Quantitative Finance, Vol. 6, No. 2, April 2006, 95–105

Routledge
Taylor & Francis Group

A simple approach for pricing equity options with Markov switching state variables

DONALD D. AINGWORTH†, SANJIV R. DAS*‡
and RAJEEV MOTWANI†

†Department of Computer Science, Gates Computer Science Building 4B,
Stanford University, Stanford, CA 94305-9045, USA
‡Dept of Finance, 208 Kenna Hall, Santa Clara University,
Santa Clara, CA 95053-0388, USA

(*Received 23 August 2004; in final form 7 December 2005*)

1. Introduction

The pricing of American options on stocks was rendered computable by the work of Cox *et al.* (1979) (CRR) and the ensuing work of Jarrow and Rudd (1983). They developed a convergent option pricing scheme in a Black–Merton–Scholes world (Black and Scholes 1973, Merton 1973) on a discrete-time lattice. In this setting stock prices $[S(t)]$ are assumed to follow a geometric Brownian motion, i.e.

$$\mathrm{d}S(t) = \mu(t)S(t)\mathrm{d}t + \sigma(t)S(t)\mathrm{d}Z(t)\,, \quad S(0) \equiv S_0. \tag{1}$$

The drift of the stock is a deterministic function $\mu(t)$ and the variance coefficient $\sigma(t)$ is also deterministic. $\mathrm{d}Z(t)$ represents a standard Wiener increment: $\mathrm{d}Z \sim N[0, \mathrm{d}t]$.

The CRR solution discretely approximates the stochastic process above on a binomial tree, using a time interval Δ. This scheme has at least two desirable properties. First, the binomial approximation to the Ito process in equation (1) renders the required terminal distribution when the discrete interval Δ becomes small. Second, the use of the geometric Brownian motion for stocks when $\sigma(t)$ is constant results in a recombining tree, making computation possible in polynomial time. Without recombination, the problem would require exponential time. Therefore the CRR model, which assumes $\sigma(t)$ to be constant, is the simplest model available for pricing American equity options. At the other end of the spectrum is the case when $\sigma(t)$ is fully stochastic, and has been solved for European options using Fourier transform methods by Heston (1993). However, the Fourier inversion is often computationally difficult and the model does not permit solutions for American options. In this paper, we are able to bridge these models at different ends of the volatility spectrum, by developing a computationally efficient algorithm that (a) has a polynomial $O(kn^{2k})$ run time, (b) can price options on stochastic volatility, (c) allows for the pricing of American options and (d) converges rapidly, thus offering one facile solution to this class of problems.

To see why recombination is achieved in the basic CRR model, examination of the integral solution for equation (1) is instructive. With constant coefficients, we have in time interval Δ:

$$S(t+\Delta) = S(t)\exp\left\{\left(\mu - \frac{V}{2}\right)\Delta + \sigma\epsilon\sqrt{\Delta}\right\}, \quad \epsilon \sim N(0,1), \tag{2}$$

where $V(t) = \sigma^2(t)$. On the binomial tree, the normal distribution for ϵ is approximated by allowing it to take values in the set $\{-1, +1\}$. In the CRR model we have

$$S(t+\Delta) = S(t)\exp\left\{\sigma\epsilon\sqrt{\Delta}\right\}, \quad \epsilon \sim N(0,1). \tag{3}$$

It is easily verified that after two periods (i.e. 2Δ), the stock price $S(t+2\Delta)$ would be the same if the shock ϵ were first positive and then negative, or vice versa, i.e. $S(t+2\Delta) = S(t)\exp[(\mu-(V/2))\Delta]$. Therefore, after n periods, the number of terminal nodes would only

*Corresponding author. Email: srdas@scu.edu

Quantitative Finance
ISSN 1469–7688 print/ISSN 1469–7696 online
http://www.tandf.co.uk/journals
DOI: 10.1080/14697680500511215

be $(n+1)$ and not 2^n as would be the case with a non-recombining scheme for the lattice. This feature of the binomial tree approach of CRR is usually unavailable when extensions of the Black–Merton–Scholes (BMS) model are undertaken. Notice that when $\sigma(t)$ is not constant, the argument outlined in the previous paragraph fails and recombination of the tree is lost.

A lattice approach such as ours allows dynamic programming implementation and hence admits the pricing of American options. We analyse the case when volatility follows a Markov-switching process. So far, there are some approaches to dealing with the problem. For one, Brigo *et al.* (2004) develop a simple formula for option prices when stock prices are modelled as a mixture of lognormals. Longstaff and Schwartz (2001) provide a simulation model in which American options may be valued quite easily; their approach would be applicable to the problem we address in this paper. The lattice model in this paper provides one method to price such options in polynomial time. It also admits the pricing of American options, where optimal stopping decisions are required.

A recent set of papers has adopted a new approach now known as the 'normal mixture diffusion' (NMD) model. Excellent expositions of this approach may be obtained in Alexander (2004), Mercurio (2002), Brigo and Mercurio (2002) and Brigo *et al.* (2004). In the NMD class of models, volatility is extended from the Black–Scholes setting to being random, i.e. drawn from a distribution of possible values but not going so far as to make volatility follow a stochastic process as in the model of Hull and White (1987) or Heston (1993). Tractable versions are obtained by limiting these values to a finite set occurring with mixing probability. This model readily develops a short-term smile. Further, as in Alexander (2004), by allowing the mixing probabilities to be stochastic, the model also admits a long-term smile. Our paper is closest in spirit to these models. We allow volatility to obey a special type of stochastic process, a regime-switching one. This is actually similar to the NMD model with stochastic probability, except in our model, the switching probabilities are state dependent and are different depending on which volatility regime the system is presently occupying. Hence, we may think of the model of this paper as one that intersects the realm of stochastic volatility and NMD models.

Extending the CRR algorithm to the case of switching volatility means that the recombination feature may be lost. No more does an up-move in ϵ followed by a down-move result in the same stock price as a down-move, then up-move in ϵ. This is because the volatility multiplying through the ϵ value varies at each point in time. In this paper, we develop an algorithm that resides on a complex, though recombining tree, resulting in a polynomial time algorithm for pricing options with Markov-switching volatility for even a large number of volatility states k.

In order to represent the joint process of $\{S, V\}$ over a time interval $[0, T]$ would require a lattice in $k+1$ dimensions (not counting the time dimension t). For example, if $k=2$, and each dimension were approximated using a binomial scheme, the entire state-space in S and V would require four branches emanating from each node on the tree (for (i) S up, V up, (ii) S up, V down, (iii) S down, V up, (iv) S down, V down). If this were not recombining, then the number of terminal nodes for n steps on the tree would be 4^n and the computing effort would grow exponentially in n.

This paper develops a computable approach to this class of problems. To the extent that the volatility stays within the same state the binomial process is still recombining. Also, a certain extent of recombination occurs in the BMS model, as a number of up and down moves in the same state will cancel each other out. Hence, Markov switching in volatility does offer a certain degree of recombination. Indeed, we show that the recombination attained is sufficient, even in the worst case, to deliver a polynomial algorithm instead of an exponential one. We will prove that the run time will be $O(kn^{2k})$, where k is the number of states in the volatility Markov chain.

