

PARTIAL HEDGING IN A STOCHASTIC VOLATILITY ENVIRONMENT

MATTIAS JONSSON

*Department of Mathematics
University of Michigan
Ann Arbor, MI 48109-1109
mattiasj@umich.edu*

K. RONNIE SIRCAR

*Department of Operations Research & Financial Engineering
Princeton University
Princeton, NJ 08544.
sircar@princeton.edu*

First version: August 25, 2000; this version: February 8, 2002.

ABSTRACT: We consider the problem of partial hedging of derivative risk in a stochastic volatility environment. It is related to state-dependent utility maximization problems in classical economics. We derive the dual problem from the Legendre transform of the associated Bellman equation and interpret the optimal strategy as the perfect hedging strategy for a modified claim.

Under the the assumption that volatility is fast mean-reverting we derive, using a singular perturbation analysis, approximate value functions and strategies that are easy to implement and study. The analysis identifies the usual mean historical volatility and the harmonically-averaged long-run volatility as important statistics for such optimization problems without further specification of a stochastic volatility model. The approximation can be improved by specifying a model and calibrated for the leverage effect from the implied volatility skew. We study the effectiveness of these strategies using simulated stock paths.

KEY WORDS: Hedging of options, stochastic volatility, asymptotic analysis, dynamic programming, utility maximization, Hamilton-Jacobi-Bellman equations.

1. INTRODUCTION

We consider the problem of computing an optimal *partial hedge* for a derivative security when the underlying asset is modeled to have uncertain volatility. Stochastic volatility models are popular in the financial industry as the inadequacy of models with constant or slowly-varying volatility, especially the Black-Scholes model, has become apparent. They have been successful in accounting for the “smile effect” of traded options prices, where the classical Black-Scholes model performs poorly.

There are, however, a number of difficulties that prevent the Black-Scholes methodology of pricing and hedging by replication from going over directly to the stochastic volatility case. In the constant volatility world, the *no arbitrage* principle leads to a unique price of any derivative security, in terms of the price of the underlying asset and, possibly, its history. Additionally, the Black-Scholes analysis gives a replicating strategy by which the risk of a short position in any derivative can be *perfectly* hedged by dynamically trading in the underlying asset. The cost of the hedge, the premium that must be put up to enter into

it, is, if the market is arbitrage free, the Black-Scholes price of the derivative. In this model, the market is said to be complete.

In a stochastic volatility world, where we always think of the volatility as having an independent random component from the asset price, *no arbitrage* is not sufficient to determine a unique price for contingent claims, and in general, they cannot be perfectly replicated (without overshooting) by a trading strategy in just the asset. This is the most common example of an incomplete market¹. Strategies which *superreplicate* might be available, and have been extensively studied (Avellaneda et al., 1995; Karatzas and Shreve, 1998; Karoui and Quenez, 1995), but the cost of these hedges is related to the worst-case volatility scenario and they tend to be very expensive (Frey, 1996). In other words, the worst-case is, under most models, pretty bad and pessimistic pricing and hedging is not satisfactory for practical purposes.

It is possible to replicate a contingent claim exactly by *dynamically* trading in other *derivatives* (“delta-sigma” hedging), but this is associated with much larger transaction costs. Some component of the risk may be absorbed by *static* positions in the other derivatives. The problem is then to hedge the remaining exposure by dynamic strategies in the *underlying*.

We are interested here in how best to hedge a contingent claim, according to some performance criterion. If an institution is not willing to pay the large amount needed to guard against the worst case, how well can it do with a given lesser premium? This question can be asked in a complete market environment too, and has been solved recently (Föllmer and Leukert, 2000; Cvitanić and Karatzas, 1999). The answer is to perfectly hedge a cheaper claim whose Black-Scholes price is equal to the premium put up.

The same problem in a general incomplete market setting has also been studied. It is closely related to the classical economics problem of maximizing expected utility, and existence and uniqueness results about an optimal strategy, exploiting the semimartingale property of asset price models that exclude arbitrage, appear in (Kramkov and Schachermayer, 1999) and also in (Föllmer and Leukert, 2000). Existence is also proven in Itô models in (Cvitanić, 2000). The answer is again to hedge a cheaper claim, but the structure of the new security is not transparent. Even in the simplest examples, it is not obvious from these results how to compute the best partial hedging strategy, since the dual problem, whose construction is a powerful tool of the convex analysis methods commonly used in these problems, is typically just as difficult.

In this article, we study the problem within a class of stochastic volatility models, analyzing the associated Hamilton-Jacobi-Bellman (or simply Bellman) partial differential equation (PDE). The main tool, convex duality and the Legendre transform are reviewed first in the context of constant and time-dependent deterministic volatility and then for stochastic volatility models.

