depicts Assumption 1 when there are only two volatility states (σ_H and σ_L) and the maximum maturity is $3\Delta t$. In this case there are $m = 8$ distinct deterministic volatility processes, marked in different colors on the figure. [14] In general, the assumption of a finite number of possible volatility states in each time interval Δt allows one to enumerate also the number of distinct total variances over each interval $n\Delta t$. For example, when there are only two possible volatility states (as in Fig. 1) the number of distinct values for $v_i(n\Delta t)$ is $n + 1$ [for $n = 1, 2, \ldots, N$]. There are only four distinct values for $v_i(3\Delta t)$, only three distinct values for $v_i(2\Delta t)$, and only two distinct values for $v_i(\Delta t)$. These values are shown in the three centre columns of Table 1.

The second assumption concerns the mixing law $[\lambda_1, \lambda_2, \ldots, \lambda_m]$ on the component normal densities $\phi_{1,t}(x), \phi_{2,t}(x), \ldots, \phi_{m,t}(x)$ in the risk neutral densities $f_t(x)$ of X:

Assumption 2. For each i, set $\lambda_i = \theta_1^{n_{i,1}} \theta_2^{n_{i,2}} \ldots \theta_d^{n_{i,d}}$ where $n_{i,j}$ = number of intervals Δt in which the ith volatility process takes value s_j. The new model parameters are such that $\theta_1 + \theta_2 + \cdots + \theta_d = 1$.

To motivate this assumption, consider again the case $d = 2$ and the 3-period model depicted in Fig. 1. For brevity, write $\theta_1 = \lambda$ and $\theta_2 = 1 - \lambda$, so the risk neutral density $f_t(x)$ of X at any point in time t has only three parameters: λ, σ_H and σ_L. The mixing law $\lambda_1, \lambda_2, \ldots, \lambda_8$ applied to the eight volatility processes $\sigma_1(t),$ $\sigma_2(t), \ldots, \sigma_8(t)$ is given in the right hand column of Table 1. Although at any point in time t the density $f_t(x)$ will be a weighted sum of the same number of normal densities, they are not all distinct. For example, in the 3-period process depicted in Fig. 1, there are only four distinct normal densities in the density at time $3\Delta t$, only three distinct normal densities in the density at time $2\Delta t$, and only two distinct normal densities in the density at time Δt. The effective weights on these densities in the mixing

---

[14] The tree drawn in Fig. 1 is useful to understand the rule of the mixing law given by Assumption 2. However, it should be emphasized that this is not a stochastic volatility model: branches of the tree do not correspond to probabilistic states for a random volatility process.
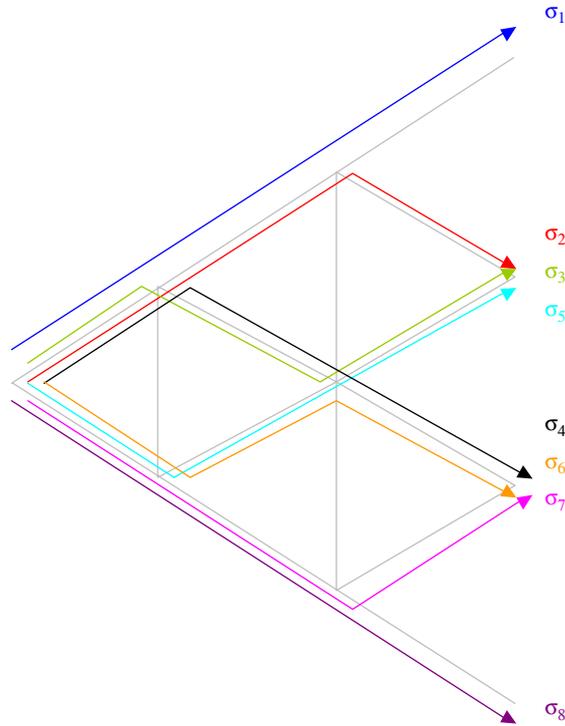
Fig. 1. Volatility tree [$d = 2$ and $N = 3$].

Table 1
Volatilities and mixing law [$d = 2$ and $N = 3$]

Volatility	$v_i \Delta(t)$	$v_i(2\Delta t)$	$v_i(3\Delta t)$	λ_i
σ_1	σ_H^2	$2\sigma_H^2$	$3\sigma_H^2$	λ^3
σ_2	σ_H^2	$2\sigma_H^2$	$2\sigma_H^2 + \sigma_L^2$	$\lambda^2(1-\lambda)$
σ_3	σ_H^2	$\sigma_H^2 + \sigma_L^2$	$2\sigma_H^2 + \sigma_L^2$	$\lambda^2(1-\lambda)$
σ_4	σ_H^2	$\sigma_H^2 + \sigma_L^2$	$\sigma_H^2 + 2\sigma_L^2$	$\lambda(1-\lambda)^2$
σ_5	σ_L^2	$\sigma_H^2 + \sigma_L^2$	$2\sigma_H^2 + \sigma_L^2$	$\lambda^2(1-\lambda)$
σ_6	σ_L^2	$\sigma_H^2 + \sigma_L^2$	$\sigma_H^2 + 2\sigma_L^2$	$\lambda(1-\lambda)^2$
σ_7	σ_L^2	$2\sigma_L^2$	$\sigma_H^2 + 2\sigma_L^2$	$\lambda(1-\lambda)^2$
σ_8	σ_L^2	$2\sigma_L^2$	$3\sigma_L^2$	$(1-\lambda)^3$

law are obtained by summing the λ_i in the right hand column of Table 1 that are rel-
evant to each variance.