There are many benefits to our computational method. First, we add to the available pricing algorithms for American options with stochastic volatility. In a recent paper, Duan *et al.* (1998) develop a model for option pricing with regime switches in the equity price. We provide a model for switching volatility, thereby offering a new lattice scheme for option pricing with stochastic volatility of a particular form. As the number of states increases, the switching model approaches stochastic volatility over a continuous state space. Second, stochastic volatility models are empirically justified (see Bates 1996), yet apart from the Heston (1993) closed form solution for a particular SV model, very few efficient pricing methods exist. The first paper in this class of models is that of Hilliard and Schwartz (1996). They show how to create a lattice for mean-reverting stochastic volatility when pricing equity options. Brigo and Mercurio (2002) present a model with a mixture of lognormal diffusions, interpreted as an approach to introducing stochastic volatility via switching processes. Bollen (1998) develops a pentanomial lattice similar in spirit to the model in this paper. He eschews a quadrinomial model such as ours since recombination is not attained as in our scheme. In addition, we are able to prove the polynomial complexity of our model. Leisen (2000) is able to show that a sequence of discrete-time models converges to the stochastic volatility case. Finally, our methods may be extended to jump-diffusions as well, by merging our Markov chain method with the techniques developed in the paper by Amin (1993). Finally, since there are several methods available for estimation of Markov chain systems (see Hamilton 1990), our model is easy to implement.

2. The Markov-switching framework

The stock price process in equation (1) will in general have a volatility process defined by k states. We denote these states $V_j, j = 1 \dots k$. The transition matrix for volatility switches (from state i to state j) is $\{p_{ij}\}, i, j = 1 \dots k$. Each probability may be written in functional form and then estimated. In general, we may write

$$\begin{aligned} p_{ij} &= f(\mathbf{x}, \Delta; \alpha), & i \neq j, \\ p_{ij} &= 1 - \sum_{j \neq i} f(\mathbf{x}, \Delta; \alpha), & i = j, \end{aligned}$$

where $\mathbf{x}$ is a vector of state variables and α is a vector of parameters for the probability function. These parameters are determined by estimation. As an illustrative example, we may employ the one-parameter logit function $p_{ij} = \exp(\alpha_{ij})/[1 + \exp(\alpha_{ij})]$. Or if we need the transition probabilities to be correlated with the stock price (S) we may estimate a model where $p_{ij} = \exp(\alpha_{ij} + \beta_{ij} S)/[1 + \exp(\alpha_{ij} + \beta_{ij} S)]$. In order to keep the exposition simple, without loss of generality, we restrict attention to a two-state model.

2.1. *A two-state model*

To fix ideas we explore a simplified version of the regime-switching volatility model in a two-state setting for volatility. Using the notation from the previous section, we write

$$\sqrt{V(t)} = \sigma(t) = \begin{cases} \sigma_{\mathrm{H}} & \text{(the high volatility state)}, \\ \sigma_{\mathrm{L}} & \text{(the low volatility state)}. \end{cases} \tag{4}$$

The transitions from one state to the other are driven by a transition probability matrix $P(\Delta)$ which is a function of the time interval between realizations of (S, V). We write the matrix as

$$P(\Delta) = \begin{bmatrix} p_{\mathrm{H}} & 1 - p_{\mathrm{H}} \\ 1 - p_{\mathrm{L}} & p_{\mathrm{L}} \end{bmatrix}.$$

Of course, the transition probability matrix estimated for a given Δ may be transformed for time intervals other than Δ using the generator matrix for the Markov chain, denoted G. As usual, $\exp(G\Delta) = P(\Delta)$. For small Δ, $P(\Delta) = I + G\Delta$, where I is the identity matrix. In general, $P(\Delta) = I + \sum_{n=1}^{\infty} \Delta^n G^n / n!$. The two-state Markov chain is representative of reality because stock volatility usually tends to stay at normal values for periods of time punctuated by moments of high variance. Furthermore, there is ample evidence of volatility persistence (see the ARCH and GARCH literature), which is also possible in the Markov-switching framework.

By increasing the number of states in the Markov chain, and choosing the transition matrix appropriately, we obtain a sequence of models that get closer to a stochastic volatility model, that is, if this is desired at all. Usually, a two- or three-state volatility process does quite well in capturing the essence of the pricing problem faced in the financial markets.

Equations (1), (2) and (4) may be discretized for embedding on the bivariate (S, V)-lattice. The equations are as follows:

$$S(t + \Delta) = S(t) \exp\left\{ \left(\mu - \frac{\sigma(t)^2}{2} \right) \Delta \pm \sigma(t) \sqrt{\Delta} \right\},$$

$$\sigma(t) = \begin{cases} \sigma_{\mathrm{H}}, \\ \sigma_{\mathrm{L}}. \end{cases}$$

If the volatility remained purely in one state or the other, we would obtain two binomial trees, one for the high volatility situation, σ_{H}, and the other for the low volatility situation, σ_{L}.

The discretization provided here to the stochastic differential equation (1) is known as the Jarrow and Rudd (1983) version of the CRR model. The basic CRR model posits a different discretization, which is as follows:

$$S(t + \Delta) = S(t) \exp\left\{ \pm \sigma(t) \sqrt{\Delta} \right\},$$

$$\sigma(t) = \begin{cases} \sigma_{\mathrm{H}}, \\ \sigma_{\mathrm{L}}. \end{cases}$$

Any version of the model is acceptable. Both forms converge to the continuous-time limit of the process. In the numerical experiments we provide, the CRR version of the model is used, as in some cases, we are able to achieve a further improvement in complexity. The complexity level $O(kn^{2k})$ (derived later) is an upper bound over all forms of lattice implementation for the model.

Since volatility is stochastic and switches between low and high states, our numerical algorithm needs to account for this behaviour. The intuitive idea is as follows. First, we generate a recombining stock price lattice for each possible volatility state. In our setting, this would result in two separate lattices. Since these are recombining, the number of nodes would be minimal, ensuring economical computational effort. Second, since volatility levels switch between low and high states, our numerical procedure would intuitively entail jumping from one lattice to the other when volatility switches. Therefore, proceeding with a more detailed explanation of the algorithm, we first discuss the most important property of the joint process for (S, V): the martingale property.

2.2. *No-arbitrage requirements*

The results of Harrison and Kreps (1979) and Harrison and Pliska (1981) state that prices of contingent claims are determined by discounting the payoffs from the security at the risk free rate of interest (r) and determining their expected value. This expected value must be taken with respect to a probability measure Q, under which the prices of discounted assets are martingales. Therefore, over time interval Δ, under the measure Q:

$$S(t) = \exp(-r\Delta) E_t^Q [S(t + \Delta)]. \tag{5}$$

Furthermore, under the well-known results of BMS, discounting is undertaken using the risk-less discount rate r, rather than the true return rate μ. In order to ensure that equation (5) holds we define the measure Q of the binomial process as follows:

$$S(t+\Delta) = \begin{cases} S(t)u, & \text{w/prob } q, \\ S(t)d, & \text{w/prob } (1-q), \end{cases}$$
$$q = \frac{\exp(r\Delta) - d}{u - d},$$
$$u = \exp\left\{+\sigma(t)\sqrt{\Delta}\right\},$$
$$d = \exp\left\{-\sigma(t)\sqrt{\Delta}\right\}.$$

Note that μ is replaced by r. It can be verified that the martingale condition (5) holds in this setting. The probability $q \approx 1/2$ and in the continuous-time limit, $q = 1/2$.

