

Hedging under Stochastic Volatility

K. Ronnie Sircar *

November 24, 1998

Abstract

We present a family of hedging strategies for a European derivative security in a stochastic volatility environment. The strategies are robust to specification of the volatility process and do not need a parametric description of it or estimation of the volatility risk premium. They allow the hedger to control the probability of hedging success according to risk aversion. The formula exploits the separation between the time scale of asset price fluctuation (ticks) and the longer time scale over which volatility fluctuates, that is, the observed “persistence” of volatility. We run simulations that demonstrate the effectiveness of the strategies over the classical Black-Scholes strategy.

1 Introduction

In this article we present a family of hedging strategies for a European derivative security that super-replicate the claim with a controllable success probability, in a stochastic volatility environment. The strategy has the following features:

- It is an approximate (asymptotic) solution to the problem, but as such, it is computable (we give an explicit formula).
- It is based on a nonparametric description of the random volatility process of the underlying asset. Thus it is robust to specific modelling of the volatility (under technical restrictions).
- It requires estimates of certain simple statistics of the volatility process that are easily obtained from historical asset price data.
- It does not need identification (or estimation) of the volatility risk premium or the market’s (equivalent martingale) pricing measure.

*Department of Mathematics, University of Michigan, Ann Arbor MI 48109-1109, *sircar@umich.edu*. The author is grateful for conversations on the subject with George Papanicolaou, to Hans Föllmer for suggesting the particular hedging problem considered here, and participants of the Courant Finance Seminar for helpful questions and comments.

The strategy is selected as follows: choose a minimum acceptable probability p_0 with which the strategy should dominate the perfect hedging strategy. Then, find the number of standard deviations ρ of a standard normal distribution whose confidence interval has this probability

$$\frac{1}{\sqrt{2\pi}} \int_{-\rho}^{\rho} e^{-z^2/2} dz = p_0. \quad (1)$$

This is easily found in tables (for example $\rho = 1$ corresponds to $p_0 = 67\%$, $\rho = 2$ to 95%). Then the hedging strategy is given by formulas (15) and (20). Note that p_0 is not the probability that the hedge is successful, but increasing p_0 increases that probability.

That the randomness of the volatility process (whose distribution is unspecified) translates into a family of hedging strategies that are distinguished simply by a *normally distributed* random variable comes from an application of the central limit theorem for Markov processes to the Black-Scholes derivative pricing PDE with a random volatility coefficient. That such a convenient characterization is a good approximation is due to the persistent nature (or burstiness) of volatility (in at least equity and F/X markets).

The next section briefly sketches the background and motivation for stochastic volatility models. In Section 3, we explain how volatility persistence is modelled here and how uncertainty in volatility translates into uncertainty in derivative prices and hedging strategies. Section 4 presents the main result which is illustrated by the simulations of Section 5. The issue of estimation, crucial for the theory to be applicable, is explained in Section 6, followed by conclusions.

2 Why Stochastic Volatility?

Several excellent survey articles, for example [7], outline the main features of stochastic volatility modelling for derivative pricing, starting from the work of Hull and White [8] in 1987. The stock price (or exchange rate) process $\{X_t, t \geq 0\}$ satisfies

$$dX_t = \mu X_t dt + \sigma_t X_t dW_t,$$

where $\{\sigma_t, t \geq 0\}$ is the stochastic volatility process. An overview of the usual approach as it relates to the work here is given in [13], and the important points are:

- Empirical studies of stock price data strongly suggest volatility is *not* constant (as assumed by the Black-Scholes theory), but has a random component. ARCH/GARCH models, whose continuous-time diffusion limits are stochastic volatility models, provide much better descriptions of the data. See [2] for details.
- Empirical studies of implied volatility data, for example [12], report frequent observation of the smile curve, a U-shaped variation of implied volatility with strike price for options with the same time-to-maturity. The minimum is at or near the current stock price.
- Any stochastic volatility model in which the volatility process is independent of the Brownian motion W_t results in predicted European option prices whose implied volatility curve smiles, with minimum at today's stock price adjusted by compound interest earned from today to expiration. See [11] or [13] for a proof.

- Most analysis and estimation is parametric: σ_t is modelled as the solution to a particular Itô SDE.
- A stochastic volatility environment is a simple example of an incomplete market. As such there is no unique pricing (or equivalent martingale) measure and this indeterminacy can be characterized by an unknown process, labelled the volatility risk premium, playing a part in derivative pricing and hedging. Most usually this process is taken to be zero, sometimes a constant or a deterministic function of present volatility, the main reasons being feasibility of estimation or to preserve the Markovian structure. Here we shall not need to make such a choice, as explained in Section 3.

We mention briefly some recent work related to the hedging problem considered here. The problem of almost sure super-replication is considered by Cvitanic *et al.* [3] (and authors referenced therein) by stochastic control methods. Such strategies that guarantee a successful hedge are usually expensive, which motivates Föllmer and Leukert [4] (and others cited there) to allow some risk of shortfall and look for a strategy that maximizes the probability of a successful hedge given some initial cash input that the hedger is willing to spend. This is explicitly computed in the constant volatility framework, and when volatility can jump by a random amount at a known time. Avellaneda *et al.* [1] construct worst-case super-replicating strategies for complex portfolios of options given that volatility lies in a known band.

3 Separation of Time-Scales

Figure 1 shows two simulated realizations of possible volatility paths over the course of a year. In the first, volatility is low (4-8%) for a large part of that year (roughly $t = 0.1$ till $t = 0.8$ - over 8 months) and then for the rest of the year it is at a higher level. In the second path, volatility fluctuates between periods of high and periods of low far more often - it seems to be low for a few weeks and then high for a few weeks, then low again, and so on. That is, often, when it is low, it stays low for a period, and similarly when it is high.

Empirical studies suggest that the latter realization is a much more typical yearly volatility pattern than the former - it exhibits volatility clustering, or the tendency of high volatility to come in bursts.

In fact, the sample paths of the second process can be obtained by simulating the first process for a much longer time (50 years) and squeezing the realization into one year. That is, the second process comes from speeding-up the first. Mathematically, if we call the first process $\{\sigma(t), t \geq 0\}$, it is convenient to denote the second process $\{\sigma\left(\frac{t}{\varepsilon}\right), t \geq 0\}$, where $\varepsilon > 0$ is a dimensionless parameter that represents the speeding-up. Because volatility clustering is really a distinguishable feature, we shall think of ε as being small.

