Hedging with Stochastic and Local Volatility

ISMA Centre Discussion Papers in Finance 2004-11
Preliminary Version: July 2004
This Version: December 2004

Carol Alexander
ISMA Centre, University of Reading
c.alexander@ismacentre.rdg.ac.uk

Leonardo M. Nogueira
ISMA Centre, University of Reading
l.nogueira@ismacentre.rdg.ac.uk

Copyright 2004 Alexander and Nogueira. All rights reserved.
Hedging with Stochastic and Local Volatility

Abstract
We derive the local volatility hedge ratios that are consistent with a stochastic instantaneous volatility and show that this ‘stochastic local volatility’ model is equivalent to the market model for implied volatilities. We also show that a common feature of all Markovian single factor stochastic volatility models, (log)normal mixture option pricing models and ‘sticky delta’ models is that they predict incorrect dynamics for implied volatility. As a result they over-hedge the Black-Scholes model in the presence of a market skew and this explains the poor delta hedging performance of these models reported in the literature. Whilst the traditional ‘sticky tree’ local volatility models do not possess this unfortunate property, they cannot be used for pricing without exogenous and ad hoc smoothing of results. However the stochastic local volatility framework allows one to extend a good pricing model into a good hedging model. The theoretical results are supported by an empirical analysis of the hedging performance of seven models, each with different volatility characteristics, on the SP500 index skew.

JEL Classification: G12, C16
Keywords: Local volatility, stochastic volatility, implied volatility, hedging, dynamic delta hedging, volatility dynamics.

Acknowledgement
We are most grateful for very useful comments on an earlier draft from Dr. Hyungsok Ahn of Commerzbank, London, Professor Emanuel Derman of Columbia University, NY, Dr. Bruno Dupire of Bloomberg, NY and Dr. Jacques Pézier of the ISMA Centre, UK.
I Introduction

Since the global equity crash in 1987 the Black and Scholes (1973) equity implied volatilities have exhibited a steep negative skew, where implied volatilities generally increase as the strike decreases. For an option pricing model to be consistent with this skew an intuitive choice is to allow the instantaneous volatility (or its square, the instantaneous variance) of the equity price process to be a function of a diffusion process itself, possibly correlated with the underlying asset. This is the approach of Hull and White (1987) and Heston (1993), among others.1 By adding a new source of uncertainty to the model it is possible to fit the observed market prices of options. But there is a cost. With two sources of uncertainty in the model, a market with only the underlying asset and a risk-free money market account is incomplete, since it is no longer possible to replicate the payoff of a simple European option and delta hedging alone is not sufficient to eliminate risk.

Then Dupire (1994), Derman and Kani (1994) and Rubinstein (1994) defined a unique instantaneous volatility that is a deterministic function of time and asset price level only and is consistent with observed market prices of options. Using ‘Dupire’s equation’ it is possible to fit any continuous market smile or skew exactly and to ‘lock-in’ a forward instantaneous volatility surface, called the ‘local volatility’ surface, assuming this surface does not change as the underlying asset price moves. Since no new source of uncertainty is introduced the market is complete and delta hedging is possible. However, several papers have tested the delta hedging performance of local volatility models and find that they perform worse than the Black-Scholes model (see Dumas, Fleming and Whaley (1998), Hagan et al. (2002) and others). A conclusion of these studies might be that the assumption of a static local volatility function is too restrictive and that stochastic volatility models may be more realistic.

Stochastic and local volatility models have been regarded as two alternative and competing approaches to the same unobservable quantity, the instantaneous volatility of the underlying asset. The former represents the instantaneous variance as a diffusion or jump-diffusion process and the latter derives instantaneous and forward volatilities that are consistent with a ‘snapshot’ of implied volatilities at a particular time. In fact, these two approaches are not inconsistent, but the few attempts to unify them into a single theory have not been developed by further research.

The heart of the hedging problem is the assumption of a deterministic instantaneous volatility that is imposed by most local volatility models. However this assumption is not actually necessary for a local volatility model. This was recognized by Dupire (1996) and Kani, Derman and Kamal (1997) who define the local variance (i.e. the square of the local volatility) as the expectation of the future instantaneous variance conditional on a given asset price level. More specifically, the local volatility at time $t_0 < t$ is the square root of

1 For a review of volatility smile consistent models, and local and stochastic volatility models in particular, see Skiadopoulos (2001).
\[
\sigma_L^2(t, S) = E^0 \{ \sigma^2(t, S(t), x(t)) \mid S(t) = S \} \tag{1}
\]

where \( E^0 \) denotes the expectation under the risk-neutral probability conditional on a filtration \( \mathcal{F}_0 \) which includes all information up to time \( t_0 \) and \( x(t) = \{x_1(t), \ldots, x_n(t)\} \) is a vector of all sources of uncertainty that may influence the instantaneous volatility process at time \( t \), other than the asset price \( S(t) \).

Therefore, even when the instantaneous volatility is stochastic, the local volatility function is a deterministic function of time \( t \) and the future asset price. In fact \( x(t) \) can be very general: it can be any ‘arbitrage-free’ set of continuous stochastic processes (Appendix B describes some explicit no-arbitrage conditions). Hence the definition (1) can be consistent with any univariate diffusion stochastic volatility model (e.g. Hull and White (1987) and Heston (1993)). For this reason, Dupire (1996) names model (1) the ‘unified theory of volatility’.

So what is the problem with local volatility models? It is precisely the residual uncertainty in \( x(t) \). In particular, by taking the expectation in (1) we ignore the variance of \( x(t) \) and its influence on the instantaneous volatility. This uncertainty is transferred to the local volatility surface itself. That is, although locally (i.e. at each calibration) the surface is indeed a deterministic function of \( t \) and \( S \), over time that surface moves in an unpredictable manner, i.e. its dynamics are stochastic. The residual uncertainty from \( x(t) \) does not just disappear from the model.

The consequence of ignoring residual uncertainty is that local volatility surfaces become very unstable. The model is incomplete because its assumptions are inconsistent with the implied volatility dynamics that we observe in the market. The assumption of a deterministic instantaneous volatility is inconsistent with any dynamics for local volatility, yet the surface does, typically, change considerably on recalibration and this is entirely unpredictable in the context of a local volatility model. Some techniques aim to increase robustness in calibration. Andersen and Andreasen (2000) overlay Dupire’s deterministic diffusion dynamics with a jump diffusion for the underlying price and Bouchouev and Isakov (1997, 1999) and Avellaneda et al. (1997) apply regularization methods. But this should not be necessary with a more complete specification of the model’s dynamics. By introducing explicit, stochastic dynamics for the parameters of the local volatility function we find that many of the limitations and criticisms of local volatility models disappear. We shall call this more general framework the ‘stochastic local volatility’ model to distinguish it from the traditional characterization of local volatility.

In this paper we first prove an interesting ‘duality’ result: that stochastic local volatility is equivalent to a multivariate stochastic volatility diffusion, called the ‘market model’ of implied volatility, that has been studied in various forms by Schönbucher (1999), Brace et al. (2001) and Ledoit et al. (2002). Hence the two approaches
will yield the same claim prices and hedge ratios, although stochastic local volatility models are much easier to calibrate than the market models. The result is interesting because it shows that the stochastic volatility and local volatility approaches can be unified within a general framework and it is only when these approaches take a restricted view of volatility dynamics that they appear to be different.

The main focus of this paper is on the hedging performance of various deterministic and stochastic volatility models. At each trading date each model is calibrated to an implied volatility smile surface from market prices of European calls and puts. We show that with stochastic local volatility (or equivalently, with the market model of implied volatilities) the delta, gamma and theta are equal to an associated deterministic local volatility hedge ratio plus an adjustment factor that depends on the degree of uncertainty in local volatility parameters and on their correlation with the underlying asset price. Our empirical results confirm that this adjustment can indeed improve the hedging performance of deterministic local volatility models.

The remainder of this paper is as follows: Section II introduces stochastic local volatility (SLV) and derives new dynamics for the local volatility model price of a contingent claim; Section III derives the implied volatility dynamics that are consistent with the new claim price dynamics and proves the duality between the SLV model and the market model of implied volatilities; Section IV examines the relationship between local volatility, stochastic local volatility, stochastic volatility and Black-Scholes (BS) hedge ratios; Section V introduces the various models used in the hedging race and compares their empirical hedging performance; Section VI summarizes and concludes.

II. Extending the Dynamics of Local Volatility

The definition of local volatility is based on the early work of Dupire (1993, 1994), Derman and Kani (1994) and Rubinstein (1994). It assumes the underlying asset price process follows a geometric Brownian motion with deterministic forward volatility \( \sigma(t, S) \) under the risk-neutral measure as:

\[
dS = (r - q)Sdt + \sigma(t, S)dW_S
\]

where \( r \) denotes the risk-free interest rate and \( q \) is the dividend yield. The celebrated ‘Dupire’s equation’ (3) shows how the local volatility \( \sigma_L(t, S) \) is uniquely determined from a surface of market prices \( f(T, K) \) of standard European options with different strikes and maturities, as:

\[
\sigma^2_{L}(t, S) \bigg|_{T=S=K} = 2 \left( \frac{\partial f}{\partial T} + (r - q)K \frac{\partial f}{\partial K} + qf \right) \bigg/ K^2 \frac{\partial^2 f}{\partial K^2}
\]

Note that \( \sigma(t, S) \) in (2) is not a function of any other (stochastic) variables \( \mathbf{x}(t) \) so (3) is consistent with the general ‘forward equation’ derived by Kani, Derman and Kamal (1997). In this case the local volatility is equal
to the instantaneous forward volatility $\sigma(t, S)$ at any future time $t$ conditional on the asset level $S$. In effect, the forward instantaneous volatilities are ‘locked-in’ by the current prices of European options, i.e. the local volatility model assumes that they will be realised with certainty.

Since Dupire’s equation requires a continuum of traded options prices, direct computation of the local volatility function using finite difference methods is problematic. The local volatility surface can be very irregular and sensitive to the interpolation methods used between quoted option prices and their extrapolation to boundary values, requiring some ‘regularization method’ to obtain the smoothest possible fit to the implied volatility surface (see e.g. Bouchouev and Isakov (1997, 1999) and Avellaneda et al. (1997)). As a result most of the recent work on local volatility has focused on the use of parametric forms for local volatility functions: see Dumas, Fleming and Whaley (1998), Brown and Randall (1999), McIntyre (2001), Brigo and Mercurio (2001, 2002), Alexander (2004), and many others. In this case the local volatility function is calibrated by changing parameters so that some distance metric between model prices and market prices is minimized. Of course, with fewer parameters than prices, the parameterized local volatility model will not fit quoted prices exactly.

In these parametric local volatility models, at any point in time $t_0$ the values $v(t_0) = \{v_1(t_0), \ldots, v_n(t_0)\}$ for the local volatility parameters are calibrated to the current implied volatility surface. The underlying asset price process assumed at time $t_0$ is then:

$$dS = (r - q)Sdt + \sigma(t, S; v(t_0))SdW^S$$ for all $t > t_0$$

and $v(t_0)$ is known at time $t_0$. Since the spot volatility is deterministic, putting $x(t) \equiv v(t_0)$ in (1) the local volatility is given by:

$$\sigma^2(t, S; v(t_0)) = \sigma^2(t, S; v(t_0))$$ for all $t > t_0$$

We have stressed that the local volatility will be sensitive to the calibration at time $t_0$. When at time $t_1 > t_0$ the model is re-calibrated, we will have:

$$\sigma^2(t_1, S; v(t_1)) = \sigma^2(t_1, S; v(t_1))$$ for all $t > t_1$$

and this can of course differ from (5) as long as $v(t_1) \neq v(t_0)$. In fact, the dynamics of the local volatility surface will be stochastic if the calibrated parameters $v(t)$ are stochastic.

Let us assume that all the uncertainty in the random variables $x(t)$ in (1) can be captured by the dynamics of the parameters $v(t) = \{v_1(t), \ldots, v_n(t)\}$ of the local volatility model. Then at time $t_0$ the spot variance $\sigma^2(t, S; v(t))$ is a function of $t$, $S$ and $v(t)$, with $v(t)$ being stochastic because these parameters are not calibrated to the market until some future time $t > t_0$. Note that only in the special case that the parameters $v(t)$ are constant...
and equal to \( v(t_0) \) do we have the local volatility model (5) above. In general, when we allow \( v(t) \) to evolve stochastically, we have the *stochastic* local volatility function:

\[
\sigma^2_{L}(t, S) = \mathbb{E}^{0} \left\{ \sigma^2(t, S(t); v(t)) \mid S(t) = S, v \right\} = \int_{\Omega} \sigma^2(t, S; v) b_{S}(v \mid S) dv \quad \text{at time } t_0
\]  

(6)

In (6), as in (1), the expectation is conditional on a filtration \( \mathcal{F}_{t_0} \), which includes all information up to time \( t_0 \). In particular the past history of \( S(t) \) and of every additional stochastic factor in the spot variance must be included in \( \mathcal{F}_{t_0} \) so that the local volatility surface is well-defined for every pair \((t, S)\) with \( t > t_0 \). The integration is over \( \Omega_t \), the space of all arbitrage-free values for \( v(t) \) and \( v \in \Omega_t \) is a realization of \( v(t) \). Finally \( b_{S}(v \mid S) \) denotes the multivariate density of \( v(t) \) conditional on a given \( S \) at time \( t \), and given \( \mathcal{F}_{t_0} \). With this definition the local volatility function \( \sigma_{L}(t, S) \) has an implicit dependence on the future parameters \( v(t) \) – and their past values, through the filtration \( \mathcal{F}_{t_0} \) – so the local volatility becomes stochastic. We call (6) the ‘stochastic local volatility’ (SLV) model.