In a bivariate system with stochastic volatility, the joint system moves $[S(t), \sigma(t)]$ forward, and one possible representation is a double-binomial model. This is described as follows.

$$\begin{bmatrix} S(t+\Delta) \\ \sigma(t+\Delta) \end{bmatrix} = \begin{cases} \begin{bmatrix} S(t)\,u_j \\ \sigma_{\mathrm{H}} \end{bmatrix} \\ \begin{bmatrix} S(t)\,u_j \\ \sigma_{\mathrm{L}} \end{bmatrix} \\ \begin{bmatrix} S(t)\,d_j \\ \sigma_{\mathrm{H}} \end{bmatrix} \\ \begin{bmatrix} S(t)\,d_j \\ \sigma_{\mathrm{L}} \end{bmatrix} \end{cases} \text{with prob} \begin{pmatrix} qp_j \\ q(1-p_j) \\ (1-q)p_j \\ (1-q)(1-p_j) \end{pmatrix}, \quad j = \{H, L\}. \tag{6}$$

In this system, it can be checked that discounted stock prices are a martingale and satisfy equation (5). Recall that p_j is the probability of remaining in state j.

2.3. Path-independence and recombination in the two-state model

There is a simple way to see that the tree recombines and does not explode, thereby enabling polynomial complexity in the model. Consider two possible paths: an up-move in the high volatility state followed by a down-move in the low volatility state, and a down-move in the low volatility state followed by an up-move in the high volatility state. In the first case, the resulting stock price will be $Su_{\mathrm{H}}d_{\mathrm{L}}$, and in the second case it will be $Sd_{\mathrm{L}}u_{\mathrm{H}}$. Since the moves are simply multiplicative, these re-factor to the same stock price, $Su_{\mathrm{H}}d_{\mathrm{L}}$. Hence, the sequence of up and down moves in each state is not relevant, just the number of each type of the four moves. This gives the stock price after n periods on the tree, the normalized form being

$$S\,(u_{\mathrm{H}})^{n_1}\,(d_{\mathrm{H}})^{n_2}\,(u_{\mathrm{L}})^{n_3}\,(d_{\mathrm{L}})^{n_4}, \quad n_1 + n_2 + n_3 + n_4 = n.$$

All paths with the same $n_j, j = 1, 2, 3, 4$, will end in the same place, which results in recombination (albeit of complex form) in the model.

2.4. Generalization to **k** *states*

The 2-state model described above is easily extended to $k > 2$ states. Note that the probability q depends only on the current volatility state. The change in $S(t)$ populates a binomial state space, and the change in volatility populates a k-nomial space. The product space of $[S(t), \sigma(t)]$ therefore results in $(2k)$ branches from each node. The probabilities p_{ij}, given current volatility state i, are read off the transition matrix $P(\Delta)$. The transition matrix is determined from the generator matrix $G(\Delta)$.

3. Algorithm complexity

In this section, we show that the run time of the algorithm is order $O(kn^{2k})$ where n is the number of periods on the lattice, and k is the number of states in the volatility Markov chain. Note that $\Delta = T/n$, where T is the maturity of the derivative security lattice. We build up our results in a series of steps.

Proposition 1: *The number of distinct underlying prices at period n is* $\binom{n+2k-1}{2k-1}$.

Proof: Due to recombination, after n steps, the stock price can be expressed as

$$S\,(u_1)^{U_1}\,(d_1)^{D_1}\,(u_2)^{U_2}\,(d_2)^{D_2}\ldots(u_k)^{U_k}\,(d_k)^{D_k},$$
$$\sum_{i=1}^{k} U_i + D_i = n,$$

where u_k and d_k are the up and down moves for volatility state k, and U_k and D_k are the total number of times those moves were taken. The number of distinct prices is in a one-to-one correspondence with $\binom{n+2k-1}{2k-1}$, the number of ways to choose $2k-1$ distinct values out of the range $1, 2, \ldots, n+2k-1$. The n unchosen numbers represent the up and down moves. The number of elements before the first one chosen is the number of up moves in the first volatility level, the number of elements between the first and second ones chosen is the number of down moves in the first volatility level, the number of elements between the second and third is the number up moves in the second state, and so on, with the number of down moves in the ultimate state being the number of elements past the last one chosen. In the two-state example, $S\,(u_{\mathrm{H}})^5\,(d_{\mathrm{H}})^3\,(u_{\mathrm{L}})^0\,(d_2)^1$ would correspond to choosing 6, 10 and 11 from $1, \ldots, 12$. Every recombinant path has precisely one representation, and every set of choices represents a valid path. □

Proposition 2: *The total number of nodes in the tree is* $\binom{n+2k}{2k}$.

Proof: Using Proposition 1 for each level we simply carry out a summation across all levels, i.e.

$$\begin{aligned}\sum_{i=0}^{n}\binom{i+2k-1}{2k-1} &= \binom{2k-1}{2k-1}+\sum_{i=1}^{n}\binom{i+2k-1}{2k-1}\\ &= \binom{2k}{2k}+\sum_{i=1}^{n}\binom{i+2k-1}{2k-1}\\ &= \binom{2k+1}{2k}+\sum_{i=2}^{n}\binom{i+2k-1}{2k-1}\\ &= \binom{2k+2}{2k}+\sum_{i=3}^{n}\binom{i+2k-1}{2k-1}\\ &\ddots\\ &= \binom{n+2k}{2k},\end{aligned}$$

which provides the result. □

Proposition 3: *The complexity for the algorithm is $O(kn^{2k})$.*

Proof: Every node in the tree has at most $2k$ outgoing edges, representing the up and down moves for each of the k volatility levels. For every node in the tree but those in the last level, we have to cross all $2k$ edges. Thus, our running time is $2k\binom{n+2k}{2k}+\binom{n+2k-1}{2k-1}$ which is $O(kn^{2k})$. □

4. Numerical results

This section explores the algorithm numerically. Option prices are computed using the approximation model. We examine the speed of the algorithm. We increase the number of time periods in the model and determine whether the model suffers serious degradation in performance. Before proceeding with numerical exercises, we describe the computer implementation of the model.

4.1. Program design

We solve our lattice in a bottom-up fashion. We begin by creating all of the leaf-nodes of the lattice, and determining their values. We concurrently compute the moments, and the American/European put/call values. We then iterate over each level of the lattice, and for each node, we push the probability-weighted values to the parent nodes. When we arrive at the root, we have solutions for all possible initial volatilities.

Algorithm 1 Valuing Options With Stochastic Volatility

```
Create terminal nodes
for all Levels ℓ = n, ..., 1 do
  for all Active nodes v in level ℓ do
    Exercise, as appropriate
    for all Incoming edge e of v do
      if w = source(e) is new then
        Create w
        Add w to active list for level ℓ − 1
      Push values from v to w weighted by e
    Deactivate v
```

We used three main data structures in the program. We had: a list of all of the nodes in the active level; a tree of the parent nodes; and a list of the parent nodes, which replaced the active list as we switched levels. We used the natural ordering to order the tree. As we iterated over the active node list, we would first check if the node should be exercised (for American options). Then we checked if each of its potential parent nodes had already been encountered as the parent of a previous node. If not, we would create it. Finally, we pushed the probability mass backward.

While testing the implementation, we found that it was more important to use a simple binary tree, created in a balanced fashion, than it was to use a balancing tree structure. This was achieved by permuting the terminal level before we began our computation. Still, the searches dominate the running time, and we expect that hashing the level could improve the running time.

Beyond the efficiency of this approach, other advantages are that we can compute the moments of the distribution, allowing us to test and fit our volatilities to real values. This approach also allows concurrent computation of multiple strike prices and gives solutions for all possible initial volatilities.

