FINANCIAL OPTIONS RESEARCH CENTRE

University of Warwick

Volatility Smile Consistent Option Models: A Survey

George Skiadopoulos

November 1999

Financial Options Research Centre
Warwick Business School
University of Warwick
Coventry
CV4 7AL
Phone: (01203) 524118

FORC Preprint: 2000/101

VOLATILITY SMILE CONSISTENT OPTION MODELS: A SURVEY [1]

George Skiadopoulos [2]

University of Warwick

November 1999

---

[1] I am grateful to Stewart Hodges and Alessandro Rossi for many helpful discussions. I would also like to thank Abhay Abhyankar, Marco Avellaneda, Les Clewlow, Peter Honore and Joao Pedro Nunes for comments. Part of this paper was written while the author benefited from the Human Capital and Mobility programme of the European Commission. Financial support from the Corporate Members of FORC is also gratefully acknowledged. Because of space limitations some contributions have not been given the attention that they deserve. Any remaining errors are my responsibility alone.

[2] Athens Derivatives Exchange Research and Development of New Products Group, and Associate Research Fellow at the Financial Options Research Centre, Chrisanthemon 76, N. Iraklio 1421, Athens, Greece, Tel: 00301-3370912, e-mail: gskiadopoulos@hotmail.com

Abstract

The developing literature on "smile consistent" no-arbitrage models has emerged from the need to price and hedge exotic options consistently with the prices of standard European options. This survey paper describes the steps through which this literature has evolved by providing a taxonomy of the various models. It highlights the main ideas behind the different models, and it outlines their advantages and limitations. Practical issues in implementing the models are also discussed.

1 Introduction

Exotic options are often hedged with European options. This hedging strategy is called static replication (see Carr, Ellis, and Gupta [22], Derman, Ergener, and Kani [34]). To improve the hedging performance, exotic and standard options need to be valued consistently. This is done by first assuming a stochastic process which describes the underlying asset price dynamics. The process is calibrated to the observed market prices of exchange-traded options. Then, the resulting (implied) process is used to price over the counter options. The need to price and hedge exotic options consistently with the prices of standard European options has led to the quite recent literature of "smile-consistent" no-arbitrage models. The aim of this paper is to survey concisely this developing literature which is of particular importance to both academics and practitioners.

The behavior of implied volatilities derived from inverting Black-Scholes [16] (BS) model, questions the validity of the model. The empirical evidence shows (see among others Derman and Kani [39], Rubinstein [84],[85] and for an extensive literature survey Mayhew [75]) that implied volatilities vary across different strikes (smiles or skews), and different times-to-expiration (term structure), while the BS model does not predict any such variation. These results suggest that implied volatilities could be viewed as a two-dimensional non-flat surface. In addition, the functional form of this surface changes over time (Gemmill [51], Jackwerth and Rubinstein [61]).

The non-flat implied volatility surface is roughly explained by either stochastic volatility (see Hull and White [59], Johnson and Shanno [68], Scott [90], Wiggins [102]), or jump models (see Bates [8] Merton [78]), or both (see Bates [9], [11], Scott [91]). The approach taken by these models consists of specifying the parameters of the processes for the underlying traded and non-traded securities (stochastic volatility and, or jump). The market price of risk of the non-traded sources of risk has also to be specified[1]. Then, option prices are derived as a function of the parameters of the processes and the prices of the underlying securities.

However, these models do not fit observed implied volatility patterns well (see Clewlow and Xu [25], Das and Sundaram [33], Taylor and Xu [99]), making it difficult to use them in practice to price and hedge exotic options. These problems have motivated the recent literature on "smile consistent" no-arbitrage models. "Smile-consistent" models reverse the approach taken by the conventional stochastic volatility, or jump models. The prices of standard European options are taken as given, and they are used to infer information about the underlying price processes.

In this paper, we survey the "smile-consistent" no-arbitrage literature by classifying the two stages through which it has been developed. First, deterministic volatility models which fit the observed European option prices were introduced (Andersen [1], Andreasen [2], Avellaneda et al. [4], Barle and Cakici [7], Derman and Kani [35], Derman, Kani and Chriss [38], Dupire [43], [45], Jackwerth [63], Rubinstein [85]). Next, stochastic volatility models were provided which allowed for smile-consistent option pricing under the no-arbitrage evolution of the volatility surface (Britten-Jones and Neuberger [20], Derman and Kani [39], Dupire [42], [44], [46], Ledoit and Santa-Clara [71]). The second class of models is more general and

[1]The market price of risk is specified by invoking equilibrium arguments. In this sense these models allow for equilibrium, and not for arbitrage pricing. For an intuitive, clarification of the difference between equilibrium and arbitrage pricing, see Dupire [44].

it nests the first class. The various models are developed in continuous, or discrete time, or both. We describe them by commenting on the key ideas behind them and we outline their advantages and limitations. Furthermore, some practical issues in implementing these models are addressed.

The remainder of the paper is structured as follows. In the second and third section, we discuss the smile consistent deterministic volatility models in continuous and discrete time, respectively. We compare them, and we discuss some practical issues in implementing them. The empirical results from the research on the validity of the smile-consistent no-arbitrage deterministic volatility models are presented in section four. Section five describes the smile-consistent no-arbitrage stochastic volatility models, and brings together the different definitions of the key concept of the forward variance. The investigation of the dynamics of volatilities, as a prerequisite for the implementation of this class of models, is also pointed out. The last section concludes and suggests some topics for future research.

2 Smile Consistent Deterministic Volatility Models in Continuous Time

In this section, we motivate the use of deterministic volatility models and we present the various smile-consistent deterministic volatility models that are developed in continuous time.

Option prices calculated from the BS model deviate from the market option prices, especially after the market crash on October 19, 1987. An alternative way of stating this is by describing the stylized characteristics of implied volatilities (see for instance Jackwerth and Rubinstein [61], Rubinstein [85]). Depending on the underlying asset and the sample period under scrutiny, implied volatilities have a term structure which is upward, or downward sloping (Derman and Kani [35]). In addition, in contrast to the BS prediction of a constant implied volatility across strikes, there is an implied volatility bias; they vary across strikes giving rise to smiles, or skews. Black [15] finds that implied volatilities decline as the strike price goes down; Macbeth and Merville [73] find that call options implied volatilities tend to be higher when the strike price declines. Rubinstein [84] finds that the implied volatility bias changes direction depending on the sample period under scrutiny. Shastri and Wethyavivorn [93] find an implied volatility smile (see also Bates [10] for an extensive survey).

A simple way of explaining the implied volatility skew which appears in Index and Futures options markets (Bates [12]), is by resorting to standard deterministic volatility models (see Cox [28], Cox and Ross [29], Emanuel and MacBeth [47], MacBeth and Merville [74]) which allow for an inverse relationship between the price of the underlying security, and the variance of the rate of return. These models specify exogenously the instantaneous volatility σ as a deterministic function of the price of the underlying asset S_t and time t, i.e.

$$\frac{dS_t}{S_t} = a(S_t, t)dt + \sigma(S_t, t)dW_t \tag{1}$$

There are three ways, to our knowledge, of explaining the systematic relationship of volatility with the underlying asset. A first informal approach is the one suggested by Derman, Kani and Zou [36]. They call $\sigma(S_t, t)$ as the local volatility. It is the volatility which prevails at the asset level S_t at time t. They think of the implied volatility as an average of local volatilities across the state space (rather than the time domain). Assuming that the local volatility varies linearly with the asset price, they show that the local volatility varies with the asset level about twice as rapidly, as implied volatility varies with the strike. They show that the smile can be explained by the variation of local volatility with the asset price and time, and other effects such as stochastic volatility and jumps are less important. The second approach, uses the negative correlation between $\sigma(S_t, t)$ and the asset price which was first observed by Black [15]. This negative relationship can be explained either as a leverage effect (Christie [23]), or by the portfolio insurance strategies that investors use (Grossman and Zhou [53]). The third approach, invokes Platen and Schweizer [88] model's which is similar in spirit to Grossman and Zhou's. They start from a microeconomic equilibrium model, where part of the demand for the underlying asset is induced by a hedging strategy. The limit of their model is a deterministic volatility diffusion, where the volatility coefficient is derived endogenously from assumptions about agents' trading behavior.

In contrast to standard deterministic volatility models, smile consistent deterministic volatility models do not specify $\sigma(S_t, t)$ in advance, but endogenously from the European option prices. Therefore, they preserve the "pricing by no-arbitrage" property of the BS model, and the option's payoff can be synthesized from a portfolio of existing assets, i.e. the markets are complete (see Dothan [40] and for a concise description Sundaram [98]). In addition, they provide us with a method for specifying $\sigma(S_t, t)$ from the market option prices, i.e. they deliver to us an implied process.

The knowledge of the process allows for the pricing and hedging of path-dependent options (Monte-Carlo methods) and American options (by dynamic programming). The hedging will be effective throughout the life of the option if the asset price behaves according to the inferred process.

2.1 Dupire (1993 and 1994)

The first "smile-consistent" deterministic volatility models developed in continuous time, were presented by Dupire. Let $C(K,T)$ be a European call option of exercise (strike) price K and maturity T. Assume that the continuum of all $(C(K,T))_{K,T}$ are traded and that their prices today are consistent with no arbitrage. Breeden and Litzenberger [18] have shown that the observed European call option prices deliver to us the conditional terminal risk-neutral density as a function of K, i.e.

$$\Phi_T(K) = e^{-r(T-t)} \frac{\partial^2 C(K,T)}{\partial K^2} \tag{2}$$

where $\Phi_T(K)$ is the terminal risk-neutral density of S_T conditional on the information at current time t, and r is the interest rate. In general, the converse is not true. From the terminal implied risk-neutral density we can not recover uniquely the asset process which generates today's option prices (see Dupire [43], Melick and Thomas [77]). However, Dupire proves that there is an exception. Under some technical regularity conditions, we can recover a *unique* diffusion process from the terminal risk-neutral implied density, if we restrict ourselves to risk-neutral deterministic volatility diffusions (see Dupire [43], [45]). This is proved by means of the forward Kolmogorov equation. Given the process

$$dx = a(x,t)dt + b(x,t)dW \tag{3}$$

the forward Kolmogorov equation is given by

$$\frac{1}{2}\frac{\partial^2(b^2 f)}{\partial x^2} - \frac{\partial(af)}{\partial x} = \frac{\partial f}{\partial T} \tag{4}$$

where $f(x,T) \equiv \Phi_T(x)$. In other words, in general the forward Kolmogorov equation is used for deriving conditional densities (or distributions) starting from a given process[2]. However, in our case we cope with the converse problem: *f is known and b is the unknown.*

Restricting ourselves to risk-neutral densities (and assuming without loss of generality that the interest rate is zero) equation (4) becomes[3]

$$\frac{1}{2}\frac{\partial^2(b^2 f)}{\partial x^2} = \frac{\partial f}{\partial T} \tag{5}$$

As f can be written as $\frac{\partial^2 C}{\partial x^2}$ (x denotes the strike price), he shows that (Dupire [43])

$$\frac{1}{2}b^2\frac{\partial^2 C}{\partial x^2} = \frac{\partial C}{\partial T} \tag{6}$$

Both derivatives are positive by arbitrage. Hence[4],

$$b(x,T) = \sqrt{\frac{2\frac{\partial C(x,T)}{\partial T}}{\frac{\partial^2 C(x,T)}{\partial x^2}}} \tag{7}$$

Combining equations (1) and (3), we obtain the instantaneous volatility by $\sigma(S,T) = \frac{b(S,T)}{S}$.