Examples

1. Suppose $\sigma(t)$ is a two-state Markov chain representing a crude model of volatility taking a high or a low state: $\sigma(t) \in \{\sigma_1, \sigma_2\}$. Then if the generator matrix Q of the process

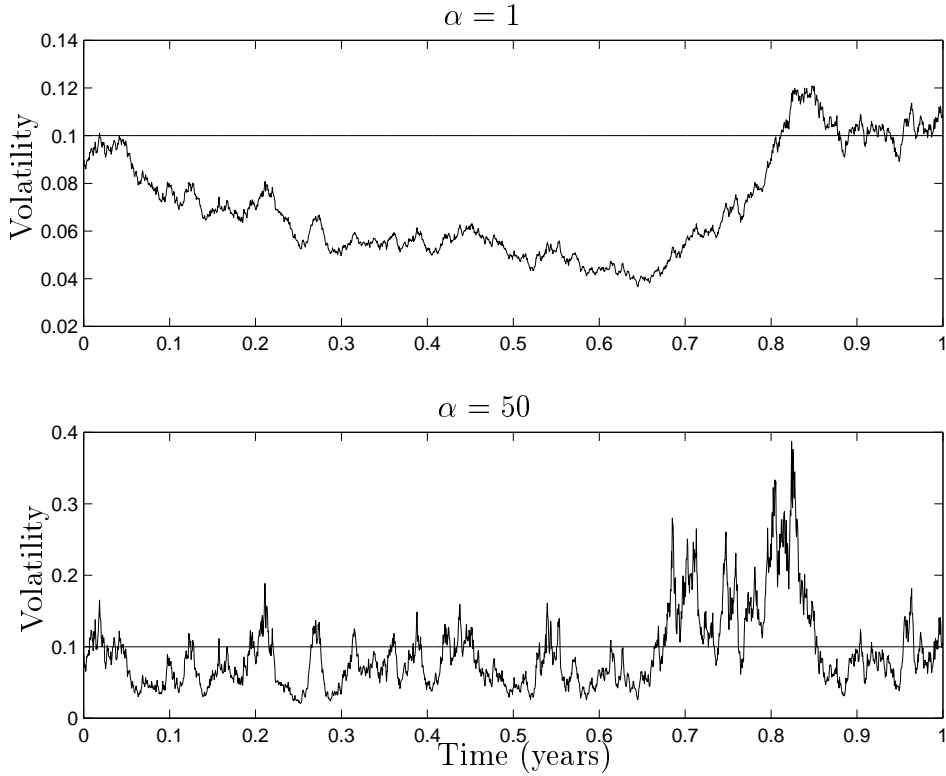


Figure 1: The top figure shows a simulated path of $\sigma_t = e^{Y_t}$, Y_t a mean-reverting Ornstein-Uhlenbeck process defined by equation (2), with $\alpha = 1$, and the bottom one shows a path with $\alpha = 50$. Note how volatility "clusters" in the latter case.

has elements that are $\mathcal{O}(1)$ in size (neither big nor small), for example

$$Q = \begin{pmatrix} -2 & 2 \\ 8 & -8 \end{pmatrix},$$

then a typical six-month sample path might look like the first graph of Figure 2. However, if we consider such a process with a generator having large entries, for example

$$Q' = \begin{pmatrix} -200 & 200 \\ 800 & -800 \end{pmatrix},$$

then a typical path, shown in the bottom graph, gives a much better description of high volatility coming in bursts of a few days or weeks, rather than months.

But notice that $Q' = \frac{1}{0.01}Q$, and that if $\sigma'(t) := \sigma\left(\frac{t}{\varepsilon}\right)$, with $\varepsilon = 0.01$, then the generator of $\sigma'(t)$ is exactly this $Q' = \frac{1}{\varepsilon}Q$. Thus the speeding-up notation $\sigma\left(\frac{t}{\varepsilon}\right)$ is a concise description of the fact that such a parametric model of volatility (and indeed *any* Markovian model) contains some parameters that are small, if the model is to reflect clustering periods of lengths that are usually observed. In this representation, the "smallness" is controlled by the ε .

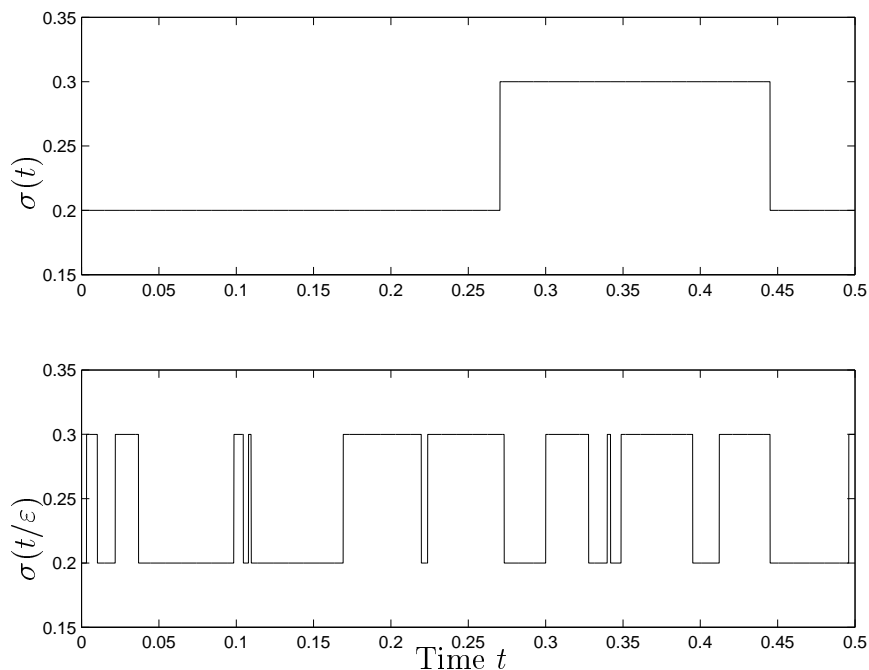


Figure 2: The top figure shows a simulated path of $\sigma(t)$ a two-state Markov chain, and the bottom one shows a path of $\sigma\left(\frac{t}{\varepsilon}\right)$, with $\varepsilon = 0.05$.