4.2. Comparing option prices

We price call and put options. We verified the accuracy using simulation. As an additional metric, we computed the price of puts using put–call parity, and then compared the prices to the value of puts computed directly. This gives another check on the algorithm and we find that the results match, as should be expected. Once again, we obtain values for a range of maturities, $T=\{0.25, 0.50, 0.75, 1.00\}$ and employ the same parameterization as in the previous section. We define the Markov chain to be (on the left, and the corresponding generator matrix on the right)

$$\begin{bmatrix}0.7 & 0.3\\ 0.2 & 0.8\end{bmatrix},\quad \begin{bmatrix}-0.415888 & 0.415888\\ 0.277259 & -0.277259\end{bmatrix}.$$

We also employ a range of moneyness: strike prices are set to be at the money forward, i.e. $K=S_0e^{rT}$ as well as $\pm 10\%$ of this value, i.e. $K=0.9\,S_0e^{rT}$ and $K=1.1\,S_0e^{rT}$. Results are presented in table 1. In each cell, we report the value of calls and puts. Since the option values tend to oscillate depending on the number of steps in the tree being odd or even, we computed option values for n and $n+1$ steps and then averaged these values. The value of the options also depends critically on whether we start from the high volatility state or the low volatility one. Hence, we report values from both states.

There are several features of the model that we notice from the the first cut numerical analysis. First, the rate at which the lattice model converges to the simulation is rapid and comes from the benefit of the polynomial run time derived earlier. Second, it is checked that put–call parity holds exactly. Third, the initial volatility state matters in the pricing because the volatility in the model is

Table 1. Call and put prices from the algorithm. The maturity for the option is $T = \{0.25, 0.50, 0.75, 1.00\}$ years, the interest rate is $r = 0.05$. The initial stock price is $S_0 = 100$. The high volatility is $\sigma_H = 0.40$ and the low volatility is $\sigma_L = 0.10$. The transition matrix used had the probability of remaining in the high regime (H) as 70% and remaining in the low regime (L) as 80%. The value of the call and put are reported for initial states high (H) and low (L), respectively. Each option value is based on the average of two computations with n and $n+1$ steps, where $n = 25$.

Maturity (yr)	Strike	Call (H)	Call (L)	Put (H)	Put (L)
0.25	91.13	13.451	10.144	3.451	0.144
0.25	101.26	7.777	2.229	7.777	2.229
0.25	111.38	4.124	0.193	14.124	10.193
0.50	92.28	15.954	10.616	5.954	0.616
0.50	102.53	10.703	3.442	10.703	3.442
0.50	112.78	6.929	0.781	16.929	10.781
0.75	93.44	17.768	11.251	7.768	1.251
0.75	103.82	12.779	4.537	12.779	4.537
0.75	114.20	8.975	1.532	18.975	11.532
1.00	94.61	19.289	11.970	9.289	1.970
1.00	105.13	14.413	5.579	14.413	5.579
1.00	115.64	10.642	2.359	20.642	12.359

Table 2. European and American put prices from the algorithm. The maturity for the option is $T = \{0.25, 0.50, 0.75, 1.00\}$ years, the interest rate is $r = 0.05$. The initial stock price is $S_0 = 100$. The high volatility is $\sigma_H = 0.40$ and the low volatility is $\sigma_L = 0.10$. The transition matrix used had the probability of remaining in the high regime (H) as 70% and remaining in the low regime (L) as 80%. The value of the call and put are reported for initial states high (H) and low (L), respectively. Each option value is based on the average of two computations with n and $n+1$ steps, where $n = 25$.

		European		American	
Maturity (yr)	Strike	Put (H)	Put (L)	Put(H)	Put(L)
0.25	91.13	3.451	0.144	7.260	0.266
0.25	101.26	7.777	2.229	14.728	3.107
0.25	111.38	14.124	10.193	23.948	12.369
0.50	92.28	5.954	0.616	12.146	1.073
0.50	102.53	10.703	3.442	20.183	5.044
0.50	112.78	16.929	10.781	29.520	14.175
0.75	93.44	7.768	1.251	15.826	2.154
0.75	103.82	12.779	4.537	24.184	6.876
0.75	114.20	18.975	11.532	33.642	15.924
1.00	94.61	9.289	1.970	18.808	3.381
1.00	105.13	14.413	5.579	27.467	8.661
1.00	115.64	20.642	12.359	37.019	17.708

persistent, i.e. transition probabilities are low, and there is a high probability of remaining with the initial volatility regime. Note that this differentiates the model from normal mixture models, where persistence is not imposed since each step in the model allows a draw from the mixture of distributions. Finally, our lattice approach enables the pricing of American options just as easily as European ones. In table 2 we show how different the prices of American and European puts are.

4.3. *Convergence and speed*

We let the number of periods (n) on the lattice increase. Retaining the parameters from the previous section, we examine option values as n becomes very large in order to assess when convergence occurs to a stable value. This analysis will also identify the nature of convergence, i.e. oscillatory versus monotonic. We plot the values of put options only in figure 1. It is clear that oscillatory convergence is achieved.

Correspondingly, we also plot the runtime taken as n becomes large in figure 2. The run time is clearly growing slowly, as expected from the polynomial run time result. We note, however, that as computers become even faster over time, the run times may become even better than portrayed here.

4.4. *Varying the parameters*

In order to understand the impact of parameters on option prices, we use the model to examine the effect of various facets of the model. There are two aspects of the Markov-switching model that bear investigation. These are (i) the difference in volatilities in the different regimes (note that in the limiting case constant volatility,

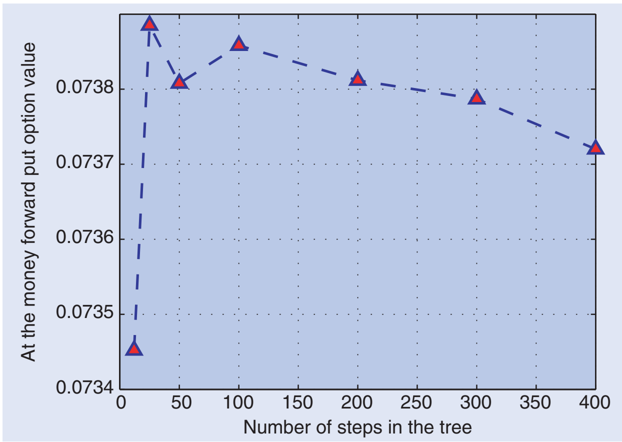

Figure 1. Plot of put values for varying time step. Prices are expressed in per dollar terms. The maturity for the option is $T = 1$ yr, the interest rate is $r = 0.10$. The initial stock price is $S_0 = 100$. The high volatility is $\sigma_H = 0.20$ and the low volatility is $\sigma_L = 0.10$. The transition matrix used had the probability of remaining in the high regime (H) as 70% and remaining in the low regime (L) as 80%. The value of the call and put are reported for initial states high (H) and low (L), respectively. The strike price is at the money forward.

we may set all $V_j = V_i$, $\forall i, j$), and (ii) variations in the degree of persistence in a state, embedded in the transition matrix.