The forward equation (6) presents the option pricing problem in a different way than the BS partial differential equation (PDE) does. It turns the option pricing problem into a problem in strikes and maturities with fixed spot and time, rather than in a problem in spot and time with fixed strike and maturity. Andreasen [2] shows that in general, the forward equations for the option prices imply a duality: the problem of pricing and hedging of European options can be solved in a dual economy, where the spot is the strike, the strike is the spot, the call is the put, the interest rate is the dividend yield, and the dividend yield is the interest rate. However, the BS PDE applies to any contingent claim, while equation (6) holds only for European call options (see Dupire [44], [45]).

---

[2]For a further discussion on the forward Kolmogorov equation and its use, see Cox and Miller [27].

[3]By restricting himself to the risk-neutral environment, he has one equation (the forward equation) and one unknown (the instantaneous volatility). Otherwise, he would have two unknowns, the drift and the instantaneous volatility and he would not be able to determine uniquely the instantaneous volatility.

[4]Equation (7) holds for every system of call prices, provided that the time derivative of the European call vanishes as the strike goes to infinity and a slow growth condition is satisfied (see Dupire [43]). For instance, the BS model for a European call on a no dividend asset with interest rate equal to zero, satisfies (7) (see Bick and Reisman [14]). Bick and Reisman also derive equation (7), independently. However, they make no distinction between the concept of local and implied volatility.

3 Smile Consistent Deterministic Volatility Models in Discrete Time

The implementation of a smile-consistent deterministic volatility model, for pricing and hedging purposes, is done in a discrete time framework. The tools used are either binomial, or trinomial implied trees, or implicit finite difference schemes. The former discretizes the asset price process, while the latter discretizes the BS type fundamental PDE (see Geske and Shastri [52]).

Binomial (or trinomial) trees are built from the known prices of European options. Such trees are called *implied trees* because they are consistent with or implied by the volatility smile. Their continuous time limit is a deterministic volatility process (see Nelson and Ramaswamy [79]). In the standard Cox, Ross, Rubinstein [30] (CRR) tree the size of the upper and down move of the underlying asset, and the respective probabilities of such moves are constant (because they depend on the volatility which is assumed to be constant). This is not any longer the case with implied trees.

In general, in order to construct a tree we need to know the way that the underlying asset price evolves and the transition probabilities corresponding to the links of the tree. The traditional way of calculating these two, is by establishing conditions under which a sequence of processes converges in distribution to the given diffusion (Nelson and Ramaswamy [79])[5]. In the case of implied trees, there is also the additional constraint that they must correctly reproduce the volatility smile. Once the tree has been built, backward induction (see Cox, Ross, Rubinstein [30]) is applied for the evaluation of the option. Alternatively, the constructed tree delivers the local volatility surface which can be used for the pricing and hedging of options via Monte Carlo simulation (see Derman, Kani, and Zou [36], and Zou and Derman [103]).

There are three different approaches to the construction of implied trees. First, binomial implied trees are constructed by using both backward and forward induction (Derman and Kani [35], Barle and Cakici [7])[6]. They fit implied volatilities in both the maturity and strike dimension. Second, binomial implied trees are constructed by using only backward induction. They fit either the strike dependence of implied volatilities (Rubinstein [85]), or both the strike and the term dependence (Jackwerth [63]). Third, trinomial trees are built by using simultaneously forward and backward induction (Dupire [43], [45], Derman, Kani and Chriss [38]). They fit both the strike and the term structure of implied volatilities. Their main difference with implied binomial trees is that the state space is fixed in advance, and the construction of the tree is reduced to the calculation of transition probabilities.

[5] For example, for the lognormal diffusion (1) we match the first two moments of the continuous process and of the discretised process (see Cox, Ross, Rubinstein [30]).

[6] Backward and forward induction are the discrete analogues of the Kolmogorov backward and forward equations, respectively. The binomial backward equation states that the price at any period n is the discounted value of the average of the prices at the two up and down nodes in the next period $n+1$. The binomial forward is the "dual" or the "adjoint" of the binomial backward equation. It states that the price of an Arrow-Debreu primitive security of any maturity $(n+1)$ is the average of the discounted at the previous two up and down nodes of the Arrow-Debreu security of maturity n (for a further description and application of the technique in the context of interest rate models see Hull and White [60], Jamshidian [64], and Rebonato [82]).

Finally, implicit finite difference schemes (Andersen [1], Andreasen [2]) and maximum entropy methods (Buchen and Kelly [19], Stutzer [97], Avellaneda et al. [4])are proposed in order to solve some of the problems encountered with implied trees.

3.1 Derman and Kani (1994)

Derman and Kani [35] build a recombining binomial implied tree by using forward and backward induction simultaneously.

Their tree has uniformly spaced levels which are Δt apart. In order to construct it, they assume that they have already implied the tree's nodes and the transition probabilities out to level n. The known price at node i and level n $S_{i,n}$ can evolve into an "up" node with price $S_{i+1,n+1}$, or into a "down" node with price $S_{i,n+1}$ at level $(n+1)$. The (unknown) probability of making a transition into the "up" node is denoted by p_i. The aim is to determine the nodes of the $(n+1)^{th}$ level at time t_{n+1} and the corresponding transition probabilities. In total, there are $2n+1$ parameters that define the transition from the n to the $(n+1)$ level of the tree. These parameters are the $n+1$ stock prices $S_{i,n+1}$, and the n transition probabilities p_i. Derman and Kani determine them by using the smile.

They find the distribution of $S_{i,n+1}$ and the transition probabilities p_i by using the theoretical values of n forwards and n *European* options, all expiring at time t_{n+1}. They require that these theoretical values match the (interpolated) market values. This provides $2n$ equations for these $2n+1$ parameters, and it ensures that they fit today's smile. They use the one remaining degree of freedom to make the centre of their tree to coincide with the centre of the standard CRR tree that has constant local volatility[7].

The above can be expressed formally as follows: The martingale condition delivers the forward price $F_{i,n}$ of the stock as

$$F_{i,n} = p_i S_{i+1,n+1} + (1-p_i) S_{i,n+1} \tag{8}$$

Let $C(S_{i,n}, t_{n+1})$ and $P(S_{i,n}, t_{n+1})$, respectively, be the known market values for a European call and put, struck today at $K = S_{i,n}$ and expiring at t_{n+1}. The values of each of these calls and puts are known from interpolating the smile curve implied from options expiring at time t_{n+1}. The theoretical binomial value of a European call struck at K and expiring at t_{n+1} in a complete market is given by:

$$C(K, t_{n+1}) = e^{-r\Delta t} \sum_{j=1}^{n} \{Q_{n,j} p_j + Q_{n,j+1}(1-p_{j+1})\} \max(S_{j+1} - K, 0) \tag{9}$$

where the sum is taken over all nodes j at the $(n+1)$ level and $Q_{n,j}$ is the price of an Arrow-Debreu security expiring at t_{n+1}. From equations (8) and (9) they get that:

$$S_{i+1,n+1} = f(r, \Delta t, S_{i,n+1}, C(S_{i,n}, t_{n+1}), \Sigma_c, Q_{i,n}, S_{i,n}, F_{i,n}) \tag{10}$$

$$p_i = f(F_{i,n}, S_{i,n+1}, S_{i+1,n+1}) \tag{11}$$

[7] For a different choice of the "centrering condition", the constructed tree would have been different. However, in the continuous time limit, where there are an infinite number of nodes at each time step, the choice of the "centrering condition" is not important (see Derman, Kani and Chriss [38]).

where $\Sigma_c = \sum_{j=i+1}^{n} Q_{j,n}(F_{j,n} - S_{i,n})$. The Arrow-Debreu prices $Q_{i,n}$ have been calculated by applying forward induction.

For all the nodes above the centre of the tree we can find iteratively $S_{i+1,n+1}$ and p_i, from equations (10) and (11) if we know $S_{i,n+1}$ at one initial node. "Centering conditions" are imposed so as to calculate $S_{i,n+1}$. If the number of nodes at the $(n+1)^{th}$ level is odd they choose the central node $S_{i,n+1}$ (for $i = \frac{n}{2} + 1$) to be today's spot price, as in the CRR tree. If the number of nodes at the $(n+1)^{th}$ level is even, they start instead by identifying as initial $S_{i,n+1}$ and $S_{i,n}$, the nodes just below and above the center of the level (i.e. $i = \frac{n+1}{2}$). This is done by making the average of the natural logarithms of the two central nodes' stock prices equal to the logarithm of today's spot price. Substituting this condition in equation (10) gives the formula for the upper of the two central nodes for even levels

$$S_{i+1,n+1} = f(r, \Delta t, S_{i,n}, C(S_{i,n}, t_{n+1}), \Sigma_c, Q_{i,n}, F_{i,n}) \tag{12}$$

for $i = \frac{n}{2} + 1$. Once we have this initial node's stock, we can continue to fix higher nodes from equation (10).

Similarly, the asset value for the nodes below the central node at level n , are calculated by using known put prices[8]. The analogous formula that determines a lower stock price from a known upper one is

$$S_{i,n+1} = f(r, \Delta t, S_{i+1,n+1}, P(S_{i,n}, t_{n+1}), \Sigma_p, Q_{i,n}, F_{i,n}) \tag{13}$$

where $\Sigma_p = \sum_{j=1}^{i-1} Q_{j,n}(S_{i,n} - F_{j,n})$.

Applying equations (10), (11) and (13) for every level and for small enough time steps between successive levels, completes the construction of the tree. These equations reveal the idea behind the "implied trees" methodology. From the current option prices, we can back out a discrete approximation to the risk-neutral stock process, and the risk-neutral transition probabilities.

The advantage of Derman and Kani's algorithm is that it provides the asset price evolution and the transition probabilities, by capturing both the term and the strike structure of implied volatilities (they interpolate across option prices for each time level). On the other hand, Barle and Cakici [7] find that Derman and Kani's algorithm fails to reproduce the smile accurately if the interest rate is high. In the next section, we demonstrate how Barle and Cakici extend Derman's and Kani's algorithm.

3.2 Barle and Cakici (1995)

In order to ensure that transition probabilities remain in the interval [0,1], Derman and Kani require that $F_{i,n} < S_{i+1,n+1} < F_{i+1,n}$. If the stock price $S_{i+1,n+1}$ violates this inequality, then they override the option price that produced it. The missing stock price is replaced by the

[8] Non-synchronous trading and the bid-ask bounce create noise in the observed option prices (see Harvey and Whaley [55], and Roll [83], respectively). By using out-of-the money calls and puts, they minimize the effect of noisy option prices on the construction of their tree. This is because the delta for these options is low.

one which keeps $\ln S_{i+1,n+1} - \ln S_{i,n+1} = \ln S_{i+1,n} - \ln S_{i,n}$. However, Barle and Cakici [7] note that Derman and Kani's algorithm fails to reproduce the smile accurately when the interest rate is high. The reason is that with higher interest rate, negative probabilities are more frequently encountered, leading to overriding the corresponding option prices. Hence, the constructed tree does not fully incorporate the information from the smile. In order to correct for this problem they propose three modifications to Derman and Kani's method.

First, they choose the option to be struck at $K = F_{i,n}$. Second, rather than fixing the center of the tree at the current stock price, they allow it to follow the evolution of the mean of the risk-neutral distribution by setting it to $Se^{rt_{n+1}}$. Third, when there is a missing stock price due to the violation of the arbitrage condition, they set $S_{i+1,n+1} = \frac{F_{i,n} + F_{i+1,n}}{2}$. Barle and Cakici's modifications are equivalent to working with the futures rather than the spot price. In a standard binomial tree this trick guarantees non-negative transition probabilities (see Hull [58]).

Even though, their modified method fits the smile accurately for very high interest rates (e.g. $r = 40\%$), it fails to do so for increasing interest rates and smile slope. Negative probabilities occur even with this modification. "These weaknesses are a consequence of the strict requirements that continuous diffusion can be modelled as a binomial process and on a recombining tree", as they state in their conclusions.