In this way we see that Assumption 2 implies that the mixing law at maturity $n\Delta t$
is the multinomial density. Again, in the case $d = 2$ we have well-known binomial
density $B(n, \lambda)$: if we denote by $v_j(n\Delta t)$ the $n + 1$ distinct variances in the density
at time $n\Delta t$, then we have, for $j = 1, 2, \ldots, n + 1$,

$$v_j(n\Delta t; \sigma_H, \sigma_L) = (n - j + 1)\sigma_H^2 + (j - 1)\sigma_L^2 \qquad (5)$$

with corresponding weight in the normal mixture

$$w_j(n\Delta t; \lambda) = (n!/((j - 1)!(n - j + 1)!))\lambda^{n-j+1}(1 - \lambda)^{j-1}. \qquad (6)$$

The binomial NMD option price for an option of maturity $n\Delta t$ with strike $K_{n,s}$ is therefore:

$$\text{NM}(x, K_{n,s}, n\Delta t; \lambda, \sigma_H, \sigma_L) = \sum_{j=1}^{n+1} w_j(n\Delta t; \lambda)\text{BS}(x, K_{n,s}, n\Delta t, \eta_i(n\Delta t; \sigma_H, \sigma_L)) \qquad (7)$$

where $\eta_j(n\Delta t; \sigma_H, \sigma_L) = \sqrt{[v_j(n\Delta t; \sigma_H, \sigma_L)/n]}$, and $v_j(n\Delta t; \sigma_H, \sigma_L)$ and $w_j(n\Delta t; \lambda)$ are given by (5) and (6).

In general, Assumption 2 implies that the mixing law, which determines the effective weight on each distinct normal density at time $n\Delta t$ is the multinomial density, from the expansion of $(\theta_1 + \theta_2 + \cdots + \theta_d)^n$. Again the number of different normal densities in the mixture will increase with maturity, giving the short-term smile effect that we seek.

The multinomial NMD parameterization has no volatility term structure. This is easy to verify, for example – ignoring terms of order Δt^2 and $(\sigma^2)^2$ – in the trinomial model the average variance at Δt is

$$\theta_1\sigma_H^2 + \theta_2\sigma_M^2 + \theta_3\sigma_L^2$$

which is the same as the average variance at $2\Delta t$:

$$\theta_1^2\sigma_H^2 + \theta_2^2\sigma_M^2 + \theta_3^2\sigma_L^2 + \theta_1\theta_2(\sigma_H^2 + \sigma_M^2) + \theta_1\theta_3(\sigma_H^2 + \sigma_L^2) + \theta_2\theta_3(\sigma_M^2 + \sigma_L^2).$$

Moreover, the model smile will always have a minimum at the ATM forward volatility, because the implemented model is a specification of the lognormal mixture dynamics of Brigo and Mercurio where all drifts are the same. The parameterization of the NMD described here therefore captures a pure smile effect, with no skew and no volatility term structure. [15]

5. Application to currency options

In this first example of calibration, the binomial NMD parameterization has been applied to the exchange traded European Euro–US dollar options of all available strikes and maturities on 28th June 2002. Thirty-three prices at three different maturities (of 2, 6 and 11 weeks) were available at eleven strikes (between 90 and 110). The spot rate was 99.07, the interbank rate for the Euro was taken as 3.86% and that

---

[15] For a volatility term structure, a time variation into the possible values $[s_1, \ldots, s_d]$ of the volatility processes could be introduced. For a skew, following Brigo and Mercurio (2001a), a displaced diffusion model could be assumed.

for the US dollar was 1.71%. [16] We chose to minimize the following objective function: [17]

$$\sum_{t=1}^{k} \sum_{i=1}^{11} \gamma_{i,t} (\sigma_{nm,i,t} - \sigma_{m,i,t})^2, \tag{8}$$

where the gamma $\gamma_{i,t}$ denotes the gamma of an option with the ith strike and tth maturity, $\sigma_{nm,i,t}$ denotes the BS implied volatility of the normal mixture model price, and $\sigma_{m,i,t}$ denotes the BS implied volatility of the market price of the option. In the following we take $\Delta t = 1$ week and $k =$ either 1 or 3. That is, we shall use (8) to calibrate to each single smile at a fixed maturity, as well as to the whole smile surface.