To keep matters simple, we examine the case where there are only two volatility levels. In this simplest case, the impact of varying volatility levels and regime persistence should be easy to infer. We price options for $T = 1$. Three different sets of volatility levels are chosen: $(V_L, V_H) = \{(0.05^2, 0.25^2), (0.10^2, 0.20^2), (0.125^2, 0.175^2)\}$. Initial volatility is chosen to be either V_L or V_H. Finally, the level of persistence of each volatility state is set to three different levels via a choice of

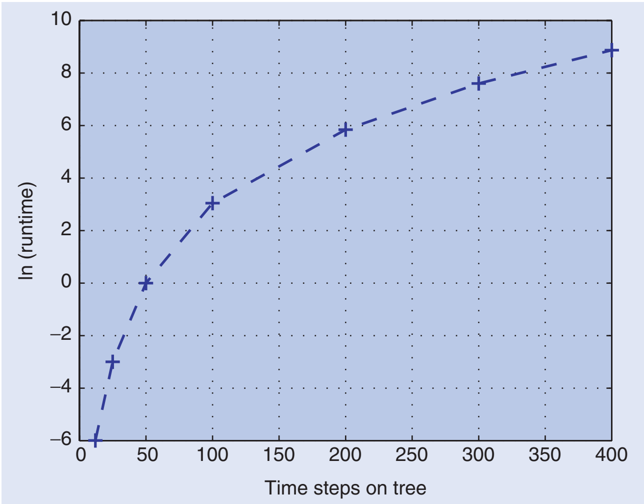

Figure 2. The natural logarithm of the runtime as a function of steps. The run time is measured in CPU seconds.

transition matrix probabilities: $(\lambda_L, \lambda_H) = \{(0.9, 0.9), (0.8, 0.8), (0.7, 0.7), (0.6, 0.6)\}$. The risk-free rate is 10%. The strike price chosen is at the money forward, i.e. $K = S_0 e^{rT}$. The number of time steps in the tree is $n = 100$. The results are presented in table 3. We report prices of call options only.

There are many interesting features of option prices which may be seen from the numerical experiments.

(a) If the initial volatility state is low, then option prices increase as the rate of transitions goes up (inferred from the lower values of $\lambda_j, j = \{H, L\}$). This occurs because the stock can migrate more easily to the high volatility state. On the other hand, when the initial volatility state is high, the converse occurs, and option prices fall with more transitioning, as migration to the low volatility state becomes more likely.

(b) The change in option prices as the transition rate increases is a function of the disparity between the high and low volatility levels. As the disparity decreases, intuitively, the change in option prices falls, as the switch in the level of volatility is correspondingly smaller.

(c) The further apart the high and low volatility levels are, it means that the volatility of volatility is higher, which should lead to higher average option prices. This can be checked on the table by averaging the option prices in high and low initial volatility columns. As we go down the table, volatility disparity declines, and so does the average option price, holding the transition rates constant for comparison across volatility disparities.

5. Option smiles in the switching model

In this section, we investigate the volatility smiles generated by this model. We also look at the ease with which the model may be calibrated using numerical examples. The parameters used in this exercise are: initial stock price $100, the number of steps in the model are $n = 25$ and the risk-free rate is 5%. The high and low volatility levels are 50% and 20%.

Table 3. European call option prices for varying parameters. In this table we vary the levels of the volatility states, and also the rates at which volatility transitions occur from one state to another. Maturity $T = 1$. Three different sets of volatility levels are chosen: $(V_L, V_H) = \{(0.05^2, 0.25^2), (0.10^2, 0.20^2), (0.125^2, 0.175^2)\}$. Initial volatility V_0 is chosen to be either V_L or V_H. The level of persistence of each volatility state is set to four different levels via a choice of transition matrix probabilities: $(\lambda_L, \lambda_H) = \{(0.9, 0.9), (0.8, 0.8), (0.7, 0.7), (0.6, 0.6)\}$. The risk-free rate is 10%. Initial stock price is $100 and the strike price chosen is at the money forward, i.e. $K = S_0 e^{rT}$. The number of time steps in the tree is $n = 25$. Option prices are the average of computations on two trees of n and $n + 1$ steps.

(V_L, V_H)	(λ_L, λ_H)	$V_0 = V_L$	$V_0 = V_H$
$(0.05^2, 0.25^2)$	$(0.9, 0.9)$	9.624	2.345
	$(0.8, 0.8)$	9.296	2.906
	$(0.7, 0.7)$	8.927	3.554
	$(0.6, 0.6)$	8.479	4.373
$(0.10^2, 0.20^2)$	$(0.9, 0.9)$	7.790	4.134
	$(0.8, 0.8)$	7.595	4.384
	$(0.7, 0.7)$	7.375	4.670
	$(0.6, 0.6)$	7.105	5.028
$(0.125^2, 0.175^2)$	$(0.9, 0.9)$	6.851	5.036
	$(0.8, 0.8)$	6.748	5.154
	$(0.7, 0.7)$	6.633	5.289
	$(0.6, 0.6)$	6.490	5.457

We varied the transition probabilities between the states and priced call options with both initial high volatility or low volatility states. Options were priced for a range of strikes, ranging from 70% of the at-the-money forward strike to 130% of the ATMF strike. Maturity was varied by a quarter year at a time to a maximum maturity of one year. The option prices were then used to back out Black–Scholes implied volatilities which are displayed in figures 3–5. Each figure has 2 plots; the one plot shows the smile when the system is initially in the high volatility state. The other plot corresponds to the initial volatility state being low.

First, notice that the smile is shallow when the system starts out in the high volatility state. On the other hand, there is a substantial smile if the system starts in the low volatility state. In figure 3, we can see this effect. In this figure, the probability of remaining in each state is 70%; figure 4 shows the case when the probability of remaining in each state is 51%. In the low volatility state, there is a possibility of a spike in volatility which leads to the exaggeration of the smile; this effect is ruled out in the high state. The mild smile in the high state comes purely from the slight tail fatness from the mixture of distributions.

Second, the slope of the term structure of implied volatilities (TSIV) is upward in the case of the initial volatility state being low. The TSIV is downward sloping when the initial volatility is high. This is intuitive, as given more time to run, the influence of the opposite state

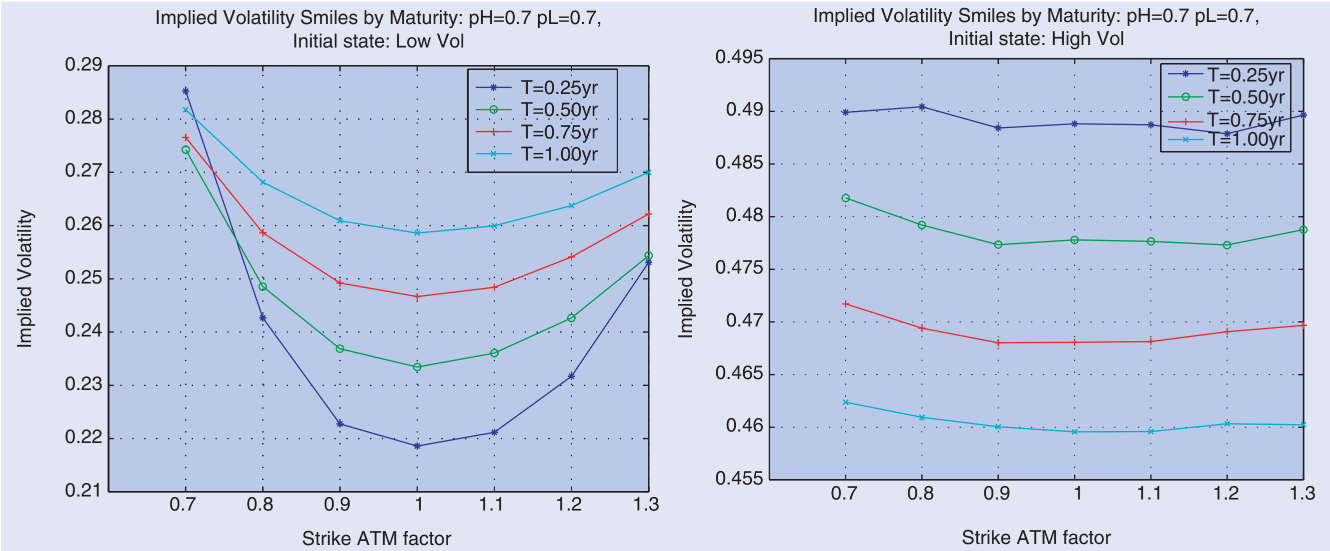

Figure 3. Option smiles. We plot smiles for four maturities: $\{0.25, 0.50, 0.75, 1.00\}$ years. The smiles are plotted for a range of strikes from 0.7 times the ATM forward strike to 1.3 times. The one-year transition probabilities are $pH = 0.7$ and $pL = 0.7$. These are the probabilities of remaining within the specified regime, H or L. The right plot is the smile if the initial volatility regime is H, and the left plot is the smile when the volatility regime is L. Parameters: initial stock price \$100, the number of steps in the model are $n = 25$ and the risk-free rate is 5%. The high and low volatility levels are 50% and 20%.