- Suppose $\sigma(t) = f(Y_t)$, where $f(\cdot)$ is a positive increasing function and Y_t is a mean-reverting Ornstein-Uhlenbeck process: it satisfies

$$dY_t = \alpha(m - Y_t)dt + \beta dZ_t. \quad (2)$$

The top graph of Figure 1 shows $\sigma(t)$ for $f(y) = e^y$ (the expOU model), with parameters $\alpha = 1$, $\beta^2 = 0.5$ and m chosen so that the RMS volatility (or long-run mean level) is 10%. With these $\mathcal{O}(1)$ values, we do not see significant volatility persistence.

Now consider the speeded-up process $\sigma\left(\frac{t}{\varepsilon}\right) = f(Y_{t/\varepsilon})$. The generator of the original Y_t is

$$\mathcal{L} := \alpha(m - y)\frac{\partial}{\partial y} + \frac{1}{2}\beta^2\frac{\partial^2}{\partial^2 y},$$

while the generator of $Y_{t/\varepsilon}$ is

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon}\mathcal{L} = \frac{\alpha}{\varepsilon}(m - y)\frac{\partial}{\partial y} + \frac{1}{2}\frac{\beta^2}{\varepsilon}\frac{\partial^2}{\partial^2 y}.$$

From this, we see that speeding-up Y_t (and hence $\sigma(t)$) is analogous to replacing α by α/ε and β by $\beta/\sqrt{\varepsilon}$. That is, the rate of mean-reversion α is scaled by $1/\varepsilon$ with the noise factor β scaled correspondingly to keep $\beta^2/2\alpha$ constant. For example, if $\varepsilon = 1/50$, we can simulate $\sigma\left(\frac{t}{\varepsilon}\right)$ by using $\alpha = 50$ and $\beta^2 = 25$ in (2). A realization of this process is shown in the bottom graph of Figure 1, and again, it better captures volatility clustering.

In this context, it is then convenient to think of speeding-up as simply the presence of a *large* mean-reversion rate: there are always Itô fluctuations in the volatility from the Brownian motion Z_t and it reverts slowly to its mean-level when looked at relative to this time-scale. But it reverts fast to the mean when looked at over the time-scale of a year. In the path shown, it crosses the mean-level over thirty times during the year.

This class of models is analyzed and estimated from market data in [5, 6].

Characterization in terms of time-scales

In summary, there are three distinct time-scales in the modelling of the underlying asset price (or exchange rate) $\{X_t^\varepsilon, t \geq 0\}$, which satisfies

$$dX_t^\varepsilon = \mu X_t^\varepsilon dt + \sigma \left(\frac{t}{\varepsilon} \right) X_t^\varepsilon dW_t. \quad (3)$$

Firstly, there is the “infinitely small” scale of the “infinitely fast” fluctuations of the Brownian motion W_t . These model the tick-by-tick fluctuations of the price. The volatility process might also have a component fluctuating on this scale (for example the Z_t in the second example above). However there is a longer time-scale representing volatility persistence which might be on the order of a few days or weeks (for example, the average mean-reversion time in the second example). These variations are slow in comparison to the tick-tick scale, but still fast when looked at over the lifetime of a derivative contract (many months) which is the long time-scale of our pricing or hedging problems.

Brownian	Volatility	Option
fluctuations	fluctuations	$[0, T]$
\ll	\ll	
\sim minutes	\sim days	\sim months.

Such a separation of scales is utilized by the asymptotic analysis of the next section.

4 Hedging

Suppose at time $t = 0$, we wish to hedge the risk of having written a European call option with strike price K and expiration date T on a stock by buying and selling only the underlying stock. The problem is to find a hedging strategy $H(t, x)$ that gives the number of units of the underlying held at time t (when its price is x) so that at time T , when we might have to sell the stock to the option holder for price K , we break even or make a profit with high probability.

In the complete market constant volatility case, a perfect (probability 1) break even hedging scheme is to hold $C_x^{BS}(t, x)$ units of stock (the delta), where C^{BS} denotes the Black-Scholes formula. In the stochastic volatility case, the additional source of randomness cannot be exactly hedged by trading in just the stock, and so we look for strategies with a high probability of success.

Illustration of Method

To explain how we find such a strategy, consider first the simplified scenario in which the stock price has a deterministic volatility $\sigma_1(t)$

$$dX_t = \mu X_t dt + \sigma_1(t) X_t dW_t,$$

but we hedge with the wrong deterministic volatility function $\sigma_2(t)$. That is, we solve the (generalized) Black-Scholes PDE

$$\begin{aligned} \overline{C}_t + \frac{1}{2} \sigma_2(t)^2 x^2 \overline{C}_{xx} &= 0, \\ \overline{C}(T, x) &= (x - K)^+, \end{aligned} \quad (4)$$

whose solution we denote $\overline{C}(t, x; [\sigma_2])$, and hold $\overline{C}_x(t, X_t; [\sigma_2])$ of the stock at time t . We take the interest rate $r = 0$ here, but our final formulas are given for the general case. At time $t = 0$, we put up the amount $\overline{C}(0, X_0; [\sigma_2])$, the cost of the hedge. The value of the hedging portfolio V_t is

$$V_t = \overline{C}(0, X_0; [\sigma_2]) + \int_0^t \overline{C}_x(s, X_s; [\sigma_2]) dX_s.$$

When the wrong volatility $\sigma_2(\cdot)$ is used, the strategy is not self-financing and V_t may be negative at certain times. We assume we are willing to put up extra cash after $t = 0$ if the strategy demands it, but our profit/loss bookkeeping is with respect to the initial cost of the hedge. The question of interest is how close is V_T to $(X_T - K)^+$, the payoff of the written claim?

