PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

A market model for stochastic implied volatility

Philipp J. Schönbucher

Phil. Trans. R. Soc. Lond. A 1999 **357**, 2071-2092
doi: 10.1098/rsta.1999.0418

THE ROYAL SOCIETY

A market model for stochastic implied volatility

By Philipp J. Schönbucher

Department of Statistics, Faculty of Economics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany
(p.schonbucher@finasto.uni-bonn.de)

In this paper a stochastic volatility model is presented that directly prescribes the stochastic development of the implied Black–Scholes volatilities of a set of given standard options. Thus the model is able to capture the stochastic movements of a full term structure of implied volatilities. Conditions are derived that have to be satisfied to ensure absence of arbitrage in the model and its numerical implementation is discussed.

Keywords: option pricing; stochastic volatility; smile effect; term structure of volatility; strike structure of volatility

1. Introduction

The aim of this paper is to provide a framework for the market-based pricing and hedging of exotic options and options on volatility indices. In addition to the usual share and bond underlying securities, the model presented here also uses the prices of liquidly traded standard options as underlying securities. The prices of the standard options are given in terms of their implied Black–Scholes volatilities, which are stochastic.

We will follow a market-based approach applied to the term structure of implied volatilities, which is similar to the market models of the term structure of interest rates by Miltersen *et al*. (1997), Brace *et al*. (1997) and Jamshidian (1997).

Using a market-based approach means that we do not model 'fundamental' quantities like, for example, the stochastic process of the volatility of the share price (as in the traditional stochastic volatility models of Hull & White (1987), Heston (1993) or Stein & Stein (1991)), or the instantaneous conditional forward volatilities (as in the effective volatility model by Derman & Kani (1998)), or forward variances (like in Dupire (1993*a*)), but we model the Black–Scholes implied volatilities *directly*. This facilitates the fitting of the model to observed option prices and gives the model a larger degree of flexibility.

As we allow stochastic dynamics for the implied volatilities we have to ensure that no arbitrage opportunities arise in the model. Therefore, sufficient conditions are derived that have to be imposed on the drift coefficients of the implied volatilities and that ensure absence of arbitrage in the model.

(*a*) *The information content of traded options prices*

There are good reasons for incorporating the prices of at least some traded options into a stochastic volatility model.

Since the advent of the famous Black & Scholes (1973) option pricing model and the introduction of exchange-traded option contracts in the same year, the volume and liquidity of traded options has increased exponentially. Simultaneously, more and more complex exotic option specifications have arisen with features ranging from American early exercise, knock-in and knock-out barriers, Asian averaging and lookbacks to combinations of these and other features with many different pay-off functions and multiple underlying securities.

While on the one end of the spectrum the development has gone towards increasingly complex specifications, there has been a significant increase in liquidity in the markets of standard European or American call and put options. For almost every major stock index or its futures contract there are liquid markets for European or American call and put options with a broad range of strike prices and maturities. These markets make the trading of a new piece of information possible: information on volatility.

The efficient markets hypothesis in its semistrong form (see, for example, Fama 1970) states that prices in liquid markets contain all information that is relevant to the pricing of the security under concern and that is publicly available in the market. The information relevant to the pricing of options is information about volatility, information that is not directly contained in the prices of the underlying security.

It has been debated whether the efficient markets hypothesis is always fully valid, but there is compelling evidence that exchange-traded options prices do contain idiosyncratic information that cannot be backed out from the price information of the underlying security alone. For example, Chiras & Manaster (1978) or Fleming *et al*. (1995) show that predictions of future stock-price volatility that are based on the implied volatility of option prices are superior to predictions that are based on information from the stock-price process alone.

That options prices contain volatility information can also be seen from the fact that options traders focus on the new information that is encoded in the securities they trade: information about volatility. Prices of standard options are usually quoted in terms of the implied volatility $\hat{\sigma}$ that has to be substituted in the Black–Scholes option pricing formula (see, for example, equation (2.2)) to reach the cash price of the option. This does not mean that the market participants assume that the Black–Scholes model with all its imperfections applies to the actual market, instead they just use the Black–Scholes formula to make their price quotations more independent from the movements of the price of the underlying security for which there is already an efficient market.

The advent of sufficiently liquid markets for standard options has several consequences.

First, the market prices will show deviations from the prices implied by the Black–Scholes formula. This is not due to any pricing errors in the market but to the inaccuracies in the assumptions of the Black–Scholes model itself which are already corrected in the market prices, and, as explained above, it is due to additional information that the Black–Scholes formula cannot reflect.

Second, because of the liquidity of the standard options market, the need to theoretically *price* these securities is diminished: a fair-price indication can be read from the market (and it is even likely to be more accurate), arbitrage opportunities will be unlikely to exist and a hedge strategy is less important because the position can be unwound quickly. Pricing models are most useful if there might be arbitrage opportunities in the markets, if the instrument to price is not well understood or if a hedging strategy is needed. This is often not the case for standard options; for exotic options on the other hand there is a need for pricing models.

Third, given that standard options markets reveal additional information about the likely dynamics of the underlying, instead of *deriving* prices for them, a pricing model should use their prices as input. This should yield an increase in accuracy over the standard Black–Scholes model. Then the standard options can also be used as additional hedge instruments.

The model presented here tries to take these points into account. It is designed to incorporate traded options prices (and thus the information that they contain) in order to improve the pricing of more exotic instruments.

(*b*) *Related literature*

The deviation of observed market prices for options from their theoretical counterparts (as given by the Black–Scholes formula) has triggered a large literature in which both academics and practitioners alike have tried to improve on the limitations of the Black–Scholes model.

One strand of the literature concentrates on the nature of the underlying asset-price process which was assumed to be a lognormal Brownian motion by Black & Scholes. Here the main focus is on *stochastic volatility* models which assume that volatility of the stock-price process is not constant but is itself stochastic. Well-known papers on this approach are by Hull & White (1987), Heston (1993) or Stein & Stein (1991).

These models can usually reproduce the typical shapes of implied volatilities observed in the markets (the 'smile') but they cannot be *fitted* easily to any given implied volatilities. Furthermore, these models have only one additional factor driving the stochastic volatility and cannot be extended to the multifactor case, and the expressions given for the prices of the standard European call and put options are very complex and cannot be considered to be in closed form (e.g. Heston's (1993) model still requires a numerical inversion of the Fourier transformation).

In another direction of research—the implied tree approach—the aim was to keep as closely as possible to the Black–Scholes set-up while exactly reproducing the option prices given in the market. This is achieved by specifying a time- and state-dependent (i.e. share price) volatility function which does not contain any additional random component. Models of this type are by Rubinstein (1995), Derman & Kani (1994), Derman *et al*. (1996) and Dupire (1993*b*, 1994).

While exactly reproducing the option prices observed in the market, the implied volatility models have the drawback that they do not allow for idiosyncratic stochastic dynamics in the option prices. This is in conflict with empirical observation† and with the continuous updating of the new information reflected in the option prices. The poor results in a hedging test performed by Dumas *et al*. (1999) are probably also due to this drawback.

Dupire (1993*a*) took a first step towards incorporating stochastic dynamics into the term structure of volatilities, but again he models *realized* volatilities (and forward contracts on it) and not implied volatilities from options prices.

† See, for example, Skiadopulous *et al*. (1998) for an analysis of the dynamics of implied volatility surface given by the S&P 500 options at the CME.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

In a recent paper, Derman & Kani (1998) have extended their implied tree approach to allow for stochastic dynamics in the full term and strike structure of implied local volatilities. They derive restrictions on the drift of the local volatilities that are necessary for absence of arbitrage, and these restrictions involve integrals over all possible share prices and times before the maturity of the forward volatility concerned. The complexity of these restrictions makes the model hard to handle and we are going to propose a slightly different approach. Furthermore it is not obvious how in Derman & Kani's (1998) model it is ensured that the implied volatilities satisfy certain no-arbitrage restrictions as expiry is approached. (These restrictions will be derived later.) The fundamental problem is that Derman & Kani (1998) specify two things that may be contradictory: the dynamics of the spot volatility and the implied volatilities for different strike prices and maturities. Nevertheless the approach taken in these two papers is closest to the approach taken here.