The optimization problem belongs to the category of non-linear, multi-dimensional constrained minimization. Such problems are difficult because of the need to keep the solution within a boundary which is determined by constraints. Several algorithms were tested which were either too slow, or lacked robustness. The fastest and most reliable algorithm that has been tested with a variety of data sets appears to be the Downhill Simplex Method. [18]

The objective function (8) will not be well-defined unless it has bounded variance, and for this we need to ensure the existence of the fourth moment of the log price density. Asymptotically valid tests for the existence of fourth moments are well known, and may be applied to historical returns distributions. Hill (1975) suggests defining a test statistic based the maximal moment exponent $a = \sup\{b > 0 : E|\varepsilon_t|^b < \infty\}$, as follows: let

$$\hat{a}_s = \left(s^{-1} \left(\sum_{j=0}^{s-1} \ln \varepsilon_{n-j} \right) - \ln \varepsilon_{n-s} \right)^{-1}, \tag{9}$$

where s is a positive integer and $\epsilon_1 \leqslant \cdots \leqslant \epsilon_n$ are the ordered historical returns. Hall (1982) showed that if n is large enough and s/n is small enough and the tails of the distribution have the asymptotic Pareto–Lévy form then $\sqrt{s}(\hat{a}_s - a)$ has a normal

---

[16] Note that we have assumed constant interest rates at this point.

[17] The weighting of the squared volatility difference by gamma has the effect of giving more weight to more certain option prices. The gamma is greatest for at-the-money short-dated options, and decreases with both maturity and moneyness, but of course, other calibration objectives could be used. In fact, if the object is to fit the smile as closely as possible, irrespective of the reliability of option prices, no weighting in the objective function may be preferred. An unweighted objective has been used, for example, for the results in Table 5.

[18] Powell's Set Direction Method transforms the problem to one dimension and then applies successive line minimizations to find local minima. The problem with this method is that it is extremely slow. Moreover the method requires as inputs a set of directions where it will search for the minima. If the set of directions defined is not a "longsighted" one, then the algorithm does not converge. The Broyden–Fletcher–Goldfarb–Schanno (BFGS) algorithm is a variable metric or "quasi-Newton" method that calculates the partial derivatives with respect to all the variables, looks for the "steepest" gradient and then employs a line minimization to that direction. This method very often gave infeasible solutions, i.e. λ greater than 1 or negative volatilities.

Table 2
Tests for the existence of the fourth moment

Exchange rate			EUR
5%	Right tail	$\hat{a}_s$	4.45
		4th moment test	1.02
	Left tail	$\hat{a}_s$	3.69
		4th moment test	−0.73
1%	Right tail	$\hat{a}_s$	5.85
		4th moment test	1.91
	Left tail	$\hat{a}_s$	4.23
		4th moment test	0.24

Note: s is defined as the percentage of number of the observations in the right (left) tail. Tests are based on 14-year period of daily close exchange rates from 2nd January 1989 to 31st December 2002 (a total of 3652 observations).

distribution with standard deviation a. Table 2 presents the estimates for the maximal moment exponent and the test results, based on daily data from January 1989 to December 2002, for the existence of the fourth moment. According to these results, and similar results on exchange rate distributions – see for example, Huisman et al. (2001, 2002), whose results support the existence of the fourth moment for all major US dollar exchange rates, based on daily data from 1979 to 1996 – we cannot reject the existence of the fourth moment.

Now consider the calibration of σ_H, σ_L and λ to all options of a fixed maturity, first calibrating a mixture of 3 normal densities to the smile at 2 weeks, then calibrating a mixture of 7 normal densities to the smile at 6 weeks, and finally calibrating a mixture of 12 normal densities to the smile at 11 weeks. The results are shown in Table 3. We see that the log price density calibrated to the 6 week smile is not the same as the density that is inferred at 6 weeks from calibration to the 2-week smile. The same remark applies to the log price density at 11 weeks. When longer term smiles are inferred from the 2-week smile parameters, the excess kurtosis in the log price densities decreases with maturity: this is exactly the short term smile effect that, by the central limit theorem, decreases fairly rapidly with maturity. The excess kurtosis that can be attributed to the short-term smile effect (that is, the excess kurtosis inferred from the 2-week price density parameters) is 1.50 at 6 weeks and 0.82 at 11 weeks. By the 11 week maturity (with 12 normal densities in the mixture) the price density inferred from the 2-week smile parameters is very near to normal.

On the other hand, the excess kurtosis in the smile, including the longer term smile effects, may be estimated by calibrating parameters directly on the individual smiles. Although a model is used, the fitted smile follows the observed market implied volatilities closely (goodness-fit results will be given in the next section). Thus it seems reasonable to infer that the model implied excess kurtosis is an accurate estimate of the total excess kurtosis in the price density. From Table 3, these estimates are 4.28 at 6 weeks and 4.05 at 11 weeks. We conclude that only a small part of the smile at longer maturities can be explained by the binomial NMD parameterization.