3.3 Rubinstein (1994)

Rubinstein [85] constructs an implied binomial tree which has T levels, by using only backward and not forward induction. In this sense his tree is an extension of the CRR tree. The key input to his algorithm is the terminal total (nodal, as opposed to the one period) risk-neutral implied probabilities of the underlying asset. He extracts them from the observed prices of European options which mature at time T, by using a nonlinear minimization method[9].

His method consists of establishing a prior guess of the terminal risk-neutral distribution; his guess is the log-normal one. Then, the implied posterior risk-neutral probabilities are those which are, in the least-squares sense, closest to the lognormal[10]. The minimization is performed subject to some constraints. The probabilities must add up to one and be non-negative. Moreover, they are calculated so as the present value of the underlying assets and all the European options calculated with these probabilities to fall between their respective bid and ask prices.

In order to proceed further, he imposes a number of assumptions: (a) binomial evolution of the asset price, (b) recombining nodes, (c) ending nodal values organized from lowest to highest, (d) constant interest rate, and (e) all paths leading to the same ending node have the same risk-neutral probability. Then, the tree is constructed through four very simple steps.

[9] In general, there is a number of approaches for estimating risk-neutral probability density functions from option prices (see also Section 3.7 for the description of the maximum entropy method). For a survey of these methods see Bahra [5], and Mayhew [76].

[10] Jackwerth and Rubinstein [61] examine alternative specifications of the minimization criterion using historically observed option prices. All of the specifications, including the quadratic one, produce similar posterior distributions (for near-the-money) which seem to be independent of the assumed prior distribution.

1. For every node j, calculate the terminal path probabilities corresponding to the T level, from the terminal nodal probabilities.

2. From the path probabilities of the T level, calculate the path probabilities for the $T-1$ level.

3. From the path probabilities of the $T-1$ level, calculate the transition probabilities of a transition from level $T-1$ to level T.

4. Uses the transition probabilities to calculate the return for the jth node at the $T-1$ level.

This exercise is repeated for every time level and completes the construction of the tree. Then, the value and the hedging parameters of any derivative instrument maturing with or before the European options can be calculated. However, the constructed tree fits only the European options with maturity T (in the sense that the model price falls within the bid and ask observed prices). This is because it uses as input only the options maturing at the T level. It does not capture the term structure of implied volatilities something which can be considered as a limitation of the technique.

3.4 Jackwerth (1997)

Rubinstein's implied binomial tree fits only the European options which expire at the terminal level of the tree. On the other hand, Derman and Kani's model fits intermediate maturity options, but the construction of the tree depends on the chosen interpolation and extrapolation method. Moreover, negative transition probabilities are frequently encountered. As a solution to these problems, Jackwerth [63] develops a generalized implied binomial tree. It is 'generalized' in the sense that the simplicity of Rubinstein's implied binomial tree is preserved, but it relaxes the assumption that all the paths which lead up to the same ending node are equally probable. This allows him to fit intermediate maturity European option prices. In addition, the transition probabilities are constrained by construction to lie within 0 and 1, as it was the case with Rubinstein's implied tree.

Let $i = 0, 1, \ldots, n$ be the time step, and $j = 0, 1, \ldots, i$, be the nodes at each time step starting with the lowest stock price at the bottom of the step. In order to fit the intermediate maturity options, Jackwerth works with nodal, rather than path probabilities, and he uses a weight function $w_{i,j}$ which has a particular interpretation. $w_{i,j}$ can be interpreted as the portion of nodal probability at the upper node going into the preceding node at the previous time step (down weight). For a standard binomial tree, $w_{i,j} = \frac{j}{i}$ (linear function). In a generalized implied binomial tree, $w_{i,j}$ is an arbitrary function; it is determined so as to fit the intermediate maturity options. Jackwerth reports that concave weight functions explain the observed European index on the S&P 500 option prices better than either linear, or convex weight functions. A concave weight function implies that a path going first down and then coming up, is more likely to be taken than a path going first up and then coming down.

Given the nodal probabilities and stock prices at time i, he can solve for the nodal probability and stock price at the preceding node in three steps:

1. $P^{nodal}_{i-1,j} = (1 - w_{i,j})P^{nodal}_{i,j} + w_{i,j+1}P^{nodal}_{i,j+1}$

2. $p = P_{i-1,j} = w_{i,j+1} \frac{P^{nodal}_{i,j+1}}{P^{nodal}_{i-1,j}}$

3. $S_{i-1,j} = [(1 - P_{i-1,j})S_{i,j} + P_{i-1,j}S_{i,j+1}]/(r/\delta)$,

where r and δ are the interest rate and dividend yield per step, and p is the transition probability. Note, that as long as the weights are between 0 and 1, the transition probabilities will also be between 0 and 1.

Jackwerth's technique recognizes that the evolution of nodal probabilities throughout any tree (standard, or implied binomial or trinomial tree) is governed by a transition probability weight. Changing the functional form for this weight, changes the nodal probabilities, i.e. changes the transition probabilities and the local volatilities. By implying the functional form of this weight from European option prices, one can change the nodal probabilities so as to price shorter term options consistently with the shorter and longer term European option prices.

3.5 Trinomial Trees

Implied trinomial trees are proposed as a solution to the problem of not-acceptable transition probabilities occurring in Derman and Kani's implied binomial tree. Moreover, trinomial trees provide a much better approximation to the continuous time process than the binomial tree for the same number of steps. This is because there are three possible future movements over each time rather than two (see Clewlow and Strickland [24]).

Trinomial trees have more parameters than binomial trees. The constraints remain the same, i.e. matching the moments of the continuous process and of the discretized process, and matching the model's forward and option prices with the market's. Inevitably, in order to match the parameters with the constraints that we have, we have to select the state space in advance (see Derman, Kani, and Chriss [38]). Then, the transition probabilities between the nodes can be easily calculated from the constraints. The "freedom" to fix the state space, enables us to come up with acceptable transition probabilities. On the other hand, we need to be careful choosing the state space so as to fit the current smile.

Derman, Kani and Chriss [38] discuss the issue of constructing the state space when volatility varies significantly with time to expiration and strike, producing a skew. In such a case, the nodal spacing has to change significantly with time and stock level. The method that they propose for constructing the state space, consists of two steps. In the first step, they assume that interest rates and dividends are zero, and they build the state space which corresponds to a trinomial tree with constant volatility (this is also suggested by Dupire [43], [45]). Then, they modify the time spacing and subsequently the nodal spacing so as to capture the basic term and skew structures of local volatility in the market. In the second step, if there are any forward price violations in any of the nodes, they multiply all node prices by the growth factor $e^{r(r-\delta)t_i}$. This is equivalent to working with the futures price, rather than the asset price (as Barle and Cakici [7] proposed), and it will remove all forward price violations.

Having set the state space in advance, the problem of the construction of the implied tree is reduced to the calculation of consistent with the smile transition probabilities (implied transition probabilities). Dupire [44], [45] sketches a way for calculating them. His technique can be summarized as follows. The state prices (Arrow-Debreu prices) are implied by the market prices of European calls and puts (implied Arrow-Debreu prices). Then, the implied transition probabilities are calculated from fitting the smile by using simultaneously backward and forward induction.

To make Dupire's description concrete, assume that we observe the market prices of European calls and puts for any strike and maturity. Then, the theoretical price of a European call with strike price K and maturity date $n\Delta t$, in a complete market is given by

$$C(n\Delta t, K) = \sum_{j=-n}^{n} Q_{n,,j} \max(S_{n,j} - K, 0) \tag{14}$$

By taking $C(n\Delta t, K)$ as observed from the market, we have to invert somehow equation (14), so as to get $Q_{n,j}$ (implied state prices). Assume that we wish to compute the implied state prices for time step n. We start at the top node n at time step n. The price of a European call with strike price $S_{n,n-1}$ and maturity date $n\Delta t$ is

$$C(n\Delta t, S_{n,n-1}) = Q_{n,n}(S_{n,n} - S_{n,n-1}) \tag{15}$$

which can be rearranged to give the state price at node (n, n). We can compute the state prices for the nodes down to the middle of the tree in a similar way, i.e. for node (n, k) we compute $Q_{n,k}$ by choosing $K = S_{n,k-1}$.

We then compute the state prices for the lower half of the tree by starting from the bottom node of the tree and using puts. This is repeated for every time step in the tree and completes the calculation of $Q_{n,,j}$ for every (n, j). Notice that for the, consistent with the smile, evaluation of European options we do not need the transition risk-neutral probabilities, but only the implied state prices. However, transition probabilities are necessary for the evaluation of more complex options. We are going to show how to calculate them by using the already derived implied state prices.

Imagine that we are at node (n, j) and we have computed the transition probabilities and the state prices for all nodes above (n, j). We want to calculate the transition probabilities $p_{u_{n,j}}, p_{m_{n,j}}, p_{d_{n,j}}$ from node (n, j) to the upward, middle, and downward node at the next time level $n+1$. We have three unknowns, and therefore we need three equations. The three equations are given by

(a) the forward induction of the state prices:

$$Q_{n+1,j+1} = e^{-r\Delta t}(p_{d_{n,j+2}} Q_{n,j+2} + p_{m_{n,j+1}} Q_{n,j+1} + p_{u_{n,j}} Q_{n,j}) \tag{16}$$

(b) the backward induction for the price of the asset:

$$S_{n,j} = e^{-r\Delta t}(p_{d_{n,j}} S_{n+1,j-1} + p_{m_{n,j}} S_{n+1,j} + p_{u_{n,j}} S_{n+1,j+1}) \tag{17}$$

(c) the forward price of a one period bond:

$$1 = (p_{d_{n,j}} + p_{m_{n,j}} + p_{u_{n,j}}) \tag{18}$$

The first equation can be rearranged to give $p_{u_{n,j}}$ directly. We then solve the second and third equations for $p_{m_{n,j}}$ and $p_{d_{n,j}}$. The above procedure is repeated for every time step in the tree. The transition probabilities are calculated by using simultaneously forward and backward induction. Hence, the constructed trinomial tree fits the observed smile.

Derman, Kani and Chriss [38] use equations which are very similar to the ones that Derman and Kani [35] use for the construction of their implied binomial tree. They calculate the transition probabilities from the known option prices, asset prices and Arrow Debreu prices. These equations have been derived by applying backward and forward induction in a way which is similar to the one that Dupire sketched.

3.6 Implicit Finite Difference Schemes

Andersen [1], and Andreasen [2] construct implicit and semi-implicit (Crack-Nicolson) schemes, which are consistent with the equity option volatility smile. They employ these schemes so as to overcome the problem of negative transition probabilities that may be encountered with binomial and trinomial trees. Avoiding negative transition probabilities is feasible within implicit finite difference schemes because these schemes have better properties, in terms of stability and convergence, than binomial and trinomial trees (see Clewlow and Strickland [24]. Andersen also shows that stability is equivalent to having acceptable probabilities). This is expected because the negative probabilities arise in the binomial and trinomial tree from restricting the asset price process into a certain kind of evolution. On the other hand, the implicit schemes discretize the BS type fundamental PDE (see Geske and Shastri [52]).

Andersen incorporates the forward induction technique into the implicit finite scheme. This enables him to estimate the risk-neutral distribution from a set of option prices. Then, he backs out the local volatilities from the risk-neutral densities of the different maturities by solving a constrained quadratic problem. Andreasen, contrary to Andersen, directly extracts the local volatilities from the input implied volatilities. He achieves this by deriving an explicit formula that relates the surface of implied volatilities to the surface of local volatilities. Moreover, in addition to Dupire's forward equation that the option price satisfies, he derives forward equations that the "Greek" hedging ratios (i.e. delta, gamma, vega, theta) must satisfy, as well.