As the dynamics of a whole term structure of security prices are to be captured, the modelling of implied volatilities is similar to the modelling of the term structure of interest rates. There one of the most elegant solutions to the problem of fitting a range of prices to a model has been proposed by Heath *et al*. (1992), which is to model the whole term structure of interest rates and then to impose restrictions on the drift of the rates to ensure absence of arbitrage.

While the Heath *et al*. (1992) approach was very successful, it still did not directly describe the dynamics of the most liquid instruments in the fixed income market: the Libor futures and the interest rate swaps. Therefore, in some recent papers (see, for example, Miltersen *et al*. 1997; Brace *et al*. 1997; Jamshidian 1997) there has been a shift to models that directly model the rates that are used in the market (i.e. forward Libor and swap rates) instead of instantaneous forward rates. These models were termed *market models* of interest rates and they were the direct inspiration for this paper: here too, the aim is to directly model the implied volatilities as they are quoted in the market and not some fundamental but unobservable quantity.

(*c*) *Structure of the paper*

The stochastic volatility model is built up in several steps, going from single options via a discrete term structure of maturities to a continuous term structure of maturities.

After the introduction of the model set-up in the next section, the main intuition is discussed using a single traded European call option. The no-arbitrage condition on the drift of the option's implied volatility is derived, and we analyse the restrictions that have to be imposed to ensure regularity of the option price at expiry. The smile and frown effects are analysed and it is shown how they can be incorporated using the volatility of volatility and the correlation of the implied volatility with the share price.

In the following section the model is extended to a discrete term structure of option prices and implied volatilities. We introduce the concept of a *forward implied volatility* and its differences to the forward volatility found in Dupire (1993*a*). Again no-arbitrage conditions and consistency conditions at expiry dates are derived.

Then this set-up is extended to a continuous term structure of implied volatilities and instantaneous forward volatilities. While this may seem more complicated in some respects, in others it is more convenient as the process of the spot volatility arises naturally from the model.

Furthermore, it is shown in each section how this model can be implemented to price a variety of exotic options, and also to price futures contracts on implied volatility. Such contracts are traded on some exchanges, e.g. the VOLAX future at the DTB. The conclusion sums up the results of this paper.

2. Model set-up

(*a*) *Traded securities*

The model is set up in a probability space $(\Omega, (\mathcal{F}_t)_{(t\geqslant 0)}, Q)$, where the filtration $(\mathcal{F}_t)_{(t\geqslant 0)}$ is generated by the $N+1$-dimensional Brownian motion $(W_0, W) = (W_0, W_1, \ldots, W_N)$ and satisfies the usual conditions. Q is the martingale measure (or pricing measure) under which discounted price processes are martingales.

There are several liquidly traded securities: the underlying of the options S (called the share price from now on), a set of European call options with strike prices and maturities† $\{(K_m, T_m)\, m \leqslant M\}$, and a risk-free investment opportunity (the bank account) with constant interest rate r. To each state variable (time, share price, implied volatilities) there is exactly one traded security (bond, share, traded options) so that the markets are complete (assuming a non-singular variance/covariance structure of the asset prices).

Market completeness distinguishes this model from the classical stochastic volatility models where the share is assumed to be the only traded risky security and therefore markets are incomplete. On the other hand, by adding one traded derivative these models can also be completed. Here the traded standard options and their prices are taken as direct input, which has the additional advantage that no market prices of risk or preferences have to be specified: the market price of risk can be implied from the observed prices.

We assume that the share-price process can be represented as a stochastic volatility lognormal Brownian motion:

$$\mathrm{d}S = rS\,\mathrm{d}t + \sigma S\,\mathrm{d}W_0, \tag{2.1}$$

where σ is *stochastic*. The drift of rS is imposed to make the discounted share price a martingale, and, given a positive price process for S, a representation like (2.1) can always be found. We will analyse the precise nature of the dynamics of σ later on.

The prices of the call options are given by the Black–Scholes formula and *option-specific implied volatilities* $\hat{\sigma}(K_m, T_m)$. The implied volatility is typically different across the traded options. We denote with *implied volatility* the Black–Scholes implied volatility as opposed to the *actual volatility*, which is the volatility of the share-price process.

As the model is not set up in a Black–Scholes world with constant share-price volatility, the Black–Scholes formula serves only as a convenient way of describing option prices via the parameter $\hat{\sigma}$. Typically, the implied Black–Scholes volatility $\hat{\sigma}$ of the options is neither equal to the actual share-price volatility σ nor to some expectation of it. There are close links between actual and implied volatilities but they are more complex and will be discussed in more detail later. For now it is only important to note that it is much easier to observe the value of the implied volatilities than the actual volatility.

† The set of options will be a continuum in later sections.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

We can restrict our attention to the call option prices as put option prices follow from put–call parity:

$$C - P = S - K\mathrm{e}^{-r(T-t)}.$$

Put–call parity follows directly from a comparison of pay-off profiles and therefore holds independently of the distribution and dynamics of the share price or of possibly stochastic volatility. Hence, call and put options with the same maturity and strikes have to have the same implied volatility.

(b) *Implied volatilities*

(i) *Definition*

The implied volatility $\hat{\sigma}$ of an option is implicitly defined as the parameter $\hat{\sigma}$ that yields the actually observed option price when it is substituted into the well-known Black–Scholes formula (together with time t, the price S of the underlying, interest rate r and the parameters K, T of the option). The Black–Scholes formula is

$$C(S, t; K, T; r, \hat{\sigma}) = SN(d_1) - \mathrm{e}^{-r(T-t)}KN(d_2), \tag{2.2}$$

where the coefficients d_1 and d_2 are given by

$$d_1 = \frac{\ln(S/K) + (r + \frac{1}{2}\hat{\sigma}^2)(T-t)}{\hat{\sigma}\sqrt{T-t}},$$
$$d_2 = \frac{\ln(S/K) + (r - \frac{1}{2}\hat{\sigma}^2)(T-t)}{\hat{\sigma}\sqrt{T-t}}.$$

($N(x)$ is the cumulative standard normal distribution function.) The Black–Scholes formula (2.2) is the solution to the well-known Black–Scholes partial differential equation

$$0 = C_t + \tfrac{1}{2}\hat{\sigma}^2 S^2 C_{SS} + rC_S - rC, \tag{2.3}$$

with the final condition $C(S, T) = (S - K)^+$ and appropriate boundary conditions at $S \to 0$ and $S \to \infty$. (To simplify notation we will write all partial derivatives as subscripts from now on.) Note that the solution of this partial differential equation equals the actually observed option price only for the *implied* volatility for this option. Similarly, implied volatilities can also be defined for more complex options (e.g. American or barrier options) which still have to satisfy the Black–Scholes partial differential equation (2.3) (unless there is strong path dependence in the option price in which case the state space will have to be extended).

It is common practice in futures markets to quote option prices not directly but in terms of the implied volatility that has to be used in the Black–Scholes formula to reach the cash price of the option. This relieves the market makers from the task of tracking every single movement in the price of the underlying asset and enables the traders to concentrate on the option-specific features.

(ii) *Forward options prices*

If, instead of the spot price of the underlying the forward price F (with the same maturity T as the option),

$$F(t) = \mathrm{e}^{r(T-t)}S(t), \quad \mathrm{d}F = \sigma F\,\mathrm{d}W_0, \tag{2.4}$$

is given, then the modified model of Black (1976) gives the forward option prices $\bar{C}$ as

$$\bar{C}(F, t; K, T; r, \hat{\sigma}) = FN(d_1) - KN(d_2), \tag{2.5}$$

where the coefficients d_1 and d_2 are now

$$d_1 = \frac{\ln(F/K) + \frac{1}{2}\hat{\sigma}^2(T-t)}{\hat{\sigma}\sqrt{T-t}}$$
$$d_2 = \frac{\ln(F/K) - \frac{1}{2}\hat{\sigma}^2(T-t)}{\hat{\sigma}\sqrt{T-t}},$$

and the forward price $\bar{C}$ of the option has to satisfy the forward version of the Black–Scholes PDE:

$$0 = \bar{C}_t + \tfrac{1}{2}\hat{\sigma}^2 F^2 \bar{C}_{SS}. \tag{2.6}$$

The Black formula remains valid for stochastic interest rates, and, as long as share price and interest rate process are independent, forward prices are martingales for all maturities under the original martingale measure. This can be used to make the analysis independent of possibly stochastic interest rates, and the forward prices of the options and the share can be approximated with the respective futures prices at very small errors.