We then exploit volatility clustering or *fast mean-reversion* to analyze the dual Bellman equation by asymptotic approximations. The existence of a fast time-scale of volatility fluctuation in financial data has been shown in the empirical analysis of S&P 500 high-frequency data in (Fouque et al., 2000c), and also in exchange rate dynamics in (Alizadeh et al., 2001), and we summarize those findings in Section 6.1. This property is used to construct two hedging strategies that well approximate the optimal ones. The first depends on the volatility model mainly through the average volatility $\bar{\sigma} := \langle \sigma_t^2 \rangle^{1/2}$ and the harmonic average volatility

¹Other examples include models with jumps such as in (Duffie et al., 2000). Here the incompleteness problem is even more difficult, but this approach is also successful in explaining the smile.

$\sigma_\star := \langle \sigma_t^{-2} \rangle^{-1/2}$, where $\langle \cdot \rangle$ denotes expectation with respect to the natural invariant measure for the underlying volatility-driving process and (σ_t) is the volatility process. This strategy is extremely robust in that it does not need full specification of a stochastic volatility model, only estimation of these statistics of volatility and the present volatility level.

The second approximate strategy, which is a better approximation to the optimal strategy for a *fully specified* volatility model, is given in terms of certain volatility group parameters that can be estimated in a relatively robust manner from traded option prices and the form of the stochastic volatility model. In fact what is needed about the rate of mean-reversion of volatility and the correlation (or leverage) effect, which are the hardest parameters to estimate from historical data, is obtained from the observed implied volatility skew.

Finally, we numerically compute the expected loss from the approximate hedging strategies and examine the behavior of the strategies on simulated stock paths.

2. NOTATION & FORMULATION

Let $(X_t)_{t \geq 0}$ denote the underlying asset price (which we shall also refer to as the stock price, though it could be an exchange rate or a commodity price), modeled as a stochastic process on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ equipped with an increasing filtration $(\mathcal{F}_t)_{t \geq 0}$ representing information on X at time t . We assume that the stock price is known at time $t = 0$ (now), so that \mathcal{F}_0 is trivial.

An investor with initial capital v holds π_t dollars worth of the stock at times $0 \leq t \leq T$, where $T < \infty$ is some fixed time horizon, the maturity of the derivative contract we are interested in hedging. We assume that he or she trades in a self-financing manner, meaning that there is no further injection or extraction of capital after time $t = 0$, so that the value V_t of his or her portfolio at time t is given by

$$(2.1) \quad V_t = v + \int_0^t \frac{\pi_s}{X_s} dX_s, \quad 0 \leq t \leq T,$$

when, as we shall henceforth assume, the interest rate is zero. The strategy process $(\pi_t)_{0 \leq t \leq T}$ is *predictable* and suitably regular so that the stochastic integral in (2.1) is well-defined.

The investor holds a short position in a contingent (European) claim that pays $h(X_T)$ on expiration date T , where h satisfies

$$\mathbb{E}\{h(X_T)^2\} < \infty.$$

The problem is to find a strategy $(\pi_t)_{0 \leq t \leq T}$ such that, starting with initial capital v , the value of the hedging portfolio V_T at time T comes as close as possible to $h(X_T)$. We want the performance of the strategy to be penalized for falling short, with the actual size of the shortfall taken into account, but when the strategy overshoots the target, the size of the overshoot has no bearing on the measure of *risk*. For the explicit computations and examples here, the penalty function we shall use is

$$\mathbb{E}\left\{\frac{1}{2}((h(X_T) - V_T)^+)^2\right\},$$

the *one-sided second-moment*, where $x^+ = \max\{x, 0\}$.

One could also consider $\mathbb{E}\{l((h(X_T) - V_T)^+)\}$, where $l(\cdot)$ is a positive convex loss function with $l(0) = 0$, (as in (Föllmer and Leukert, 2000)). This allows control of one's risk aversion. Typical examples are $l(x) = x^p/p$ for $p > 1$. The analysis is identical up to changing terminal conditions.

We shall also insist that $V_t \geq 0$ for all $0 \leq t \leq T$ almost surely for the strategy to be admissible. That is, the hedging portfolio is bounded below by zero, a portfolio constraint.

The problem is to solve

$$(2.2) \quad \min_{(\pi_t)_{0 \leq t \leq T}} \mathbb{E} \left\{ \frac{1}{2} ((h(X_T) - V_T)^+)^2 \right\},$$

where V is defined in (2.1), given the initial premium v , subject to

$$(2.3) \quad V_t \geq 0, \text{ for all } 0 \leq t \leq T.$$

The problem (2.2) is reminiscent of *utility maximization* and we shall reformulate it as such, following (Föllmer and Leukert, 2000). Namely, we consider the (state-dependent) *utility function*

$$U(x, v) = \frac{1}{2} [h(x)^2 - ((h(x) - v)^+)^2].$$

For fixed x , this is concave in v on $(0, \infty)$, strictly concave on $(0, h(x))$ and satisfies $U(x, 0) = 0$, $U(x, v) = \frac{1}{2}h(x)^2$ for $v \geq h(x)$ (see Figure 1 for an illustration).