Table 3
Comparison of log price densities inferred from two-week fitted smile and fitted to each smile

	2 weeks		6 weeks				11 weeks			
λ	7.89%		6.99%		7.89%		5.66%		7.89%	
σ_H	41.24%		58.02%		41.24%		68.62%		41.24%	
σ_L	11.01%		7.56%		11.01%		3.99%		11.01%	
	Weight	Variance	Weight	Variance	Weight	Variance	Weight	Variance	Weight	Variance
$w1$	0.006225	0.170058	0.000000	0.336686	0.000000	0.170058	0.000000	0.470854	0.000000	0.170058
$w2$	0.145353	0.091088	0.000009	0.281525	0.000017	0.143734	0.000000	0.428194	0.000000	0.155700
$w3$	0.848422	0.012118	0.000309	0.226364	0.000493	0.117411	0.000000	0.385534	0.000000	0.141341
$w4$			0.005486	0.171203	0.007677	0.091088	0.000000	0.342873	0.000000	0.126983
$w5$			0.054787	0.116042	0.067218	0.064764	0.000000	0.300213	0.000005	0.112625
$w6$			0.291809	0.060881	0.313883	0.038441	0.000011	0.257553	0.000074	0.098267
$w7$			0.647599	0.005720	0.610711	0.012118	0.000189	0.214893	0.000863	0.083909
$w8$							0.002254	0.172233	0.007195	0.069550
$w9$							0.018778	0.129573	0.041994	0.055192
$w10$							0.104315	0.086913	0.163414	0.040834
$w11$							0.347690	0.044253	0.381539	0.026476
$w12$							0.526763	0.001593	0.404917	0.012118
Annual vol (%)	15.68		16.98		15.68		16.78		15.68	
XS kurtosis	4.50		4.28		1.50		4.05		0.82	

6. Modelling longer term smile effects with uncertain volatility

In practical terms, the multinomial NMD parameterization is limited because the model and market prices of longer term OTM options can be very different. In an attempt to improve this fit, we now introduce uncertainty about which local volatility surface to apply to the price process. In this case, the price density may no longer converge to a normal density as the maturity increases.

In this section, volatility uncertainty is introduced to the binomial NMD parameterization by making λ stochastic. We assume that λ is a Bernoulli variable, independent of the Brownian motion driving X, with probability p on λ_H and probability $(1 - p)$ on λ_L. Thus at time $t = 0$, the trader perceives two possible local volatility surfaces; a high volatility surface with probability p and a low volatility surface with probability $(1 - p)$. The Bernoulli distribution over the two local volatility surfaces may be interpreted as a representation of the traders' state of knowledge, at any point in time. As this additional uncertainty introduces a market incompleteness, there is no unique 'risk neutral' option price. However, for liquid options the price differences arising from differences in risk premia should be small, so market prices should be close to the 'risk neutral' option prices, which in this model are given by

$$p\mathrm{NM}(x, K_{n,s}, n\Delta t; \lambda_H, \sigma_H, \sigma_L) + (1 - p)\mathrm{NM}(x, K_{n,s}, n\Delta t; \lambda_L, \sigma_H, \sigma_L)$$

with $\mathrm{NM}(x, K_{n,s}, n\Delta t; \lambda, \sigma_H, \sigma_L)$ given by (5)–(7). As before, the model calibration will equate model prices to the observed market prices, using the favoured calibration objective. The option sensitivities can be taken as simple weighted averages of the BS sensitivities, but it is important to note that there will be residual hedging uncertainty because of the market incompleteness arising from the uncertainty surrounding the local volatility surface.

Table 4 reports the results of calibrating the stochastic λ model simultaneously to all 33 currency option implied volatilities observed on 28th June 2002. At each maturity the log price densities are a weighted average of two binomial NMD log price densities, one fitted with λ_H, σ_H, and σ_L and the other fitted with λ_L, σ_H, and σ_L.

Table 4
Stochastic λ model parameter estimates

Parameters		Maturity (weeks)	λ	Deterministic λ models		Stochastic λ model	
				Volatility	XS kurtosis	Volatility	XS kurtosis
p	28.12%	2	84.71%	26.87%	0.22		
σ_H	28.98%	2	2.45%	9.25%	2.92		4.19
σ_L	8.17%	6	84.71%	26.87%	0.07	$\sqrt{E}$ (variance) $= 16.27\%$	
λ_H	84.71%	6	2.45%	9.25%	0.97		3.74
λ_L	2.45%	11	84.71%	26.87%	0.04	E (volatility) $= 14.21\%$	
		11	2.45%	9.25%	0.53		3.63

These two log price densities have quite different volatilities. For example, at 2 weeks the annual volatility of the density calibrated with λ_H is 26.87%, and the annual volatility of the density calibrated with λ_L is 9.25%. The difference in standard deviations of these densities, which becomes more pronounced with maturity, implies that when the two densities are mixed in the log price density there is considerable excess kurtosis. Even at the longer maturities where the excess kurtosis in each of the binomial NMD parameterizations has virtually disappeared (so we are taking a mixture of two almost normal densities), because of the difference in their standard deviations, there is still a substantial excess kurtosis in the mixture log price density. The excess kurtosis estimated from the stochastic λ model compares well with the estimated excess kurtosis from the smile that was evaluated in Table 3 by fitting a deterministic λ model to each maturity separately. Table 4 gives an estimated excess kurtosis with the stochastic λ binomial NMD parameterization of 3.74 at 6 weeks and 3.63 at 11 weeks.