(iii) *Stochastic implied volatilities*

In actual markets the implied volatility for a traded option is by no means constant. We therefore specify the following dynamics for the implied volatility of an option with maturity T and strike K:

$$\mathrm{d}\hat{\sigma}(T, K) = u(T, K)\,\mathrm{d}t + \gamma(T, K)\,\mathrm{d}W_0 + \sum_{n=1}^{N} v_n(T, K)\,\mathrm{d}W_n. \tag{2.7}$$

The implied volatility is driven by the Brownian motions $W_1, \ldots, W_N$ and a term $\gamma\,\mathrm{d}W_0$ which is driven by the same Brownian motion that is driving the share price. This can be used to model the correlation between implied volatility and share-price movements. Negative correlation of this type is frequently observed, especially at large down movements of the share price where there is an increase in the implied volatility.

To simplify notation we will write the N-dimensional Brownian motion $W = (W_1, \ldots, W_N)^{\mathrm{T}}$ and the volatility vector $v = (v_1, \ldots, v_N)$ in vector notation such that equation (2.7) becomes

$$\mathrm{d}\hat{\sigma}(T, K) = u(T, K)\,\mathrm{d}t + \gamma(T, K)\,\mathrm{d}W_0 + v(T, K)\,\mathrm{d}W. \tag{2.8}$$

The implied volatility $\hat{\sigma}$ and the diffusion parameters γ, u and v are predictable stochastic processes which can depend on the full state vector $(S, t, \hat{\sigma})$, consisting of share price S, time t and all implied volatilities $\hat{\sigma}$. To keep the notation clear, only the dependence on the maturity T and the strike K are shown explicitly.

To ensure existence and uniqueness of the process of implied stochastic volatilities the diffusion parameters must satisfy certain regularity conditions which are given in the following well-known theorem (see, for example, Karatzas & Shreve 1991, p. 284).

Theorem 2.1 (Existence and uniqueness). *Let M be the number of traded options and let $X = (S, \hat{\sigma}_1, \hat{\sigma}_2, \ldots, \hat{\sigma}_M) \in \mathbb{R}^{M+1}$ be the state vector. Let $T > 0$, and $u(\cdot,\cdot) : [0,T] \times \mathbb{R}^{M+1} \to \mathbb{R}^{M+1}$ and $v(\cdot,\cdot) : [0,T] \times \mathbb{R}^{M+1} \to \mathbb{R}^{M+1 \times N+1}$ be measurable functions satisfying*

$$|u(t,x)| + |v(t,x)| \leqslant C(1 + |x|) \tag{2.9}$$

for all $x \in \mathbb{R}^{M+1}$, $t \in [0,T]$ and some constant C, and

$$|u(t,x) - u(t,y)| + |v(t,x) - v(t,y)| \leqslant D|x - y| \tag{2.10}$$

for all $x, y \in \mathbb{R}^{M+1}$, $t \in [0,T]$ and some constant D.

Then the stochastic differential equation $X(0) = X_0$ and

$$\mathrm{d}X = u(t,X)\,\mathrm{d}t + \sum_{i=0}^{N} v(t,X)\,\mathrm{d}W_i \tag{2.11}$$

has a unique t-continuous solution $X(t;\omega) = (S(t;\omega), \hat{\sigma}_1(t;\omega), \hat{\sigma}_2(t;\omega), \ldots, \hat{\sigma}_M(t;\omega))$, each component of which is measurable, adapted and square-integrable.

This solution is called a strong *solution.*

Because we absorbed the share-price process into the state vector X, the Lipschitz growth condition will also have to be satisfied for the diffusion coefficients of S. Specifically, the regularity conditions of this theorem are satisfied if the diffusion coefficients of the implied volatilities given in equation (2.7) are Lipschitz continuous, and if the spot volatility process σ is regular and some Lipschitz continuous function of the other state variables.

In this model the crucial problem will be to ensure regularity for the volatility σ of the share price and regularity of the drift coefficients u_n of the implied volatilities $\hat{\sigma}$. The other coefficients will not pose any problems: the drift of the share price is given by rS under the martingale measure, and the volatility of the implied volatilities can be specified by the user.

3. Modelling one implied volatility

(*a*) *No-arbitrage conditions*

So far there are no provisions in the model to ensure that there are no arbitrage opportunities. The situation is similar to that in the Heath *et al*. (1992) model for interest rates: in both cases we have an over-specified model with more securities than sources of randomness: in Heath *et al*. (1992) there is a continuum of bond prices (which are specified in terms of forward rates) and only a finite number of Brownian motions driving the model; here we have a possibly large number of option prices (in terms of implied volatilities) with again only a finite number of Brownian motions.

The solution to this problem is in both cases to impose restrictions on the dynamics of the factors, which ensure that the discounted security prices are martingales under the pricing measure. We will do this now for the case of only one traded option (and thus also only one implied volatility).

(b) Dynamics

To describe the dynamics of the option prices we will need the partial derivatives of the option prices. These are, for the Black–Scholes formula,

$$C_S = N(d_1), \qquad C_{SS} = \frac{n(d_1)}{S\hat{\sigma}\sqrt{T-t}},$$
$$C_{\hat{\sigma}} = S\sqrt{T-t}\,n(d_1), \qquad C_{\hat{\sigma}\hat{\sigma}} = S\sqrt{T-t}\,n(d_1)\frac{d_1 d_2}{\hat{\sigma}},$$
$$C_t = -\frac{Sn(d_1)\hat{\sigma}}{2\sqrt{T-t}} - rK\mathrm{e}^{-r(T-t)}N(d_2), \qquad C_{S\hat{\sigma}} = -\frac{1}{\hat{\sigma}}d_2 n(d_1).$$

And also $C_K = -C_S$, $C_{KK} = C_{SS}$, $C_T = -C_t$. In forward prices the partial derivatives are

$$\bar{C}_F = N(d_1), \qquad \bar{C}_{FF} = \frac{n(d_1)}{F\hat{\sigma}\sqrt{T-t}},$$
$$\bar{C}_{\hat{\sigma}} = F\sqrt{T-t}\,n(d_1), \qquad \bar{C}_{\hat{\sigma}\hat{\sigma}} = \frac{1}{\hat{\sigma}}F\sqrt{T-t}\,n(d_1)d_1 d_2,$$
$$\bar{C}_t = -\frac{Fn(d_1)\hat{\sigma}}{2\sqrt{T-t}}, \qquad \bar{C}_{F\hat{\sigma}} = \frac{1}{\hat{\sigma}}d_2 n(d_1),$$

again with $\bar{C}_K = -\bar{C}_F$, $\bar{C}_{KK} = \bar{C}_{FF}$, $\bar{C}_T = -\bar{C}_t$. Here $N(x)$ is the cumulative standard normal distribution function and

$$n(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$$

is its density.