When \( v(t) \) is stochastic the local volatility and the forward instantaneous volatility are no longer the same. However, in Appendix A we show that, if the calibration of a deterministic local volatility model fits market prices for standard European options precisely, then the same local volatility surface will be given by both (5) and (6). That is:

\[
\sigma^2(t, S; v(t_0)) = \mathbb{E}^{0} \left\{ \sigma^2(t, S(t); v(t)) \mid S(t) = S \right\} \quad \text{for all } t > t_0
\]  

(7)

This is an important constraint on the permissible dynamics for \( v(t) \). In fact, using a Taylor series expansion of the forward variance, Appendix A also shows that:

\[
\sigma^2(t, S; v(t_0)) = \sigma^2(t, S; \tilde{v}(t_0)) + \frac{1}{2} \sum_{i} \sum_{j} \frac{\partial^2 \sigma^2(t, S; \tilde{v}(t_0))}{\partial v_i \partial v_j} \text{Cov}^{0} \left\{ v_i(t), v_j(t) \mid S(t) = S \right\}
\]  

(8)

where \( \tilde{v}(t_0) = \mathbb{E}^{0} \left\{ v(t) \mid S(t) = S \right\} \). Therefore \( \sigma^2(t, S; v(t_0)) \neq \mathbb{E}^{0} \left\{ v(t) \mid S(t) = S \right\} \).

Hence there is no obvious relationship between the calibrated parameters at time \( t_0 \) and the future parameters \( v(t) \) unless the spot variance is either deterministic or a linear function of the parameters \( v(t) \), when the second derivative term in (8) will be zero.

---

2 In practice, it is hard to validate equation (7). It requires the calibrated model prices for standard European options to match a continuum of market prices, but clearly these are not available in the market. This poses the problem of uniqueness, as we could find several continuums that match the market prices. But in practice this is often ignored, at least when many options for a variety of strikes and maturities are available.
Although the local volatility surface can fit the current smile, the assumption of a deterministic instantaneous volatility yields unrealistic model dynamics. In the same way that stochastic volatility models extend the Black-Scholes assumptions to allow more realistic variance dynamics, local volatility models can be extended to account for the uncertainty in the future values of their parameters. We can now formalize the stochastic local volatility model and derive the corresponding dynamics for the price of a contingent claim. These results are necessary to derive the appropriate hedge ratios, and to prove the duality between the SLV model and the ‘market model’ of stochastic implied volatility.

Assume that the asset price process follows a geometric Brownian motion under the risk-neutral measure:

\[ dS = (r - q)Sdt + \sigma(t, S; \nu(t))SdW_S \]  

(9)

in which the instantaneous volatility \( \sigma(t, S; \nu(t)) \) is continuous and satisfies the no-arbitrage conditions in Appendix B. Assume that the continuously-compounded risk-free rate \( r \) and dividend yield \( q \) are constant and \( \nu(t) = \{ \nu_1(t), \nu_2(t), \ldots, \nu_n(t) \} \) is a vector of stochastic parameters that are correlated with the asset price \( S(t) \) and with each other. Suppose the risk-neutral dynamics for each parameter \( \nu_i \) in \( \nu(t) \) are as follows:\n
\[ d\nu_i = \alpha_i(t, S, \nu)dW_i + \beta_i(t, S, \nu)dZ_i \]

with \n
\[ dZ_i = \Omega_{i,S}(t, S, \nu)dW_S + \sqrt{1 - \rho_i^2(t, S, \nu)}dW_i \]

\[ dZ_i, dZ_j \xrightarrow{d.s.a} 0 \quad \text{and} \quad dW_i, dW_j \xrightarrow{d.s.a} 0 \quad \text{for} \ i, j \in \{1, 2 \ldots n\} \]

(10)

satisfying, almost surely, the usual regularity conditions and for all \( T > t_0 \):

\[ \int_{t_0}^{T} |\alpha_i(t, S, \nu)| dt < \infty \quad \text{and} \quad \int_{t_0}^{T} |\beta_i^2(t, S, \nu)| dt < \infty \]

Here \( \rho_{ij} \in [-1, 1] \) is the correlation between variations in \( \nu_i \) and \( \nu_j \) and \( \rho_{i,S} \in [-1, 1] \) is the correlation between variations in \( \nu_i \) and \( S \). As before, \( \nu \in \Omega_i \) is a realization of \( \nu(t) \), where \( \Omega_i \) is the space of arbitrage-free values for \( \nu(t) \). Together (9) and (10) provide the full specification of the SLV model.

---

3 Kani, Derman and Kamal (1997) note the analogy with the Heath-Jarrow-Morton (HJM) model for interest rates (Heath, Jarrow and Morton, 1992): “forward rates are the arbitrage-free interest rates at future times that can be locked in by trading bonds today. Similarly, local volatilities are the arbitrage-free volatilities at future times and market levels that can be locked in by trading options today”. Indeed, spot and local volatilities are analogous to spot and forward interest rates, with the stochastic local volatility surface being the analogue of the forward yield curve, which is assumed to be stochastic in the HJM framework.

4 In (10) we allow all coefficients to depend on all variables in the model so that the parameter dynamics can be as general as possible, including a variety of reasonable implementations for each parameter \( \nu_i \), e.g. arithmetic or geometric Brownian motions, mean-reverting, etc. There is also an implicit dependence on the filtration \( \mathcal{F}_t \) at time \( t_0 \).
Denote the local volatility price, calibrated at time \( t_0 \), of a contingent claim by \( f_L = f_L(t, S(t); v(t) | \mathcal{F}_0) \). So \( v(t) \) is included in the filtration \( \mathcal{F}_0 \). Since \( v(t) \) contains the future parameters of a deterministic local volatility model the claim price \( f_L \) must satisfy the following partial differential equation at each time \( t > t_0 \):

\[
\frac{\partial f_L(t, S; v)}{\partial t} + (r - q)S \frac{\partial f_L(t, S; v)}{\partial S} + \frac{1}{2} \sigma^2(t, S; v) S^2 \frac{\partial^2 f_L(t, S; v)}{\partial S^2} = r f_L(t, S; v)
\]

in which the filtration \( \mathcal{F}_0 \) has been omitted for convenience. Equation (11) only holds locally, i.e. assuming the DLV model is re-calibrated at each time \( t \). But since in general the calibrated parameters change at each re-calibration we assume \( v(t) \) is stochastic and defined as in (10) above. Then, under assumptions (9) and (10), the risk-neutral dynamics of the model price \( f_L(t, S(t); v(t) | \mathcal{F}_0) \) are given by the following theorem.

**Theorem 1**

Under assumptions (9) and (10) and assuming all no-arbitrage conditions in Appendix B are satisfied, the risk-neutral dynamics of the model price \( f_L(t, S(t); v(t) | \mathcal{F}_0) \) are given by:

\[
df_L = rf_L dt + \left( \sigma S \frac{\partial f_L}{\partial S} + \sum_i \beta_i \frac{\partial f_L}{\partial v_i} \right) dW_S + \sum_i \beta_i \sqrt{1 - \gamma_i^2} \frac{\partial f_L}{\partial v_i} dW_i \tag{12}
\]

and the coefficients in (10) must satisfy the following drift condition for every \( t \in (t_0, T) \):

\[
\sum_i \left( \alpha_i \frac{\partial f_L}{\partial v_i} + \sigma S \beta_i \frac{\partial f_L}{\partial S} + \gamma_i \frac{\partial f_L}{\partial v_i} \right) = 0 \tag{13}
\]

**Proof:** This follows from Ito’s lemma using (9), (10) and the standard risk-neutrality argument (Appendix A).

The dynamics (12) contrast with the typical option price dynamics from deterministic local volatility models, in which no stochastic behavior for the model parameters is taken into account. The drift condition (13) implies that if \( \beta_i = 0 \) for every parameter \( v \) then also \( \alpha_i = 0 \) for every parameter \( v \) because (13) must hold for any claim. This places a strong constraint on the dynamics of any local volatility surface: if the volatility surface moves at all, it does so stochastically. In accordance with this, if in (12) we put \( \beta_i = 0 \) for every parameter \( v \) in the SLV model we have the standard local volatility option price risk-neutral dynamics, in which the volatility surface is fixed, i.e.

\[
df_L = rf_L dt + \sigma S \frac{\partial f_L}{\partial S} dW_S \tag{14}
\]

---

5 Within a deterministic local volatility model, \( v(t) \) is assumed constant, hence \( f_L \) can be expressed as a function of \( t \) and \( S \) only. Then (11) follows from application of Ito’s lemma and the standard risk-neutrality argument.
Otherwise, if $\beta_i \neq 0$, there are two special cases to consider according as (a) $\rho_{i,S} = 1$ and (b) $\rho_{i,S} = 0$. When
instantaneous volatility parameters are perfectly correlated with the underlying, no new source of uncertainty is introduced so (12) becomes:

$$df_L = \sigma S \left( \frac{\partial f_L}{\partial S} + \sum_i \beta_i \frac{\partial f_L}{\partial \psi_i} \right) dW^v + \sum_i \beta_i \frac{\partial f_L}{\partial \psi_i} dW^i$$

(15)

where the diffusion coefficient is modified but the instantaneous volatility is still ‘deterministic’. In (b), the claim price dynamics will be driven by a multi-factor model:

$$df_L = \sigma S \frac{\partial f_L}{\partial S} dW^v + \sum_i \beta_i \frac{\partial f_L}{\partial \psi_i} dW^i$$

(16)

When the parameters $\psi(t)$ of the local volatility model are stochastic with less than perfect correlation with asset price movements, the claim price has multi-factor dynamics with one Brownian motion from the underlying asset price dynamics (9) and another Brownian motion for each stochastic parameter in the model. This is very different from the standard local volatility model dynamics (14). Of course different dynamics imply different sensitivities and different hedge ratios: these will be derived in section IV.

III. Implied Volatility Dynamics

Recent work of Dupire (2003) derives a general relationship between local volatilities and Black-Scholes implied volatilities, in which implied volatilities are gamma-weighted averages of local volatilities. This raises the question of duality between the stochastic local volatility model and a stochastic implied volatility model. This section formalizes the duality result by deriving an explicit relationship between the stochastic local volatility price dynamics for a standard European option and the evolution of the associated implied volatility.

For a vanilla European option with strike $K$ and maturity $T$, the local volatility price of this option at time $t$ when the asset price is $S$ is denoted by $f_L(K, T; t, S, \psi)$. Note that we term $f_L$ the ‘local volatility’ price, with no distinction between the standard local volatility model price and the price obtained when its dynamics are adjusted for stochastic instantaneous volatility. Each time the local volatility model is recalibrated the claim prices will be identical, it is just their dynamics are different.

When the implied volatility is 0, we denote the BS price of this option at time $t$ when the asset price is $S$ by $f_{BS}(K, T; t, S, 0)$. We define the market implied volatility $\theta_M = \theta_M(K, T; t, S)$ as that $\theta$ such that the BS model price equals the observed market option price. Since market prices are observable, market implied volatilities are observable. Now assume that the local volatility model is calibrated to an implied volatility surface at each
time \( t \). Then the *local* implied volatility \( \theta = \theta(K, T; t, S, v) \) is defined by equating the local volatility price to the BS price conditional on the filtration \( \mathcal{F}_0 \):

\[
f_L(K, T; t, S, v) = f_{BS}(K, T; t, S, \theta) \tag{17}
\]

The following results are derived using local implied volatilities, not market implied volatilities. That is, we derive the relationship between a stochastic local volatility function and the associated local implied volatilities on the assumption that the parametric local volatility model can fit market options prices on any day with acceptable accuracy.

To prove the theorem of this section, we first need two lemmas that focus on the sensitivities of the local implied volatility surface \( \theta(K, T; t, S, v) \) to changes in \( t, S \) and \( v \). While the sensitivities to \( K \) and \( T \) can be derived from a ‘snapshot’ of the surface, the sensitivities to \( t, S \) and \( v \) depend on the dynamics of the implied volatility surface. There is a large empirical literature on the implied volatility sensitivity to \( S \) (see Derman and Kamal (1997), Skiadopoulos et al. (1999), Alexander (2001), Cont and da Fonseca (2002), Fengler et al. (2003) and others) but we shall take a theoretical approach here. We assume that the parameters of a local volatility model evolve stochastically as specified in (10) and we derive the implied volatility dynamics that are consistent with this.