3.7 Maximum Entropy Methods

Another method for calculating the volatility coefficient of equation (1) from the observed standard European option prices, is based on the maximum entropy concept. The use of this idea is motivated by Rubinstein's [85] paper. There, he extracts the risk-neutral distribution by using a least-squares criterion in order to minimize the distance between an assumed distribution and the distribution to be derived from option prices. However, other measures of this distance can be used, as well.

Buchen and Kelly [19] first apply the method of Maximum Entropy in order to measure this distance. The entropy of a probability density p can be interpreted as a measure of "missing information" and it is defined as

$$S(p) = -\int_0^\infty p(x)\ln p(x)dx \tag{19}$$

The maximum entropy density is the one which maximizes $S(p)$ subject to the constraints that the derived density (a) integrates to one, and (b) prices correctly the observed set of

European options. Stutzer [97] also uses the maximum entropy idea to derive the risk-neutral distribution from the historical distribution of the asset price without using observed option prices.

Avellaneda et al. [4] derive the local volatility surface by generalizing the method to minimize the Kullback-Leibler relative entropy (or information distance), under the above constraints.. The Kullback Leibler relative entropy is defined as

$$S(p,q) = \int_0^\infty p(x) \ln[\frac{p(x)}{q(x)}]dx \tag{20}$$

where $q(x)$ is a prior density for the random variable X, representing prior knowledge about X. However, their volatility surface is not smooth, but it has spikes at the data points. This is because the maximum entropy method in general, does not impose any smoothness constraints.

4 Testing the Validity of the Deterministic Volatility Assumption

Smile consistent deterministic volatility models have theoretical and practical advantages. They are a simple extension of the BS model preserving its arbitrage pricing property. In addition, they achieve a cross-sectional fit of the observed for the different strikes and maturities option prices. Finally, they are easy to implement. However, the validity of the deterministic volatility assumption has to be investigated empirically, before concluding that these models are appropriate for option pricing and hedging purposes. We are aware of three papers which address this issue: Dumas, Fleming and Whaley's [41], Jackwerth and Rubinstein's [62], and Buraschi and Jackwerth[21] [11].

Dumas, Fleming and Whaley [41] assess the stability of the deterministic volatility function for the S&P 500 Index, by examining how well it predicts future option prices. They estimate, every week, various polynomial specifications of the volatility function. Their estimation is performed by minimizing the sum of squared deviations of theoretical option prices from the observed market option prices. Then, they examine the price deviations from theoretical values one week later. The deterministic volatility model can always fit the cross-section of observed option prices, as long as the volatility function is complex enough. However, their out-of-sample results indicate that the instantaneous volatility function is not stable over time.

Jackwerth and Rubinstein [62] compare the out-of-sample empirical performance of alternative models, including Jackwerth's [63] generalized implied binomial tree, in terms of the pricing and hedging errors. They find that both generalized binomial trees and stochastic volatility models outperform the BS model. However, the size of the standard deviation of the pricing errors makes it difficult to conclude whether generalized binomial trees are better, or worse than stochastic volatility models.

[11] Related empirical studies by Bates [12], and Bakshi, Cao, and Chen [6] compare different option models, without looking explicitly at smile-consistent deterministic volatility models.

Buraschi and Jackwerth [21] rather than exploring the empirical performance of deterministic volatility models being based on the size in dollars of the pricing and hedging errors, they provide a general statistical test of deterministic volatility models versus stochastic volatility ones. They test directly the implication of deterministic volatility models that options are redundant securities. Their null hypothesis is that the payoff of any asset can be replicated with a dynamic trading strategy that involves two primitive assets, such as the underlying asset. They construct their tests from the properties of the implied risk-neutral density. Their tests reject the null hypothesis. The results suggest that the returns of the in-and out-of-the-money options are needed for spanning purposes. This finding is even stronger in the postcrash period.

The above-mentioned empirical studies indicate that the instantaneous volatility is not a deterministic function of the asset price and of time. Therefore, the asset price does not behave according to the inferred, from the implied deterministic volatility models, process. In this case the hedging will not be effective. As a solution to this Dupire [43] proposes a method of hedging which is robust, no matter what the dynamics of the asset price look like. He hedges a claim against movements in the volatility by using a portfolio of European options. This portfolio is rebalanced periodically, so as a change in the volatility surface will change the value of both the targeted claim and of the portfolio by the same amount. Bates [13] also provides a simple non-parametric method for inferring the deltas and gammas from the implied volatility patterns. His method is based on the assumption that the underlying asset price follows a stochastic process with constant returns to scale, so as option prices are homogeneous of degree one in the underlying asset price and strike.

Despite these suggestions, the empirical evidence implies that deterministic volatility models have to recalibrated every day, so as to fit the smile (Gemmill [51], Jackwerth and Rubinstein [61]). Therefore, they do not offer a unified theory of volatility which can be used for the pricing and hedging of exotic options. S*mile-consistent stochastic volatility models* have been developed in order to provide such a theory.

5 Smile Consistent No Arbitrage Stochastic Volatility Models

The development of the stochastic volatility literature is similar to the evolution of the interest-rate literature. In the latter, there was a transition from equilibrium considerations (e.g. Vasicek [101], Cox, Ingersoll, Ross [32]) to arbitrage arguments (e.g. Ho and Lee [57], Heath, Jarrow, Morton [56], (HJM)) (for a survey of the continuous time interest rate literature, see Strickland [96]). In that context, the aim is not to explain the yield curve, but taking it for granted and along with evolution assumptions to obtain arbitrage prices for derivative securities. In the stochastic volatility literature, first models which start by assuming a stochastic process for the evolution of the instantaneous volatility were introduced. Then, models which start from today's European option prices and achieve option pricing and hedging of more complicated claims under a consistent with no-arbitrage evolution of the volatility surface, were developed.

5.1 Dupire (1992)

Dupire's [42] approach is similar to HJM's. He starts from today's European option prices, and he derives the process of the instantaneous volatility "endogenously" from the process of the "forward" volatility (see below for his definition of forward volatility).

Let the financial instrument f_T delivering $f(S_T)$ at time T. Without loss of generality, he assumes the interest rate to be zero at all times and he also assumes the absence of any arbitrage opportunities in the market. Let (Ω, I, I_t, P), be a filtered probability space, where (I_t) is a right continuous filtration, and P is the objective probability measure. Let now P_T be the P-equivalent probability measure. Then, the value at time 0 of f_T is given by (see Dothan [40])

$$f_T(0) = \int f(S_T)\Phi_T(S_T)dS_T \equiv E^{P_T}[f(S_T) \mid I_0] \tag{21}$$

where $\Phi_T(S_T)$ has been extracted from the known prices of European call options (equation (2)) and the expectation is taken with respect to P_T (risk-neutral measure). Hence, in order to price any instrument, we have to get the risk-neutral process of the underlying asset. In order to get the risk-neutral process for the spot he assumes that

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_{1,t} \tag{22}$$

where W_1 is a P-Brownian motion adapted to F_t and μ_t and σ_t can be measurable processes themselves adapted to F_t. By defining $d\overline{W_{1,t}} \equiv dW_{1,t} + \frac{\mu_t}{\sigma_t}dt$, Dupire gets

$$\frac{dS_t}{S_t} = \sigma_t d\overline{W_{1,t}} \tag{23}$$

where $\overline{W_1}$ is a Brownian motion under P_T, obtained by Girsanov's theorem (see Oksendal [81] for a description of Girsanov's theorem). Equation (23) is the risk-neutral process for the Spot. We can use it to price options by Monte-Carlo simulation (see Boyle [17]). However, in order to price options consistently with the market, the process for instantaneous volatility σ_t must incorporate information from today's option prices. Moreover, it should respect the evolution of the implied volatility surface. For pricing purposes, we need the risk-neutral process for such an evolution of the volatility (this point is explained in Hull [58]). Then, we will be able to plug it into equation (23) and achieve smile-consistent pricing.

Applying Ito's lemma to $lnSt$, using equation (22), integrating the result between T_1 and T_2, and rearranging terms we get

$$\int_{T_1}^{T_2} \sigma_t^2 dt = 2\int_{T_1}^{T_2} \frac{dS_t}{S_t} - 2(\ln S_{T_2} - \ln S_{T_1}) \tag{24}$$

The left hand side (LHS) of equation (24) can be thought of as being the payoff at time T_2 of a forward contract which is traded at time t. Assuming that $T_2 - T_1 = \varepsilon$, Dupire defines as V_T this forward contract which delivers the instantaneous variance to be observed at time

T^{12}. By interpreting the stochastic integral on the right hand side of equation (24) as the gain or loss from a self financing strategy, results in the integral to vanish[13]. Let L_T be a claim delivering the logarithm of S_T at time T (log-contract). Its price at time 0 is[14]

$$L_T(0) = E^{P_T}[\ln S_T \mid I_0] \tag{25}$$

So, the payoff of a contract traded on the future volatility, is equal to the payoff of a self-financing strategy of buying and selling today's log-contracts of maturities T_1 and $T_1+\varepsilon$. Then, their prices must be the same, if no arbitrage is to exist. By dividing both sides by ε, and taking the limits, so as $\varepsilon \to 0$, we get by arbitrage pricing that the value of the forward contract V_T at any time $t < T$ is

$$V_T(t) = -2\frac{\partial L_T(t)}{\partial T} \tag{26}$$

This is the point where Dupire defines implicitly as $V_T(t)$ the instantaneous forward variance (IFV) to be observed at time T. The IFV $V_T(t)$, is defined as the value at time t of a forward contract which will deliver the instantaneous variance to be observed at some time T in the future[15].

From the values $(L_T(0))_T$, we can deduce the initial instantaneous forward variance curve as

$$V_T(0) = -2\frac{\partial L_T(0)}{\partial T} \tag{27}$$

Next, he models the forward variance which automatically ensures compatibility with $(L_T(0))_t$ as equation (27) shows. Consequently, this will ensure compatibility with the current volatility surface, as equations (21) and (25) show. He makes among other possible choices, the assumption that $V_T(t)$ is lognormal i.e.

$$\frac{dV_T(t)}{V_T(t)} = adt + bdW_{2,t} \tag{28}$$

where a and b are constant or a deterministic function of time and W_2 is another Brownian motion adapted to F_t, *possibly correlated with* W_1. Defining $d\overline{W_{2,t}} = dW_{2,t} + \frac{a}{b}dt$, equation (28) can be rewritten as

$$\frac{dV_T(t)}{V_T(t)} = bd\overline{W_{2,t}} \tag{29}$$

[12]The assumption of the existence of a forward contract which is traded on volatility is not unrealistic. Over-the-counter futures and options contracts on foreign-currency and interest-rate volatility indexes, are currently being developed by a number of investment banking firms in the U.S. and Europe (see Grunbichler and Longstaff [54]).

[13]For the definition of a self-financing strategy, see Dothan [40].

[14]Notice that in order to price the log-contract consistently with the market, we need the implied risk-neutral density function. This explains why Dupire needs to take the observed option prices as granted.