Following [9], we know that

$$\begin{aligned} \overline{C}(t, X_t; [\sigma_2]) &= \overline{C}(0, X_0; [\sigma_2]) + \int_0^t \overline{C}_x(s, X_s; [\sigma_2]) dX_s \\ &\quad + \int_0^t \left(\overline{C}_t(s, X_s; [\sigma_2]) + \frac{1}{2} \sigma_1(s)^2 X_s^2 \overline{C}_{xx}(s, X_s; [\sigma_2]) \right) ds, \end{aligned}$$

by Itô's lemma, and the last integral can be re-written as

$$\frac{1}{2} \int_0^t \left(\sigma_1(s)^2 - \sigma_2(s)^2 \right) X_s^2 \overline{C}_{xx}(s, X_s; [\sigma_2]) ds, \quad (5)$$

using (4). This gives

$$V_T = (X_T - K)^+ + \frac{1}{2} \int_0^T \left(\sigma_2(s)^2 - \sigma_1(s)^2 \right) X_s^2 \overline{C}_{xx}(s, X_s; [\sigma_2]) ds.$$

Clearly, taking $\sigma_2 \equiv \sigma_1$ hedges the claim perfectly. When $\sigma_1(t)$ is a random process, we want to find a $\sigma_2(\cdot)$ so that $V_T \geq (X_T - K)^+$ with a high probability. This reduces to finding $\sigma_2(\cdot)$ such that

$$\frac{1}{2} \int_0^T \sigma_2(s)^2 X_s^2 \overline{C}_{xx}(s, X_s; [\sigma_2]) ds \geq \frac{1}{2} \int_0^T \sigma_1(s)^2 X_s^2 \overline{C}_{xx}(s, X_s; [\sigma_2]) ds \quad (6)$$

in some “best” way. This is a stochastic control type problem to find a volatility path that maximizes a probability and this is the approach of [3, 4]. To our knowledge, the exact solution is not easily computable, so we will look for a dominating solution: find $\sigma_2(\cdot)$ such that (6) holds with high probability.

The remaining ingredient is to characterize possible hedging strategies in a convenient way that translates uncertainty in the volatility into uncertainty in the final value V_T . This is achieved by an asymptotic approximation that exploits the separation of time-scales.

Asymptotic Approximation

We define the stochastic option price $C^\varepsilon(t, x)$ as the solution to the Black-Scholes PDE with the random speeded-up volatility coefficient $\sigma\left(\frac{t}{\varepsilon}\right)$ from (3):

$$\begin{aligned} C_t^\varepsilon + \frac{1}{2}\sigma^2\left(\frac{t}{\varepsilon}\right)x^2C_{xx}^\varepsilon &= 0, \\ C^\varepsilon(T, x) &= (x - K)^+. \end{aligned} \tag{7}$$

This can be thought of as a conditional Black-Scholes PDE with each realization of the process $\{C^\varepsilon(t, x), 0 \leq t \leq T\}$ a call option pricing function given the path of the volatility. As $\sigma\left(\frac{t}{\varepsilon}\right)$ and W_t are independent, the derivation of this equation exactly follows the derivation of the classical Black-Scholes PDE.

We are interested in the asymptotic behaviour of C^ε as $\varepsilon \downarrow 0$: this approximation will tell us how to deal with the risk from the randomness of the volatility. We shall make the following assumptions on the process $\{\sigma(t), t \geq 0\}$ (which also hold after speeding-up):

1. $\{\sigma^2(t), t \geq 0\}$ is wide-sense stationary: it has time-independent mean $\overline{\sigma^2} := E\{\sigma^2(t)\}$ and its autocorrelation $E\{(\sigma^2(t) - \overline{\sigma^2})(\sigma^2(s) - \overline{\sigma^2})\}$ is a function of $|t - s|$ only.
2. It is ergodic and Markov.

As ε becomes smaller and smaller, the distinction between the time-scales disappears and $C^\varepsilon(t, x)$ looks more and more like the Black-Scholes formula with a constant *averaged* volatility. This is a standard averaging principle result following from the ergodic theorem. It says that $C^\varepsilon(t, x)$ converges in probability to $C^{BS}(t, x; \sqrt{\overline{\sigma^2}})$, the Black-Scholes formula with volatility $\sqrt{\overline{\sigma^2}}$, which satisfies the (averaged) PDE

$$\begin{aligned} C_t^{BS} + \frac{1}{2}\overline{\sigma^2}x^2C_{xx}^{BS} &= 0, \\ C^{BS}(T, x) &= (x - K)^+. \end{aligned} \tag{8}$$

That is,

$$P\left(\sup_{0 \leq t \leq T} \sup_{x > 0} \left|C^\varepsilon(t, x) - C^{BS}(t, x; \sqrt{\overline{\sigma^2}})\right| > \delta\right) \longrightarrow 0$$

as $\varepsilon \downarrow 0$, for any $\delta > 0$. A proof is given in [13].

So far we have a crude approximation to possible prices under stochastic volatility by the Black-Scholes formula. What is of use is the next correction term, valid for small $\varepsilon > 0$, that quantifies volatility risk.

Let us write

$$C^\varepsilon(t, x) = C^{BS}(t, x; \sqrt{\sigma^2}) + \sqrt{\varepsilon}Z^\varepsilon(t, x),$$

which defines the error term $Z^\varepsilon(t, x)$. Then subtracting (8) from (7), we find that $Z^\varepsilon(t, x)$ satisfies

$$\begin{aligned} Z_t^\varepsilon + \frac{1}{2}\sigma^2\left(\frac{t}{\varepsilon}\right)x^2Z_{xx}^\varepsilon + \frac{1}{2}\left(\frac{\sigma^2\left(\frac{t}{\varepsilon}\right) - \overline{\sigma^2}}{\sqrt{\varepsilon}}\right)x^2C_{xx}^{BS} &= 0, \\ Z^\varepsilon(T, x) &= 0. \end{aligned} \quad (9)$$

We shall use two convergence results: the first is **weak averaging for stochastic differential equations** which says that as $\varepsilon \downarrow 0$, the solution to (3) with initial condition $X_0^\varepsilon = x$ converges weakly to the solution of the analogous SDE with the averaged volatility coefficient $\sqrt{\overline{\sigma^2}}$:

$$d\overline{X}_t = \mu\overline{X}_t dt + \sqrt{\overline{\sigma^2}}\overline{X}_t d\overline{W}_t, \quad (10)$$

with the same starting value $\overline{X}_0 = x$, where $\{\overline{W}_t, t \geq 0\}$ is a standard Brownian motion. See, for example, [14] for a proof.