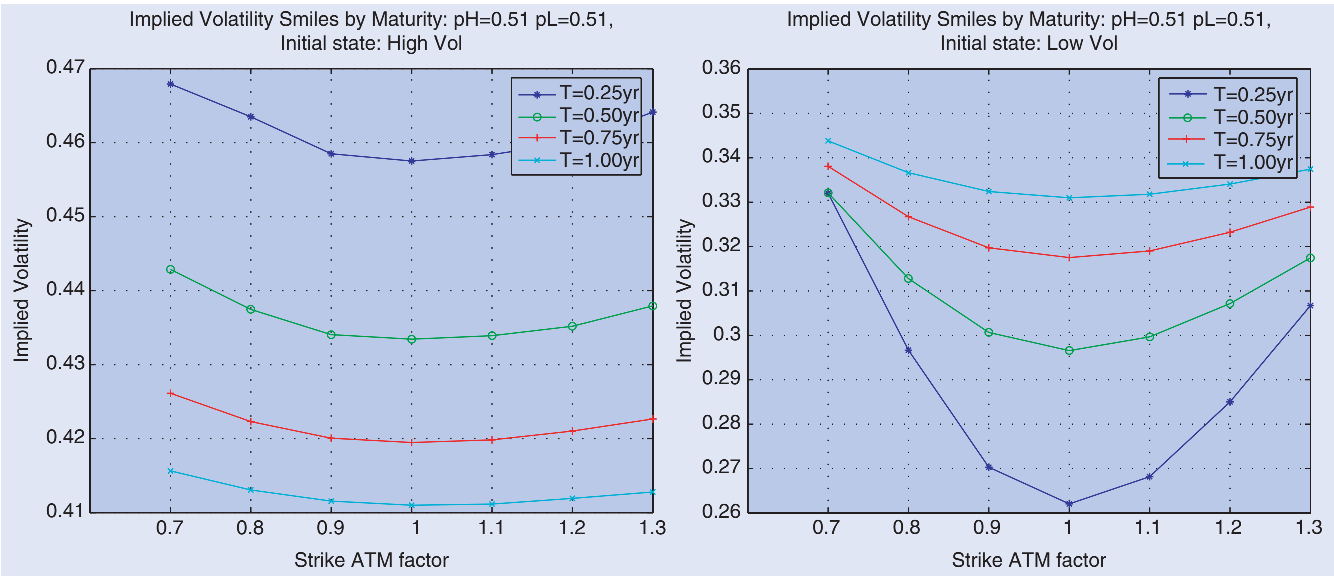

Figure 4. Option smiles. We plot smiles for four maturities: $\{0.25, 0.50, 0.75, 1.00\}$ years. The smiles are plotted for a range of strikes from 0.7 times the ATM forward strike to 1.3 times. The one-year transition probabilities are $pH = 0.51$ and $pL = 0.51$. These are the probabilities of remaining within the specified regime, H or L. The left plot is the smile if the initial volatility regime is H, and the right plot is the smile when the volatility regime is L. Parameters: initial stock price \$100, the number of steps in the model are $n = 25$ and the risk-free rate is 5%. The high and low volatility levels are 50% and 20%.

becomes more prevalent, and hence, volatility drifts upward in expectation with time if it starts in the low state; conversely for the high state.

Third, the smile becomes asymmetric at short maturities when the transition probabilities between states are not balanced. In figure 5 this may be seen for the shortest maturity of 0.25 years where there is an appreciable skew for the initial low volatility case.

Fourth, the level of the smile depends on both the initial state and the persistence of the states. From a comparison of figures 3 and 4 we can see that the steepness of the smile is lower in the former when the initial state is low. This is because there is higher persistence (70%) in the former figure as compared to the latter figure (51% persistence). Likewise, the smile level is higher when the initial state is high and persistence is high.

5.1. *Fitting the skew*

In this section we examine a shortcoming of this class of models and propose and implement a solution. Preceding numerical examples of the model generated smiles centred

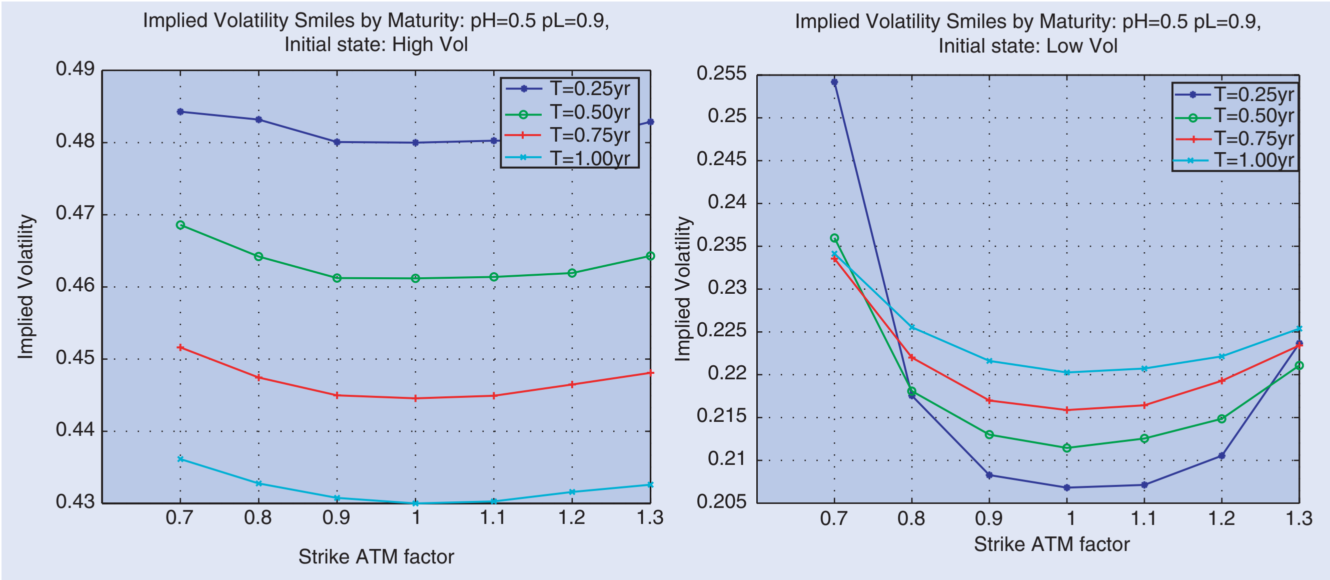

Figure 5. Option smiles. We plot smiles for four maturities: $\{0.25, 0.50, 0.75, 1.00\}$ years. The smiles are plotted for a range of strikes from 0.7 times the ATM forward strike to 1.3 times. The one-year transition probabilities are $pH = 0.5$ and $pL = 0.9$. These are the probabilities of remaining within the specified regime, H or L. The left plot is the smile if the initial volatility regime is H, and the right plot is the smile when the volatility regime is L. Parameters: initial stock price \$100, the number of steps in the model are $n = 25$ and the risk-free rate is 5%. The high and low volatility levels are 50% and 20%.