[15]Stewart Hodges and Alessandro Rossi, have pointed out that by definition, $V_T(t) = V_T(t, S_t)$. This is because

$$\begin{aligned} V_T(t) &= E^{Q_2}[-2(\ln S_{T+\varepsilon} - \ln S_T) \mid I_t] \\ &= E^{Q_2}[-2(\ln S_{T+\varepsilon} - \ln S_T)] \end{aligned}$$

since the increments of a Brownian motion are orthogonal and this holds true for the increments of any function of the Brownian motion.

where $\overline{W_2}$ is a Brownian motion under Q_2, the P-equivalent probability measure obtained by Girsanov's theorem. Equation (29) is the risk-neutral process for the instantaneous forward variance. Using equation (29), he derives the risk-neutral process for the instantaneous volatility. That is

$$\frac{d\sigma_t}{\sigma_t} = (\frac{1}{2}\frac{\partial \ln V_t(0)}{\partial t} - \frac{b^2}{8})dt + \frac{b}{2}d\overline{W_{2,t}} \tag{30}$$

Equation (30) is probably the most important contribution of this paper. The first term of the drift of equation (30) can be estimated from the prices of the log-contracts, as equation (27) shows, and the second term can be estimated from the initial forward volatility surface. Therefore, the drift of (30) ensures consistency with the initial (implied) volatility term structure. In addition, $\overline{W_{2,t}}$ moves stochastically the volatility surface. In this way he hopes that his model will fit better the smile every day[16]. Solving equation (30) and plugging it into equation (23) delivers the risk-neutral process for the spot.

As a final step, he assumes that the filtration associated with $\overline{W_1}$ and $\overline{W_2}$ conveys the same information, and therefore is the same. This is a reasonable assumption because the filtration associated with $\overline{W_1}$ is used for the pricing of the log-contract; the log-contract is used for the construction of the forward variance and hence of $\overline{W_2}$. Define Q (through Girsanov's Theorem) as the P-equivalent measure, under which $\overline{W_1}$ and $\overline{W_2}$ are Q-Brownian motions. Then, the price at time t, $h(t)$ of an instrument h that delivers at time T a payoff $h(S_T)$, is

$$h(t) = E^Q[h(S_T) \mid F_t] \tag{31}$$

where F_t is the natural augmented filtration associated with W_1 and W_2. The expectation in equation (31) can be evaluated consistently with the smile by "joint" Monte Carlo (see Boyle [17]) of the risk-neutral process of the spot (equation (23)) and of the instantaneous volatility (equation (30)). Monte Carlo simulation allows us also to compute the hedging parameters through a small shift of the paths.

To summarize, Dupire starts by assuming that European Calls of all strikes and maturities are traded, and that their market prices are consistent with no arbitrage. From these prices, he deduces the arbitrage price $L_T(t)$ of contingent claims (log-contract) that promise lnS_T at date T (the log-contract is not traded in the listed market, but it can be approximated by a combination of standard listed options (see Derman et al. [37]). These prices $L_T(t)$ permit to synthesize the value of a forward market on the instantaneous variances to be observed at any maturity, as equation (26) shows[17]. Next, he assumes a one factor model for the forward variance and he derives the risk neutral process for the instantaneous variance. Using the risk-neutral processes for the spot and the instantaneous volatility, he obtains arbitrage free prices that do not depend on any risk premia, nor on a volatility drift. They do depend on the term structure of the IFV, on the correlation between spot and IFV, and on the volatility of the latter, as equation (23) combined with equation (30) show.

---

[16] This is analogous to the form of the drift of the short-term interest rate that Heath,Jarrow,Morton [56] derived by starting from the process for the instantaneous forward variance. That drift depends among other things and on the initial instantaneous forward rate which reflects the current yield curve.

[17] Neuberger [80] shows that the log-contract is very useful in order to hedge against volatility. He does this by delta-hedging log-contracts on futures. Then, he regresses the percentage hedging error on the squared outcome volatility, and he gets a correlation coefficient of 99.99%. That is to say that by trading a portfolio of log-contracts, we can replicate (and therefore eliminate) the variance.

A serious limitation of the model is that it recognizes the strike and term structure of implied volatilities for only the current time, and not for future times, as the volatility surface evolves stochastically. The two dimensional information $(C_{K,T}(0))_{K,T}$ has been compacted at each time into a 1-dimensional $(L_T(0))_T$ in the process. In order to conclude about the performance of Dupire's model over the "traditional" stochastic volatility models (e.g. Hull and White [59]), the models have to be implemented and compared.

5.2 The Definition of the Forward Variance Revisited

Even though Dupire [42] introduced the idea of the forward variance (some people call it local), a formal treatment of the forward volatility concept was not done until 1996 in Kani, Derman and Kamal's [70] paper. There, they assume that the risk-neutral index evolution is governed by the following equation

$$\frac{dS_t}{S_t} = (r_t - \delta)dt + \sigma_t d\overline{Z_t} \tag{32}$$

where r_t is the riskless rate of return at time t, assumed to be a deterministic function of time, δ is a continuous compounded dividend yield, and σ_t is the instantaneous index volatility at time t, assumed to follow some unspecified stochastic process.

The forward (local) variance $\sigma^2_{K,T}$, corresponding to level K, and maturity T, is defined as the conditional expectation of the instantaneous variance of index return at the future time T, contingent on index level S_T being equal to K, i.e.

$$\sigma^2_{K,T}(t,S) = E^*_t(\sigma^2_T \mid S_T = K) \tag{33}$$

where the expectation has been taken with respect to the risk-neutral measure and the subscript t indicates that the expectation is based on information at time t.

The local volatility $\sigma_{K,T}$, is defined as the square root of the local variance, $\sigma_{K,T} = (\sigma^2_{K,T})^{\frac{1}{2}}$. Hence, the local volatility is defined as the forecast (estimate) of index volatility at a particular future time and market level, so as to make current option prices fair. It is worth noting that the concept of the local volatility is different to the concept of the implied volatility $\Sigma_{K,T}$ which can be thought of as the market's estimate of the expected average future index volatility during the life of the option[18].

[18]The statement that implied volatilities are the estimates of the (expected) average future index volatility is rather loose. It can be proved formally only for the case that (a) we are in a Hull-White world of stochastic volatility, and (b) we deal with at-the-money options. To make things more concrete, Hull and White [59] show that when the volatility risk is not priced and the correlation between the underlying asset and the volatility is zero, then the price P of an option is given by

$$P = \int BS(\overline{V_t})h(\overline{V_t} \mid I_t)d\overline{V} = E[BS(\overline{V_t}) \mid I_t]$$

where $BS(.)$ is the BS pricing formula, $\overline{V} = \frac{1}{T-t}\int_t^T V_s ds$. $h(\overline{V_t} \mid I_t)$ is the density of $\overline{V_t}$ conditional on the the information set at time t I_t, and T is the expiration date of the option. Moreover, Cox and Rubinstein [31] show that the BS formula is approximately a linear function of the standard deviation for at-the-money options, so that $E[BS(\overline{V_t}) \mid I_t] = BS[E(\overline{V_t}) \mid I_t]$.

Note, that in the case where the instantaneous index volatility is assumed to be deterministic as a function of the index level and time (deterministic volatility case) i.e. $\sigma_T = \sigma(S_T, T)$, equation (33) becomes

$$\sigma^2_{K,T}(t,S) = E_t(\sigma^2_T \mid S_T = K) = E_t[\sigma^2(S_T,T) \mid S_T = K] = \sigma^2(K,T) \tag{34}$$

Equation (34) shows that the forward variance equals the instantaneous variance, when the latter is assumed to be deterministic. That is to say that the smile-consistent stochastic volatility case, nests the deterministic volatility one. In this case we are left with a static local volatility surface, whose shape remains unchanged as time evolves since the right hand side is independent of t and S. This is analogous to the result that in a deterministic interest rates economy, the forward rates do not change over time (see Rebonato [82]).

Derman and Kani [39] show that if the asset price evolves according to equation (32), then the forward variance is given by

$$\sigma^2_{K,T}(t,S) = \frac{2\{\frac{\partial C_{K,T}}{\partial T} + (r-\delta)K\frac{\partial C_{K,T}}{\partial K} + \delta C_{K,T}\}}{K^2\frac{\partial^2 C_{K,T}}{\partial K^2}} \tag{35}$$

Equation (35) shows that the local volatility can be locked in by trading portfolios of currently available options. Hence, it can be thought as the market price of volatility we can lock in today, in order to obtain volatility exposure over some specific range of future times and market levels. Local volatility is the volatility analog of the forward rate. The forward rate is the future rate prevailing in a time interval, so as the current yields to maturity are justified, and it can be locked in by trading current bond portfolios[19].

The existence of the risk neutral measure is used for the expectation definition of local volatility, and it also allows for an alternative definition of the local volatility. In order to show this, Derman, Kamal and Kani [37] assume that under the risk neutral measure, the local volatility evolves according to the following process

$$\frac{d\sigma^2_{K,T}}{\sigma^2_{K,T}} = a_{K,T}dt + \beta_{K,T}d\overline{Z_t} + \sum_i \vartheta^i_{K,T}d\overline{W^i_t} \tag{36}$$

where the instantaneous volatility σ_t is equal to the instantaneous local volatility at time t and level S_t, i.e. $\sigma_t = \sigma_{S,t}(t,S)$. Notice also as $\sigma^2_{K,T}$ depends on the shock of the asset process, i.e. the stochastic variations of the local volatility surface may depend on the prevailing level.

Let $P_{K,T} = P(t, S_t, K, T)$ be the total probability that the index level at time T arrives at $S_T = K$ given that the stock price at time t is S. Both the stock price and volatility are assumed to evolve stochastically. Knowing that under the risk neutral measure $P_{K,T}$ evolves as a martingale, i.e.

$$\frac{dP_{K,T}}{P_{K,T}} = \phi_{K,T}d\overline{Z_t} + \sum_i X^i_{K,T}d\overline{W^i_t} \tag{37}$$

[19] Dupire [46] has also derived equation (35) independently, for the case that $r = \delta = 0$. He defines the forward variance $V_{K,T}(S_0, t_0)$ as the price of a forward contract $V_{K,T}$ introduced at time t_0. The difference with his earlier definition (Dupire [42]) is that the contract will exchange at time T, the instantaneous variance observed at time T $v(S_T, T)$, against an agreed at time t_0 amount $V_{K,T}(S_0, t_0)$ if and only if $S_T = K$. If $S_T \neq K$, no exchange takes place. His definition is an alternative way of interpreting equation (33).

they show that

$$\frac{d\sigma_{K,T}^2}{\sigma_{K,T}^2} = \beta_{K,T}\widehat{dZ_t} + \sum_i \vartheta_{K,T}^i \widehat{dW_t^i} \tag{38}$$

where $\widehat{dZ_t} = dZ_t - \phi_{K,T}dt$ and $\widehat{dW_t^i} = dW^i - X_{K,T}^i dt$ are the Brownian motions with respect to the new measures. In other words, under these measures, the local variance is a martingale. They call the new measure *the K-level, T-maturity forward risk-adjusted measure.* Letting $E_t^{K,T}(.)$ to denote expectations with respect to this measure, conditional on the information at time t, equation (33) can be rewritten as

$$\sigma_{K,T}^2 = E_t^{K,T}(\sigma_T^2) \tag{39}$$

Hence, in the $K-T$ forward risk-adjusted measure, the local variance $\sigma_{K,T}^2$ is the conditional expectation of the future instantaneous variance σ_T^2[20].

5.3 Derman and Kani (1998)

Derman and Kani [39] use the concept of local volatility, as this is defined by Kani, Derman and Kamal [70], to present a method for option pricing and hedging based on it. They follow a methodology similar to Dupire's [42], by starting from the initial set of index option prices and their associated local volatility surface. They assume that the asset price and the forward volatility evolve according to equations (32) and (36), respectively.

The drift coefficients $a_{K,T}(t,S)$ in equation (36) must satisfy mild measurability and integrability conditions, as the factor volatility $\vartheta_{K,T}^i$. Moreover, they must also be restricted in such a way, so as the stochastic theory described by equations (32) and (36), precludes any arbitrage opportunities among the standard options, forwards and their underlying stock. This is similar to the HJM approach, where the drift of the instantaneous forward rate had to be constrained by a no-arbitrage condition, so as forward rates to evolve in a no-arbitrage fashion.