Using the dynamics of the implied volatilities and of the share price we can now derive the dynamics of the option prices that are implied by these dynamics using Itô's lemma:

$$\mathrm{d}C = C_t\,\mathrm{d}t + C_S\,\mathrm{d}S + \tfrac{1}{2}\sigma^2 S^2 C_{SS}\,\mathrm{d}t + C_{\hat{\sigma}}\,\mathrm{d}\hat{\sigma} + \tfrac{1}{2}C_{\hat{\sigma}\hat{\sigma}}\,\mathrm{d}\langle\hat{\sigma}\rangle + C_{S\hat{\sigma}}\,\mathrm{d}\langle\hat{\sigma}, S\rangle. \tag{3.1}$$

(c) Drift restrictions

For absence of arbitrage we need the discounted option-price process to have zero drift under the martingale measure, or equivalently, it has drift $rC\,\mathrm{d}t$ if it is not discounted. The drift component of C is, according to (3.1), given by

$$rC\,\mathrm{d}t = \boldsymbol{E}[\mathrm{d}C] = C_t\,\mathrm{d}t + rSC_S\,\mathrm{d}t + \tfrac{1}{2}\sigma^2 S^2 C_{SS}\,\mathrm{d}t + C_{\hat{\sigma}}u\,\mathrm{d}t + \tfrac{1}{2}C_{\hat{\sigma}\hat{\sigma}}v^2\,\mathrm{d}t + C_{S\hat{\sigma}}\gamma\sigma S\,\mathrm{d}t. \tag{3.2}$$

To simplify notation all dependencies on $(S, t, \hat{\sigma}; K, T)$ in $\hat{\sigma}, u, v$ and γ have been dropped. Furthermore, we write

$$v^2 := \sum_{n=1}^{N} v_n^2 \tag{3.3}$$

for the volatility of the implied volatility, and

$$f := \ln(F/K), \quad s := \ln(S/K), \quad \tau := T - t. \tag{3.4}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Equation (3.2) can be reduced using the Black–Scholes partial differential equation (2.3):

$$0 = (C_t + rSC_S + \tfrac{1}{2}\hat{\sigma}^2 S^2 C_{SS} - rC_S)\,\mathrm{d}t + (\tfrac{1}{2}(\sigma^2 - \hat{\sigma}^2)S^2 C_{SS} + C_{\hat{\sigma}}u + \tfrac{1}{2}C_{\hat{\sigma}\hat{\sigma}}v^2 + \gamma\sigma SC_{S\hat{\sigma}})\,\mathrm{d}t, \tag{3.5}$$

whence we can derive the no-arbitrage drift of the implied volatility of the option:

$$u = \frac{1}{2C_{\hat{\sigma}}}((\hat{\sigma}^2 - \sigma^2)S^2 C_{SS} - C_{\hat{\sigma}\hat{\sigma}}v^2 - 2\gamma\sigma SC_{S\hat{\sigma}}). \tag{3.6}$$

For European call options the no-arbitrage drift restriction is expanded to

$$\hat{\sigma}u = \frac{1}{2\tau}(\hat{\sigma}^2 - \sigma^2) - \tfrac{1}{2}d_1 d_2 v^2 + \frac{d_2}{\sqrt{\tau}}\sigma\gamma. \tag{3.7}$$

It is necessary for absence of arbitrage that this restriction is satisfied for all options at all times ($Q-$ almost surely). Note that the restriction (3.7) has to hold for *each implied volatility* and its diffusion parameters individually. If we have several traded options we will also have a set of restrictions, one for each drift parameter. Not surprisingly, we also get the same restriction when the derivation is taken via the forward options prices $\bar{C}$. Here we have to impose that $\bar{C}$ has no drift under the martingale measure, which will yield equation (3.7).

(*d*) *Volatility bubbles at* $t = T$

Examination of equation (3.7) shows several interesting features of the stochastic implied volatility.

(i) If the implied volatility is constant (i.e. $\gamma = u = v = 0$), the drift restriction is only satisfied if $\hat{\sigma} = \sigma$. In this case the option price must be given by the standard Black–Scholes equation with the correct implied volatility.

(ii) The implied volatility $\hat{\sigma}$ has a *mean fleeing* behaviour which is shown in the term with $(\hat{\sigma}^2 - \sigma^2)$ ('mean fleeing' as opposed to 'mean reversion'). The further it is away from the spot volatility σ the more it is pushed away from it.†

(iii) The speed of the mean-fleeing behaviour seems to go to infinity as $t \to T$. This means that the solution to the stochastic volatility equation will blow up as $t \to T$ unless there is another force counteracting it.

(i) *The case of constant* σ

The reason for the mean-fleeing behaviour becomes clearest in the situation of constant σ (i.e. the classical Black–Scholes world) but with $\hat{\sigma}$ different from σ, and possibly stochastic. This is clearly a situation with arbitrage opportunities, because by Black–Scholes the option prices should exhibit a constant implied volatility of σ. This arbitrage manifests itself in the form of a *volatility bubble*.

If the implied volatility is too large ($\hat{\sigma} > \sigma$), the option is too expensive compared to its Black–Scholes price, and the dS-component will have a negative contribution

† As d_1 and d_2 also contain terms in $\hat{\sigma}$ this statement will have to be modified slightly.

to the expected growth rate. This must be compensated, and the only possible compensation would be through a locally expected *increase* in the implied volatility. This would push it even further away from the correct level of $\hat{\sigma} = \sigma$.

The whole mechanism is very similar to a price bubble in general equilibrium theory. There prices are moved further and further away from their fundamental value because the agents expect them to do exactly this. As long as agents expect this to go on, the wrong valuation can be sustained.

Here we have a bubble in the option price which is driven by the dynamics of the implied volatility. The initial option price is wrong but it does not revert to its correct value because the implied volatility grows. This pushes the option price even further away from its correct value, which in turn requires an even larger drift in the implied volatility to sustain it.

Because we have a finite time-horizon the bubble has to burst at the maturity of the option. This is where the drift of the implied volatility explodes, but even an infinite implied volatility cannot support the incorrect option price. The discounted option price loses its martingale property at this point and the solution to the SDE for the implied volatility ceases to exist. (The Lipschitz growth condition (2.9) in theorem 2.1 is not satisfied.) Thus the only specification of $\hat{\sigma}$ that prevents volatility bubbles is to set $\hat{\sigma} = \sigma = \text{const}$.

(ii) *The case of time-dependent $\sigma(t)$*

As a further example let us consider the case of a time-dependent (but non-stochastic) spot volatility function $\sigma(t)$. Here it is well known that the correct specification of the implied volatility for any option with maturity T is given by the average future volatility

$$\hat{\sigma}^2 = \frac{1}{T-t}\int_t^T \sigma^2(s)\,\mathrm{d}s. \tag{3.8}$$

The dynamics of the implied volatility can be inferred from equation (3.8) by taking the time-derivative of $\sqrt{\hat{\sigma}^2}$:

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}t} = \frac{1}{2\hat{\sigma}}\frac{1}{T-t}(\hat{\sigma}^2 - \sigma^2), \tag{3.9}$$

which is in exact accordance with the drift restriction (3.7). Conversely, the average future volatility (3.8) is the only solution of (3.7) that remains finite at T. A specification of implied and spot volatilities should therefore obey some relationship that ensures that the implied volatility is (close to) the expected average spot volatility over the remaining lifetime of the option.

(*e*) *No-bubbles restrictions*

Bubbles in the implied volatilities are an undesirable feature of any pricing model and comparable to the presence of arbitrage opportunities. Therefore, restrictions have to be imposed to prevent bubbles from occurring. This can be achieved by using the last degree of freedom that is left in the model: the stochastic process of the spot volatility σ.

The explosion at time T of the drift of the implied volatility is caused by the terms in $1/(T-t)$ in equation (3.7). Noting that d_1 and d_2 also contain terms of $1/\sqrt{T-t}$ we must therefore require that

$$\lim_{t\to T}\left\{\frac{1}{2\hat{\sigma}}\frac{1}{T-t}(\hat{\sigma}^2-\sigma^2)-\frac{1}{2\hat{\sigma}}d_1d_2v^2+\frac{d_2}{\hat{\sigma}\sqrt{T-t}}\sigma\gamma\right\}<\infty, \quad \forall\hat{\sigma}, \tag{3.10}$$

and that $\hat{\sigma}$ remains bounded a.s., too. Then the SDE for the implied volatilities $\hat{\sigma}$ still has a unique and bounded solution and all price processes are well specified and arbitrage-free.