**Lemma 1**

Denote the BS model price sensitivities by \( \delta_{BS} = \frac{\partial f_{BS}}{\partial S}; \gamma_{BS} = \frac{\partial^2 f_{BS}}{\partial S^2}; \Theta_{BS} = \frac{\partial f_{BS}}{\partial t}; \nu_{BS} = \frac{\partial f_{BS}}{\partial \theta}; \kappa_{BS} = \frac{\partial^2 f_{BS}}{\partial \theta^2}; \) and \( \Omega_{BS} = \frac{\partial^2 f_{BS}}{\partial S \partial \theta} \). Denote the local volatility sensitivities to \( S \) and \( t \) by \( \delta_L = \frac{\partial f_L}{\partial S}; \gamma_L = \frac{\partial^2 f_L}{\partial S^2}; \) and \( \Theta_L = \frac{\partial f_L}{\partial t} \). Then we can derive the local implied volatility function \( \theta(K, T; t, S, v) \) sensitivities to \( t, S \) and \( v \) as:

\[
\begin{align*}
\frac{\partial \theta(K, T; t, S, v)}{\partial t} &= \frac{\Theta_L(K, T; t, S, v) - \Theta_{BS}(K, T; t, S, \theta)}{\nu_{BS}(K, T; t, S, \theta)} \\
\frac{\partial \theta(K, T; t, S, v)}{\partial S} &= \frac{\delta_L(K, T; t, S, v) - \delta_{BS}(K, T; t, S, \theta)}{\nu_{BS}(K, T; t, S, \theta)} \\
\frac{\partial \theta(K, T; t, S, v)}{\partial v} &= \frac{1}{\nu_{BS}(K, T; t, S, \theta)} \frac{\partial f_L(K, T; t, S, v)}{\partial v} \\
\frac{\partial^2 \theta(K, T; t, S, v)}{\partial S^2} &= \frac{1}{\nu_{BS}} \left[ \gamma_L - \gamma_{BS} - \kappa_{BS} \left( \frac{\partial \theta}{\partial S} \right)^2 - 2\Omega_{BS} \frac{\partial \theta}{\partial S} \right] \\
\frac{\partial^3 \theta(K, T; t, S, v)}{\partial S^2 \partial v} &= \frac{1}{\nu_{BS}} \left[ \frac{\partial^2 f_L}{\partial S^2} \frac{\partial \theta}{\partial v} + \frac{\partial \theta}{\partial S} \frac{\partial^2 f_L}{\partial S \partial v} \left( \Omega_{BS} + \frac{\partial \theta}{\partial S} \right) \right]
\end{align*}
\]
\[
\frac{\partial^2 \Theta(K, T; t, S, \mathbf{v})}{\partial v_j \partial v_j} = \frac{1}{\nu_{BS}} \left[ \frac{\partial^2 f_{L_{i}},}{\partial v_i \partial v_j} - \nu_{BS} \frac{\partial \nu_{L_{i}}}{\partial v_j} \right]
\]

**Proof:** Differentiate (17) with respect to \(t\), \(S\) and each \(v_i\) and apply the chain rule in the right-hand side whenever necessary. For instance:

\[
\frac{\partial f_{L_{i}}}{\partial S} = \frac{\partial f_{BS}}{\partial S} + \frac{\partial f_{BS}}{\partial \nu} \frac{\partial \nu}{\partial S} \Rightarrow \frac{\partial \nu}{\partial S} = \frac{\delta_{L_{i}} - \delta_{BS}}{\nu_{BS}} \tag{18}
\]

\[
\frac{\partial^2 \Theta}{\partial S^2} = \frac{\partial}{\partial S} \left( \frac{\delta_{L_{i}} - \delta_{BS}}{\nu_{BS}} \right) \tag{19}
\]

and so forth.

**Lemma 2**

Any local implied volatility \(\Theta(K, T; t, S, \mathbf{v})\) for an European option with strike \(K\) and maturity \(T\), must satisfy the following:

\[
\frac{\partial \nu}{\partial t} + \left( r - \sigma^2 \frac{d_2}{2 \sqrt{\tau}} \right) \frac{\partial \nu}{\partial S} + \frac{\sigma^2 S^2}{2} \left( \frac{\partial^2 \Theta}{\partial S^2} + 
\frac{d_1 d_2}{0} \left( \frac{\partial \Theta}{\partial S} \right)^2 \right) + \frac{1}{2 \tau} \left( \sigma^2 - \nu^2 \right) = 0
\]

where \(\tau = T - t > 0\) and \(d_1\) and \(d_2\) are such as in the Black-Scholes formula:

\[
d_1 = \frac{\ln M}{0 \sqrt{\tau}} \quad d_2 = d_1 - 0 \sqrt{\tau} \quad M = \frac{S e^{-\tau}}{K e^{-\tau}}.
\]

**Proof:** Subtract the Black-Scholes PDE from equation (11), apply (17) and Lemma 1, and use the following relationships between the Black-Scholes sensitivity:

\[
\nu_{BS} = \frac{1}{0 \tau S^2} \nu_{BS} \quad \kappa_{BS} = \frac{d_1 d_2}{0} \nu_{BS} \quad \text{and} \quad \Omega_{BS} = \frac{d_2}{0 \tau S^2 \sqrt{\tau}} \nu_{BS}
\]

Lemma 2 describes the dynamics of implied volatility that are consistent with any parametric local volatility model. Note that the differential equation in Lemma 2 has no partial derivative on the elements of \(\mathbf{v}\), so the implied volatility surface does move over time even when the local volatility surface is static, as assumed in standard local volatility models. However, whilst these models are not inconsistent with movement in implied volatilities over time, problems may arise because the permissible movements in implied volatility are too restricted. The main result of this section shows how the dynamics of the entire implied volatility surface will be governed by the same stochastic factors as those driving the local volatility and the option price:

---

6 But of course \(\Theta(K, T; t, S, \mathbf{v})\) is not independent of \(\mathbf{v}\) because it depends on the spot volatility \(\sigma(t, S; \mathbf{v})\).
Theorem 2

Under assumptions (9) and (10) the dynamics of the local implied volatility \( \theta = \theta(K, T; t, S, v) \) for a European option with strike \( K \) and maturity \( T \), are given by:

\[
\theta dt = \xi dt + \sigma S \frac{\partial \theta}{\partial S} dW_S + \sum_j \frac{\partial \theta}{\partial v_j} dZ_j
\]

(20)
in which the drift \( \xi \) is:

\[
\xi = \frac{1}{2} \frac{d_2}{\sigma \sqrt{T - t}} \left( \theta^2 - \sigma^2 \right) + \sigma \frac{d_2}{\sigma \sqrt{T - t}} \psi - \frac{1}{2} \frac{d_1 d_2}{\sigma^2} \eta^2
\]

satisfying \( \int_t^T \xi dt < \infty \) and \( \psi \) is related to the covariance between implied volatility and asset price movements:

\[
\psi dt = d\theta dW_S = \left( \sigma S \frac{\partial \theta}{\partial S} + \sum_j \beta_j \frac{\partial \theta}{\partial v_j} \right) dt
\]

and \( \eta^2 \) is the variance of the implied volatility process:

\[
\eta^2 dt = d\theta d\theta = \left( \psi^2 + \sum_i \sum_j \beta_i \beta_j \left( \rho_{i,j} - \rho_{i,S} \rho_{j,S} \right) \frac{\partial \theta}{\partial v_i} \frac{\partial \theta}{\partial v_j} \right) dt
\]

and all partial derivatives of \( \theta \) are as in Lemma 1.

Proof: This follows from an application of Ito’s lemma to the dynamics of \( \theta(K, T; t, S, v) \) with respect to \( t, S \) and \( v \), and using lemmas 1 and 2. See Appendix A.

Note that the option prices that are consistent with the implied volatility dynamics (20) must also satisfy the no arbitrage conditions mentioned in Appendix B.\(^7\) The following corollary simply re-writes (20) using only uncorrelated Brownian motions. But it is interesting because it shows that the SLV model is equivalent to the dynamic model for implied volatilities introduced by Schönbucher (1999):

Corollary 1

Assuming the vector \( dW = [dW_1, dW_2, \ldots dW_n] \) has positive definite correlation matrix \( \Sigma \), the dynamics of the local implied volatility from Theorem 2 can also be expressed in terms of uncorrelated Brownian motions \( W_j^* \) as:

\(^7\) Besides this, there is an interesting singularity on the drift \( \xi \) as \( t \to T \), when it can explode, as reported by Schönbucher (1999). However, from Theorem 2, this is not a problem as long as \( \int_t^T \xi dt < \infty \) for all \( T > t \). In effect, Dupire (2003) points out that the density of the implied volatility is equivalent to the density of the Brownian bridge from \( (S, t) \) to \( (K, T) \), which concords with our findings. More information about the asymptotic relationship between local volatility and Black-Scholes implied volatility can be found in Berestycki, Busca and Florent (2002).
\begin{align}
d\mathbf{d} & = \xi dt + \psi dW_S + \sum_j \omega_j dW_j^* \tag{21}
\end{align}

with \( \xi \) and \( \psi \) defined as in Theorem 2, \( dW_j^* dW_i^* = dW_j^* = 0 \) for all \( i \neq j \) almost surely, and:

\begin{equation}
\omega_j = \sum_{i=j}^6 \beta_i \sqrt{1-q_{i,j}^2} A_{i,j} \frac{\partial \theta}{\partial W_j} \tag{22}
\end{equation}

where \( A_{i,j} \) are the elements of the Cholesky decomposition \( A \) of the correlation matrix \( \Sigma \) with:

\begin{align}
\Sigma_{i,j} &= \frac{q_{i,j} - q_{i,S} q_{j,S}}{\sqrt{(1-q_{i,S}^2)(1-q_{j,S}^2)}} \quad \text{and} \quad \eta_j^2 = \psi^2 + \sum_j \omega_j^2.
\end{align}

**Proof:** This follows from a standard application of Cholesky decomposition and from:

\begin{equation}
dZ_i dZ_j = \left( q_{i,S} q_{j,S} + \sqrt{1-q_{i,S}^2} \sqrt{1-q_{j,S}^2} \Sigma_{i,j} \right) dt = q_{i,j} dt \Rightarrow \Sigma_{i,j} = \frac{q_{i,j} - q_{i,S} q_{j,S}}{\sqrt{(1-q_{i,S}^2)(1-q_{j,S}^2)}}.
\end{equation}

Apart from minor differences in notation, (21) is precisely the same as equation (2.7) from Schönbucher (1999) for the dynamics of a stochastic implied volatility with the drift term given by equation (3.7). The corollary is interesting because Schönbucher (1999) models the stochastic implied volatility for a given strike \( K \) and maturity \( T \), whilst we begin with a stochastic local volatility model for which the dynamics (21) hold for all strikes and maturities simultaneously.

In the ‘market model’ approach an implied volatility (or implied variance) diffusion is defined for each strike \( K \) and maturity \( T \). So if there are options for \( k \) strikes and \( m \) maturities in the market, the market model specifies \( mk \) diffusions, one for each traded option. In the SLV approach the smile surface is parameterized, with the number of parameters \( n \ll mk \). Hence the SLV model reduces the probability space from \( mk + 1 \) random variables to only \( n + 1 \), including the asset price \( S \).

**Corollary 2**

The correlation between the local implied volatility and the asset price changes is given by:

\begin{equation}
Q_{0,S}(K, T; t; S) = \frac{\psi}{\sqrt{\eta^2 + \sum_j \omega_j^2}} = \frac{\psi}{|\eta|} \tag{23}
\end{equation}

**Proof:** The correlation follows from (9) and (21):

\begin{equation}
Q_{0,S} = \frac{\text{Cov}(d\mathbf{d}, dS)}{\sqrt{\text{Var}(d\mathbf{d})\text{Var}(dS)}} = \frac{\psi \sigma S dt}{\sqrt{(\psi^2 + \sum_j \omega_j^2) dt (\sigma^2 S^2 dt)}} = \frac{\psi}{|\eta|}.
\end{equation}

---

8 We believe there was a typo in equation (3.3) from Schönbucher (1999) for the variance of implied volatility, where the term \( \gamma^2 \) appears to be missing. Many thanks to Hyungsok Ahn of Commerzbank, London for drawing our attention to this.
In Corollary 2, the absolute value $|\eta|$ stresses that the denominator of (23) is strictly positive. The $\omega$ are defined in (22) and are non-zero unless the parameters are non-stochastic. This result shows that the local implied volatility and price movements will have perfect correlation (of $\pm 1$, depending on the sign of the covariance $\psi$) if and only if $\omega = 0$ for all $j$, i.e. when local volatility surface is fixed. In other words, the instantaneous volatility is deterministic if and only if variations in implied volatility and the asset price are perfectly correlated.

IV. Hedging with Local and Stochastic Volatility

One well-known application of local volatility models is to hedge over-the-counter options consistently with the observed prices of standard European options through static replication (see e.g. Derman, Ergener and Kani (1995) and Carr, Elis and Gupta (1998)). Perfect hedging is possible if the assumptions of the local volatility model are valid, because the only source of uncertainty in the future claim price is the underlying asset. But if the assumptions need to be extended, as we have done above, perfect hedging becomes very complex since it involves new sources of randomness. Recognizing this problem, Kani, Derman and Kamal (1997) propose using 'volatility gadgets', i.e. small portfolios of traded options combined in such a way that it is possible to hedge any specific region from the local volatility surface. Then, by combining these gadgets, a multitude of hedging possibilities is available to the volatility trader.

In this section we do not focus on perfect hedging. Instead we compare the hedge ratios obtained from local volatility models, with and without the assumption of stochastic instantaneous forward volatilities, and the hedge ratios from stochastic volatility models. First we show that if the instantaneous volatility is stochastic and correlated with the asset price the hedge ratios derived from 'standard' local volatility models will be incorrect. In fact, the delta, gamma and theta of any European option require an adjustment to the local volatility hedge ratios delta: $\delta_{L} = \frac{\partial f_{L}}{\partial S}$, gamma: $\gamma_{L} = \frac{\partial^{2} f_{L}}{\partial S^{2}}$ and theta: $\Theta_{L} = \frac{\partial f_{L}}{\partial t}$ that are calculated at time $t_{0} < t$ using a calibrated local volatility surface, i.e. for a given $\nu(\mathcal{F}_{0})$.