This weak approximation of X^ε by \overline{X} can then be combined with the **central limit theorem for Markov processes** to show that the scaled fluctuation of the volatility process behaves weakly like the increment of a Brownian motion $\{B_t, t \geq 0\}$:

$$\int_t^T g(s, X_s^\varepsilon) \left(\frac{\sigma^2\left(\frac{s}{\varepsilon}\right) - \overline{\sigma^2}}{\sqrt{\varepsilon}} \right) ds \longrightarrow \gamma \int_t^T g(s, \overline{X}_s) dB_s, \quad (11)$$

for bounded non-anticipating functions g ; see [10], for example. The Brownian motion B_t is standard and γ contains the remaining trace of the original volatility process through the integral of its correlation function:

$$\gamma^2 = 2 \int_0^\infty E\{(\sigma^2(s) - \overline{\sigma^2})(\sigma^2(0) - \overline{\sigma^2})\} ds. \quad (12)$$

It is shown in [13] that $Z^\varepsilon(t, x)$ converges weakly to the Gaussian process $Z(t, x)$ that satisfies the linear stochastic PDE

$$dZ_t + \frac{1}{2}\overline{\sigma^2}x^2Z_{xx}dt = -\frac{1}{2}\gamma x^2C_{xx}^{BS}dB_t, \quad (13)$$

the limit equation of (9), with $Z(T, x) = 0$. Thus we have the approximation

$$C^\varepsilon(t, x) = C^{BS}(t, x; \sqrt{\overline{\sigma^2}}) + \sqrt{\varepsilon}Z(t, x) + \mathcal{O}(\varepsilon), \quad (14)$$

for small $\varepsilon > 0$: possible option prices are decomposed as the sum of a Black-Scholes price and a normally distributed random function $Z(t, x)$ no matter what the original volatility distribution. All that is left is γ^2 (known as the power spectral density at zero frequency of the volatility), a statistic whose estimation from data is considered in Section 6.

Similar approximations are derived in [13] in the more general situation that volatility is of the form $\sigma(t, x)$, a function-space-valued random process.

Hedging Strategy

How now does this representation help with the hedging problem? Motivated by the simple form of (14), let us look for a strategy $H(t, x)$ that is the delta of a correction to the Black-Scholes delta with the averaged volatility coefficient:

$$H(t, x) = C_x^{BS}(t, x; \sqrt{\sigma^2}) + \sqrt{\varepsilon} F_x(t, x). \quad (15)$$

To measure the performance of such strategies, it is convenient to define the *effective volatility function* $E^\varepsilon(t, x)$. We want to write $C^{BS}(t, x; \sqrt{\sigma^2}) + \sqrt{\varepsilon} F(t, x)$ as $\overline{C}(t, x; [E^\varepsilon])$, the solution to (4) with $E^\varepsilon(t, x)$ instead of $\sigma_2(t)$. Expanding $E^\varepsilon(t, x)$ in the form $\sqrt{\sigma^2}$ plus correction, substituting into (4) and comparing powers of ε gives

$$E^\varepsilon(t, x) = \sqrt{\sigma^2} - \sqrt{\varepsilon} \frac{\mathcal{L}^{BS} F(t, x)}{x^2 \sqrt{\sigma^2} C_{xx}^{BS}(t, x)} + \mathcal{O}(\varepsilon),$$

where

$$\mathcal{L}^{BS} := \frac{\partial}{\partial t} + \frac{1}{2} \overline{\sigma^2} x^2 \frac{\partial^2}{\partial x^2}.$$

Now $H(t, x) = \overline{C}_x(t, x; [E^\varepsilon])$ and, given an initial cash input V_0 (to be determined), the value of our hedging portfolio is

$$\begin{aligned} V_t &= V_0 + \int_0^t \overline{C}_x(s, X_s^\varepsilon; [E^\varepsilon]) dX_s^\varepsilon \\ &= V_0 - \overline{C}(0, X_0^\varepsilon; [E^\varepsilon]) + \overline{C}(t, X_t^\varepsilon; [E^\varepsilon]) \\ &\quad - \frac{1}{2} \int_0^t \left(\sigma^2 \left(\frac{s}{\varepsilon} \right) - E^\varepsilon(s, X_s^\varepsilon)^2 \right) (X_s^\varepsilon)^2 \overline{C}_{xx}(s, X_s^\varepsilon; [E^\varepsilon]) ds, \end{aligned}$$

analogous to (5). Therefore, the final value is

$$\begin{aligned} V_T &= V_0 - \left(C^{BS}(0, X_0^\varepsilon; \sqrt{\sigma^2}) + \sqrt{\varepsilon} F(0, X_0^\varepsilon) \right) + (X_T^\varepsilon - K)^+ \\ &\quad - \frac{1}{2} \int_0^T \left(\sigma^2 \left(\frac{s}{\varepsilon} \right) - \overline{\sigma^2} + 2\sqrt{\varepsilon} \frac{\mathcal{L}^{BS} F(s, X_s^\varepsilon)}{(X_s^\varepsilon)^2 C_{xx}^{BS}(s, X_s^\varepsilon)} \right) (X_s^\varepsilon)^2 \overline{C}_{xx}(s, X_s^\varepsilon) ds + \mathcal{O}(\varepsilon). \end{aligned}$$

So the replication error, which determines whether the strategy yields a profit or a loss is weakly approximated by

$$\begin{aligned} V_T - (X_T^\varepsilon - K)^+ &= V_0 - \left(C^{BS}(0, x; \sqrt{\sigma^2}) + \sqrt{\varepsilon} F(0, x) \right) \\ &\quad - \sqrt{\varepsilon} \left(\int_0^T \mathcal{L}^{BS} F(s, \overline{X}_s) ds + \frac{1}{2} \gamma \int_0^T (\overline{X}_s)^2 C_{xx}^{BS}(s, \overline{X}_s) dB_s \right) + \mathcal{O}(\varepsilon), \end{aligned} \quad (16)$$

where $x = X_0^\varepsilon$, the observed current stock price, and we have used (11) and the weak approximation of X^ε by \overline{X} , the solution of (10), and that $\overline{C}_{xx} = C_{xx}^{BS} + \mathcal{O}(\sqrt{\varepsilon})$ within our region of asymptoticity.

Let us choose $V_0 = C^{BS}(0, x; \sqrt{\sigma^2}) + \sqrt{\varepsilon} F(0, x)$, defined to be the cost of the hedge, and find an F that, with respect to the probability measure defined by the Brownian motion

$\{B_t, t \geq 0\}$ (on some abstract space to which we do not make specific reference) makes the combined last two terms positive, given the path of \bar{X} .