As equation (3.10) contains only terms in $1/(T-t)$ it is sufficient to have linear convergence to zero of the term in the curly brackets, i.e.

$$\{(\hat{\sigma}^2-\sigma^2)-d_1d_2(T-t)v^2+2d_2\sqrt{T-t}\sigma\gamma\}=O(T-t) \quad \text{as } t\to T. \tag{3.11}$$

Noting that

$$\lim_{t\to T} d_1\sqrt{T-t}=\lim_{t\to T} d_2\sqrt{T-t}=\frac{1}{\hat{\sigma}}\ln\left(\frac{S}{K}\right),$$

this simplifies in the limit of $t\to T$ to

$$\hat{\sigma}^2\sigma^2-2\gamma f\hat{\sigma}\sigma-\hat{\sigma}^4+f^2v^2=0. \tag{3.12}$$

Equation (3.12) can be viewed as a quadratic equation for the spot volatility σ or as a fourth-order polynomial equation for the implied volatility $\hat{\sigma}$. The equation for the spot volatility has the roots

$$\sigma=\frac{\gamma f}{\hat{\sigma}}\pm\sqrt{\hat{\sigma}^2-\frac{f^2}{\hat{\sigma}^2}(v^2-\gamma^2)}. \tag{3.13}$$

Here the positive root has to be taken to ensure a positive relationship between $\hat{\sigma}$ and σ.

Although there are closed-form solutions for the full fourth-order polynomial equation (3.12) for $\hat{\sigma}$, here we only consider the case of $\gamma=0$. Then (3.12) has the root

$$\hat{\sigma}^2=\tfrac{1}{2}\sigma^2+\sqrt{\tfrac{1}{4}\sigma^4+f^2v^2}. \tag{3.14}$$

(The other root would yield a negative value for $\hat{\sigma}^2$.)

Equations (3.13) and (3.14) have several consequences.

(i) For non-stochastic $\hat{\sigma}$, i.e. $v=\gamma=0$, the spot volatility equals the implied volatility $\sigma=\hat{\sigma}$ in the limit as maturity approaches. For times before maturity, the process of the spot volatility follows from the rate of change of the implied volatility via equation (3.7).

(ii) Stochastic implied volatilities require stochastic spot volatilities. Otherwise the spot volatility σ could not converge to its limit as is required in equation (3.13).

(iii) For a given maturity T and different strike prices K and K', the limits of the implied volatilities $\hat{\sigma}(K)$ and $\hat{\sigma}(K')$ as $t\to T$ are linked by (3.14). This yields an implied volatility structure that exhibits the *smile effect* as shown in figure 1.

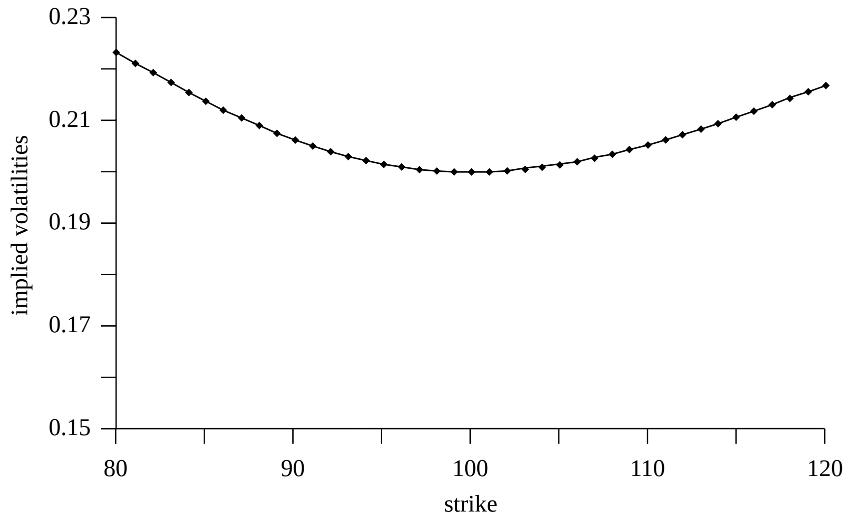

Figure 1. Implied volatilities as $t \to T$. Parameter values: $S = 100$, $\sigma = 0.2$, $v = 0.1$, $\gamma = 0$.

(iv) The extent of the smile effect is directly related to the volatility of volatility, v^2. Thus v^2 need not necessarily be estimated from historical data, it can also be fitted to an observed market smile.

(v) Similar to the smile effect, a 'sneer' (i.e. asymmetry in the implied volatilities) can be incorporated using γ.

(*f*) *Specification of the spot volatility process*

There is considerable freedom in the choice of the specification of the spot volatility process provided equation (3.13) is satisfied as maturity is approached, $t \to T$. Apart from this there is a unique specification of the spot volatility process for any given drift of the implied volatility *under the martingale measure*. Here it should be noted that the specification of the drift of the implied volatility under the martingale measure need not necessarily agree with empirical observations, as this drift will typically contain a risk premium.

Assuming one would like to specify the drift of the implied volatility as a function $u^*(\hat{\sigma}, \dots)$ of the implied volatility (and possibly other parameters) the process of the spot volatility has to be chosen such that equation (3.7) is satisfied, i.e.

$$\hat{\sigma} u^* = \frac{1}{2\tau}(\hat{\sigma}^2 - \sigma^2) - \tfrac{1}{2} d_1 d_2 v^2 + \frac{d_2}{\sqrt{\tau}} \sigma \gamma.$$

This is a quadratic polynomial in σ and its solution is

$$\left.\begin{aligned} \sigma &= \gamma d_2 \sqrt{\tau} + \{\hat{\sigma}^2 - 2\tau\hat{\sigma}u^* + \tau d_2(\gamma^2 d_2 - v^2 d_1)\}^{1/2}, \\ \sigma &= \frac{\gamma}{\hat{\sigma}}(f - \tfrac{1}{2}\hat{\sigma}^2\tau) \\ &\quad + \left\{\hat{\sigma}^2 - 2\tau\hat{\sigma}u^* + \frac{1}{\hat{\sigma}^2}(f - \tfrac{1}{2}\hat{\sigma}^2\tau)[(\gamma^2 - v^2)f - (\gamma^2 + v^2)\tfrac{1}{2}\hat{\sigma}^2\tau)]\right\}^{1/2}. \end{aligned}\right\} \tag{3.15}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

So, given a specific u^*, the spot volatility process that is consistent with u^* is given by equation (3.15). The approach taken here is exactly the opposite of the classical stochastic volatility models. There, after specifying the dynamics of the spot volatility process, the option prices and implied volatilities are derived. Here we specify the process of the implied volatility and derive a consistent spot volatility process. As the spot volatility cannot be observed directly this seems to be a more pragmatic approach.

One natural specification would be to choose $u^* = 0$, i.e. the implied volatility is a martingale. This, and zero correlation $\gamma = 0$, would yield

$$\sigma^2 = \hat{\sigma}^2 - \frac{v^2}{\hat{\sigma}^2}(f^2 - \tfrac{1}{4}\tau^2\hat{\sigma}^4). \tag{3.16}$$

Here again we can see that σ and $\hat{\sigma}$ can be constant only if both are equal. Which of the specifications for the drift of the implied volatility to choose remains part of the modelling problem.

(g) *Implementation of the one-factor model and the pricing of other derivative securities*

By a suitable choice of the spot volatility process $\sigma(\tau, \hat{\sigma}, S)$ as a function of the state variables time (as time to expiry τ), implied volatility $\hat{\sigma}$ and share price (as log-moneyness f), any drift can be supported for the implied volatility process, and the model even keeps its Markovian structure. This makes the implementation of the model in a tree- or finite-difference-based algorithm possible, without needing to recourse to Monte Carlo methods.