**Theorem 3**

Under the SLV model (9) and (10) the first and second order sensitivities of the claim price $f_{L}(t, S; \nu | \mathcal{F}_{0})$ at time $t_{0}$ with respect to $S$, and the first order sensitivity to time $t$, are given by:

$$\delta_{SLV}(t, S; \nu) = \delta_{L}(t, S; \nu) + \sum_{i} \frac{\beta_{i} \varphi_{i}}{\sigma S} \frac{\partial f_{L}}{\partial v_{i}}$$

$$\gamma_{SLV}(t, S; \nu) = \gamma_{L}(t, S; \nu) + \sum_{i} \frac{\beta_{i} \varphi_{i}}{\sigma S} \left( 2 \frac{\partial^{2} f_{L}}{\partial S^{2}} - \frac{1}{S} \frac{\partial f_{L}}{\partial v_{i}} + \sum_{j} \frac{\beta_{j} \varphi_{j}}{\sigma S} \frac{\partial^{2} f_{L}}{\partial v_{i} \partial v_{j}} \right)$$
\[ \Theta_{SLV}(t,S;\nu) = \Theta_{L}(t,S;\nu) + \sum_i \frac{\partial f_L}{\partial \nu_i} \left( \alpha_i - (r - q) \frac{\beta_i Q_{i,S}}{\sigma} + \frac{1}{2} \sigma \partial^2 Q_{i,S} \right) \] 

(26)

**Proof:** This follows from the chain rule, the dynamics (10), and Ito’s lemma. See Appendix A.

Since the market model of stochastic implied volatilities is equivalent to the SLV model, they produce the same hedge ratios for standard European options. For instance, the ‘market model’ delta is related to the BS delta by the chain rule:

\[ \delta_{MM} = \delta_{BS} + \nu_{BS} \frac{d\theta}{dS} \]

and from (9) and (21) we have:

\[ \frac{d\theta}{dS} = \frac{\text{Cov}(d\theta, dS)}{\text{Var}(dS)} = \frac{\psi \sigma S d\theta}{\sigma^2 S^2 dt} = \frac{\partial \theta}{\partial S} + \sum_i \frac{\beta_i Q_{i,S}}{\sigma S} \frac{\partial \theta}{\partial \nu_i} \]

so that, using Lemma 1:

\[ \delta_{MM} = \delta_{BS} + \nu_{BS} \left( \frac{\delta_L - \delta_{BS}}{\nu_{BS}} + \sum_i \frac{\beta_i Q_{i,S}}{\sigma S} \frac{1}{\nu_{BS}} \frac{\partial f_L}{\partial \nu_i} \right) = \delta_L + \sum_i \frac{\beta_i Q_{i,S}}{\sigma S} \frac{\partial f_L}{\partial \nu_i} \]

which is identical to (24).

Theorem 3 shows that the traditional delta hedge is \( \delta_L \) imperfect. The second term on the right hand side of the above equations is an adjustment factor which depends on the correlation between movements of each \( \nu_i \) and the asset price \( S \). In effect we can split each hedge ratio into two parts: a sensitivity \( \delta_L \) derived from the standard view (i.e. calibrated to the smile at a fixed point in time) and an adjustment factor that depends on the dynamics of the stochastic parameters \( \nu(t) \).

In Appendix A we prove that the delta hedging error from using \( \delta_L \) rather than \( \delta_{SLV} \) is:

\[ \Lambda_{DLV} = \sum_i \int_0^T \beta_i \frac{\partial f_L}{\partial \nu_i} dZ_i \] 

(27)

It is the sum of Ito’s stochastic integrals, each of them with zero expected value but non-zero variance. So \( E\{\Lambda_{DLV}\} = 0 \) but

\[ \text{Var}\{\Lambda_{DLV}\} = \sum_i \sum_j \text{Cov}\left[ \int_0^T \beta_i \frac{\partial f_L}{\partial \nu_i} dZ_i, \int_0^T \beta_j \frac{\partial f_L}{\partial \nu_j} dZ_j \right] \]

and the only possibility for \( \text{Var}\{\Lambda_{DLV}\} = 0 \) (i.e. for perfect hedging) is that either \( \beta_i = 0 \) or \( \partial f_L/\partial \nu_i = 0 \) (or both) for all \( i \). This clearly requires a deterministic instantaneous volatility process. Hence, whilst the traditional delta hedge strategy is unbiased, it is not a perfect hedge when the instantaneous volatility is stochastic.
It is also interesting to calculate the hedging error from using the correct delta \( \delta_{SLV} \). Appendix A shows that this is:

\[
\Lambda_{SLV} = \sum_{i} \int_{0}^{T} \rho_{i,s} \sqrt{1 - \rho_{i,s}^2} \frac{\partial f_{L}}{\partial v_{i}} dW_{i}
\]

(28)

Again, this is the sum of Ito’s stochastic integrals and so \( E\{\Lambda_{SLV}\} = 0 \), but of course the delta hedge strategy is still not perfect. In fact, it follows that:

\[
Var\{\Lambda_{DLV}\} = Var\{\Lambda_{SLV}\} + Var\left\{\sum_{i} \int_{0}^{T} \frac{\partial f_{L}}{\partial v_{i}} \rho_{i,s} dW_{i}\right\}
\]

(29)

Therefore, although using the adjusted delta (24) does not resolve all uncertainty in the model, it should at least reduce the total variance of the hedging error, improving the overall hedging performance. All the above results assume that (9) and (10) are a good approximation of reality. But if, for instance, the price process were discontinuous the expressions for the hedging error above could not hold and the delta hedge strategy could be even biased, having non-zero expected value for the hedging error.

Now consider any Markov process such that at time \( t > t_0 \):

\[
\frac{dS(t)}{S(t)} = (\mu(t) - q(t)) dt + \sigma(t) dW(t)
\]

(30)

where \( \mu(t) \) and \( q(t) \) are deterministic functions of time and \( W \) defines a measure under which the discounted \( S(t) \) is a martingale. Here \( \sigma(t) \) is only required to be a continuous and predictable process with a known value at time \( t_0 \), denoted \( \sigma(t_0) \). Hence the instantaneous volatility may be either deterministic or stochastic. Appendix C proves that for any such model the implied volatility takes the form \( \theta = g(Y, T; t_0, \sigma(t_0)) \) i.e. it is a function of moneyness \( Y = S(t_0)/K \), the maturity of the option \( T \), and the initial volatility \( \sigma(t_0) \) at time \( t_0 \) only. This result is already known (for instance see Fouque et al. (2000)), but its important implications for the delta and gamma hedging of vanilla options have thus far been ignored.

Since \( \theta \) only depends on \( Y \), differentiating with respect to \( S(t_0) \) and \( K \) gives:

\[
\frac{\partial \theta(K, T; t_0, S(t_0), \sigma(t_0))}{\partial S(t_0)} = -\frac{K}{S(t_0)} \frac{\partial \theta(K, T; t_0, S(t_0), \sigma(t_0))}{\partial K}
\]

(31)

so that the change in the model’s implied volatility after a small variation in the underlying asset price is of opposite sign (and is proportional to) the slope of the implied volatility curve in the strike metric.
The delta of this type of model has an unfortunate property. From (31):

$$\delta = \frac{\partial f}{\partial S(t_0)} = \delta_{BS} + \nu_{BS} \frac{\partial \theta}{\partial S(t_0)} = \delta_{BS} - \frac{K}{S(t_0)} \nu_{BS} \frac{\partial \theta}{\partial K}$$

(32)

where the Black-Scholes delta and vega are calculated using the market implied volatility for each vanilla option. Since $\nu_{BS} > 0$ we have:

$$\delta > \delta_{BS} \text{ if } \frac{\partial \theta}{\partial K} < 0 \text{ and } \delta < \delta_{BS} \text{ if } \frac{\partial \theta}{\partial K} > 0$$

(33)

That is, the model delta will be greater than the Black-Scholes delta whenever the slope of the implied volatility curve is negative and vice-versa. For instance, in the case of equity indices where $\partial \theta / \partial K$ is typically negative (except perhaps at very high strikes) the model deltas will be greater than the corresponding Black-Scholes deltas. In short, any model of the form (30) will over hedge in the face of a skew.

Moreover, it is shown in Appendix C that these models have the following ‘additively separable’ property:

$$f(K, T; t_0, S(t_0), \sigma(t_0)) = S(t_0) \frac{\partial f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial S(t_0)} + K \frac{\partial f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial K}$$

(34)

So the model delta may be obtained directly from market data, as

$$\delta = \frac{\partial f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial S(t_0)} = \left( f(K, T; t_0, S(t_0), \sigma(t_0)) - K \frac{\partial f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial K} \right) / S(t_0)$$

(35)

Similarly, the model gamma is:

$$\gamma(K, T; t_0, S(t_0), \sigma(t_0)) = \frac{\partial^2 f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial S^2(t_0)} = \left( \frac{K}{S(t_0)} \right)^2 \frac{\partial^2 f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial K^2}$$

(36)

We have shown that the option price for any model of the form (30) obeys an ‘additively separable’ property in which we can decompose the option price in terms of its partial derivatives with respect to $S(t_0)$ and $K$. Likewise, the gamma of the option is proportional to the convexity of the option price with respect to the exercise price. Now, if prices for vanilla options are observable and sufficiently smooth (35) and (36) allow one to calculate the delta and gamma of vanilla options directly regardless of the actual specification of the volatility process! That is, all models of the form (30) should give almost identical deltas and gammas if they fit the market implied surface with the same degree of accuracy.
The additively separable property (34) is an important one to consider. It can easily be shown that it implies (31), and hence also (32) and (33). So any model that has this property will over hedge in the face of a skew. But (34) is a very common property. For instance any mixture model, where the option price is a weighted average of BS prices, also satisfies (34).

Hence it is clear that these results are very general. Under the umbrella of the model (30) we have not only the vast majority of univariate stochastic volatility models including the Hull and White (1987) and Heston (1993) model (see Appendix C) but also any model where the instantaneous volatility is a function of $S(t)/S(\theta)$ and not of $S(t)$ separately. Moreover any model with the additively separable property (34), and this includes normal mixture option pricing models, will have larger deltas than BS deltas in the presence of the skew.

An unfortunate misconception that all local volatility models have larger deltas than BS deltas in the presence of the skew has crept into the literature, for instance in Hagan et al (2002). But we have shown that this delta property only applies in special cases, such as when the instantaneous volatility is a function of $S(t)/S(\theta)$ only, or more generally when (34) holds. The problem is that the term ‘local volatility’ has been used for any model where the forward instantaneous volatility is a deterministic function of $S$ and $t$. However, the models that have this delta property are not ‘true’ local volatility models – they are ‘sticky delta’ models – yet the intention of Dupire (1993, 1994) and Derman and Kani (1994) was that a local volatility model should be a ‘sticky tree’ model, not a ‘sticky delta’ model. The ‘sticky tree’ can be used to price all options and does not move as the underlying moves. But sticky delta models, like stochastic volatility models, carry the implied surface with them as the underlying moves and the dynamics when it moves are very odd, as shown by Hagan et al (2002). The implied volatility, which can be written as a function of moneyness in these models, needs not change when $S$ changes but if it does change we would have $\partial \theta / \partial S > 0$ in equity markets, which is contrary to observed market behaviour.

For ‘true’ local volatility models, i.e. those where there is unique local volatility tree that can be use to price all options now and in the future, the forward instantaneous volatilities must be expressed more generally as: $\sigma(t,S(t);t_0,S(t_0))$ so that the spot volatility is clearly a function of the current level of the asset price, $S(\theta)$. Then, it can be shown that the option price does not obey (34). Instead

$$f_L = S(t_0) \frac{\partial f_L}{\partial S(t_0)} + K \frac{\partial f_L}{\partial K} - S(t_0) \frac{\partial \sigma(t_0)}{\partial \sigma(\theta)} \frac{\partial \sigma(\theta)}{\partial S(\theta)}$$

so there is an additional term in the right-hand side that is proportional to $\partial \sigma(\theta) / \partial S(\theta)$, which vanishes only if $\sigma(\theta)$ is independent of $S(\theta)$. Likewise,
So (33) does not follow, i.e. the model does not necessarily over hedge in the presence of a skew. In summary we have only shown that ‘sticky delta’ models, normal mixture models and stochastic volatility models have the wrong sign for $\frac{\partial \theta}{\partial S}$, and true local volatility models, being based on the ‘sticky tree’, can predict the correct sign for $\frac{\partial \theta}{\partial S}$.

V Empirical Analysis

In this section we consider how effective the delta and gamma adjustments in Theorem 3 are. Starting with clearly the wrong delta, i.e. from a ‘sticky delta’ model we use a short history of calibrated parameters to adjust the models delta and gamma using (24) and (25). The parameter adjustment process has the added benefit of being determined by the stability of the calibration. For hedging options, as opposed to pricing them, it is not just the closeness of the fit to the market data today that matters. Of course the fit is important but we also require a model that has relatively constant parameters if we plan to use the model for pricing as well as hedging. Besides, if the parameters are not constant then the basic assumption of a local volatility model is obviously violated and the model should not be used on these grounds alone.

Surprisingly few studies of hedging performance have been reported in the academic literature, even though this is of paramount importance for banks. Dumas, Fleming and Whaley (1998) test several different parametric and semi-parametric forms of local volatility function. They calibrate the parameters to SP500 index options prices on a particular date, repeating this on a weekly basis, and subsequently compare the hedging performance of the local volatility models with that of the Black and Scholes model. Their conclusion is that the Black-Scholes deltas appear to be more reliable than any of the deltas from the local volatility models that they tested. McIntyre (2001) reaches a similar conclusion, although Coleman et al. (2001) claim that over long hedging periods the local volatility deltas do improve somewhat.

However, in the light of our results these findings should be treated with caution. The local volatility deltas that are applied in these papers do not account for the impact of the dynamics of local volatility surfaces. In fact, the complete market model on which these tests are based assumes perfect delta hedging is possible because the local volatility surface is static. This may be the main reason for their disappointing conclusions.
V.1 The models

The pricing and hedging performance of the following models will be assessed:

(i) The Black-Scholes model (specified in Appendix C)
(ii) The Heston (1993) “SQRT” model:

\[
dS / S = (r - q) dt + \sqrt{V(t)} dB(t)
\]

\[
dV = \tilde{\psi}(w - V) dt + \xi \sqrt{V} dW(t)
\]

\[
\tilde{\psi} = (1 - \lambda) \varphi \xi + \sqrt{\varphi^2 - \lambda (1 - \lambda) \xi^2}
\]

where $\lambda$ is the volatility risk premium and $\varphi$ is the price-variance correlation.

(iii) The Brigo and Mercurio (2001) normal mixture ‘NM’ model with marginal price density:

\[
f(x; t) = \lambda \varphi (x; (r - \frac{1}{2} \sigma^2_1 + s) t; \sigma^2_1 t) + (1 - \lambda) \varphi (x; (r - \frac{1}{2} \sigma^2_2 + s_L) t; \sigma^2_2 t)
\]

where $x = \ln S$, $\varphi(x)$ is the normal density and $s_L = \frac{1}{\tau} \ln \left( \frac{1 - \lambda \exp(s \tau)}{1 - \lambda} \right)$ with $\tau = T - t$.