Suppose we knew the path of the average (Black-Scholes) stock price $\bar{X}_s, 0 \leq s \leq T$. Since

$$\int_0^T \mathcal{L}^{BS} F(s, \bar{X}_s) ds + \frac{1}{2} \gamma \int_0^T (\bar{X}_s)^2 C_{xx}^{BS}(s, \bar{X}_s) dB_s$$

is, given this path, a Gaussian random variable with mean $M := \int_0^T \mathcal{L}^{BS} F(s, \bar{X}_s) ds$ and variance

$$S^2 := \gamma^2 \int_0^T \frac{1}{4} (X_s^\varepsilon)^4 C_{xx}^{BS}(s, X_s^\varepsilon)^2 ds,$$

our choice of F is based on a quantity that makes M negative (so that the average profit in (16) is positive). We also want $-M$ to be a number ρ times the standard deviation S . We can then choose $\rho > 0$ depending on how much risk we are prepared to allow of the normal random variable exceeding that number of its standard deviations.

First we solve

$$\begin{aligned} \mathcal{L}^{BS} \zeta(t, x) &= -\frac{1}{2} \gamma x^2 C_{xx}^{BS}(t, x) \\ \zeta(T, x) &= 0, \end{aligned} \tag{17}$$

and then taking (at time $t = 0$),

$$F(t, x) = \frac{\rho}{\sqrt{T}} \zeta(t, x) \tag{18}$$

for some $\rho > 0$ (the number of standard deviations we want), we have

$$\begin{aligned} \int_0^T \mathcal{L}^{BS} F(s, \bar{X}_s) ds &= \frac{\rho \gamma}{2\sqrt{T}} \int_0^T (\bar{X}_s)^2 C_{xx}^{BS}(s, \bar{X}_s) ds \\ &\geq \rho S. \end{aligned} \tag{19}$$

The last inequality follows from the Cauchy-Schwarz inequality

$$\left(\int_0^T f(s) ds \right)^2 \geq T \int_0^T f(s)^2 ds$$

for nonnegative functions $f(\cdot)$.

Note that from (17) and (18), $\mathcal{L}^{BS} F \leq 0$ so that $M \leq 0$ and the average replication error is positive. Thus, with this choice of F , the replication error is weakly approximated by a random variable that, conditional on the path $\bar{X}_s, 0 \leq s \leq T$, is normal with mean ρ times the standard deviation. Because the convergence is weak, this cannot be translated into a result along almost all paths of the Brownian motion $\{W_t, t \geq 0\}$. Thus we can say the chosen hedging strategy dominates the perfect hedging strategy with B -probability (probability with respect to the limiting volatility fluctuation measure defined by B_t) along almost all paths of the *average* stock price \bar{X} , but as the individual paths of \bar{X} may not be close to the paths of X^ε , we cannot make a precise quantification in terms of the joint law of X^ε . The simulations of the next section will demonstrate the effectiveness of the strategy.

In practice, ρ controls the lower bound on the risk: taking more standard deviations increases the probability of a successful hedge. Choosing $\rho = 2$ means the hedging strategy dominates the perfect strategy with B -probability 95%. Again, we stress that this is probability on the space of paths of the limiting Brownian motion B_t . We expect that the hedge success probability will also be high although we cannot quantify it exactly. A higher ρ also increases the cost of the hedge, as seen from the choice of V_0 above.

It remains to compute $F(t, x)$ by solving (17). Using the Green's function for the Black-Scholes PDE (see, for example [13, Appendix C]), we find

$$F(t, x) = \rho\gamma\sqrt{\frac{\tau}{2\pi\sigma^2T}}\frac{K^{a+1}}{x^a}\exp\left(-\nu\tau - \frac{L^2}{2\sigma^2\tau}\right),$$

where $L = \log(x/K)$, $\tau = T - t$, $a = r/\sigma^2 - 1/2$. This is the formula incorporating a nonzero interest rate r which was omitted from the equations so far for simplicity of presentation.

The hedging strategy is given by (15), where

$$F_x = -\rho\gamma\sqrt{\frac{\tau}{2\pi\sigma^2T}}\left(\frac{K}{x}\right)^{a+1}\left(\frac{\log(x/K)}{\sigma^2\tau} + \frac{r}{\sigma^2} - \frac{1}{2}\right)\exp\left(-\nu\tau - \frac{L^2}{2\sigma^2\tau}\right). \quad (20)$$

5 Simulations

In this section, we demonstrate the effectiveness of the hedging strategy derived above by simulating many stock price paths in a stochastic volatility environment, and presenting profit/loss histograms with respect to these realizations. We simulate (3) in which $\sigma(t/\varepsilon)$ is a rapidly fluctuating two-state Markov chain volatility process. This is not proposed as a realistic model of volatility, but we use it here to illustrate the performance of the asymptotic theory.

We compare two alternative hedging strategies in the underlying stock, the first with the Black-Scholes strategy using the averaged volatility $\bar{\sigma}$, in which the option writer holds $C_x^{BS}(t, X_t^\varepsilon)$ units of the stock at time t . The cost of the strategy is $C^{BS}(0, x)$, where x is the observed stock price at $t = 0$.

The second strategy incorporates the asymptotic correction for the randomly fluctuating volatility: it involves holding $H(t, X_t^\varepsilon)$ of the stock, where H is defined by (15), and we choose $\rho = 1$ in (18). That is, the strategy dominates the perfect hedging strategy with B -probability 67%. The cost of this hedge is larger than the Black-Scholes hedge: it is $C^{BS}(0, x) + \sqrt{\varepsilon}F(0, x)$.

In Figure 3, we show the stock, volatility and hedging processes along a typical realization, and in Figure 4 the profit-loss histograms from 3000 runs implementing the two strategies over the length of a twelve-month ($T = 1$) contract with 200 equally spaced re-hedgings. The profits/losses are with respect to the different costs of each strategy.

We see that the conservative second strategy yields profits much more often. On average, the profit is \$3.19 with respect to the initial cost of \$16.38. The Black-Scholes strategy produces an average profit of \$0.28 with respect to the lower cost of the strategy of \$13.32. An even more successful strategy would be to take $\rho = 2$, for example, though of course the cost will be much higher.