There are two state variables: share price S and implied volatility $\hat{\sigma}$. The dynamics of these two state variables under the martingale measure is

$$\mathrm{d}S = rS\,\mathrm{d}t + \sigma(\tau, \hat{\sigma}, S)\,\mathrm{d}W_0, \tag{3.17}$$

$$\mathrm{d}\hat{\sigma} = u^*\,\mathrm{d}t + \gamma\,\mathrm{d}W_0 + v\,\mathrm{d}W_1. \tag{3.18}$$

(Without loss of generality we can collapse the N Brownian motions $W_1, \ldots, W_N$ to one.) Similar to the argument used in the derivation of the drift restriction on the implied volatility, we can derive the restriction on the drift of the price $P(S, \hat{\sigma}, t)$ of any derivative security that is not strongly path dependent. This price can be expanded using Itô's lemma and it turns out that it must satisfy the following partial differential equation:

$$0 = P_t - rP + rSP_S + \tfrac{1}{2}\sigma^2(\tau, \hat{\sigma}, S)S^2P_{SS} + \gamma\sigma SP_{\hat{\sigma}S} + u^*P_{\hat{\sigma}} + \tfrac{1}{2}v^2P_{\hat{\sigma}\hat{\sigma}}, \tag{3.19}$$

with appropriate final and boundary conditions. The price of the security could be written as a function of the state variables because of the Markovian nature of the model set-up.

The correct specification of the boundary conditions and the solution of partial differential equations like (3.19) with finite-difference methods is now standard in options pricing theory.†

The pricing becomes particularly simple if the pay-off does not depend on the share price, like in the case of the implied volatility futures contract VOLAX at the

† See Wilmott *et al.* (1993) for an applied introduction.

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

DTB. Here the pay-off at time T_1 is proportional to the weighted average of the implied volatilities of a basket of options at the money with maturity $T_2 > T_1$. If one simplified the basket of options at the money to one prespecified option, then the pay-off is simply the value of $\hat{\sigma}(T_1)$ at time T_1.

The price of this security is the expected discounted value of the pay-off, thus

$$P(t) = \mathrm{e}^{-r(T_1-t)}\boldsymbol{E}[\hat{\sigma}(T_1)] = \mathrm{e}^{-r(T_1-t)}\left(\hat{\sigma}_0 + \boldsymbol{E}\left[\int_t^{T_1} u^* \,\mathrm{d}s\right]\right). \tag{3.20}$$

The price of the volatility future depends directly on the specification of the drift of the implied volatility under the martingale measure. Thus it is undetermined as long as the process of $\hat{\sigma}$ is not determined. This can be done either by specifying a process for the spot volatility σ and then deriving the process that the implied volatility $\hat{\sigma}$ has to follow, or by directly specifying the process of the implied volatility $\hat{\sigma}$. Furthermore, equation (3.20) can be used to fit u^* to the price of a volatility futures contract.

4. Stochastic forward volatilities

If implied volatilities are given for several options with increasing maturities, T_1, $T_2, \dots, T_n$, then the structure of the implied volatilities can be analysed more clearly if *forward implied volatilities* are used.

(*a*) *Change to local variances*

It is more convenient for the following sections to change the set-up from modelling the implied *volatility* $\hat{\sigma}$ to the modelling of its square, the implied *variance* $\hat{\sigma}^2$. With the definition

$$\mathrm{d}\hat{\sigma} := u\,\mathrm{d}t + v\,\mathrm{d}W + \gamma\,\mathrm{d}W_0, \tag{4.1}$$

Itô's lemma yields

$$\mathrm{d}\hat{\sigma}^2 = (2\hat{\sigma}u + v^2 + \gamma^2)\,\mathrm{d}t + 2\hat{\sigma}v\,\mathrm{d}W + 2\hat{\sigma}\gamma\,\mathrm{d}W_0, \tag{4.2}$$

which can be expanded to

$$\mathrm{d}\hat{\sigma}^2 = U\,\mathrm{d}t + 2\hat{\sigma}v\,\mathrm{d}W + 2\hat{\sigma}\gamma\,\mathrm{d}W_0, \tag{4.3}$$

$$U = \frac{1}{T-t}(\hat{\sigma}^2 - \sigma^2) + (1 - d_1 d_2)v^2 + \frac{2d_2}{\sqrt{T-t}}\gamma\sigma + \gamma^2. \tag{4.4}$$

Equation (4.4) is the no-arbitrage restriction on the drift U of $\hat{\sigma}^2$.

(*b*) *Forward volatilities*

Using the implied variance the concept of a forward implied volatility is easily explained: if the share price S follows a geometric Brownian motion, the variance of the log of the share price at time T is

$$\boldsymbol{E}[(\ln S(T) - \ln S(t))^2] = (T-t)\sigma^2. \tag{4.5}$$

If there are two implied volatilities $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ for the maturities $T_1 < T_2$, the *forward volatility* $\hat{\sigma}_{12}$ over the time-interval $(T_1, T_2]$ is defined as

$$(T_2 - t)\hat{\sigma}_2^2 = (T_1 - t)\hat{\sigma}_1^2 + (T_2 - T_1)\hat{\sigma}_{12}^2. \tag{4.6}$$

This can be rearranged to yield

$$\hat{\sigma}_{12}^2 = \frac{1}{T_2 - T_1}((T_2 - t)\hat{\sigma}_2^2 - (T_1 - t)\hat{\sigma}_1^2). \tag{4.7}$$

Just like the implied volatility $\hat{\sigma}_1$ gives an indication of the market's expectation of the average volatility of the share-price process until T_1, the forward volatility $\hat{\sigma}_{12}$ gives an indication of the expected volatility in the time interval $[T_1, T_2]$. If the volatility of S jumps from $\hat{\sigma}_1$ to $\hat{\sigma}_{12}$ at T_1, then the log share price

$$\ln S(T_2)/S(t)$$

has the variance $(T_2 - t)\sigma_2^2$, which is consistent with the second implied volatility $\hat{\sigma}_2$.

While this interpretation with time-dependent volatility gives a good intuition of the workings of the model, it is not exactly true for *stochastic* volatility.

Furthermore, the forward *implied* volatility is different from the *forward volatility* and the *forward contract on realized volatility* introduced in Dupire (1993*a*, *b*). Dupire's forward contract on volatility captures the market expectation of the realized volatility of S over the relevant interval, while the definition above uses *implied* volatilities from options prices. In a stochastic volatility environment implied volatilities and expected realized average volatilities do not coincide, and forward implied volatilities and forward contracts on realized volatility also differ.

The difference is easily seen for the implied volatilities: the option price is $\boldsymbol{E}[C(\sigma)]$, the expectation of the price of the option as a function of some stochastic spot volatility σ. By definition of the implied volatility this must be equal to $C(\hat{\sigma})$. Because C is a nonlinear function, $\boldsymbol{E}[C(\sigma)]$ is *not* equal to $C(\boldsymbol{E}[\sigma])$. Therefore, the expected average volatility $\boldsymbol{E}[\bar{\sigma}]$ is not equal to the implied volatility $\hat{\sigma}$. The same argument also applies to forward volatilities.

(*c*) *Forward volatility model set-up*

We assume that we are given a set of option maturities $T_1, T_2, \dots, T_M$, the implied volatility $\hat{\sigma}_1$ of the first option (i.e. the option with maturity T_1) and the forward volatilities $\hat{\sigma}_{i,i+1}$ for the later intervals $[T_i, T_{i+1}]$, $i \in \{1, \dots, M-1\}$.