Fig. 2 compares the market smiles and the model smiles that are fitted with the calibrated stochastic λ model. Clearly the model does not explain all the observed excess kurtosis in this smile – stochastic interest rates, liquidity, hedging costs and market imperfections also affect currency option prices at longer maturities – but the entire fitted smile surface is based on only five parameters, and should therefore be relatively stable over time.

To investigate the parameter stability of the stochastic λ model, it has been calibrated to European currency option smiles for the Euro–US dollar exchange rate on three consecutive Fridays: 23rd May, 30th May, and 6th June 2003. [19] The four option maturities were for the June, July, August and September expires. Compared with the previous data the option volatilities are lower: for example, the ATM volatility was approximately 10% but it was about 14% in June 2002 (see Fig. 2). There is also a pronounced skew in the implied volatilities for the June option. This is probably a result of the steep rise in the exchange rate during the first five months of 2003. The lowest strike options have very little liquidity and the Bloomberg call option prices are almost definitely stale quotes, left over from a time when the market was near to that level. [20] The exceptionally high price for the June ITM options imparts a very significant skew, which is not apparent in the July, August and September options of this strike. Including the June options in the calibration data leads to extremely high model kurtosis at not only the June maturities, but also at subsequent maturities. Since even small errors in the ITM call prices would lead to large errors in the implied volatility, the very near term implied volatility data were not used in the calibration. A better test of the parameter stability and goodness of fit of the model is to calibrate it to the July, August and September smiles. Seven strikes at three

---

[19] The contemporaneous spot rate was 117, 117.8 and 118.4 respectively, the inter-bank rate for the Euro was taken as 2.5% and that for the US dollar was 1.31%. Only the Ask option prices (obtained from Bloomberg) were used because there were very few Bid prices available on these days.

[20] Indeed, there were no put prices for the June ITM options, since their value was zero; similarly the call prices for the far ITM options were unreliable.
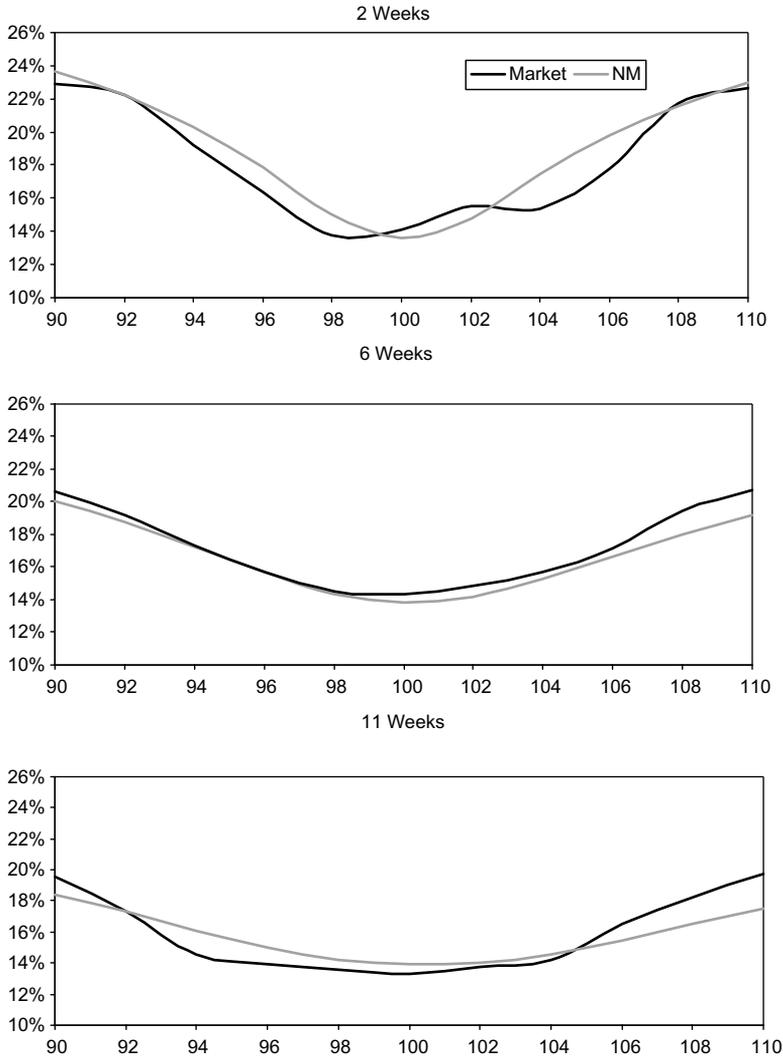
Fig. 2. Market smiles and fitted smiles with the stochastic λ model.

maturities were used to calibrate the model on each of the three Fridays in question and the results – including extrapolated results for the June maturity – are summarized in Table 5.