In order to derive this no-arbitrage restriction, they work with the total transition probability $P_{K,T}(t,S)$[21]. They show that the drift functions $a_{K,T}(t,S)$ have to satisfy the following no-arbitrage condition

$$\begin{aligned} a_{K,T}(t,S) &= -\sum_{i=0}^{n} \vartheta_{K,T}^i(t,S)\{\frac{1}{P(t,S,T,K)} \\ &\int_t^T \int_0^\infty \vartheta_{K',T'}^i(t,S)P(t,S,T',K')K'^2\frac{\partial^2}{\partial K'^2}P(T',K',T,K)dK'dT' - \Pi^i\} \end{aligned} \tag{40}$$

where the quantities Π^i denote the market prices of risk associated with the volatility risk factors W^i, $i = 1,\ldots,n$ and $W^0 = Z$. The double integral of equation (40) reflects the two dimensional dependence of the local volatility on K and T and it makes the continuous

[20] This is analogous to the relationship between the forward and the future short rate. The forward rate f_T is the T-maturity forward risk-adjusted expectation of the future short rate at time T (see Jamshidian [65]).

[21] Working with total transition probabilities is equivalent to working with option prices, as equation (2) shows.

time implementation of their model very difficult. Therefore, they prefer implementing their model in discrete time by building a trinomial *stochastic implied tree.*

A stochastic implied trinomial tree is an extension of the implied trinomial tree. The local volatilities and the transition probabilities which correspond to the future nodes, vary stochastically as time elapses and the index level moves. The stochastic tree is the output of an algorithm which combines a standard trinomial implied tree and Monte Carlo simulation. Four steps are necessary for the construction of the stochastic tree. First an implied trinomial tree, like Derman, Kani and Chriss's [38] is built. This provides the initial inputs to the algorithm. Second, the extracted from the implied trinomial tree initial local volatility surface is perturbed. This is done by simulating a discretized version of equation (36) which can be written as

$$\Delta\sigma^2_{m,n}(i,j) = \sigma^2_{m,n}(i,j)[a_{m,n}(i,j)\Delta t_i + \sum_{l=0}^{n} \vartheta^l_{m,n}(i,j)\Delta W^l_i] \tag{41}$$

where (i,j) corresponds to the node (t_i, S_j), describing the current location of the stock at the *ith* step of the simulation. (n,m) labels the future node (t_n, S_m) corresponding to the future time and level (T,K). The evolution of the local volatility surface has to be consistent with no-arbitrage. This is achieved by determining the drift parameters $a_{m,n}(i,j)$ from a "martingale condition" (the volatility parameters $\vartheta_{m,n}(i,j)$ are pre-specified by applying for example Principal Components Analysis). This condition is that the total probabilities $P_{m,n}(i,j)$ of arriving at the future node (n,m) from the (fixed) initial node (i,j) must be jointly martingales for all future nodes (n,m). Once $a_{m,n}(i,j)$ are determined by $P_{m,n}(i,j)$ a random vector $(\Delta W^0_i, \Delta W^1_i, \ldots, \Delta W^n_i)$ is drawn from the distribution of the increments of W^i at time t_i. This vector is used to arrive at new values for all future local volatilities $\sigma^2_{m,n}(i+1,j)$.

In the third step, $\sigma^2_{m,n}(i+1,j)$ are used for the calculation of the one period transition probabilities. These probabilities describe the transition from node (i,j) to the up, middle and down nodes at time t_{i+1}. The transition probabilities are calculated from the condition that they must add up to one, and from the conditions which match the first two moments of the discretized process with those of the continuous process. In the fourth step, a new location S_j for the asset price at time t_{i+1}, is determined. This is done by comparing the random draw ΔW^l_i to the calculated from the third step one period transition probabilities. The third and fourth steps establish the dependence of the evolution of the asset price on the evolution of the local volatilities.

Therefore, the key idea for the construction of the stochastic tree, is to perturb the local volatility surface, so as to exclude arbitrage. Once one of the possible positions of the surface is determined, the stochastic volatility problem is reduced to a deterministic one. Then, the consistent with the new surface, asset price can be traced. Notice that since the state space is fixed for every step of the simulation, it may be the case that large simulated local volatilities produce unacceptable transition probabilities at certain nodes. In this case, the unacceptable probabilities are overwritten and current option prices may not be fitted exactly, as discussed in Derman and Kani [35], and Derman, Kani and Chriss [38]. However, Derman and Kani find that for their stochastic implied tree the overwriting is rarely encountered.

In general, the stochastic implied tree model is a very promising one, and future research should (a) explore the empirical performance of the model, and (b) investigate the number and the nature of the shocks appearing in the local volatility process.

5.4 Extracting the Local Volatility Surface

The local volatility surface implied from today's option prices is necessary for initializing Derman and Kani's algorithm. They extract it by building an implied trinomial tree. We have already seen that maximum entropy methods can also be used for this purpose. In this section we are going to present alternative methods for inferring the local volatilities.

Equation (35) shows that in principle, we can extract the local volatilities from the option prices. This requires a smooth surface of option prices, so as the partial derivatives appearing in equation (35) to be evaluated. In other words, we need to interpolate between the observed strikes and maturities option prices, and to extrapolate beyond them. However, it is always easier to interpolate in the space of implied volatilities, rather in the space of option prices, because the former is smoother than the latter. One way that this can be carried out is by using Shimko's [94] method. The method can be described as follows:

1. Convert the observed option prices to implied volatilities.
2. Use a least squares regression to estimate a quadratic volatility smile.
3. Convert the estimated implied volatilities back into option prices.

To be more precise, Shimko's method is an estimation, rather than an interpolation and the estimated function of option prices does not necessarily go through the original prices. Hence, the extracted local volatilities may not be consistent with today's smile.

Andersen [1] and Andreasen [2] interpolate and extrapolate in the space of implied volatilities, as well, by using a different method from Shimko's. Andersen derives the local volatility surface by using the implicit finite difference scheme that we have already described. Andreasen derives a formula which explicitly expresses the local volatilities in terms of the strike and maturity partial derivatives of implied volatilities. Then, he approximates these derivatives by performing a bicubic spline procedure, so as to achieve smoothness in both directions of the implied volatility surface.

Derman, Kani and Zou [36], and Zou and Derman [103] point out that the calculation of the second derivative is very sensitive to the interpolation method (see also Andreasen [2]). Rather than interpolating, they recognize the property that the second derivative of a European call option with respect to the strike price is a probability density function and they approximate it with an Edgeworth expansion. This is similar to Rubinstein [87], who approximates the expiration-date risk neutral distribution by means of an Edgeworth expansion in a discrete-space framework, assuming that the investor knows its skewness and kurtosis. The idea behind an Edgeworth expansion is that the probability distribution function can be expressed in terms of another known distribution plus its derivatives, provided that all moments of both distributions exist (it is similar to the Taylor expansion for analytical functions)[22]. This expansion extracts the asset risk-neutral probability distribution from prices of options that expire at a fixed time. The time derivative is calculated by interpolating the volatility term structure between options' expiration dates by means of a cubic

[22] Jarrow and Rudd [66] were the first to apply the Edgeworth expansion method to the problem of option valuation.

spline. However, it is not certain that even their method is robust because the Edgeworth expansion does not always represent a proper probability density function; there are many intervals for which it could take negative values (see Johnson and Kotz [67]).

5.5 Britten-Jones and Neuberger (1999)

A limitation of Derman and Kani's [39] algorithm is that it is quite computer intensive. As a solution to this, Britten-Jones and Neuberger [20] provide a trinomial lattice procedure which is simple and fast. Their aim remains the same: smile consistent arbitrage pricing under stochastic volatility. In order to construct their trinomial tree they establish a sufficient and necessary condition for a process to be consistent with a set of initial option prices.

To make the above more concrete, let time $t = 0, h, 2h, \ldots, T$. They define a grid consisting of nodes of time-price events (t, K) where $K = \{k : k = S_0 u^i, i = 0, \pm 1, \pm 2, \ldots, \pm T/h\}$, and $u > 1$. The single risky underlying asset has initial price S_0, and can move on this grid. Moreover, define

$$\lambda(k,t) \equiv \frac{1}{h}\frac{(1+u)(C(t+h,k) - C(t,k))}{C(t,ku) - (1+u)C(t,k) + uC(t,k/u)}$$

and

$$\pi(k,t) \equiv \frac{C(t,ku) - (1+u)C(t,k) + uC(t,k/u)}{k(u-1)}$$

Their first proposition states that the terminal probability that the asset price takes the value k, is given by

$$\Pr(S_t = k) = \pi(k,t)$$

for $0 \leq t \leq T$, $k \in K$. Their second proposition states that if the transition probability for a continuous martingale process is given by

$$\Pr[S_{t+h} \neq S_t \mid S_t = k] = h\lambda(K,t)$$

then this is a sufficient and necessary condition for this process to fit today's option prices. Furthermore, they show that their second proposition is equivalent to extracting the local volatility from equation (35).

Consider now a martingale process in which the probability of stock price moves depends not only on time and the stock price, but also on the volatility state Z, where $z = 1, 2, \ldots, N$. Z is assumed to evolve as a time homogeneous Markov process with transition matrix $P = p_{jk}$, where $\Pr[Z_{t+h} = j \mid Z_t = k, F_t] = p_{jk}$. The transition probabilities are chosen exogenously to reflect the type of volatility behavior we think appropriate. Define then the asset transition probabilities as

$$\Pr[S_{t+h} \neq S_t \mid S_t = k, Z_t = z] \equiv h\lambda(k,z,t)$$

We need to calculate $\lambda(k,z,t)$ in order to apply the backward induction for option pricing and hedging (recall that we do not have to worry about the determination of the state space, because it is set exogenously in a trinomial tree). In order to ensure that the model's option prices fit the market option prices, $\lambda(k,z,t)$ has to be calculated consistently with the European option market prices, i.e. their second proposition has to be satisfied. They show that the proposition holds true if the $\lambda(k, z, t)$ satisfy

$$\lambda(k,t)\pi(k,t) = \sum_{z=1}^{N} \lambda(k,z,t)\pi(k,z,t) \tag{42}$$

where $\pi(k, z, t) \equiv \Pr(S_t = k \text{ and } Z_t = z)$ is the joint probability of a particular stock price and volatility state. Notice that the left-hand-side of equation (42) can be calculated from today's option prices, while the right-hand side is unobserved.

They calculate $\pi(k, z, t)$ by using the forward Kolmogorov equation. Then, in order to calculate $\lambda(k, z, t)$ they assume that

$$\lambda(k,z,t) = q(t,k)v(z)$$

This is a separability assumption, where the function $v(z)$ is chosen exogenously. The form of the volatility process is defined by the transition probabilities p_{jk}, and the $v(z)$ function which serves for moving local volatilities in a parallel fashion. Then,

$$q(t,k) = \frac{\lambda(k,t)\pi(k,t)}{\sum_{z=1}^{N} v(z)\pi(k,z,t)}$$

This completes the calculation of $\lambda(k, z, t)$.

Britten-Jones and Neuberger's model provides a simple and fast implementable algorithm for smile-consistent pricing under stochastic volatility. It does not have to specify a priori any process for the forward volatility. On the other hand, the separability assumption is crucial for the development of the algorithm, and its implications should be explored further. Finally, the determination of the transition probabilities p_{jk} and of the $v(z)$ function may pose problems to the practitioner.

5.6 Ledoit and Santa-Clara (1998)

In general, option pricing and hedging can be done by Monte Carlo simulation if we have the risk-neutral processes for the underlying (traded and non-traded) assets. In Derman and Kani's model [39] this was not possible because of the very complicated no-arbitrage condition that the drift of the local volatility process had to satisfy. Ledoit and Santa-Clara's [71] model overcomes this problem by modeling implied rather than local volatilities.