The dynamics of these variables and of the share price S is

$$\left.\begin{aligned} \mathrm{d}S &= rS\,\mathrm{d}t + \sigma S\,\mathrm{d}W_0, \\ \mathrm{d}\hat{\sigma}_1^2 &= U_1\,\mathrm{d}t + \gamma_1\,\mathrm{d}W_0 + v_1\,\mathrm{d}W, \\ \mathrm{d}\hat{\sigma}_{i,i+1}^2 &= U_{i,i+1}\,\mathrm{d}t + \gamma_{i,i+1}\,\mathrm{d}W_0 + v_{i,i+1}\,\mathrm{d}W, \end{aligned}\right\} \tag{4.8}$$

for all $i \in \{1, \dots, M-1\}$, where again $\mathrm{d}W$ is the increment of an N-dimensional Brownian motion and the volatility parameters v are in vector form. Given this specification the values and dynamics of the (direct) implied volatilities can be derived using the following relationships:

$$\hat{\sigma}_n^2 = \frac{1}{T_n - t}\left[(T_1 - t)\hat{\sigma}_1^2 + \sum_{i=1}^{n-1}(T_{i+1} - T_i)\hat{\sigma}_{i,i+1}^2\right], \tag{4.9}$$

$$\gamma_n = \frac{1}{2\hat{\sigma}_n(T_n - t)}\left[2(T_1 - t)\hat{\sigma}_1\gamma_1 + \sum_{i=1}^{n-1}(T_{i+1} - T_i)\gamma_{i,i+1}\right], \tag{4.10}$$

$$v_n = \frac{1}{2\hat{\sigma}_n(T_n - t)}\left[2(T_1 - t)\hat{\sigma}_1 v_1 + \sum_{i=1}^{n-1}(T_{i+1} - T_i)v_{i,i+1}\right]. \tag{4.11}$$

These parameters will be needed in the drift restrictions on the forward volatilities in the next section.

(*d*) *No-arbitrage dynamics of the forward volatilities*

From definition (4.7) it is now easy to derive the arbitrage-free dynamics of the forward volatilities. The dynamics must satisfy

$$\mathrm{d}(\hat{\sigma}_{12}^2) = \frac{1}{T_2 - T_1}[\mathrm{d}((T_2 - t)\hat{\sigma}_2^2) - \mathrm{d}((T_1 - t)\hat{\sigma}_1^2)], \tag{4.12}$$

where the $\hat{\sigma}^2$ follow the arbitrage-free dynamics of equation (4.4). This is expanded to

$$\begin{aligned} \mathrm{d}\hat{\sigma}_{12}^2 &:= U_{12}\,\mathrm{d}t + v_{12}\,\mathrm{d}W + \gamma_{12}\,\mathrm{d}W_0 \\ &= \frac{1}{T_2 - T_1}[(T_2 - t)U_2 - (T_1 - t)U_1 - (\hat{\sigma}_2^2 - \hat{\sigma}_1^2)]\,\mathrm{d}t \\ &\quad + \frac{1}{T_2 - T_1}[(T_2 - t)v_2 - (T_1 - t)v_1]\,\mathrm{d}W \\ &\quad + \frac{1}{T_2 - T_1}[(T_2 - t)\gamma_2 - (T_1 - t)\gamma_1]\,\mathrm{d}W_0. \end{aligned} \tag{4.13}$$

Substitution from the arbitrage-free dynamics of the plain implied volatilities (4.4) yields the drift U_{12} of the forward volatility $\hat{\sigma}_{12}^2$:

$$\begin{aligned} U_{12} &= \frac{1}{T_2 - T_1}[(T_2 - t)U_2 - (T_1 - t)U_1 - (\hat{\sigma}_2^2 - \hat{\sigma}_1^2)] \\ &= \frac{1}{T_2 - T_1}[\tau_2 v_2^2(1 - d_{21}d_{22}) - \tau_1 v_1^2(1 - d_{11}d_{12}) \\ &\qquad + 2\sigma(d_{22}\sqrt{\tau_2}\gamma_2 - d_{12}\sqrt{\tau_1}\gamma_1) + \tau_2\gamma_2^2 - \tau_1\gamma_1^2]. \end{aligned} \tag{4.14}$$

For the absence of direct correlation between volatility and share price $\gamma_1 = 0 = \gamma_2$, it takes a particularly simple form:

$$U_{12} = \frac{1}{T_2 - T_1}[(T_2 - t)v_2^2(1 - d_{21}d_{22}) - (T_1 - t)v_1^2(1 - d_{11}d_{12})]. \tag{4.15}$$

This restriction must be applied to the drift coefficients of the forward volatilities, and in addition to this the restriction in equation (3.7) must be satisfied by the first volatility $\hat{\sigma}_1$. The remarks and the modelling strategy of the previous section still apply to the first implied volatility $\hat{\sigma}_1$; we have just extended the model using the forward volatilities for later maturities.

Note that there are no regularity problems in equations (4.14) and (4.15). The implied volatilities for time-intervals in the future are not directly connected to today's spot volatility and therefore there is no need to achieve direct consistency between both.

(i) *Implementation strategy*

For the implementation of a stochastic implied volatility model with forward volatilities, a Monte Carlo (MC) simulation is the appropriate method. First, the problem will be in at least three dimensions (share price and two implied volatilities) and MC methods are superior for higher-dimensional models, and second there will be path dependence in the model (e.g. through the summation terms in (4.9)) which makes tree- and PDE-based methods unfeasible.

The implementation will have to be done in several steps. First, the dynamics for the spot volatility σ has to be derived using the implied volatility $\hat{\sigma}_1$ with the shortest time to maturity T_1 and its no-arbitrage drift u^*. This is done exactly as in the previous section. This will yield the dynamics of the short end of the term structure of volatilities $\mathrm{d}\hat{\sigma}_1$, of the spot volatility $\mathrm{d}\sigma$ and of the share price $\mathrm{d}S$.

Given this specification and the volatilities $\gamma_{i,i+1}, v_{i,i+1}$ of the forward implied volatilities $\hat{\sigma}_{i,i+1}$, the drifts $U_{i,i+1}$ of the forward implied volatilities can be derived using equations (4.14) and (4.15) in conjunction with equation (4.8).

Now all dynamics are specified and the MC simulation can be run until maturity T_1 of the shortest option. At T_1 the first option disappears and the role of the shortest option $\hat{\sigma}_1$ in the specification of the spot volatility is taken over by the next maturity T_2 and the respective implied volatility $\hat{\sigma}_2$ (which coincides with $\hat{\sigma}_{12}$ at this point).

(*e*) *A continuum of forward volatilities*

In practical applications there will only be a discrete set of available options (and thus of forward volatilities). Nevertheless it has some advantages for analysing the case when there is a full term structure of options. The regularity problems will be resolved very elegantly and the relationship between spot and implied forward volatilities will be uniquely determined, thus removing some potential for misspecification of the drift of the first implied volatility u_1^* in the case of discrete options maturities.

Given a continuous set of implied volatilities $\hat{\sigma}^2(t,T)$ for all maturities $T > t$, we can define the forward volatilities $\hat{\sigma}_f^2(t,T)$ with

$$\hat{\sigma}^2(t,T) = \frac{1}{T-t}\int_t^T \hat{\sigma}_f^2(t,s)\,\mathrm{d}s \tag{4.16}$$

and equivalently

$$\hat{\sigma}_f^2(t,T) = \frac{\partial}{\partial T}[(T-t)\hat{\sigma}^2(t,T)], \tag{4.17}$$

which is the continuous analogue to equations (4.9) and (4.7). The spot volatility at time t (for which we had to make assumptions earlier on) is now given directly by the limit of the implied volatilities as $T \searrow t$ through equation (3.13). We define the short implied volatility

$$\hat{\sigma}(t) := \lim_{T \searrow t} \hat{\sigma}_f(t, T) \tag{4.18}$$

and define $f(t), \gamma(t)$ and $v(t)$ analogously. Substituting into (3.13) yields the spot volatility

$$\sigma(t) := \frac{\gamma f(t)}{\hat{\sigma}(t)} + \sqrt{\hat{\sigma}(t)^2 - \frac{f(t)^2}{\hat{\sigma}(t)^2}(v^2(t) - \gamma^2(t))}. \tag{4.19}$$

The spot volatility must assume this value to ensure absence of arbitrage in the limit of the very short maturity option. For options at the money ($f = 0$) this reduces to $\sigma(t) = \hat{\sigma}(t)$.