The first rows of Table 5 give the calibrated parameter values: σ_H is a little over 30% and σ_L is a about 4% – both are growing, though not rapidly, as the date of the samples progresses; λ_H and λ_L are also fairly stable, and it should be noted that their difference is far less pronounced than it was one year earlier, in June 2002. In the earlier calibration of the smile model (in Table 4) we calibrated $\lambda_H = 84.71\%$ and

Table 5
Parameter stability and goodness-of-fit

Parameter estimates	23rd May	30th May	6th June
σ_H	31.42%	32.61%	33.12%
σ_L	3.96%	3.55%	4.21%
λ_H	13.62%	14.92%	11.82%
λ_L	11.19%	10.72%	11.35%
P	39.35%	47.45%	70.23%
High volatility	12.17%	13.01%	12.05%
Low volatility	11.15%	11.19%	11.84%
June option			
Maturity (weeks)	4	3	2
XS kurtosis	3.9	5.63	8.72
July option			
Maturity (weeks)	8	7	6
XS kurtosis	2.13	2.52	2.91
August option			
Maturity (weeks)	12	11	10
XS kurtosis	1.43	1.63	1.74
September option			
Maturity (weeks)	17	16	15
XS kurtosis	1.02	1.14	1.16
RMSE (%)	0.7847	0.5193	0.6456

$\lambda_L = 2.45\%$ but now (in Table 5) the calibrated λ's have more similar values. This is particularly noticeable on 6th June 2003, where the two λ's are almost identical: 11.82% and 11.35% respectively. Consequently, there is only a small difference between the average volatilities in the two normal log price densities – these are given in the rows labeled 'high volatility' and 'low volatility' respectively – so the modelled smile effect does not persist into longer maturities as much as it did when the model was calibrated to the 2002 data. In fact, the model excess kurtosis, given for the different options in the lower rows of Table 5, drops from a high of 8.72 for the June option on 6th June, to a low of 1.02 for the September option on 23rd May. Clearly there is only a small smile effect in the September option.

An important point to note about the results of Table 5 is that the diminishing smile effect in the calibrated model is not necessarily due to a poor fit to the market data (the goodness-of fit is examined below). Although this was shown to be the case for the simple binomial NMD parameterization calibrated on the June 2002 data (see Table 3), in June 2003 the smile in the market implied volatilities was also diminishing with maturity – far more than it did in 2002. The stochastic λ model calibrates similar values for both λ_H and λ_L when this best describes the current data, that is, when there is relatively little uncertainty about the exchange rate volatility. It calibrates quite different values for λ_H and λ_L when there is more uncertainty about the exchange rate volatility, which is when the smile effect persists into longer

maturities. It is therefore reasonable to suppose that there was considerably more uncertainty about the Euro–US dollar exchange rate volatility in June 2002 than there was in June 2003. The Euro had been stable and upwards trending during the interim between the two periods, and the dollar exchange rate volatility was lower and more stable during the year leading up to June 2003 than during the year leading up to June 2002. [21]

Another significant factor that contributes to currency option smiles at longer maturities is the uncertainty in interest rates. The stochastic λ binomial NMD parameterization may use volatility uncertainty as a proxy for interest rate uncertainty. For the Euro–US dollar exchange rate, it is likely that there was less uncertainty in interest rates in 2003 than there was in 2002. Indeed absolute volatility in interest rates has dropped significantly, given that both Euro and US dollar interest rates were falling continually since the year 2000 until June 2003. Thus the similarity of the estimates for the two values of λ in June 2003 may be reflecting the reduction in interest rate uncertainty, as well as the reduction in volatility uncertainty.

To investigate the goodness-of-fit of the model, the last row of Table 5 reports the root mean square error (RMSE) between the fitted smile surface and the market smile surface – including the June maturity – for each of the sample dates. The RMSE can be regarded as a measure of the average error in volatility over the whole surface, and it is well below 1%. The errors are smallest for ATM options so most of the contribution to the error comes from the less liquid options. For simple ATM options, which are approximately linear in volatility, an error on volatility translates directly to a pricing error. The average errors would therefore translate into between 0.5% and 1% error on simple ATM option prices, and this is well within the bid–ask spread for ATM options on liquid instruments (which is normally about 4%). Due to their convexity as a function of volatility, and the slightly larger volatility errors, there will be higher pricing errors for ITM and OTM options – but then, the bid–ask spread is also greater.

We may conclude that the stochastic λ model provides a reasonable fit to the entire smile surface for currency options, even though it is based on only five parameters. Moreover, the calibrated parameter values are robust and reasonably stable over time, their values reflecting changes in the market conditions. In particular, the difference between the two λ values will capture the degree of uncertainty in volatility, and this is reflected in the persistence in the market smile through to longer maturities.

7. Summary and conclusions

This paper began by placing additional structure on the NMD of Brigo and Mercurio (2000, 2001a,b, 2002) and Brigo et al. (2002) by assuming the finite number of

---

[21] GARCH(1,1) volatility models give short term Euro–US dollar volatility forecasts of approximately: 12% (in December 2001); 8% (in June 2002); 6% (in December 2002); and 6% (in June 2003).

deterministic volatility processes are piecewise constant over time and are limited to a small number of values. The simplest parameterization, with only two volatility values, has a mixing law at time $n\Delta t$ that is binomial $B(n, \lambda)$. More generally, with several possible states for volatility in each Δt, the mixing law was shown to be multinomial. Only short-term smile effects are explained by the multinomial NMD parameterization, and since marked smile effects that do not decrease with maturity are often observed in implied volatilities, in order to model a longer term smile effect within the same general framework we have assumed that traders have uncertainty about the volatility surface.