(iv) The constant elasticity of variance ‘CEV’ model:

\[
dS / S = (r - q) dt + \sigma(t, S) dB(t)
\]

where $\sigma(t, S) = \sigma_0 \left( \frac{S}{S_0} \right)^{(\beta - 2)/2}$ with $\sigma_0$ and $\beta$ constant.

(v) The CEV local volatility model

\[
dS / S = (r - q) dt + \sigma(t, S) dB(t)
\]

where $\sigma(t, S) = \xi S^{(\beta - 2)/2}$ with $\xi$ and $\beta$ constant.

(vi) SLV versions of models (iii) and (iv) with adjustments (24) and (25) that allow the forward instantaneous volatilities to be stochastic.

V.2 Data and the Model’s Calibrations

We have obtained data from Bloomberg on the June 2004 European options on the SP500 index: i.e. daily close prices from 02 Jan 2004 to 15 June 2004 (111 business days) for 34 different strikes (from 1005 to 1200). Only the strikes within ±10% of the current index level were used for the model’s calibration each day but all strikes were used for the hedging strategies described below.
Each model was calibrated daily by minimizing the root mean square error (RMSE) between the model implied volatilities and the market implied volatilities of the options used in the calibration set. Yet, for the BS model the deltas and gammas are obtained directly from the market data and there is no need for model calibrations. For the Heston (1993) model we used the closed form price based on Fourier transforms (see for instance Lewis, 2000), chose a volatility risk premium of zero and set the long-term volatility at 12%. The calculation of the CEV option price is based on the non-centered chi-square distribution result of Schroder (1989).

Exhibits 1 - 3 show the values of the calibrated parameters of each model. These are quite stable until after the 2nd June. At the end of the period the ‘vol of vol’ and mean-reversion parameters in the Heston model become very large indeed, as does the high volatility in the mixture model and the power parameter in the CEV model. The only parameter missing from these exhibits is the coefficient $\zeta$ in the CEV local volatility model (v); it was impossible to plot this as a time series because after about one month of trading it displayed huge fluctuations from day to day that became increasingly variable as the options matured. Exhibit 4 shows that the RMSE deteriorates in all models after 2nd June, and as a result we decided to use only the option prices from 2nd January to 2nd June in the hedging race.

The adjustments (24) and (25) to the NM and CEV models require a calibration method that uses a combination of cross-sectional market data and a time series of data on the calibrated parameters, so that we can estimate the correlation between the parameters and the underlying price. The method used to calibrate these SLV models is described in detail in Appendix D. We have used dynamic correlations based on the previous two weeks of calibrated parameters, as shown in exhibit 5, to adjust the deltas and gammas for these models to take into account the movements in the skew as the underlying moves. However the $\zeta$ parameter in the CEV local volatility model changes so much from day to day that the correlation with the underlying was virtually zero. Hence no SLV adjustment is made for this model. For the other two models it is interesting to notice a sharp increase in the correlations on 11th March 2004, when the SP500 index had just fallen about 4.5% in the space of a few days.

Since we need two weeks of data to obtain the first calibrations of the SLV models, the data period used in the hedging race is 16th January 2004 to 2nd June 2004.

V.3 Hedging Strategies, Deltas and Gammas

The delta hedge strategy consists of one delta-hedged short call in each option, rebalanced daily. That is, one call on each of the 34 strikes from 1005 to 1200 is sold on 16th January (or when the option is issued, if later than this) and hedged by buying an amount $\delta$ of the underlying asset, where $\delta$ is determined by both the
model and the strike of the option. The portfolio is rebalanced daily, stopping on 2nd June for the reasons stated above.

The delta-gamma hedge strategy again consists of a short call in each option, but this time an amount $\gamma$ is bought of the 1125 option, which is closest to ATM in general, over the period. Thus the gamma on each option has been set to zero and then we delta hedge the portfolio as above.

This option-by-option strategy on a large and complete database of liquid options allows one to assess the effectiveness of hedging by strike or moneyness of the option, and day by day as well as over the whole period. A data set of P&L with 1324 observations is obtained which allows an in-depth investigation of the hedging effectiveness of each model.

The deltas and gammas of each model, whilst changing daily, exhibit some strong patterns. When they are plotted, by strike or by moneyness, on any particular day the same shapes emerge day after day. In Exhibits 6 and 7 we compare the deltas and gammas from the different models on 21st May 2004. As expected, given our theoretical results in section III, the stochastic volatility model (Heston’s SQRT model), the mixture model (Brigo and Mercurio’s NM model) and the ‘sticky delta’ model (the CEV with parameterization (iv)) all have deltas that are greater than the BS deltas – they all obey the additively separable property (34), hence their deltas must be similar. Only the CEV local volatility model (CEV_LV) and the two SLV models have deltas that are ever less than the BS delta, and for low strike options only the CEV local volatility model gives deltas that are less than the BS deltas.

The gammas in Exhibit 7 also demonstrate strong patterns. In particular, the gammas for the CEV, NM and SQRT models follow the convexity property (36) and so they should be similar since these models have been calibrated to the same options. However, this can hardly be verified empirically in Exhibit 7, although these three gammas do tend to move together at least. Nevertheless, there is a good reason for that. For any two models that obey properties (34) and (36) the only way their deltas and gammas can differ is when the models cannot produce the same fitted smile curve. And this is exactly the case, as shown in Exhibit 8. Both NM and SQRT models fit the smile curve on 21 May 2004 with good accuracy, whilst the CEV model – severely restricted by the small number of parameters – could not fit the data properly (the market implied volatilities on that date are shown in black). However, even for the NM and SQRT models, there is still a substantial difference in the convexity of the local implied volatility $\frac{\partial^2 \theta}{\partial K^2}$, which explains why their gammas were so different despite the good fit. On the other hand, the NM_SLV, CEV_LV and CEV_SLV gammas are closer to each other. This suggests that not only the SLV correction is consistent across different models, but also it minimizes the impact from the fact that ‘sticky delta’ models imply the wrong dynamics for implied volatility.
V.4 Performance Criteria and Results

Table 1 reports for each model the sample statistics of the aggregate daily P&L over all options and over all days in the hedging period. In each part of the table the models are ordered by the standard deviation of the daily P&L, since to minimize this is a prime objective of hedging. Small skewness and excess kurtosis in the P&L distribution is also desirable — high values for these sample statistics indicate that the model was spectacularly wrong on a few days in the sample. Another important performance criterion is that the P&L be uncorrelated with the underlying asset. In our case over hedging would result in a significant positive correlation between the hedge portfolio and the SP500 index return. We have therefore performed a regression, based on all 1324 P&L data points, where the average P&L for the day is explained by a quadratic function of the SP500 returns. The $R^2$ from this regression should be zero: it is reported in the last column of the table.

In the delta hedge strategy, shown in the upper part of Table 1, both the standard deviation and the $R^2$ criterion agree on the model’s ranking. According to both these criteria the best hedging model is the CEV local volatility, closely followed by the two SLV models. The CEV also has P&L that is closest to being normally distributed according to the skewness and excess kurtosis. The BS model is ranked in the middle, with the NM and SQRT performing worse than the BS. Apart from this the positive mean P&L is a result of gamma effects, since we have only rebalanced daily.

The delta-gamma hedge strategy results show a mean P&L that is close to zero. It is remarkable that the BS model clearly performs best of all models according to all criteria. The other models ranked more or less as before, an exception being that the ranking by $R^2$ changes and there is a particularly low $R^2$ for the Heston’s SQRT model.

Table 2 reports the delta-gamma hedge portfolio P&L standard deviation by strike, averaged over all days in the sample period. It shows that the BS model only performs best for low strike options and that mid to high strike options are better hedged using the CEV or the two SLV models. Clearly the over hedging of the BS delta is compensated by an equivalent over hedging by the BS gamma, but only for low strike options. Bakshi, Cao and Chen (1997) also find that BS performs well for low strike options, but in contrast to our results here they find that even a single factor stochastic volatility model performs better than BS for delta hedging.

It is possible that the apparent superiority of the BS model for delta-gamma hedging low strike options is a result of the gamma hedging strategy chosen. If we were to gamma hedge with a set of options on the SP500
with a different maturity date then results may be different. Also, no hedging costs have been included in the analysis and these cost would be greater for the over hedging strategies like BS.

VI Summary and Conclusions

Two separate strands of literature, on stochastic volatility and on local volatility models, have been developed quite separately although Dupire (1996) and Kani, Derman and Kamal (1997) identified potential links between many years ago. Most research on stochastic volatility has specified a single factor diffusion or jump-diffusion for the instantaneous variance or volatility of the underlying asset. Most research on local volatility models has assumed a deterministic instantaneous volatility function for the underlying asset price diffusion, with no reference to the dynamic evolution of volatility. Both approaches are incomplete, the former capturing the dynamic properties of volatility but only in a one-dimensional space, the latter focusing on the multi-dimensional aspects of volatility but ignoring its time-evolution. However recent developments of multivariate diffusions for implied volatility have extended the stochastic volatility approach to be consistent with the cross-section of implied volatilities as well as their dynamics. To concord with this view, the deterministic local volatility model, which implies only a deterministic evolution for implied volatility, requires generalization.

Following Dupire (1996) and Kani, Derman and Kamal (1997) we regard the deterministic local volatility model as merely a special case of a more general stochastic local volatility model. That is, we define local volatility as the square root of the conditional expectation of a future instantaneous variance that depends on \( n + 1 \) stochastic risk factors, viz. the underlying price plus \( n \) parameters of the local volatility function. Hence we provide an explicit model of the stochastic evolution of a locally deterministic volatility surface over time. We have proved that this general stochastic local volatility model is equivalent to the market model for implied volatilities that was introduced by Schönbucher (1999).

Several results on the behavior of implied volatility, which have previously been proved only in the context of specific models, are here proved within this general framework. More importantly, we provide the correct derivation of local volatility hedge ratios. Deterministic local volatility models fail to capture the proper dynamics for implied volatilities and as a result the hedge ratios derived from these models are incorrect. Hence the standard critique of the hedging performance of local volatility models no longer applies. Indeed, from the equivalence of the implied volatility market model and the general stochastic local volatility model, we show that these models have identical hedge ratios.

Both theoretical and empirical analysis shows that the hedging performance of stochastic local volatility models is superior to that of single factor stochastic volatility models, and to a large class of models based on
deterministic instantaneous forward volatility functions. These include the ‘sticky delta’ models that have been already criticized by Hagan et al. (2002) for predicting incorrect dynamics for implied volatility, and any normal mixture model such as that of Brigo and Mercurio (2001). We have shown that single factor stochastic volatility models and normal mixture models also predict incorrect dynamics for implied volatility and this is the reason why their hedging performance is so poor. The CEV local volatility model, whilst hedging very well, cannot be used for pricing due to parameter instability. However the stochastic local volatility models are clearly useful for both pricing and hedging: variation of their parameters is endogenous to the model and, perversely, our empirical results show them to be much more stable than the ‘sticky tree’ model parameters, which are supposed to be fixed over time.

References


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


Appendix A: Proofs of Main Results

Proof of Equation (7)

Suppose a deterministic local volatility (DLV) model has been calibrated at time $t_0$ assuming:

$$dS = (r - q)Sdt + \sigma(t, S; v(t_0))SdW_t$$ for all $t > t_0$$

while the underlying asset process actually follows:

$$dS = (r - q)Sdt + \sigma(t, S; v(t))SdW_t$$ for all $t > t_0$$

with $v(t)$ stochastic. Now define the delta-hedged portfolio $\Pi = f_L - \delta_LS$ where $f_L = f_L(t, S; v(t))$ is the value of a standard European option and $\delta_L$ is the option delta consistent with the DLV model. Then, from a standard application of Ito’s lemma and the PDE (11), we have:

$$d\Pi = df'_{L} - \delta_L (dS + qSdt) = r\Pi dt + \frac{1}{2}\left(\sigma^2(t, S; v(t)) - \sigma^2(t, S; v(t_0))\right)S^2\gamma_L dt$$

where $\gamma_L$ is the option gamma consistent with the DLV model. Next, integrating over $t \in [t_0, T]$, the total hedging error (hence the total pricing error) is:

$$\Lambda = \int_{t_0}^{T} \gamma_L \left(\sigma^2(t, S; v(t)) - \sigma^2(t, S; v(t_0))\right)S^2 dt$$

which is stochastic since $S$ and $v(t)$ are stochastic. Thus, conditioning on $S$ and taking expectation we have:

$$E^0\{\Lambda\} = \int_{t_0}^{T} \gamma_L \left(\sigma^2(t, S; v(t)) - \sigma^2(t, S; v(t_0))\right)S^2 dSdt$$

which must be zero for any arbitrary $T > t_0$ if options are fairly priced by the model, thus:

$$\sigma^2(t, S; v(t)) = E^0\left\{ \sigma^2(t, S; v(t))\right\}$$ \hspace{1cm} (A-1)

Finally, since the expectation in (A-1) is precisely the general definition (6) for the local volatility, we conclude that the local volatility surface $\sigma^2(t, S; v(t))$ calibrated by a DLV model is correct if options prices are fitted properly, i.e. the expected pricing error $E^0\{\Lambda\}$ is zero. ■
Proof of Equation (8)