They start from the observed at time t implied volatilities $V(t, T, X)$ corresponding to an option with time to maturity $T \equiv s - t$, and moneyness $X \equiv \frac{S_t}{K}$, and they allow them to evolve stochastically. In order to simulate jointly the processes for the underlying asset and the implied volatility, they establish a relationship between the stock's instantaneous volatility σ_t and the implied volatility V. Under the assumption that the stock price evolves according to equation (32), they show that the Black-Scholes implied volatility of an at-the-money call option converges to the stock price volatility when the time to maturity goes to zero, i.e. $\sigma_t = V(t, 0, 1)$. Then, they deduce the dynamics of implied volatilities and of the call option price, by applying Ito's lemma on $V(t, T, X)$ and $C(t, S_t, V)$. The evolution of the call option prices has to preclude any arbitrage opportunities, i.e. call option prices have to evolve as martingales. Imposing this martingale constraint they get the risk-neutral process for the implied volatility, i.e.

$$dV(t,0,1) = [\sigma_{V1}(t,0,1)(r-q-\frac{1}{2}V(t,0,1)^2) - \frac{\partial \sigma_{V1}}{\partial X}(t,0,1)V(t,0,1)^2$$
$$+2\frac{\partial V}{\partial T}(t,0,1) - \frac{\partial V}{\partial X}(t,0,1)V(t,0,1)(\sigma_{V1}(t,0,1)+\frac{1}{2}V(t,0,1))$$
$$-\frac{1}{2}\frac{\partial^2 V}{\partial X^2}(t,0,1)V(t,0,1)^2]dt$$
$$+V(t,0,1)\sigma_{V1}(t,0,1)d\overline{W_{1,t}} + V(t,0,1)\sigma_{V2}(t,0,1)d\overline{W_{2,t}} \tag{43}$$

where σ_{V1}, σ_{V2} are the volatilities of the implied volatilities, and $\overline{W_{1,t}}, \overline{W_{2,t}}$ are the, under the risk-neutral measure, two shocks driving implied volatilities.

Even though the risk-neutral process for the implied volatility seems to be relatively simple for a continuous time implementation, there are three issues that needs to be addressed in order to implement their model. First, they assume that the implied volatility process is driven by two shocks. This is an issue that it has to be explored empirically. Second, the coefficients $\sigma_{V1}(t,0,1)$ and $\sigma_{V2}(t,0,1)$ have to be estimated for every t, so as to be used in every step of the simulation. Third, in order to calculate the partial derivatives which appear in equation (43), we have to interpolate the estimated volatilities of implied volatilities and the implied volatilities themselves, across moneyness and the time to maturity. Given that the functional form of the implied volatility surface (and probably of the shocks, as well) changes unpredictably over time, the choice of an appropriate interpolation scheme is a difficult task.

5.7 The Dynamics of Volatilities

In order to implement the smile-consistent no-arbitrage stochastic volatility class of models, we need to understand the dynamics of the implied volatility surface, as we have already pointed out. More specifically, three questions should be explored: (1) how many factors are needed to explain the dynamics of the volatility surface? (2) what do these factors look like? and (3) how are these factors correlated with the innovation in the underlying asset's process[23]?

Kamal and Derman [69] analyze the dynamics of the differences of implied volatilities over time, by applying Principal Components Analysis (PCA) to over the counter (OTC) S&P 500 and Nikkei 225 Index options. The changes in implied volatilities are indexed with the delta of the option (delta metric). They find that three Principal Components (PCs) explain about 95% of the variance of the volatility surface. Their interpretation is a level of volatilities for the first PC, a term structure of volatilities for the second PC and a skew for the third. Their results suggest that a four factor model for pricing and hedging options under a stochastic volatility "smile consistent no-arbitrage pricing" type model is an

[23]This also parallels work on implementing the HJM term structure model where Litterman and Scheinkman [72] and others, have applied Principal Components Analysis to analyze innovations in the yield curve.

appropriate one. One factor is needed for the underlying asset's process and three factors for the implied volatility process. In addition, they find that the first PC is negatively correlated with the indices returns, while the other two PCs are uncorrelated. The negative correlation can be said that it is consistent with the leverage effect (see Christie [23]), and the related regularity that implied volatilities are negatively correlated with the market index returns (Franks and Schwartz [48], Rubinstein [85], Schmalensee and Trippi [92]).

Skiadopoulos, Hodges and Clewlow [95] apply PCA to the implied volatility smiles and surfaces of the Futures options on the S&P 500. They use a different methodology than Kamal and Derman and they find markedly different results. They index the changes in implied volatilities in the moneyness metric. In both the smile and the surface analysis, two factors, which have similar shape, are identified. These explain about 60% of the surface variation. The first factor is interpreted as an essentially parallel shift and the second as a Z-shaped twist. Their results suggest that in order to implement a "smile-consistent no-arbitrage" stochastic volatility model for the pricing and hedging of futures options we need three factors. They caution though that the factor structure may have higher dimensionality since the factors explain only 60% of the surface variation. Portfolios of options modelled as riskless under a three or four factor models, may in fact exhibit substantial market risk.

Gallant, Chien and Tauchen [50] do a time series analysis of returns, and they find similar results to Skiadopoulos, Hodges and Clewlow. Using daily data on close-to-close price movements and the high/low spread, they find that a model with two stochastic volatility shocks plus the underlying asset component, fits the data very well, and in particular it mimics the long-memory feature of volatility.

The differences in the number and shape of shocks driving the dynamics of implied volatilities could be attributed to either one (or all) of the following: (a) the market futures option data have a noisier volatility structure than that estimated from the quotations for OTC index options, (b) the variables on which the PCA has been performed have been indexed with different metrics, (c) the dynamics of the volatility surface may simply depend on the choice of the underlying asset. This is analogous to the documented fact that the magnitude of observed implied volatility smiles, depends on the underlying asset (Fung and Hsieh [49], Tompkins [100]). Future research should apply PCA to different data sets under different metrics, so as to solve these issues.

6 Concluding Remarks and Suggestions for Future Research

This paper has reviewed the fast developing literature of smile consistent models. We have recognized the two classes of models which constitute this literature, so far. This allows several issues and models to be presented, compared and discussed. We classify the various smile consistent models in deterministic, and stochastic volatility ones. We note that the former is nested within the latter, and we bring together the several definitions of the key concept of forward volatility.

Both classes of models start from the observed European option prices, but they differ on the assumption about how the underlying asset evolves. Part of the popularity of deterministic volatility models, can be explained by their theoretical and practical advantages.

They preserve the Black-Scholes arbitrage-pricing property and they are tractable enough for implementation purposes. However, the empirical evidence undermines the deterministic volatility assumption. On the other hand, stochastic volatility models allow for the evolution of the (implied or forward) volatility surface, so as to preclude any arbitrage opportunities.

Future research should empirically test the stochastic volatility implied models. In particular, their out-of-sample hedging performance should be investigated, and compared against the pricing and hedging errors of deterministic volatility implied models. This will allow us to conclude whether the use of a more complex class of models is necessary.

It is very probable that a jump component should be included (see Das and Sundaram [33]), and researchers should look at how to develop a *smile-consistent stochastic volatility-jump model*. The papers by Andersen and Andreasen [3], and Pappalardo [89] that add Poisson jumps to the deterministic volatility stock price diffusion process, provide a start towards this direction.

A broader topic of research is whether implied volatility smiles can be explained significantly by transaction costs. If this is the case, then the performance of smile-consistent models should be weak. Jackwerth and Rubinstein [61] point out that a potential trading-cost theory of the smile needs to explain why, given the extreme shift in the option smile, these costs were apparently of much less importance before the 1987 crash than after. We are aware of only Constandinides [26] transaction costs model which implies that transaction costs can not account for the volatility smile. However, the extent to which implied volatility patterns are affected by transaction costs, should be explored further.

Finally, even though the smile consistent approach is a promising one for the pricing and hedging of exotic options, it does not solve the problem of pricing standard options correctly, since their prices are taken as given. Future research should investigate this issue.

References

[1] L. Andersen, *The Equity Option Volatility Smile: an Implicit Finite Difference Approach, Five Essays on the Pricing of Contingent Claims*, Ph.D. Thesis, The Aarhus School of Business (1996).

[2] J. Andreasen, *Implied Modelling: Stable Implementation, Hedging and Duality*, Working Paper, The Aarhus School of Business (1996).

[3] L. Andersen and J. Andreasen *Jump-Diffusion Processes: Volatility Smile Fitting and Numerical Methods for Pricing*, Working Paper, General Re Financial Products (1999).

[4] M. Avellaneda, C. Friedman, R. Holmes and D. Samperi, *Calibrating Volatility Surfaces via Relative-Entropy Minimization*, App. Math. Finance **4** (1997) 37-64.

[5] B. Bahra, *Implied Risk-Neutral Probability Density Functions From Option Prices: Theory and Application*, Bank of England Working Paper Series (1997).

[6] G. Bakshi, C. Cao and Z. Chen, *Empirical Performance of Alternative Option Pricing Models*, J. Finance **52** (1997) 2003-2049.

[7] S. Barle and N. Cakici, *Growing a Smiling Tree*, Risk **8** (1995) 76-81.

[8] D. Bates, *Pricing Options under Jump-Diffusion Processes*, Working Paper, Rodney L. White Center (1988).

[9] D. Bates, *Dollar Jump Fears, 1984-1992: Distributional Abnormalities Implicit in Currency Futures Options*, J. International Money and Finance **15** (1996) 65-93.

[10] D. Bates, *Testing Option Pricing Models*, in *Handbook of Statistics* Vol. 14, eds. G. Maddala and C. Rao, Elsevier Science B.V (1996).

[11] D. Bates, *Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options*, Rev. Financial Studies **9** (1996) 69-107.

[12] D. Bates, *Post-87 Crash Fears in the S&P 500 Futures Option Market*, J. Econometrics, Forthcoming (1997).

[13] D. Bates, *Hedging the Smirk*, Working Paper, Univ. of Iowa (1998).

[14] A. Bick and H. Reisman, *Generalized Implied Volatility*, Working Paper, Simon Fraser Univ., Canada (1993).

[15] F. Black, *Fact and Fantasy in the Use of Options*, Financial Analysts J. **31** (1975) 36-72.

[16] F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, J. Political Econ. **81** (1973) 637-654.

[17] P. Boyle, *Options: A Monte Carlo Approach*, J. Financial Econ. **4** (1977) 323-338.

[18] D. Breeden and R. Litzenberger, *Prices of State-Contingent Claims Implicit in Option prices*, J. Business **51** (1978) 621-651.

[19] P. Buchen and M. Kelly, *The maximum entropy distribution of an asset inferred from option prices*, J. Financial and Quantitative Analysis **31** (1996) 143-159.

[20] M. Britten-Jones and A. Neuberger *Option Prices, Implied Price Processes, and Stochastic Volatility*, J. Finance, Forthcoming (1999).

[21] A. Buraschi and J. Jackwerth, *Explaining Option Prices: Deterministic vs. Stochastic Models*, Working Paper, London Business School (1998).

[22] P. Carr, K. Ellis and V. Gupta, *Static Hedging of Exotic Options*, J. Finance **53** (1998) 1165-1190.

[23] A. Christie, *The Stochastic Behavior of Common Stock Variances, Value, Leverage and Interest Rate Effects*, J. Financial Econ. **10** (1982) 407-432.

[24] L. Clewlow and C. Strickland, *Implementing Derivatives Models*, John Wiley and Sons (1998).