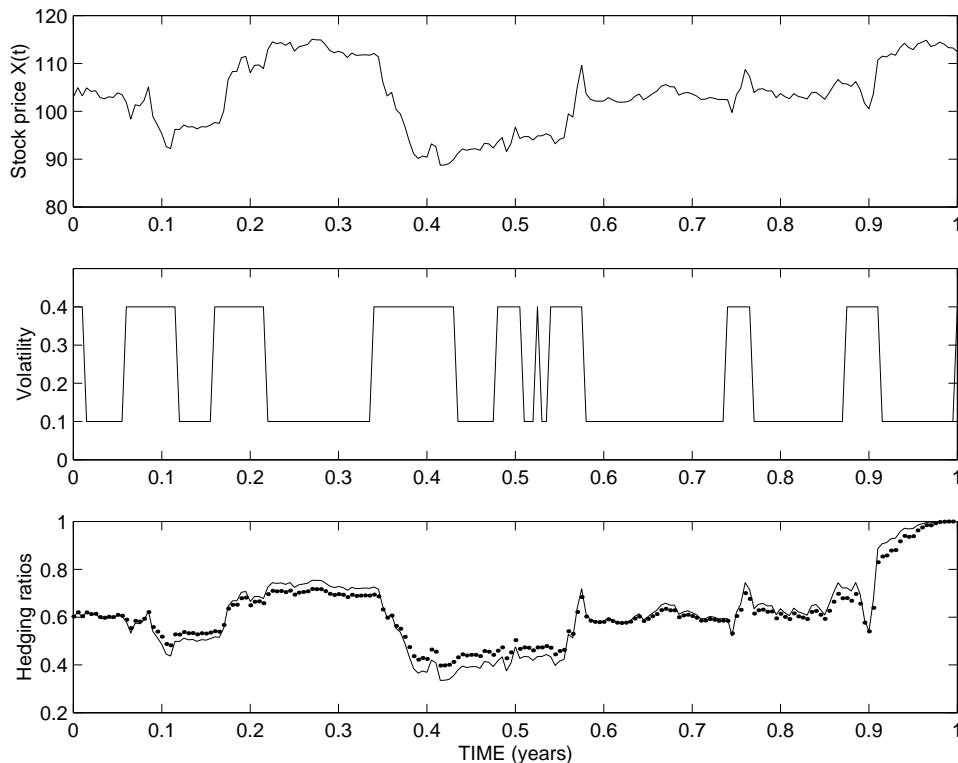


Figure 3: Stock price X_t^ε , volatility $\sigma(t/\varepsilon)$ and hedging ratios along one path. The parameter values are $\varepsilon = 0.0005$, $\rho = 1$, $K = 100$, $T = 1$. Volatility is a two-state Markov chain with values 0.1 or 0.4 and $\gamma = 1$. In the bottom graph, the dotted line shows the asymptotics-adjusted hedging ratio $H(t, X_t^\varepsilon)$ and the solid line is the Black-Scholes strategy $C_x^{BS}(t, X_t)$.

6 Estimation of Parameters

We now present a simple algorithm to estimate $\overline{\sigma^2}$ and $\varepsilon\gamma^2$ using long-run historical stock price data. The method exploits the conditional lognormal distribution of X_t^ε in the model (3). These are the only parameters needed in the asymptotic theory.

Suppose we have discrete observations $X^\varepsilon(t_n)$ of the stock price at evenly-spaced times $t_n = n\Delta t$, $n = 0, \dots, N$. Then, as $Y^\varepsilon(t) := \log X^\varepsilon(t)$ satisfies

$$dY^\varepsilon(t) = \left(\mu - \frac{1}{2}\sigma^2 \left(\frac{t}{\varepsilon} \right) \right) dt + \sigma \left(\frac{t}{\varepsilon} \right) dW_t,$$

the discrete increments of the logs of the observations satisfy

$$D_n := Y_n - Y_{n-1} = \sigma \left(\frac{t_n}{\varepsilon} \right) \Delta W_n + \left(\mu - \frac{1}{2}\sigma^2 \left(\frac{t_n}{\varepsilon} \right) \right) \Delta t,$$

where ΔW_n is a $\mathcal{N}(0, \Delta t)$ random variable.

Then the quantities

$$M_k := \frac{1}{N-k} \sum_{n=1}^{N-k} (D_n - D_{n+k})^2,$$

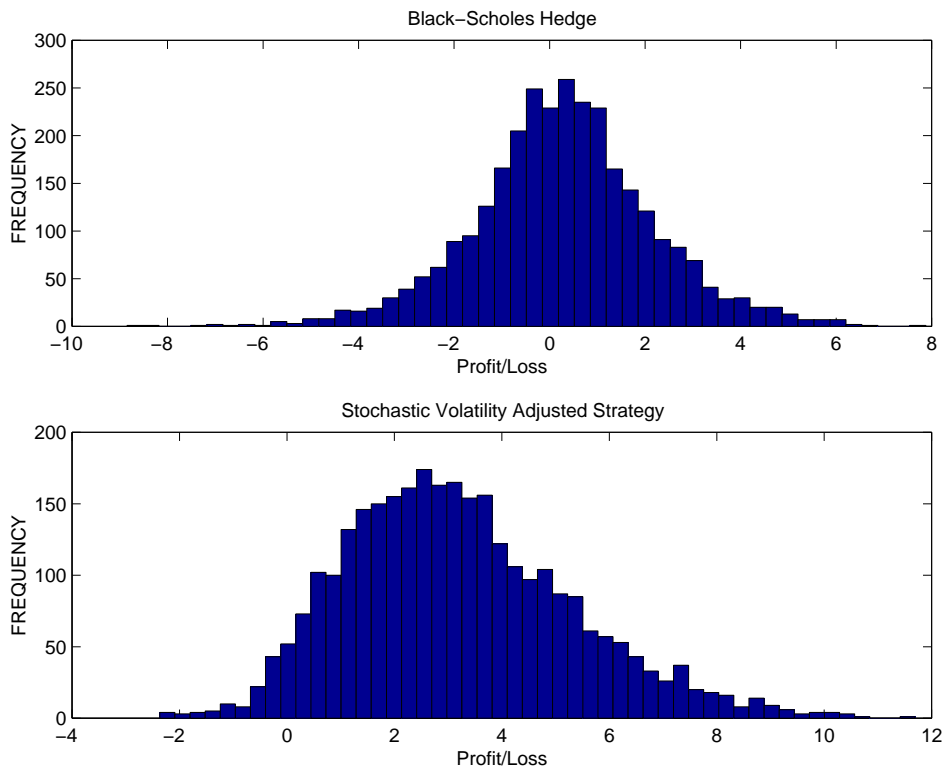


Figure 4: Profit/loss statistics of the two strategies from 3000 simulations using the parameter values listed in Figure 3. The adjusted hedge is successful more often. Note that this does not imply an arbitrage as the cost of the second hedge is greater.

for $k = 1, \dots, N - 1$, can be used to estimate $\overline{\sigma^2}$ because

$$E\{M_k\} = 2\overline{\sigma^2}\Delta t + \mathcal{O}(\Delta t^{3/2}),$$

where we have used the stationarity of $\sigma^2(\cdot)$.