Define \( \bar{v}(t_0) = E^0\{v(t)|S\} \). Then, a standard Taylor’s series expansion of \( \sigma^2(t, S; v(t)) \) gives:

\[
\sigma^2(t, S; v(t)) = \sigma^2(t, S; \bar{v}(t_0)) + \sum_i \frac{\partial \sigma^2(t, S; \bar{v}(t_0))}{\partial v_i} (v_i(t) - \bar{v}_i(t_0)) + \frac{1}{2} \sum_{i,j} \frac{\partial^2 \sigma^2(t, S; \bar{v}(t_0))}{\partial v_i \partial v_j} (v_i(t) - \bar{v}_i(t_0))(v_j(t) - \bar{v}_j(t_0))
\]

and taking the expectation at time \( t_0 \) conditional on \( S \) we have:

\[
E^0\{\sigma^2(t, S; v(t))|S\} = \sigma^2(t, S; \bar{v}(t_0)) + \frac{1}{2} \sum_{i,j} \frac{\partial^2 \sigma^2(t, S; \bar{v}(t_0))}{\partial v_i \partial v_j} \text{Cov}^0\{v_i(t), v_j(t)|S\}
\]

where the first order term cancels out in the expectation. Finally, replacing (A-1) we conclude the proof. Note that the covariance above refers to the portion of \( v(t) \) that is uncorrelated with \( S \). Hence if \( v(t) \) is deterministic the second order term above also cancels out.

\[\blacksquare\]

Proof of Equations (27) and (28)

This proof of the hedging error is similar to the proof of equation (7) above, except that it uses the dynamics (12) for the claim price from Theorem 1. The dynamics of the delta-hedged portfolio \( \Pi = f_L - \delta_L S \) is:

\[
d\Pi = df_L - \delta_L (dS + qSdt) = \Pi dt + \left( \sigma S \left( \frac{\partial f_L}{\partial S} - \delta_L \right) + \sum_i \beta_i q_{i,S} \frac{\partial f_L}{\partial v_i} \right) dW_1 + \sum_i \beta_i \sqrt{1 - q_{i,S}^2} \frac{\partial f_L}{\partial v_i} dW_i
\]

so that, replacing \( \delta_L = \frac{\partial f_L}{\partial S} \) (i.e. the deterministic local volatility delta), the total hedging error is:

\[
\Lambda_{DL} = \sum_i \left[ \int_{t_0}^T \frac{\partial f_L}{\partial v_i} dW_1 + \beta_i \sqrt{1 - q_{i,S}^2} \frac{\partial f_L}{\partial v_i} dW_i \right] = \sum_i \left[ \int_{t_0}^T \beta_i \frac{\partial f_L}{\partial v_i} dZ_i \right] \quad \text{for some } T > t_0
\]

where we have used the definition for \( dZ_i \) from (10). Therefore, under the assumption of a deterministic instantaneous volatility, the hedging error is the sum of stochastic integrals related to all the uncertainty around the local volatility parameters.

Instead, if we had used the correct delta from (24) and followed the same argument as above, the total hedging error associated with the delta-hedged portfolio \( \Pi = f_L - \delta_{SL} S \) would be:

\[
\Lambda_{SL} = \sum_i \left[ \int_{t_0}^T \beta_i \sqrt{1 - q_{i,S}^2} \frac{\partial f_L}{\partial v_i} dW_i \right] \quad \text{for some } T > t_0
\]
which is stochastic if the correlation between S and at least one parameter \( v_i \) is less than perfect, i.e. \( \rho_{i,S} \neq \pm 1 \). That is, the delta hedge will not be perfect.

**Proof of Theorem 1**

From Ito’s lemma, the dynamic of the claim price \( f_i(t, S; v) \), defined as a function of \( t, S \) and a set of parameters \( v \), under the risk-neutral measure, is:

\[
df_i = \frac{\partial f_i}{\partial t} dt + \frac{\partial f_i}{\partial S} dS + \frac{1}{2} \frac{\partial^2 f_i}{\partial S^2} dS^2 + \sum_j \left( \frac{\partial f_i}{\partial v_j} \right) dv_j + \frac{1}{2} \sum_{j,k} \frac{\partial^2 f_i}{\partial v_j \partial v_k} dt_j dt_k
\]

Using (9) and (10):

\[
df_i = \Xi dt + \left( \alpha_i \frac{\partial f_i}{\partial S} + \sum_j \beta_{i,j} \frac{\partial f_i}{\partial v_j} \right) dW_S + \sum_j \beta_{i,j} \sqrt{1 - \rho_{i,S}^2} \frac{\partial f_i}{\partial v_j} dW_j
\]

with

\[
\Xi = \frac{\partial f_i}{\partial t} + (r-q) \frac{\partial f_i}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f_i}{\partial S^2} + \sum_j \left( \alpha_j \frac{\partial f_i}{\partial v_j} + \alpha \beta_{i,j} \frac{\partial f_i}{\partial S} \right)
\]

Then, using the PDE (11) and since under the risk-neutral probability the drift of \( f_i \) must be the risk-free rate, the following drift condition must hold:

\[
\sum_i \left( \alpha_i \frac{\partial f_i}{\partial v_i} + \alpha \beta_{i,i} \frac{\partial f_i}{\partial S} + \frac{1}{2} \sum_{j,k} \beta_{i,j} \beta_{i,k} \frac{\partial^2 f_i}{\partial v_i \partial v_j} \right) = 0.
\]

**Proof of Theorem 2:**

From Ito’s lemma and using (9) and (10), the dynamics of \( \theta(K, T; t, S, v) \) are given by:

\[
\theta = \left( \frac{\partial \theta}{\partial t} + (r-q) \frac{\partial \theta}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 \theta}{\partial S^2} + \sum_i \left( \alpha_i \frac{\partial \theta}{\partial v_i} + \alpha \beta_{i,i} \frac{\partial \theta}{\partial S} + \frac{1}{2} \sum_{j,k} \beta_{i,j} \beta_{i,k} \frac{\partial^2 \theta}{\partial v_i \partial v_j} \right) \right) dt + \alpha S \frac{\partial \theta}{\partial S} dW_S + \sum_i \frac{\partial \theta}{\partial v_i} dZ_i
\]

Using Lemma 1, the drift expands to:

\[
\frac{\partial \theta}{\partial t} + (r-q) \frac{\partial \theta}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 \theta}{\partial S^2} + \sum_i \left( \alpha_i \frac{\partial f_i}{\partial v_i} + \alpha \beta_{i,i} \frac{\partial f_i}{\partial S} + \frac{1}{2} \sum_{j,k} \beta_{i,j} \beta_{i,k} \frac{\partial^2 f_i}{\partial v_i \partial v_j} \right) - \sum_i \left( \Omega_{BS} \frac{\partial \theta}{\partial v_i} + \kappa_{BS} \frac{\partial \theta}{\partial S} \right) \alpha \beta_{i,i} \frac{\partial \theta}{\partial S} = 0
\]
Then, using Theorem 1 and lemmas 1 and 2, this re-arranges to:
\[
\xi = \frac{1}{2} \frac{d_2 - d_1}{\sqrt{\tau}} \psi - \frac{1}{2} \frac{d_1^2 + d_2^2}{\tau} \eta^2
\]
with \( \tau > 0 \), \( \theta > 0 \), and \( \psi \) and \( \eta^2 \) defined as:
\[
\psi = \alpha S \frac{\partial \theta}{\partial S} + \sum_i \beta_i Q_{i,S} \frac{\partial \theta}{\partial p_i}, \quad \text{and} \quad \eta^2 = \psi^2 + \sum_i \sum_j \beta_i \beta_j (Q_{i,j} - Q_{i,S} Q_{j,S}) \frac{\partial \theta}{\partial p_i} \frac{\partial \theta}{\partial p_j}
\]
Finally, if \( \theta \) is a proper Ito’s process, then \( \int_t^T \xi \, dt < \infty \), among other regularity conditions.

**Proof of Theorem 3:**
When movements in \( S(t) \) and \( v(t) \) are correlated, we can express each \( v_i(t) \) as a function of \( t, S \) and \( W_i \) so that from Ito’s lemma:
\[
dv_i = \left( \frac{\partial v_i}{\partial t} + (r - q) S \frac{\partial v_i}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 v_i}{\partial S^2} + \frac{\sigma^2 v_i}{W_i^2} \frac{\partial^2 v_i}{\partial W_i^2} \right) \, dt + \sigma S \frac{\partial v_i}{\partial S} \, dW_i + \frac{\partial v_i}{\partial W_i} \, dW_i
\]
and equating coefficients with (10):
\[
\frac{\partial v_i}{\partial S} = \frac{\beta_i Q_{i,S}}{\sigma S} \Rightarrow \frac{\partial^2 v_i}{\partial S^2} = -\frac{\beta_i Q_{i,S}}{\sigma S^2}
\]
\[
\frac{\partial v_i}{\partial W_i} = \beta_i \sqrt{1 - \frac{Q_{i,S}^2}{\sigma^2}} \Rightarrow \frac{\partial^2 v_i}{\partial W_i^2} = 0
\]
\[
\frac{\partial v_i}{\partial t} = \alpha_i - (r - q) \frac{\beta_i Q_{i,S}}{\sigma} + \frac{\sigma^2 \beta_i Q_{i,S}}{\sigma}
\]
Then, the chain rule gives the first order price sensitivity, delta, as:
\[
\delta_{S,L,V} = \frac{d}{dS} (f_L(t,S,v)) = \frac{\partial f_L}{\partial S} + \sum_i \frac{\partial f_L}{\partial p_i} \frac{\partial p_i}{\partial S} \Rightarrow \delta_{S,L,V} = \delta_L + \sum_i \frac{\beta_i Q_{i,S} \delta f_L}{\sigma S} \frac{\partial f_L}{\partial p_i}
\]
Note that using the claim price dynamics of Theorem 1 it can be readily verified that the same delta can be derived using a delta-hedged portfolio \( \Pi = f_L - \delta_{S,L,V} S \), with dynamics \( d\Pi = df_L - \delta_{S,L,V} (dS + qSdt) \). Similarly it is easy to show that (15-b) and (15-c) hold for gamma and theta using:
\[
\gamma_{S,L,V} = \frac{d}{dS} \left( \delta_{S,L,V}(t,S,v) \right) \quad \text{and} \quad \Theta_{S,L,V} = \frac{d}{dt} \left( f_L(t,S,v) \right)
\]
Appendix B: No Arbitrage Conditions

Throughout this paper we assume that both the asset price and the instantaneous volatility are continuous processes, adapted to the filtration $\mathcal{F}_t$ at time $t_0$. Besides, whatever the functional form for the instantaneous volatility $\sigma(t, S; v)$ in (9), a risk-neutral density $g_{L,t}(S)$ for the underlying that is consistent with (9) must satisfy the following conditions:

$$g_{L,t}(S) \geq 0 \text{ for all } S \geq 0 \quad \text{and} \quad \int_0^\infty g_{L,t}(S) dS = 1 \quad (B-1a)$$

$$f_L = e^{-r(t-t_0)} \int_0^\infty G(t, S) g_{L,t}(S) dS \quad \text{for every } t > t_0 \quad (B-1b)$$

where $G(t, S)$ is the value of any tradable asset at some time $t > t_0$. Conditions (B-1a) define $g_{L,t}(S)$ as a proper probability density function of $S(t)$, while (B-1b) defines a martingale measure. Although rather obvious, these conditions add an important constraint when pricing options. For instance, if $C(K,T; t, S)$ is the price of a vanilla European call at time $t$ with $K \geq 0$ and $T > t$, then Carr (2001) and Brunner and Hafner (2003) show that these conditions imply:

$$S \geq C(K, T; t, S) \geq \max\{S - Ke^{-(T-t)}; 0\} \quad (B-2a)$$

$$C(0, T; t, S) = S \quad \text{and} \quad \lim_{K \to 0} C(K, T; t, S) = 0 \quad (B-2b)$$

$$-1 \leq \frac{\partial C(K, T; t, S)}{\partial K} \leq 0 \quad \text{and} \quad \frac{\partial^2 C(K, T; t, S)}{\partial K^2} \geq 0 \quad (B-2c)$$

Whilst (B-2a) and (B-2b) are intuitive, (B-2c) tells an interesting story. It requires the option price to be a convex and monotonically decreasing function of $K$, otherwise there is an arbitrage opportunity. Brunner and Hafner (2003) also prove two more necessary but not sufficient conditions on the term structure of call prices. For $T_1, T_2 \in (t, T)$, $T_1 < T_2$, they require:

$$C(K, T_2; t, S) \geq C(Ke^{-(T_2-T_1)}, T_1; t, S) \quad (B-3a)$$

$$\int_0^\infty \max\{S - K; 0\} \left[ e^{-r(T_2-T_1)} g_{L,T_1}(S) - g_{L,T_1}(Ske^{-(T_2-T_1)}) \right] dS \geq 0 \quad (B-3b)$$

In effect, even when $g_{r,t}(S)$ is the density of a martingale measure satisfying (B-1a) and (B-1b), there can be an arbitrage opportunity between different maturities if either (B-3a) or (B-3b) is violated. Note that we have been careful to distinguish between a model risk-neutral density $g_{r,t}(S)$ (consistent with a certain local volatility model) and the market risk-neutral density $g(S)$ (consistent with observed market options prices). Clearly whilst we expect these two densities to share similar properties, they are unlikely to be the same, since parametric local volatility models can only approximate observed options prices in general. Nevertheless, a valid calibration of the local volatility surface must satisfy all conditions outlined above.
Appendix C: Stochastic Volatility & ‘Sticky Delta’ Local Volatility Models

We consider volatility models in which implied volatility can be expressed as a function of moneyness \( S/K \) alone, instead of \( K \) and \( S \) separately. This is the case for all stochastic volatility models with constant starting variance, such as those of Hull and White (1987) and Heston (1993). It is also the case for the normal mixture models introduced by Brigo and Mercurio (2001) and for any other ‘sticky delta’ volatility model. We then examine the implications of using the moneyness metric for the model delta and gamma.