The dynamics of the $\hat{\sigma}_f^2(t, T)$ are defined in analogy to equation (4.13) as

$$\mathrm{d}\hat{\sigma}_f^2(t, T) = U_f(t, T)\,\mathrm{d}t + v_f(t, T)\,\mathrm{d}W + \gamma_f(t, T)\,\mathrm{d}W_0, \tag{4.20}$$

and again (given sufficient regularity to interchange the order of integration) we can recover the dynamics of the implied volatilities $\hat{\sigma}^2(t, T)$ in analogy to equations (4.11) and (4.16):

$$\gamma(t, T) = \frac{1}{2(T - t)\hat{\sigma}(t, T)} \int_t^T \gamma_f(t, s)\,\mathrm{d}s, \tag{4.21}$$

$$v(t, T) = \frac{1}{2(T - t)\hat{\sigma}(t, T)} \int_t^T v_f(t, s)\,\mathrm{d}s. \tag{4.22}$$

(*f*) *No-arbitrage dynamics of the forward volatilities*

The restriction on the drift $U_f(t, T)$ of the continuous forward volatility $\hat{\sigma}(t, T)$ follows from the discrete case. In the restriction (4.14) on the drift U_{12} of the discrete forward volatility $\hat{\sigma}_{12}$ we let maturity T_2 approach T_1, i.e. the limit as $T_2 \searrow T_1 =: T$. This yields

$$\begin{aligned} U_{12} &= \frac{1}{T_2 - T_1}[(T_2 - t)U_2 - (T_1 - t)U_1 - (\hat{\sigma}_2^2 - \hat{\sigma}_1^2)] \\ U_f(t, T) &= \lim_{T_2 \searrow T_1} U_{12} \\ &= \frac{\mathrm{d}}{\mathrm{d}T}[(T - t)U(t, T) - \hat{\sigma}^2(t, T)]. \end{aligned} \tag{4.23}$$

The no-arbitrage drift $U(t, T)$ of the T-maturity implied volatility is given in equation (4.4), which makes the evaluation of (4.23) a matter of straightforward but tedious algebra. Alternatively, one could perform the differentiation in (4.23) numerically when the model is implemented. This can be done without major losses in accuracy as most of the parameters have to be evaluated numerically anyway.

Although the expression (4.23) for the drift restriction looks rather complicated, it is still preferable to the drift restrictions derived by Derman & Kani (1998) in a similar context. Derman & Kani (1998) derive drift restrictions on the conditional local volatility of the share price, and the restrictions involve a double integral of which

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

one is infinite. Here we have an (admittedly complicated) expression in elementary functions with only finite integrals of the relevant parameters from t to T as they appear in the interest-rate model of Heath *et al*. (1992), which was the inspiration for this model.

(*g*) *Implementation*

The implementation of the continuous-maturity version of the model is very similar to the MC implementation of the discrete-maturity model of the previous section. The only difference is that we are now relieved from the task of specifying a drift u^* for the first implied volatility to recover the spot volatility σ. Now the spot volatility is given directly by equation (4.19), which in turn defines the share-price process $\mathrm{d}S = rS\,\mathrm{d}t + \sigma S\,\mathrm{d}W_0$.

For the forward volatilities $\hat{\sigma}_f(t,T)$ we have to specify the initial values $\hat{\sigma}_f(0,T)$, their volatilities $v(t,T)$ and their correlations $\gamma(t,T)$ with the share price. The drifts follow from equation (4.23).

Now the model dynamics can be simulated using standard MC techniques. Here after each time-step the new drift restrictions have to be calculated.

5. Conclusion

In this paper a class of stochastic volatility models is presented that is based on implied volatilities that are observed in the prices of liquidly traded options. It is shown how to derive a consistent spot volatility process and which restrictions have to be satisfied to ensure absence of arbitrage in the model.

The approach taken here is fundamentally different from classical stochastic volatility models where the spot volatility is taken as a fundamental variable, and we believe it has several advantages.

First, for the implementation of the model the estimation of the relevant parameters (the volatility of the implied volatility) is much facilitated because implied volatilities are directly observable in market prices.

Second, the model will be automatically fitted to the fundamental options prices, and the additional information that is reflected in their implied volatilities is also incorporated in the model. This ability to fit is only comparable to models of the implied-tree class, but this model incorporates stochastic dynamics which most implied-tree models do not.

Third, the extension of the model to a multifactor setting has been demonstrated. In its multifactor versions (either with discrete or with continuous sets of implied volatilities) the model is capable of reproducing much richer dynamics than one-factor models.

The stochastic implied-tree model by Derman & Kani (1998) is the model that is closest in scope and philosophy to this model. Nevertheless, the reader will have realized by now that there are fundamental differences between both approaches, most importantly the no-bubbles restrictions (which are not in Derman & Kani (1998)) in this model and the market-based approach (as opposed to Derman & Kani's (1998) 'effective volatility' approach).

Although the analysis in this model is based on European call options, the methods presented can also be used with the implied volatilities of other options (e.g. options

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

of American type) as underlying factors. Then, the partial derivatives of the options are needed to derive the no-arbitrage drift restrictions (see, for example, § 3 *c*), but qualitatively the model would not change.

Another interesting extension of the paper would be the incorporation of independent dynamics for options of the same maturities but different strike prices. The problem here is that the no-bubbles restrictions still must be satisfied as maturity approaches. Thus the final value for the implied volatilities would be predetermined (via the value of the spot volatility and equation (3.13)). Further research will have to show whether there is a sufficiently simple way to ensure the final condition while still allowing richer dynamics within the smile.

Financial support by the Deutsche Forschungsgemeinschaft (DFG) through SFB 303 at the University of Bonn is gratefully acknowledged.

References

Black, F. 1976 The pricing of commodity contracts. *J Financial Econ.* **3**, 167–179.

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

Brace, A., Gatarek, D. & Musiela, M. 1997 The market model of interest rate dynamics. *Math. Finance* **7**, 127–155.

Chiras, D. P. & Manaster, S. 1978 The information content of option prices and a test of market efficiency. *J. Financial Econ.* **6**, 213–234.

Derman, E. & Kani, I. 1994 Riding on the smile. *Risk* **7**, 32–29.

Derman, E. & Kani, I. 1998 Stochastic implied trees: arbitrage pricing with stochastic term and strike structure of volatility. *Int. J. Theor. Appl. Finance* **1**, 61–110.

Derman, E., Kani, I. & Chriss, N. 1996 Implied trinomial trees of the volatility smile. *J. Derivatives* **4**, 7–22.

Dumas, B., Fleming, J. & Whaley, R. E. 1999 Implied volatility functions: empirical tests. *J. Finance* **53**, 2059–2106.

Dupire, B. 1993*a* Arbitrage pricing with stochastic volatility. Working paper, Paribas Capital Markets, Swaps and Options Research Team.

Dupire, B. 1993*b* Pricing and hedging with smiles. Working paper, Paribas Capital Markets, Swaps and Options Research Team.

Dupire, B. 1994 Pricing with a smile. *Risk* **7**, 18–20.

Fama, E. F. 1970 Efficient capital markets: a review of theory and empirical work. *J. Finance* **25**, 383–420.

Fleming, J., Ostdiek, B. & Whaley, R. E. 1995 Predicting stock market volatility: a new measure. *J. Futures Markets* **15**, 265–302.

Heath, D., Jarrow, R. & Morton, A. 1992 Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica* **60**, 77–105.

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

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

Jamshidian, F. 1997 LIBOR and swap market models and measures. *Finance Stochastics* **1**, 293–330.

Karatzas, I. & Shreve, S. E. 1991 *Brownian motion and stochastic calculus*. Springer.

Miltersen, K. R., Sandmann, K. & Sondermann, D. 1997 Closed form solutions for term structure derivatives with log-normal interest rates. *J. Finance*, **52**, 409–430.

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

Rubinstein, M. 1995 As simple as one, two, three. *Risk* **8**, 44–47.

Skiadopulous, G., Hodges, S. & Clewlow, L. 1998 The dynamics of smiles. Working paper, Financial Options Research Centre, Warwick Business School, University of Warwick, UK.

Stein, E. & Stein, J. 1991 Stock price distributions with stochastic volatility: an analytic approach. *Rev. Financial Stud.* **4**, 727–752.

Wilmott, P., Howison, S. & Dewynne, J. 1993 *Option pricing: mathematical models and computation*. Oxford: Oxford Financial Press.