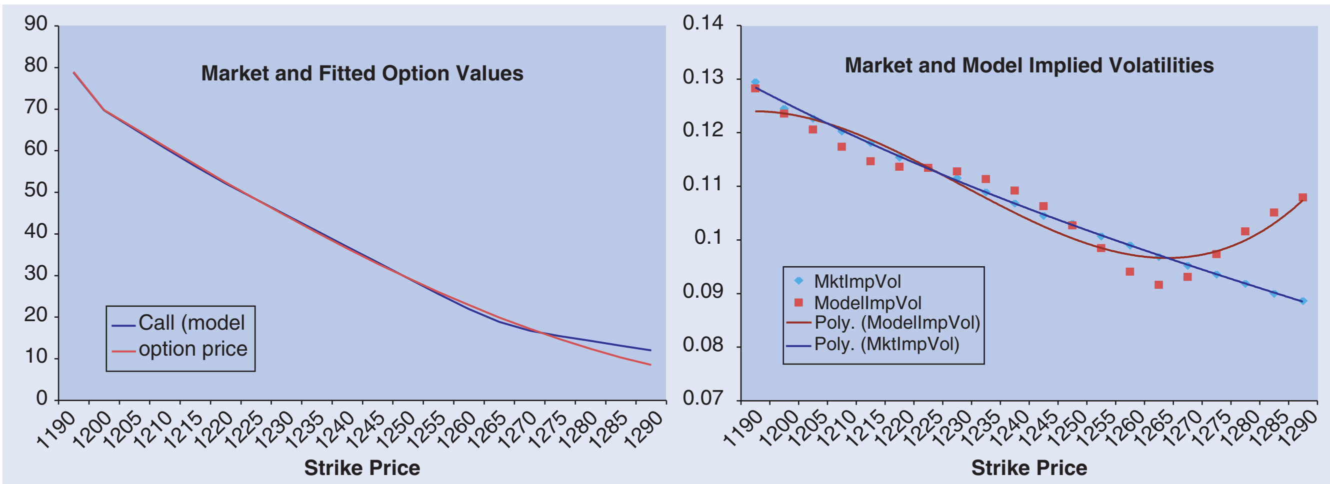

Figure 6. Calibrating S&P500 option prices. We accessed the options data for SPX on 22 November 2005, and downloaded the prices for calls maturing on 20 January 2006. The index value is 1261.23. The Libor rate for the same maturity is approximated to be 4.2%. We fitted the model to the cross-section of all calls that exceeded \$0.05 in value. Only 4 parameters need fitting, values are: $\{\sigma_{\mathrm{H}} = 0.3759, \sigma_{\mathrm{L}} = 0.0101, \lambda_{\mathrm{H}} = 0.5214, \lambda_{\mathrm{L}} = 0.9417\}$. The transition probabilities are for 1/15 of a year. After fitting we regenerated option prices from the model and plotted them against the prices in the market. The left plot shows the calibration to the market prices when the initial volatility state is low (the best fit), and the right plot shows the same fit in terms of implied volatility. The actual implied volatilities are plotted as dots, and the smiles are shown as third-order polynomial fits to the dots. We used call options since puts are priced by put–call parity. Options less than 0.5% of the value of the index are dropped from the analysis.

at-the-money, and this had been pointed out in previous work by Brigo *et al.* (2004). In our model, there is a tendency towards this, though this is not always the case, as we can see in the following example, where we calibrate the model to option on the S&P500 index. We chose this index for its well-known implied volatility skew.

We fitted the model to S&P500 options on 22 November 2005. Figure 6 shows the results of the calibration, both in terms of option prices and implied volatility. The fit to option prices is quite good, and the implied volatility skew fits well, except for out-of-the-money options where the model is not good at fitting the skew. This is a feature of the model, in that it generates smiles more easily than skews, and hence may be better calibrated to the FX options market as in Brigo *et al.* (2004); symmetry follows from the fact that the volatility state does not depend on the stock

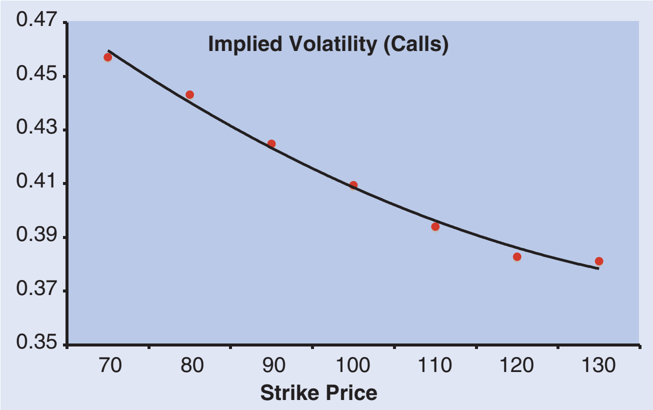

Figure 7. Calibrating an option skew. In this plot, we show the option smile from an extended model in which switching probabilities depend on the level of the stock price. We use the following parameters for the transition matrices: $\{\lambda_{\mathrm{H}}^{\mathrm{high}}=0.75, \lambda_{\mathrm{L}}^{\mathrm{high}}=0.9, \lambda_{\mathrm{H}}^{\mathrm{low}}=0.9, \lambda_{\mathrm{L}}^{\mathrm{low}}=0.75\}$. The other parameters are: $S=100$, $\sigma_{\mathrm{H}}=0.6$, $\sigma_{\mathrm{L}}=0.2$, $r_f=0.05$ and $T=0.5$.

price. The model may be extended to making volatility negatively correlated with the stock price, in which case skews will become easier to generate. However, this will result in a loss of the recombination feature of the tree and make computation more difficult. Extension to additional diffusion mixtures and a skew adjustment may be undertaken as in the work of Brigo *et al.* (2004). The ideas in this paper offer the lattice complement to simple versions of their work, allowing for the valuing of American-style options on mixture lattices.

Brigo *et al.* (2004) proposed an extended version of the mixture class of models with asymmetric volatility, which generates skews. Here we show that we can also achieve this result by making an adjustment in the transition probabilities and not in volatility, as suggested in Brigo *et al.* (2004). Our idea is to have one applicable transition matrix when stock prices are high and another one when stock prices are low. This is a new approach to generating skews, not proposed earlier. This is done by allowing the transition matrix to be state-dependent, where it favours remaining in the high volatility state when the stock price on the lattice is below its starting value (the 'low' stock state), and favours remaining in the low volatility state when the stock price is above its initial price (the 'high' stock state). Hence, we have two additional parameters in the model, and in total we need to fit four transition probabilities: $\{\lambda_{\mathrm{H}}^{\mathrm{high}}, \lambda_{\mathrm{L}}^{\mathrm{high}}, \lambda_{\mathrm{H}}^{\mathrm{low}}, \lambda_{\mathrm{L}}^{\mathrm{low}}\}$ where the superscripts depend on whether the stock is above or below its initial value. The advantage of this approach over that of Brigo *et al.* (2004) is that the lattice is still recombining, which we cannot achieve if we make an adjustment in volatilities. Whether this provides better fitting than the model of Brigo *et al.* (2004) is an open empirical question, and would require a larger empirical effort than in this paper. Figure 7 shows the smile when the transition matrix is dependent on 'high' and 'low' states.