A binomial NMD parameterization has been implemented in this paper, where the mixing law coefficient λ is (a) fixed and (b) stochastic with λ being a simple Bernoulli variable. In case (a) the very parsimonious model has just three parameters but it only explains a short-term smile effect; in case (b) the five-parameter explains the persistence of smile effects into longer maturities. We have calibrated the fixed and stochastic λ binomial NMD parameterizations to a market smile for the Euro–US dollar exchange rate in June 2002, and the stochastic λ binomial NMD parameterization to three consecutive smiles in May/June 2003. The calibrated models for June 2002 point to considerable uncertainty in exchange rate volatility, which is only being captured by the stochastic λ model. The three consecutive calibrated models in May/June 2003 indicate a substantial reduction in volatility uncertainty in the market one year later, because the calibrated values of the two λ's are similar on all three dates in 2003. The calibrated parameters are quite stable over time and the fitted smile surfaces, which are based only five parameters, have root mean square errors which are less than 1% in each case. This translates to pricing errors that are well within the bid–ask spread.

With the stochastic λ binomial NMD parameterization the fit can be close but, since it has only five parameters, it will not provide an exact fit to market prices. Moreover, the model has two obvious limitations. Firstly, unless the values of the volatility processes are allowed to vary over time, there will be no volatility term structure in the model. Secondly, since the model framework is one of lognormal mixture diffusions with equal drifts, the minimum of the smile must always occur at the ATM forward volatility. Thus it will not be suitable for highly skewed structures, for example, where there are large risk reversals in the foreign exchange market.

Of course, local volatility models can fit market prices exactly but, in contrast to much other research on local volatility, this paper has focused more on the interpretation of local volatility than the accuracy of option pricing. Econometric 'stylised facts' – and the central limit theorem – suggest that a term structure in excess kurtosis is a feature of deterministic volatility models. Perhaps, with increasingly elaborate parameterizations that are designed to fit prices perfectly, we have been missing the essence of the local volatility framework? Just as a term structure in excess kurtosis is always evident in ex-post returns, as distributions converge to the normal distribution when the sampling frequency decreases (and in accordance with the central limit theorem) so the log price density always converges to a normal density as the maturity increases, under the multinomial NMD parameterization. This

local volatility model has been designed to capture only the short-term smile effect that is associated with decreasing leptokurtosis in the price density with term. The results in this paper show that its extension to stochastic volatility, when it is natural to use the uncertain volatility framework introduced by Avellaneda et al. (1995), can capture the more persistent smile effects.

References

Andersen, L., Brotherton-Radcliffe, R., 1997. The equity option volatility smile: An implicit finite difference approach. Journal of Computational Finance 1, 5–37.

Avellaneda, M., Friedman, C., Holmes, R., Samperi, D., 1997. Calibrating volatility surfaces via relative entropy minimization. Applied Mathematical Finance 4, 37–64.

Avellaneda, M., Levy, A., Paras, A., 1995. Pricing and hedging derivative securities in markets with uncertain volatility. Applied Mathematical Finance 2, 73–88.

Baillie, R.T., Bollerslev, T., 1989. The message in daily exchange rates: A conditional-variance tale. Journal of Business and Economic Statistics 7 (3), 297–305.

Baillie, R.T., Bollerslev, T., 1990. Intra-day and inter-market volatility in foreign exchange rates. Review of Economic Studies 58, 565–585.

Ball, C.A., Torous, W.N., 1983. A simplified jump process for common stock returns. Journal of Financial and Quantitative Analysis 18 (1), 53–65.

Beaglehole, D., Chebanier, A., 2002. Mean reversion with a smile. Risk 15 (4), 95–98.

Bingham, N.H., Kiesel, R., 2002. Semi-parametric modelling in finance: Theoretical foundations. Quantitative Finance 2 (4), 241–250.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Bollerslev, T., 1986. Generalised autoregressive conditional heteroskedasticy. Journal of Econometrics 31, 307–327.

Bollerslev, T., 1987. A conditional heteroskedasticity time series model for security prices and rates of return data. Review of Economics and Statistics 69, 542–547.

Bollerslev, T., Chou, R., Kroner, K., 1992. ARCH modelling in finance. Journal of Econometrics 52, 5–59.

Bollerslev, T., Engle, R.F., Nelson, D., 1994. ARCH models. In: Engle, R.F., McFadden, D.L. (Eds.), Handbook of Econometrics, vol. 1. North-Holland, Amsterdam.

Bouchouev, I., Isakov, V., 1997. The inverse problem of option pricing. Inverse Problems 13, 11–17.

Bouchouev, I., Isakov, V., 1999. Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. Inverse Problems 15, 95–116.