[25] L. Clewlow and X. Xu, *A Review of Option Pricing with Stochastic Volatility*, Working Paper, Financial Options Research Centre, Univ. of Warwick (1992).

[26] G. Constandinides, *Transaction Costs and the Volatility Implied by Option Prices*, Working Paper, Univ. of Chicago (1998).

[27] D. Cox and H. Miller, *Theory of Stochastic Processes*, Chapman and Hall (1965).

[28] J. Cox, *The Constant Elasticity of Variance Option Pricing Model*, J. Portfolio Management, Special Issue (1996) 15-17.

[29] J. Cox and S. Ross *The Valuation of Options for Alternative Stochastic Processes*, J. Financial Econ. **3** (1976) 145-166.

[30] J. Cox, S. Ross and M. Rubinstein, *Option Pricing: A Simplified Approach*, J. Financial Econ. **7** (1979) 229-263.

[31] J. Cox, and M. Rubinstein, *Options Markets*, Prentice-Hall, Englewood Cliffs, N.J. (1985).

[32] J. Cox, J. Ingersoll and S. Ross, *A Theory of the Term Structure of Interest Rates*, Econometrica **53** (1985) 385-407.

[33] S. Das and R. Sundaram, *Of Smiles and Smirks: A Term-Structure Perspective*, J. Financial and Quantitative Analysis, forthcoming (1998).

[34] E. Derman, D. Ergener and I. Kani, *Static Options Replication*, J. Derivatives **2** (1995) 78-95.

[35] E. Derman and I. Kani, *Riding on a Smile*, Risk **7** (1994) 32-39.

[36] E. Derman, I. Kani and J. Zou, *The Local Volatility Surface: Unlocking the Information in Index Option Prices*, Financial Analysts J. **52** (1996) 25-36.

[37] E. Derman, M. Kamal, I. Kani, J. McClure, C. Pirasteh and J. Zou, *Investing in Volatility*, Goldman Sachs Quantitative Strategies Research Notes (1996).

[38] E. Derman, I. Kani and N. Chriss, *Implied Trinomial Trees of the Volatility Smile*, J. Derivatives **3** (1996) 7-22.

[39] E. Derman and I. Kani, *Stochastic Implied Trees: Arbitrage Pricing With Stochastic Term and Strike Structure of Volatility*, International J. of Theoretical and Applied Finance **1** (1997) 61-110.

[40] M. Dothan, *Prices in Financial Markets,* Oxford Univ. Press (1990).

[41] B. Dumas, J. Fleming, and R. Whaley, *Implied Volatility functions: Empirical Tests,* J. Finance **53** (1998) 2059-2106.

[42] B. Dupire, *Arbitrage Pricing with Stochastic Volatility*, Proc. of French Finance Association Meetings, Paris 1992.

[43] B. Dupire, *Pricing and Hedging with Smiles*, Proc. of French Finance Association Meetings, La Baule 1993.

[44] B. Dupire, *Model Art*, Risk **6** (1993) 118-121.

[45] B. Dupire, *Pricing with a Smile*, Risk **7** (1994) 18-20.

[46] B. Dupire, *A Unified Theory of Volatility*, Working Paper, Paribas Capital Markets (1994).

[47] D. Emanuel and J. MacBeth, *Further Results on the Constant Elasticity of Variance Option Pricing Model*, J. Financial and Quantitative Analysis **17** (1982) 533-554.

[48] J. Franks and E. Schwartz, *The Stochastic Behavior of Market Variance Implied in the Prices of Index Options*, Economic J. 101 (1991) 1460-1475.

[49] K. Fung and D. Hsieh, *Empirical Analysis of Implied Volatility: Stocks, Bonds and Currencies*, Working Paper, Fuqua School of Business, Duke Univ. (1991).

[50] A. Gallant, H. Chien and G. Tauchen *Calibrating Volatility Diffusions and Extracting Integrated Volatility*, Working Paper, Duke Univ. (1998).

[51] G. Gemmill, *Did Option Traders Anticipate the Crash? Evidence from Volatility Smiles in the U.K., with U.S. Comparisons*, J. Futures Markets **16** (1996) 881-897.

[52] R. Geske and K. Shastri, *Valuation of Approximation: a Comparison of Alternative Approaches*, J. Financial and Quantitative Analysis **20** (1985) 45-72.

[53] S. Grossman and Z. Zhou, *Equilibrium Analysis of Portfolio Insurance*, J. Finance **51** (1996) 1379-1403.

[54] A. Grunbichler and F. Longstaff, *Valuing Futures and Options on Volatility*, J. Banking and Finance **20** (1994) 985-1001.

[55] C. Harvey and R. Whaley, *S&P 100 Index Option Volatility*, J. Finance **46** (1991) 1551-1561.

[56] D. Heath, R. Jarrow and A. Morton, *Bond Pricing and the Term Structure of Interest Rates: A New Methodology For Contingent Claims Valuation*, Econometrica **60** (1992) 77-105.

[57] L. Ho and S. Lee, *Term Structure Movements and Pricing Interest Rate Contingent Claims*, J. Finance **41** (1986) 1011-1029.

[58] J. Hull, *Options, Futures, and Other Derivative Securities*, Prentice Hall, Second Edition (1991).

[59] J. Hull and A. White, *The Pricing of Options on Assets with Stochastic Volatilities*, J. Finance **3** (1987) 281-300.

[60] J. Hull and A. White, *One-Factor Interest-Rate Models and the Valuation of Interest-Rate Contingent Claims*, J. Financial and Quantitative Analysis **28** (1992) 235-254.

[61] J. Jackwerth and M. Rubinstein, *Recovering Probability Distributions from Option Prices*, J. Finance **51** (1996) 1611-1631.

[62] J. Jackwerth and M. Rubinstein, *Recovering Stochastic Processes from Options Prices*, Working Paper, Univ. of California, Berkeley (1996).

[63] J. Jackwerth, *Generalized Binomial Trees*, J. Derivatives **5** (1997) 7-17.

[64] F. Jamshidian, *Forward Induction and Construction of Yield Curve Diffusion Models*, J. Fixed Income **1** (1991) 62-74.

[65] F. Jamshidian, *Option and Futures Evaluation with Deterministic Volatilities*, Math. Finance **3** (1993) 149-159.

[66] R. Jarrow and A. Rudd, *Approximate Option Valuation for Arbitrary Stochastic Processes*, J. Financial Econ. **10** (1982) 347-369.

[67] N. Johnson and S. Kotz, *Continuous Univariate Distributions*, Vol 1 and 2. New York: John Wiley & Sons (1970).

[68] H., Johnson and D. Shanno, *Option Pricing when the Variance is Changing*, J. Financial and Quantitative Analysis **22** (1987) 143-151.

[69] M. Kamal and E. Derman, *The Patterns of Change in Implied Index Volatilities*, Goldman Sachs, Quantitative Strategies Notes (1997).

[70] I. Kani, E. Derman and M. Kamal, *Trading and Hedging Local Volatility*, J. Financial Engineering **6** (1997) 1233-1268.

[71] O. Ledoit and P. Santa-Clara, *Relative Pricing of Options with Stochastic Volatility*, Working Paper, Univ. of California, Los Angeles (1998).

[72] R. Litterman and J. Scheinkman, *Common Factors Affecting Bond Returns*, Goldman Sachs, Financial Strategies Group (1988).

[73] J. MacBeth and L. Merville, *An Empirical Estimation of the Black-Scholes Call Option Pricing Model*, J. Finance **34** (1979) 1173-1186.

[74] J. MacBeth and L. Merville, *Tests of the Black-Scholes and Cox Call Option Valuation Models*, J. Finance **35** (1980) 285-301.

[75] S. Mayhew, *Implied Volatility*, Financial Analysts J. **51** (1995) 8-13.

[76] S. Mayhew, *On Estimating The Risk-Neutral Probability Distribution Implied by Option Prices*, Working Paper, Univ. of California, Berkeley (1995).

[77] W. Melick and C. Thomas, *Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis*, J. Financial and Quantitative Analysis **32** (1997) 91-115.

[78] R. Merton, *Option Pricing when Underlying Stock Returns are Discontinuous*, J. Financial Econ. **3** (1976) 125-144.

[79] D. Nelson and K. Ramaswamy, *Simple Binomial Processes as Diffusion Approximations in Financial Models*, Rev. Financial Studies **3** (1990) 393-430.

[80] A. Neuberger, *The Log Contract: A New Instrument to Hedge Volatility*, J. Portfolio Management **75** (1994) 74-80.

[81] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, Springer-Verlag (1992).

[82] R. Rebonato, *Interest-Rate Option Models: Understanding, Analysing and Using Models for Exotic Interest-Rate Options*, John Wiley and Sons (1996).

[83] R. Roll, *A Simple Implicit Measure of the Effective Bid-Ask Spread*, J. Finance **39** (1984) 1127-1139.

[84] M. Rubinstein, *Non-Parametric Tests of Alternative Option Pricing Models using all Reported Trades and Quotes on the 30 most Active CBOE Option Classes from August 23, 1976 through August 31, 1978*, J. Finance **40** (1985) 455-480.

[85] M. Rubinstein, *Implied Binomial Trees*, J. Finance **69** (1994) 771-818.

[86] M. Rubinstein, *As Simple as One, Two, Three*, Risk **8** (1995) 44-47.

[87] M. Rubinstein, *Edgeworth Binomial Trees*, J. Derivatives (1998) 20-27.

[88] E. Platen and M. Schweitzer, *On Smile and Skewness*, Working Paper, Australian National Univ. (1995).

[89] L. Pappalardo, *Option Pricing and Smile Effect when Underlying Stock Prices are Driven by a Jump Process*, Working Paper, Financial Options Research Centre, Univ. of Warwick (1996).

[90] L. Scott, *Option Pricing when the Variance changes Randomly: Theory, Estimation and an Application*, J. Financial and Quantitative Analysis **22** (1987) 419-438.

[91] L. Scott, *Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods*, Math. Finance **7** (1997) 413-426.

[92] R. Schmalensee and R. Trippi, *Common Stock Volatility Expectations Implied By Option Premia*, J. Finance **33** (1978) 129-147.

[93] K. Shastri and K. Wethyavivorn, *The valuation of currency options for alternate stochastic processes*, J. Financial Research **10** (1987) 283-293.

[94] D. Shimko, *Bounds of Probability*, Risk **6** (1993) 33-37.

[95] G. Skiadopoulos, S. Hodges and L. Clewlow, *The Dynamics of the S&P 500 Implied Volatility Surface*, Rev. Derivatives Research (1999), forthcoming.

[96] C. Strickland, *A Comparison of Diffusion Models of the Term Structure*, European J. Finance **2** (1996) 103-123.

[97] M. Stutzer, *A simple nonparametric approach to derivative security valuation*, J. Finance **101** (1996) 1633-1652.

[98] R. Sundaram, *Equivalent Martingale Measures and Risk-Neutral Pricing: An Expository Note*, J. Derivatives **5** (1997) 85-98.

[99] S. Taylor and X. Xu, *The Magnitude of Implied Volatility Smiles: Theory and Empirical Evidence for Exchange Rates*, Rev. Futures Markets **13** (1994) 355-380.

[100] R. Tompkins, *Implied Volatility Surfaces: Uncovering Regularities for Options on Financial Futures*, Working Paper, Financial Options Research Centre, Univ. of Warwick (1998).

[101] O. Vasicek, *An Equilibrium Characterization of the Term Structure*, J. Financial Econ. **5** (1977) 177-88.

[102] J. Wiggins, *Option Values under Stochastic Volatility*, J. Financial Econ. **19** (1987) 351-372.

[103] J. Zou and E. Derman, *Monte Carlo Valuation of Path-Dependent Options on Indexes with a Volatility Smile*, J. Financial Engineering **6** (1996) 149-168.