In continuous time (i.e. without jumps in the underlying asset price process), we assume a standard Markov process at any time \( t > t_0 \) such as the geometric Brownian motion with instantaneous volatility \( \sigma(t) \) and continuous dividend yield \( q(t) \):

\[
dS(t) = (\mu(t) - q(t))S(t)dt + \sigma(t)S(t)dW(t)
\]  

(C-1)

where \( \mu(t) \) and \( q(t) \) are deterministic functions of time and \( \sigma(t) \) is only required to be a continuous and predictable process, with known value at time \( t_0, \sigma(t_0) \). That is, thus far we do not restrict the instantaneous volatility to be either deterministic or stochastic. Following Girsanov’s theorem, we know that if the discounted asset price is a martingale under the risk-neutral measure, then the option price is given by the discounted expectation of the payoff at maturity under the same measure:

\[
f(K,T;t_0,S(t_0),\sigma(t_0)) = B(t_0,T)E\left[ \max(w(S(T) - K),0) \big| S(t_0),\sigma(t_0) \right] \]  

(C-2)

where \( w \) is 1 or -1 for calls or puts respectively, and \( B(t, T) \) is the value of a riskless bond with deterministic interest rate \( r(t) \) for simplicity:

\[
B(t,T) = \exp \left( -\int_t^T r(u)du \right)
\]

We follow a procedure similar to Fouque et al (2000, section 2.6) and define a new variable \( X(t) = S(t)/S(t_0) \). Then, \( X(t_0) = 1 \) and the dynamics of \( X(t) \) are given by:

\[
dX(t) = \frac{dS(t)}{S(t_0)} = (\mu(t) - q(t))X(t)dt + \sigma(t)X(t)dW(t)
\]

so that the future values of \( X(t) \) for \( t > t_0 \) are independent of the initial asset price \( S(t_0) \). Thus, the inner part of the expectation in (C-2) can be re-arranged to:

\[
\text{This is true even when the risk-neutral measure is not unique, such as in stochastic volatility models. See e.g. Fouque et al (2000, section 2.5).}
\]
\[
\max(\text{w}((S(T) - K), 0) = \max(\text{w}(S(t_0)X(T) - K), 0) = K \max\left(\text{w}\left(\frac{S(t_0)}{K}X(T) - 1\right), 0\right)
\]

and the option price (C-2) becomes:

\[
f(K, T; t_0, S(t_0), \sigma(t_0)) = KB(t_0, T)E\left[\max\left(\text{w}\left(\frac{S(t_0)}{K}X(T) - 1\right), 0\right)\right]S(t_0), \sigma(t_0)\]

(C-3)

Next, since \(X(T)\) does not depend on \(S(t_0)\) or \(K\), the option price can be expressed more generally in terms of a certain function \(Q\) of \(S(t_0)/K\) directly, rather than \(S(t_0)\) and \(K\) separately:

\[
f(K, T; t_0, S(t_0), \sigma(t_0)) = KQ\left(\frac{S(t_0)}{K}, T; t_0, \sigma(t_0)\right)
\]

(C-4)

where the actual specification of \(Q\) depends on the assumption about the volatility process \(\sigma(t)\).

In the Black-Scholes framework where \(\sigma(t) \equiv \theta\) for all \(t \geq t_0\) it is easy to see that:

\[
f_{BS}(K, T; t_0, S(t_0), \theta) = KQ_{BS}\left(\frac{S(t_0)}{K}, T; t_0, \theta\right) = Kb\left(\frac{S(t_0)}{K}e^{-\theta(T-t_0)}\Phi(\text{wd}_1) - e^{-\theta(T-t_0)}\Phi(\text{wd}_2)\right)
\]

(C-5)

where \(w \in \{-1, 1\}\) as before and \(\Phi(\cdot)\) is the cumulative standard normal distribution and:

\[
d_1 = \frac{\ln\left(\frac{S(t_0)}{K}\right) + (r - q + \frac{1}{2} \theta^2)(T - t_0)}{\theta\sqrt{T - t_0}} \quad \text{and} \quad d_2 = d_1 - \theta\sqrt{T - t_0}
\]

where \(r, q\) and \(\theta\) are assumed constant.

We assume that the model (C-4) has been calibrated to a given market smile surface and equate the prices (C-4) and (C-5) to define the model’s implied volatility \(\theta\) by inverting the identity:

\[
f(K, T; t_0, S(t_0), \sigma(t_0)) = f_{BS}(K, T; t_0, S(t_0), \theta) \Leftrightarrow Q\left(\frac{S(t_0)}{K}, T; t_0, \sigma(t_0)\right) = Q_{BS}\left(\frac{S(t_0)}{K}, T; t_0, \theta\right)
\]

Thus the model’s implied volatility, i.e. the average volatility over the remaining life of the option that sets the model price equal to the BS price of a European put or call, can be parameterized as:

\[
\theta = g(Y, T; t_0, \sigma(t_0))
\]
where \( g(.) \) is a certain function of moneyness \( Y = S(t_0)/K \), the maturity of the option \( T \), and the initial volatility \( \sigma(h_0) \) at time \( h_0 \). This result is already known, for instance see Fouque et al (2000), but in Section IV we use it to derive important implications for the delta and gamma hedging of vanilla options.

From (C-4) we write the delta and gamma of the model price in the moneyness metric as:

\[
\delta = \frac{\partial f}{\partial S(t_0)}(K, T; t_0, S(t_0), \sigma(t_0)) = K \frac{\partial Q}{\partial Y} \left( Y, T; t_0, \sigma(t_0) \right) = \frac{1}{K} \frac{\partial Q}{\partial Y} \left( Y, T; t_0, \sigma(t_0) \right)
\]

\[
\gamma = \frac{\partial^2 f}{\partial S^2(t_0)}(K, T; t_0, S(t_0), \sigma(t_0)) = \frac{\partial}{\partial S(t_0)} \left( \frac{\partial Q}{\partial Y} \right) = \frac{1}{K} \frac{\partial^2 Q}{\partial Y^2} \left( Y, T; t_0, \sigma(t_0) \right)
\]

But also:

\[
\frac{\partial f}{\partial K} = Q + K \frac{\partial Q}{\partial Y} \left( -\frac{S(t_0)}{K^2} \right) = Q - Y \frac{\partial Q}{\partial Y}
\]

\[
\frac{\partial^2 f}{\partial K^2} = \frac{\partial}{\partial K} \left( Q - Y \frac{\partial Q}{\partial Y} \right) = \frac{\partial Q}{\partial Y} \left( -\frac{S(t_0)}{K^2} \right) \left( -\frac{S(t_0)}{K^2} \right) + Y \frac{\partial^2 Q}{\partial Y^2} \left( -\frac{S(t_0)}{K^2} \right) = \frac{Y^2}{K} \frac{\partial^2 Q}{\partial Y^2}
\]

Hence

\[
\frac{\partial f}{\partial K} = Q - Y \frac{\partial f}{\partial S(t_0)} \Leftrightarrow Q = Y \frac{\partial f}{\partial S(t_0)} + \frac{\partial f}{\partial K}
\]

so after multiplying by \( K \) we conclude that

\[
f(K, T; t_0, S(t_0), \sigma(t_0)) = S(t_0) \frac{\partial f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial S(t_0)} + K \frac{\partial f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial K} \quad (C-6)
\]

and that

\[
\gamma(K, T; t_0, S(t_0), \sigma(t_0)) = \frac{\partial^2 f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial S^2(t_0)} = \left( \frac{K}{S(t_0)} \right)^2 \frac{\partial^2 f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial K^2} \quad (C-7)
\]

Any two models of the form (1) should give identical deltas and gammas for all option. We now show that both (standard) stochastic volatility models and ‘sticky delta’ volatility models fall into this category. For the former, let \( \sigma(t) = h(U(t)) \) where \( h(.) \) is a well-behaved function and \( U(t) \) is a Markov process with risk-neutral dynamics given by:

\[
dU(t) = \alpha(t, U(t))dt + \beta(t, U(t))dZ(t) \quad (C-8)
\]
possibly correlated with the underlying asset price process (30) so that \(dW(t)dZ(t) = \rho dt\) almost surely. When \(h(u) = \sqrt{u}\), the definition above includes well-known models such as Hull and White (1987), with \(\rho = 0\), and Heston (1993), with \(\rho \neq 0\). But other specifications are also possible such as the process mentioned by Fouque et al (2000) when \(h(u) = \exp(u)\) and \(U(t)\) is a mean-reverting Ornstein-Uhlenbeck process. A key property of all these stochastic volatility models is that the option price can be written as an expectation of Black-Scholes prices:

\[
f_{SV}(K, T; t_0, S(t_0), \sigma(t_0)) = E[f_{BS}(K, T; t_0, S(t_0)\xi, \sqrt{\sigma^2})]
\]

where:

\[
\xi = \exp\left(\frac{T}{t_0} \sigma(s) dZ(s) - \frac{\rho^2}{2} \int_{t_0}^{T} \sigma^2(s) ds\right) \quad \text{and} \quad \sigma^2 = \frac{1}{T - t_0} \int_{t_0}^{T} (1 - \rho^2) \sigma^2(s) ds
\]

In particular, since \(\xi\) and \(\sigma^2\) are not functions of \(S(t_0)\) and \(K\), it is quite straightforward to obtain that:

\[
\frac{\partial f}{\partial S(t_0)} = E\left[\xi \frac{\partial f_{BS}}{\partial S(t_0)}(K, T; t_0, S(t_0)\xi, \sqrt{\sigma^2})\right] \quad (C-10)
\]

\[
\frac{\partial f}{\partial K} = E\left[\xi \frac{\partial f_{BS}}{\partial K}(K, T; t_0, S(t_0)\xi, \sqrt{\sigma^2})\right] \quad (C-11)
\]

Next, by multiplying (C-10) by \(S(t_0)\) and (C-11) by \(K\) and summing, we have in the right-hand side:

\[
E\left[S(t_0)\xi \frac{\partial f_{BS}}{\partial S(t_0)}(K, T; t_0, S(t_0)\xi, \sqrt{\sigma^2}) + K \frac{\partial f_{BS}}{\partial K}(K, T; t_0, S(t_0)\xi, \sqrt{\sigma^2})\right] = E[f_{BS}(K, T; t_0, S(t_0)\xi, \sqrt{\sigma^2})]
\]

Therefore:

\[
f_{SV}(K, T; t_0, S(t_0), \sigma(t_0)) = S(t_0) \frac{\partial f_{SV}}{\partial S(t_0)}(K, T; t_0, S(t_0), \sigma(t_0)) + K \frac{\partial f_{SV}}{\partial K}(K, T; t_0, S(t_0), \sigma(t_0))
\]

so that the additively separable property has been verified. Likewise, we have:

\[\text{See e.g. Fouque et al (2000, section 2.8).}\]
\[
\frac{\partial^2 f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial S^2(t_0)} = E \left[ \xi^2 \frac{\partial^2 f_{BS}(K, T; t_0, S(t_0), \sqrt{\sigma^2})}{\partial S^2(t_0)} \right] \tag{C-12}
\]

\[
\frac{\partial^2 f(K, T; t_0, S(t_0), \sigma(t_0))}{\partial K^2} = E \left[ \frac{\partial^2 f_{BS}(K, T; t_0, S(t_0), \sqrt{\sigma^2})}{\partial K^2} \right] \tag{C-13}
\]

so that using:

\[
\frac{\partial^2 f_{BS}}{\partial S^2(t_0)} = \left( \frac{K}{S(t_0)} \right)^2 \frac{\partial^2 f_{BS}}{\partial K}
\]

it follows that:

\[
\frac{\partial^2 f}{\partial S^2(t_0)} = \xi^2 E \left[ \left( \frac{K}{S(t_0), \xi} \right)^2 \frac{\partial^2 f_{BS}(K, T; t_0, S(t_0), \sqrt{\sigma^2})}{\partial K} \right] = \left( \frac{K}{S(t_0)} \right)^2 \frac{\partial^2 f}{\partial K}
\]

and hence the convexity property also holds.
Appendix D: Calibration of Stochastic Local Volatility

Here we can split the calibration problem into two parts: the calibration of the local volatility model on a snapshot of option prices for each day separately, and a discrete time series analysis of calibrated parameters. That is, we assume a parametric local volatility model has been calibrated at \( m \) points in time \( \{ t_1, \ldots, t_m \} \) prior to time \( t_0 \). Then the adjustment factors for the delta and gamma in Theorem 3 may be approximated as:

\[
\delta_{SLV}(t; S; \nu) \approx \delta_L(t; S; \nu) + \sum_i \frac{\text{Cov}(\Delta S_i, \Delta \nu_i)}{\text{Var}(\Delta S)} \frac{\partial f_i}{\partial \nu_i}
\]

(D-1)

\[
\gamma_{SLV}(t; S; \nu) \approx \gamma_L(t; S; \nu) + \sum_i \frac{\text{Cov}(\Delta S_i, \Delta \nu_i)}{\text{Var}(\Delta S)} \left( 2 \frac{\partial^2 f_i}{\partial \Delta S \partial \nu_i} - \frac{1}{S} \frac{\partial f_i}{\partial \nu_i} + \sum_j \frac{\text{Cov}(\Delta S_i, \Delta \nu_j)}{\text{Var}(\Delta S)} \frac{\partial^2 f_i}{\partial \nu_i \partial \nu_j} \right)
\]

(D-2)

when the \( \alpha \)'s, \( \beta \)'s and \( \rho \)'s in (10) are assumed constant.

These approximations are exact if the following no-arbitrage condition is satisfied:

\[
\sum_i \left( \frac{1}{m} \sum_{t=1}^m \frac{\partial f_i}{\partial \nu_i} - \frac{\text{Cov}(\Delta S_i, \Delta \nu_i)}{\sqrt{\text{Var}(\Delta S)}} \lambda \sqrt{\Delta t} + \frac{\text{Cov}(\Delta S_i, \Delta \nu_i)}{\text{Var}(\Delta S)} \frac{\partial^2 f_i}{\partial \nu_i \partial \nu_i} + \frac{1}{2} \sum_j \frac{\text{Cov}(\Delta \nu_i, \Delta \nu_j)}{\text{Var}(\Delta S)} \frac{\partial^2 f_i}{\partial \nu_i \partial \nu_j} \right) = 0
\]

where \( \lambda \) is the market price of risk.