Breeden, D.T., Litzenberger, R.H., 1978. Prices of state contingent claims implicit in option prices. Journal of Business 51, 621–651.

Brigo, D., 2002. The general mixture-diffusion SDE and its relationship with an uncertain-volatility option model with volatility–asset decorrelation. Available from <www.damianobrigo.it>.

Brigo, D., Mercurio, F., 2000. A mixed-up smile. Risk 13 (9), 123–126.

Brigo, D., Mercurio, F., 2001a. Displaced and mixture diffusions for analytically-tractable smile models. In: Geman, H., Madan, D.B., Pliska, S.R., Vorst, A.C.F. (Eds.), Mathematical Finance, Bachelier Congress 2000. Springer, Berlin.

Brigo, D., Mercurio, F., 2001b. Interest-Rate Models: Theory and Practice. Springer-Verlag, Berlin.

Brigo, D., Mercurio, F., 2002. Lognormal-mixture dynamics and calibration to market volatility smiles. International Journal of Theoretical and Applied Finance 5 (4), 427–446, Extended version with F. Rapisarda including FX surface calibration. Available from <www.damianobrigo.it>.

Brigo, D., Mercurio, F., Sartorelli, G., 2002. Alternative asset price dynamics and volatility smiles. Banca IMI report available from <www.damianobrigo.it>.

Brown, G., Randall, C., 1999. If the skew fits. Risk 12 (4), 62–65.

Cai, J., 1994. A markov model of switching-regime ARCH. Journal of Business and Economic Statistics 12, 309–316.

Coleman, T.F., Yuying L., Verma, A. 1998. Reconstructing the Unknown Local Volatility Function. In: Proceedings of Computational and Quantitative Finance 98, New York.

Cox, J.C., Ross, S.A., 1976. The valuation of options for alternative stochastic processes. Journal of Financial Economics 3, 145–166.

Dacorogna, M., Muller, U., Olsen, R., Pictet, O., 1998. Modelling short-term volatility with GARCH and HARCH models. In: Dunis, C., Zhou, B. (Eds.), Non-Linear Modelling of High Frequency Financial Time Series. John Wiley, Chichester.

Derman, E., Kani, I., Chriss, N., 1996. Implied trinomial trees of the volatility smile. Journal of Derivatives 3 (4), 7–22.

Dumas, B., Fleming, F., Whaley, R., 1998. Implied volatility functions: Empirical tests. Journal of Finance 53 (6), 2059–2106.

Dupire, B., 1994. Pricing with a smile. Risk 7 (1), 18–20.

Dupire, B., 1997. Pricing and Hedging with Smiles. In: Dempster, M.A.H., Pliska, S.R. (Eds.), Mathematics of Derivatives Securities. Cambridge University Press, Cambridge, pp. 103–111.

Epps, T.W., Epps, M.L., 1976. The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis. Econometrica 44 (2), 305–321.

Gencay, R., Dacorogna, M., Muller, U., Olsen, R., Pictet, O., 2001. An Introduction to High Frequency Finance Theory and Applications. Academic Press, New York.

Goodhart, C.A.E., O'Hara, M., 1997. High frequency data in financial markets: Issues and applications. Journal of Empirical Finance 4, 73–114.

Gray, S.F., 1996. Modeling the conditional distribution of interest rates as a regime-switching process. Journal of Financial Economics 42, 27–62.

Hall, P., 1982. On some simple estimates of an exponent of regular variation. Journal of the Royal Statistical Society, Series B 44 (1), 37–42.

Hamilton, J.D., Susmel, R., 1994. Autoregressive conditional heteroscedasticity and changes in regime. Journal of Econometrics 64, 307–333.

Heston, S., 1993. A closed form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6, 327–343.

Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. The Annals of Statistics 3 (5), 1163–1174.

Huisman, R., Koedijk, K., Kool, C., Palm, F., 2001. Tail-index estimator in small sample. Journal of Business and Economic Statistics 19, 208–216.

Huisman, R., Koedijk, K., Kool, C., Palm, F., 2002. The tail-fatness of FX returns reconsidered. De Economist 150, 299–312.

Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. Journal of Finance 42, 281–300.

Klaassen, F., 1998. Improving GARCH volatility forecasts. SSRN Discussion Paper No. 1998-52.

McFarland, J.W., Petit, R.R., Sung, S.K., 1982. The distribution of foreign exchange price changes: Trading day effects and risk measurement. Journal of Finance 37 (3), 693–715.

McIntyre, M., 2001. Performance of dupires implied diffusion approach under sparse and incomplete data. Journal of Computational Finance 4 (4), 33–84.

Melick, W.R., Thomas, C.P., 1997. Recovering an assets implied pdf from option prices: An application to crude oil during the gulf crisis. Journal of Financial and Quantitative Analysis 32, 91–115.

Merton, R., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125–144.

Ritchey, R.J., 1990. Call option valuation for discrete normal mixtures. Journal of Financial Research 13, 285–296.

Rubinstein, M., 1994. Implied binomial trees. Journal of Finance 49, 771–818.