Similarly, the quantities

$$T_k := \frac{1}{N - k} \sum_{n=1}^{N-k} D_n^2 D_{n+k}^2$$

can be used to estimate the (non-centred) autocorrelation of σ because

$$E\{T_k\} = E\left\{\sigma^2\left(\frac{k\Delta t}{\varepsilon}\right)\sigma^2(0)\right\}\Delta t^2 + \mathcal{O}(\Delta t^{5/2}),$$

where we have used the fact that $E\{\sigma^2(t+h)\sigma^2(t)\}$ depends only on h (second-order stationarity).

From our observations, we calculate the empirical autocorrelation

$$R_k := \frac{1}{\Delta t^2} \left[T_k - \frac{1}{4} \left(\frac{1}{k} \sum_{l=1}^k M_l \right)^2 \right].$$

The expected value of each R_k approximates the autocorrelation:

$$\begin{aligned} E\{R_k\} &= E\left\{\sigma^2\left(\frac{k\Delta t}{\varepsilon}\right)\sigma^2(0)\right\} - (\overline{\sigma^2})^2 + \mathcal{O}(\Delta t^{1/2}) \\ &= E\left\{\left(\sigma^2\left(\frac{k\Delta t}{\varepsilon}\right) - \overline{\sigma^2}\right)(\sigma^2(0) - \overline{\sigma^2})\right\} + \mathcal{O}(\Delta t^{1/2}). \end{aligned}$$

Then as

$$\int_0^\infty E\left\{\left(\sigma^2\left(\frac{s}{\varepsilon}\right) - \overline{\sigma^2}\right)(\sigma^2(0) - \overline{\sigma^2})\right\} ds = \varepsilon \int_0^\infty E\left\{\left(\sigma^2(s) - \overline{\sigma^2}\right)(\sigma^2(0) - \overline{\sigma^2})\right\} ds$$

by a change of variable, it follows that twice the area under the curve obtained by interpolating the empirical autocorrelation $\{R_k\}$ is an estimate of $\varepsilon\gamma^2$.

This procedure is tested on simulated data in [13] and practical issues (such as adaptation to nonevenly-spaced high-frequency data) are addressed there. We are presently working with real market data.

7 Conclusions

The family of hedging strategies computed here (distinguished by the number $\rho(p_0)$) dominate the perfect hedging strategy (which depends on the realized volatility path) with B -probability p_0 . This approximate bound comes from two modelling features: separation of time-scales and the generality of studying the Black-Scholes pricing PDE with a random volatility coefficient, allowing the results to apply for a large class of possible stochastic volatility processes, with unspecified distribution. The result is restricted to uncorrelated volatility (no skew or leverage effect) which is realistic in F/X and some equity markets. The simulations demonstrate the effectiveness of the $\rho = 1$ strategy, although a precise analytical quantification of the success probability is not possible because the asymptotic approximations converge weakly and not pathwise.

Greater precision in the quantification of the risk of an unsuccessful hedge would require a more detailed (parametric) description of volatility and the market pricing measure. This is studied in [6] using an additional feature of volatility that it is mean-reverting. The separation of scales analysis then aids estimation of the additional parameters of the model from S&P 500 index data.

The major benefit of the asymptotic analysis is computability, and it is a topic of future work to apply it to the stochastic control problems that arise when one wants to maximize a probability over possible hedging strategies.

References

- [1] M. Avellaneda, A. Levy, and A. Paras. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 1(2):73–88, 1995.
- [2] D. Bates. Testing option pricing models. In G. Maddala and C. Rao, editors, *Statistical Methods in Finance*, volume 14 of *Handbook of Statistics*, chapter 20, pages 567–611. North Holland, Amsterdam, 1996.

- [3] J. Cvitanić, H. Pham, and N. Touzi. Super-replication in stochastic volatility models under portfolio constraints. *Preprint*, October 1997.
- [4] H. Föllmer and P. Leukert. Quantile hedging. *Preprint*, June 1998.
- [5] J.-P. Fouque, G. Papanicolaou, and K.R. Sircar. Asymptotics of a two-scale stochastic volatility model. In *Equations aux derivees partielles et applications, in honour of Jacques-Louis Lions*, pages 517–525. Gauthier-Villars, May 1998.
- [6] J.-P. Fouque, G. Papanicolaou, and K.R. Sircar. Mean-Reverting Stochastic Volatility. *Preprint*, November 1998.
- [7] R. Frey. Derivative asset analysis in models with level-dependent and stochastic volatility. Technical report, Department of Mathematics, ETH Zürich, November 1996.
- [8] J. Hull and A. White. The Pricing of Options on Assets with Stochastic Volatilities. *J. Finance*, XLII(2):281–300, June 1987.
- [9] N. El Karoui, M. Jeanblanc-Picqué, and S. Shreve. Robustness of the Black and Scholes formula. *Math. Finance*, 8(2):93–126, April 1998.
- [10] G. C. Papanicolaou. Asymptotic analysis of stochastic equations. In *Murray Rosenblatt (ed.), Studies in Probability Theory, Mathematical Association of America*, 18:111–179, 1978.
- [11] E. Renault and N. Touzi. Option Hedging and Implied Volatilities in a Stochastic Volatility Model. *Mathematical Finance*, 6(3):279–302, July 1996.
- [12] M. Rubinstein. Nonparametric Tests of Alternative Option Pricing Models. *J. Finance*, XL(2):455–480, June 1985.
- [13] K.R. Sircar and G.C. Papanicolaou. Stochastic volatility, smile and asymptotics. *Applied Mathematical Finance*, to appear, 1998.
- [14] A. Skorohod. *Asymptotic Methods in the Theory of Stochastic Differential Equations*. Translations of Mathematical Monographs, American Mathematical Society, 1989.