To prove (D-1) and (D-2), note that in the physical measure the dynamics (9) can be written as:

\[
dS = (\mu - r)Sdt + \sigma(t, S; \nu)SdW_s^P
\]

(D-3)

with the associated Girsanov transformation:

\[
dW_s^P = dW_s - \lambda dt = dW_s - \frac{\mu - r}{\sigma(t, S; \nu)} dt
\]

(D-4)

in which \( \lambda = \frac{\mu - r}{\sigma(t, S; \nu)} \) is the market price of risk, \( \sigma(t, S; \nu) \) is assumed constant over the infinitesimal time-step \( dt \), and the superscript \( P \) indicates the physical measure. Now if we assume (D-3) and (D-4) also hold over a small time-step \( \Delta t \), the (observable) discrete price process under the physical measure can be described as:

\[
\Delta S = (\mu - q)S\Delta t + \sigma(t, S; \nu)S\Delta W_s^P
\]

(D-5)

\[
\Delta \nu_i = \alpha_i \Delta t + \beta_i \Delta Z_i^P
\]

(D-6)

\[
\Delta Z_i^P = \omega_i \Delta W_s^P + \sqrt{1 - \omega_i^2} \Delta W_j = \Delta Z_i - \omega_i \lambda \Delta t
\]

satisfying
\[
E[\Delta Z_i, \Delta Z_j] \xrightarrow{\text{a.s.}} Q_{i,j} \Delta t \quad \text{and} \quad E[\Delta W_i, \Delta W_j] \xrightarrow{\text{a.s.}} 0 \quad \text{for } i,j \in \{1, 2, \ldots, n\}
\]
so that
\[
\alpha^p_i = \alpha_i + \beta_i Q_{i,i} \lambda .
\]

Then, using (D-5) and (D-6), we assume, for \(1 \leq i \leq n\):
\[
\begin{align*}
Var(\Delta S) & \approx \sigma^2 S^2 \Delta t \\
Var(\Delta v_i) & \approx \sigma^2_i (\Delta t_i^2 + (1 - \Delta t_i^2) \Delta t) = \beta_i^2 \Delta t \\
Cov(\Delta S, \Delta v_i) & \approx \sigma S_i \Delta t \\
Cov(\Delta v_i, \Delta v_j) & \approx \beta_i \beta_j Q_{i,j} \Delta t
\end{align*}
\]
so it follows that:
\[
\begin{align*}
\beta_i & \approx \sqrt{\frac{\text{Var}(\Delta v_i)}{\Delta t}} \\
Q_{i,i} & \approx \sqrt{\frac{\text{Cov}(\Delta S, \Delta v_i)}{\text{Var}(\Delta S) \text{Var}(\Delta v_i)}} \\
Q_{i,j} & \approx \sqrt{\frac{\text{Cov}(\Delta v_i, \Delta v_j)}{\text{Var}(\Delta v_i) \text{Var}(\Delta v_j)}}
\end{align*}
\]

Now, consider the expected value of \(\Delta v_i\). From (D-5):
\[
E[\Delta v_i] = \alpha^p_i \Delta t \Rightarrow \alpha_i \approx \frac{1}{m \Delta t} \sum_{j=1}^{n} \Delta v_{i,j} - \beta_i Q_{i,i} \lambda
\]
Finally, replacing each approximation into Theorems 1 and 3, we derive the approximations (D-1) and (D-2) for the delta and gamma and for the no-arbitrage condition.

Hence we have a pragmatic method to adjust local volatility hedge ratios so that they account for the uncertainty about the future calibrated parameters. It is only an approximation but the main issue here is that the traditional local volatility and stochastic volatility models never take any account of the uncertainty on the calibrated parameters. To calibrate the SLV model at time \(t_0\) we will use the sample covariance matrix \(X^T X / m\), where \(X = [\Delta S, \Delta v_1, \Delta v_2, \ldots, \Delta v_n]_m\) is the \(m \times (n + 1)\) matrix of variations in each of the risk factors. That sample moments approximate population moments is a strong assumption. Also, while it is quite standard to assume the instantaneous variance is constant over a small time-step \(\Delta t\), we approximate this variance by the historical variance over a sample of daily observations. Clearly any application should be considered with care and justified only when its no-arbitrage condition is at least approximately satisfied.
Table 1: Aggregate performance over all dates and strikes

<table>
<thead>
<tr>
<th>Delta Hedge</th>
<th>Average</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>XS Kurtosis</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEV_LV</td>
<td>0.1462</td>
<td>0.5847</td>
<td>-0.3424</td>
<td>0.7820</td>
<td>0.1132</td>
</tr>
<tr>
<td>CEV_SLV</td>
<td>0.1393</td>
<td>0.6280</td>
<td>-0.5701</td>
<td>1.6041</td>
<td>0.2105</td>
</tr>
<tr>
<td>NM_SLV</td>
<td>0.1399</td>
<td>0.6329</td>
<td>-0.5224</td>
<td>1.2351</td>
<td>0.2229</td>
</tr>
<tr>
<td>BS</td>
<td>0.1401</td>
<td>0.7451</td>
<td>-0.7029</td>
<td>2.0370</td>
<td>0.4119</td>
</tr>
<tr>
<td>CEV</td>
<td>0.1367</td>
<td>1.1035</td>
<td>-0.6525</td>
<td>1.6691</td>
<td>0.6781</td>
</tr>
<tr>
<td>NM</td>
<td>0.1373</td>
<td>1.1788</td>
<td>-0.5928</td>
<td>1.4834</td>
<td>0.6995</td>
</tr>
<tr>
<td>SQRT</td>
<td>0.1375</td>
<td>1.1828</td>
<td>-0.5749</td>
<td>1.4906</td>
<td>0.6980</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Delta + Gamma Hedge</th>
<th>Average</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>XS Kurtosis</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS</td>
<td>-0.0014</td>
<td>0.2612</td>
<td>-0.4353</td>
<td>2.5297</td>
<td>0.0202</td>
</tr>
<tr>
<td>CEV_LV</td>
<td>0.0098</td>
<td>0.2691</td>
<td>-0.0291</td>
<td>3.0850</td>
<td>0.0511</td>
</tr>
<tr>
<td>NM_SLV</td>
<td>0.0182</td>
<td>0.2773</td>
<td>0.1673</td>
<td>3.2069</td>
<td>0.0477</td>
</tr>
<tr>
<td>CEV_SLV</td>
<td>0.0154</td>
<td>0.2832</td>
<td>0.1304</td>
<td>4.1414</td>
<td>0.0660</td>
</tr>
<tr>
<td>CEV</td>
<td>0.0382</td>
<td>0.4102</td>
<td>0.2501</td>
<td>4.0340</td>
<td>0.0733</td>
</tr>
<tr>
<td>SQRT</td>
<td>0.0329</td>
<td>0.4520</td>
<td>0.7635</td>
<td>5.7402</td>
<td>0.0201</td>
</tr>
<tr>
<td>NM</td>
<td>0.0428</td>
<td>0.4548</td>
<td>0.0208</td>
<td>4.0123</td>
<td>0.0986</td>
</tr>
</tbody>
</table>
Table 2: Standard deviation of P&L over sample period, by strike of option

<table>
<thead>
<tr>
<th>Strike</th>
<th>BS</th>
<th>CEV</th>
<th>NM_SLV</th>
<th>CEV_SLV</th>
<th>SQRT</th>
<th>NM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1025</td>
<td>0.3775</td>
<td>0.3693</td>
<td>0.3885</td>
<td>0.3900</td>
<td>0.5690</td>
<td>0.6900</td>
</tr>
<tr>
<td>1040</td>
<td>0.2538</td>
<td>0.5141</td>
<td>0.4110</td>
<td>0.5702</td>
<td>0.5378</td>
<td>0.5042</td>
</tr>
<tr>
<td>1045</td>
<td>0.2899</td>
<td>0.3156</td>
<td>0.3114</td>
<td>0.3213</td>
<td>0.3129</td>
<td>0.3186</td>
</tr>
<tr>
<td>1050</td>
<td>0.3207</td>
<td>0.3389</td>
<td>0.3528</td>
<td>0.3587</td>
<td>0.5047</td>
<td>0.6860</td>
</tr>
<tr>
<td>1060</td>
<td>0.3084</td>
<td>0.5438</td>
<td>0.5012</td>
<td>0.5856</td>
<td>0.5641</td>
<td>0.6030</td>
</tr>
<tr>
<td>1065</td>
<td>0.3920</td>
<td>0.5496</td>
<td>0.5469</td>
<td>0.6001</td>
<td>0.6454</td>
<td>0.7086</td>
</tr>
<tr>
<td>1070</td>
<td>0.3384</td>
<td>0.5912</td>
<td>0.5991</td>
<td>0.6317</td>
<td>0.6454</td>
<td>0.7078</td>
</tr>
<tr>
<td>1075</td>
<td>0.2466</td>
<td>0.2575</td>
<td>0.2930</td>
<td>0.2770</td>
<td>0.4137</td>
<td>0.6058</td>
</tr>
<tr>
<td>1080</td>
<td>0.3350</td>
<td>0.4933</td>
<td>0.5337</td>
<td>0.5348</td>
<td>0.5878</td>
<td>0.6671</td>
</tr>
<tr>
<td>1085</td>
<td>0.2428</td>
<td>0.2513</td>
<td>0.2814</td>
<td>0.2682</td>
<td>0.3820</td>
<td>0.5434</td>
</tr>
<tr>
<td>1090</td>
<td>0.4105</td>
<td>0.4917</td>
<td>0.5219</td>
<td>0.5102</td>
<td>0.5625</td>
<td>0.6372</td>
</tr>
<tr>
<td>1095</td>
<td>0.2962</td>
<td>0.3189</td>
<td>0.3427</td>
<td>0.3342</td>
<td>0.4059</td>
<td>0.4644</td>
</tr>
<tr>
<td>1100</td>
<td>0.2141</td>
<td>0.2087</td>
<td>0.2187</td>
<td>0.2117</td>
<td>0.2766</td>
<td>0.3723</td>
</tr>
<tr>
<td>1105</td>
<td>0.4103</td>
<td>0.3883</td>
<td>0.3863</td>
<td>0.3884</td>
<td>0.3968</td>
<td>0.3679</td>
</tr>
<tr>
<td>1110</td>
<td>0.2207</td>
<td>0.2355</td>
<td>0.2468</td>
<td>0.2414</td>
<td>0.2754</td>
<td>0.3212</td>
</tr>
<tr>
<td>1115</td>
<td>0.2053</td>
<td>0.2084</td>
<td>0.2129</td>
<td>0.2100</td>
<td>0.2347</td>
<td>0.2477</td>
</tr>
<tr>
<td>1120</td>
<td>0.2204</td>
<td>0.2192</td>
<td>0.2190</td>
<td>0.2195</td>
<td>0.2240</td>
<td>0.2281</td>
</tr>
<tr>
<td>1125</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>1130</td>
<td>0.1975</td>
<td>0.1885</td>
<td>0.1915</td>
<td>0.1888</td>
<td>0.2082</td>
<td>0.2121</td>
</tr>
<tr>
<td>1135</td>
<td>0.2216</td>
<td>0.1922</td>
<td>0.1930</td>
<td>0.1906</td>
<td>0.2242</td>
<td>0.2230</td>
</tr>
<tr>
<td>1140</td>
<td>0.2278</td>
<td>0.2042</td>
<td>0.2023</td>
<td>0.2029</td>
<td>0.2332</td>
<td>0.2476</td>
</tr>
<tr>
<td>1145</td>
<td>0.2069</td>
<td>0.1916</td>
<td>0.1902</td>
<td>0.1864</td>
<td>0.2619</td>
<td>0.2291</td>
</tr>
<tr>
<td>1150</td>
<td>0.2772</td>
<td>0.2632</td>
<td>0.2646</td>
<td>0.2689</td>
<td>0.3983</td>
<td>0.3592</td>
</tr>
<tr>
<td>1155</td>
<td>0.2024</td>
<td>0.1736</td>
<td>0.1554</td>
<td>0.1718</td>
<td>0.2374</td>
<td>0.1830</td>
</tr>
<tr>
<td>1160</td>
<td>0.1861</td>
<td>0.1741</td>
<td>0.1767</td>
<td>0.1799</td>
<td>0.3649</td>
<td>0.3286</td>
</tr>
<tr>
<td>1165</td>
<td>0.0950</td>
<td>0.0773</td>
<td>0.0818</td>
<td>0.0725</td>
<td>0.2103</td>
<td>0.1149</td>
</tr>
<tr>
<td>1170</td>
<td>0.1620</td>
<td>0.1621</td>
<td>0.1676</td>
<td>0.1772</td>
<td>0.4295</td>
<td>0.3191</td>
</tr>
<tr>
<td>1175</td>
<td>0.2961</td>
<td>0.2555</td>
<td>0.2399</td>
<td>0.2619</td>
<td>0.6701</td>
<td>0.4519</td>
</tr>
<tr>
<td>1180</td>
<td>0.1253</td>
<td>0.1034</td>
<td>0.0865</td>
<td>0.1284</td>
<td>0.0908</td>
<td>0.0849</td>
</tr>
<tr>
<td>1190</td>
<td>0.1132</td>
<td>0.0899</td>
<td>0.0815</td>
<td>0.1028</td>
<td>0.0906</td>
<td>0.0759</td>
</tr>
<tr>
<td>1200</td>
<td>0.3421</td>
<td>0.2957</td>
<td>0.2943</td>
<td>0.3158</td>
<td>0.8760</td>
<td>0.4936</td>
</tr>
</tbody>
</table>
Exhibit 1: Calibration of Heston (1993) model
Exhibit 2: Calibration of Brigo and Mercurio (2001) model

Exhibit 3: Calibration of CEV model
Exhibit 4: RMSE of model calibrations

Exhibit 5: Correlation of parameters in SLV models

(i) NM Model
(ii) CEV Model

Exhibit 6: Comparison of model deltas
Exhibit 7: Comparison of model gammas

Exhibit 8: Quality of fit