Thus, having shown that this facile modification to the model accommodates the skew, we may now refit

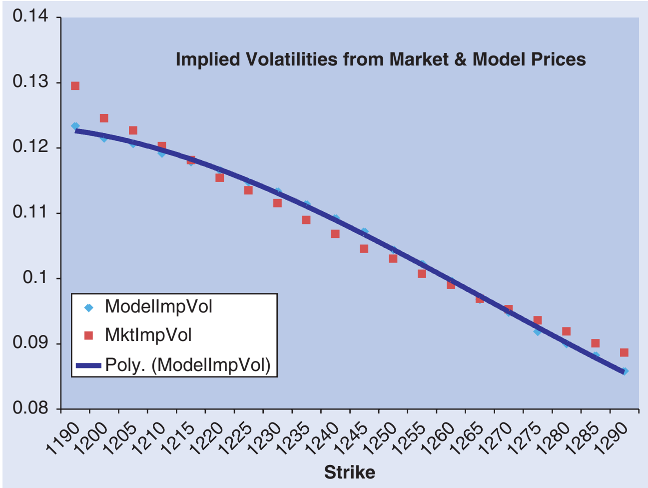

Figure 8. Calibrating an option skew to the S&P500 market. We accessed the options data for SPX on 22 November 2005, and downloaded the prices for calls maturing on 20 January 2006. The index value is 1261.23. The Libor rate for the same maturity is approximated to be 4.2%. We fitted the model to the cross-section of all calls that exceeded \$0.05 in value. Only 4 parameters need fitting, and the fitted values are: $\sigma_{\mathrm{H}}=0.1498$, $\sigma_{\mathrm{L}}=0.0462$, $\lambda_{\mathrm{H}}^{\mathrm{high}}=\lambda_{\mathrm{L}}^{\mathrm{low}}=0.5214$, $\lambda_{\mathrm{L}}^{\mathrm{high}}=\lambda_{\mathrm{H}}^{\mathrm{low}}=0.9417$. The transition probabilities are for 1/15 of a year. The actual implied volatilities are plotted as dots, and the smiles are shown as third-order polynomial fits to the dots. We used call options since puts are priced by put–call parity.

the model to the options on the S&P500. The implied volatility plot for this fit is presented in figure 8, which may be compared with the inferior fit of the model in figure 6. Further, we imposed the following condition on the fitting procedure:

$$\lambda_{\mathrm{H}}^{\mathrm{high}}=\lambda_{\mathrm{L}}^{\mathrm{low}}, \quad \lambda_{\mathrm{H}}^{\mathrm{low}}=\lambda_{\mathrm{L}}^{\mathrm{high}}.$$

Hence, the number of parameters to be fit remains exactly four as before. Despite this, we achieve an extremely good fit, and now there is a skew (smirk) instead of a smile.

6. Concluding comments

We provide a polynomial time lattice algorithm for pricing options with Markov-switching volatility. This parsimonious model bridges the computational gap between constant volatility models of Cox *et al.* (1979), which are easy to compute, and fully stochastic volatility models (e.g. Heston 1993), which are hard to compute and do not handle American options. By increasing the number of states in the Markov chain, we can move from the former class of models to the latter. It complements the literature on trees with stochastic volatility (see e.g. Bollen 1998) as well as the normal mixture diffusion class of models (see e.g. Alexander (2004), Mercurio (2002), Brigo and Mercurio (2002) and Brigo *et al.* (2004)). By making an adjustment to transition probabilities, we are able to implement skew models whereas the standard class of mixture models tends to have smiles with lowest implied volatilities at the money. The algorithm is $O(kn^{2k})$, where k is the number of states volatility can take and n is the number of periods on the lattice. We show that the model runs very fast and also converges to the exact price when n becomes large, without a substantial increase in run time. The model has been implemented differently in subsequent work by Balsara (2004) using a finite-difference approach, and binomial trees by Florescu and Viens (2004).

Acknowledgments

Thanks to Emanuel Derman and Steve Figlewski for comments. We thank two anonymous referees for several constructive suggestions. Donald Aingworth is at D.E. Shaw & Co., and was a doctoral candidate at Stanford University when this work was begun. Sanjiv Das is an Associate Professor at Santa Clara University, and Rajeev Motwani is a Professor at Stanford University. The second author gratefully acknowledges support from the Dean Witter Foundation, Breetwor Fellowship and a Research Grant from Santa Clara University.

References

Alexander, C., Normal mixture diffusion with uncertain volatility: modeling short and long-term smile effects. *J. Bank. Finance*, 2004, **28**, 2957–2980.

Amin, K., Jump-diffusion option valuation in discrete-time. *J. Finance*, 1993, **48**(5), 1833–1863.

Balsara, N., A finite difference approach to valuing options with markov switching variables. *Working Paper*, 2004 (Northeastern Illinois University).

Bates, D., Jumps and stochastic volatility: exchange rate processes implicit in deutsch mark options. *Rev. Financ. Stud.*, 1996, **9**(1), 69–107.

Black, F., and Scholes, M., The pricing of options and corporate liabilities. *J. Polit. Econ.*, 1973, **81**, 637–654.

Bollen, N., Valuing options in regime-switching models. *J. Deriv.*, 1998, **6**, 38–49.

Brigo, D. and Mercurio, F., Lognormal-mixture dynamics and calibration to market volatility smiles. *Int. J. Theoret. Appl. Finance*, 2002, **5**(4), 427–446.

Brigo, D., Mercurio, F. and Rapisarda, F., Smile at uncertainty. *Risk*, 2004, **17**(5), 97–101.

Cox, J., Ross, S. and Rubinstein, M., Option pricing: a simplified approach. *J. Financ. Econ.*, 1979, **3**, 145–166.

Duan, J.C., Popova, I. and Ritchken, P., Option pricing under regime switching. *Working Paper*, 1998 (Hong Kong University of Science and Technology).

Florescu, I. and Viens, F., A binomial tree approach to stochastic volatility driven model of the stock price. *Ann. Univ. Craiova Math. Comp. Sci. Ser.*, 2004, **31**(2), 1–17.

Hamilton, J., Analysis of time series subject to changes in regime. *J. Econometrics*, 1990, **45**, 39–70.

Harrison, J., and Kreps, D., Martingales and arbitrage in multiperiod securities markets. *J. Econ. Theory*, 1979, **20**, 381–408.

Harrison, J. and Pliska, S., Martingales and stochastic integrals in the theory of continuous trading. *Stochast. Process. Appl.*, 1981, **11**, 215–260.

Heston, S., A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.*, 1993, **6**, 327–343.

Hilliard, J.E., and Schwartz, A., Binomial option pricing under stochastic volatility and correlated state variables. *J. Deriv.*, 1996, **4**(1), 23–39.

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

Jarrow, R.A. and Rudd, A., *Option Pricing*, 1983 (Dow Jones-Irwin: Homewood, IL).

Leisen, D., Stock evolution under stochastic volatility: a discrete approach. *Discussion Paper No. B-407*, 2000 (Stanford University).

Longstaff, F.A. and Schwartz, E.S., Valuing American options by simulation: a simple least-squares approach. *Rev. Financ. Stud.*, 2001, **14**, 113–147.

Mercurio, F., A multi-stage uncertain volatility model. *Working Paper*, 2002 (Banca IMI: Milano).

Merton, R.C., The theory of rational option pricing. *Bell J. Econ. Manage. Sci.*, 1973, **4**, 141–183.