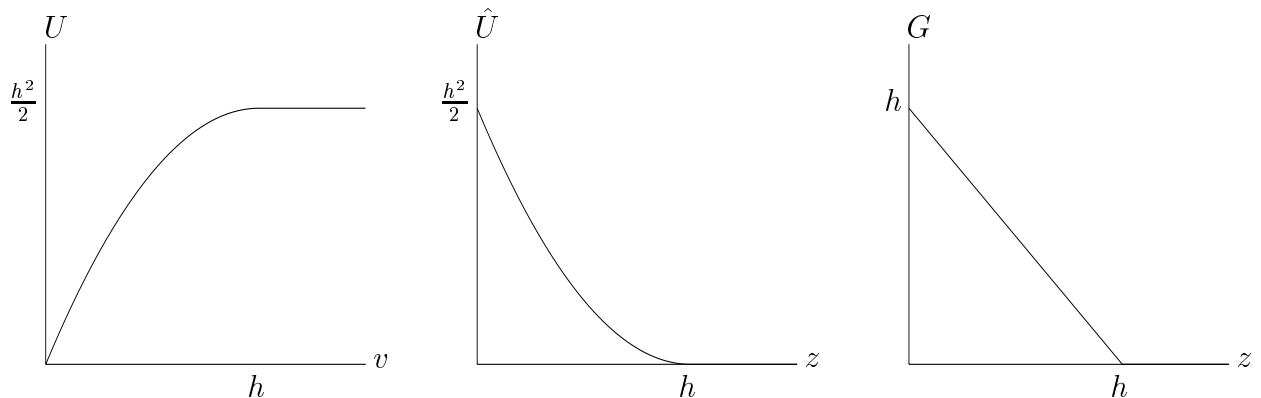


FIGURE 1. *Terminal conditions: if $h(X_T)$ is the European claim we want to hedge, for the one-sided second-moment loss function, these are typical graphs of the functions*

$$U(x, v) = \frac{1}{2}h(x)^2 - \frac{1}{2}((h(x) - v)^+)^2 \quad G(x, z) = (h(x) - z)^+ \\ \hat{U}(x, z) = \frac{1}{2}((h(x) - z)^+)^2$$

as functions of v (or z) for fixed $X_T = x$.

It is then easy to see that (2.2) is equivalent to

$$(2.4) \quad \sup_{(\pi_t)_{0 \leq t \leq T}} \mathbb{E} \{U(X_T, V_T)\},$$

subject to (2.3).

3. CONSTANT VOLATILITY

In this section we assume that the stock price is lognormal, which is an example of a complete market model. We rederive, from the Bellman PDE, the optimal partial hedging strategy as found in (Föllmer and Leukert, 2000) and (Cvitanic and Karatzas, 1999).

3.1. Bellman PDE. We assume that we are in the Black-Scholes world, that is the stock price (X_t) and the value process (V_t) satisfy:

$$(3.1) \quad \begin{aligned} \frac{dX_t}{X_t} &= \mu dt + \sigma dW_t \\ dV_t &= \pi_t \mu dt + \pi_t \sigma dW_t, \end{aligned}$$

for a strategy π_t . Here μ and σ are constants and (W_t) is a Brownian motion. We shall often use for brevity the following notation for the infinitesimal generators:

$$\begin{aligned} \mathcal{L}_x &= \mu x \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2}{\partial x^2} \\ \mathcal{L}_{x,v} &= \mathcal{L}_x + \pi \mu \frac{\partial}{\partial v} + \frac{1}{2} \pi^2 \sigma^2 \frac{\partial^2}{\partial v^2} + \pi \sigma^2 x \frac{\partial^2}{\partial x \partial v}. \end{aligned}$$

By the dynamic programming principle, the maximum expected utility

$$H(t, x, v) = \sup_{\pi} \mathbb{E}_{t,x,v} \{U(X_T, V_T)\}$$

is conjectured to satisfy the Bellman PDE

$$H_t + \sup_{\pi} \mathcal{L}_{x,v} H = 0,$$

i.e.

$$(3.2) \quad H_t + \mathcal{L}_x H + \sup_{\pi} \left(\pi (\mu H_v + \sigma^2 x H_{xv}) + \frac{1}{2} \pi^2 \sigma^2 H_{vv} \right) = 0,$$

in $x > 0, v > 0, t < T$. The terminal condition is

$$H(T, x, v) = U(x, v),$$

and the boundary condition is

$$H(t, x, 0) = 0.$$

If one can find a classical solution of this equation to which Itô's lemma can be applied, then proof of optimality follows from a verification theorem. See (Fleming and Soner, 1993), for example.

The maximum in (3.2) occurs for

$$(3.3) \quad \pi^*(t, x, v) = - \frac{(\mu H_v + \sigma^2 x H_{xv})}{\sigma^2 H_{vv}}$$

and so we get the fully nonlinear PDE

$$(3.4) \quad H_t + \mathcal{L}_x H - \frac{(\mu H_v + \sigma^2 x H_{xv})^2}{2\sigma^2 H_{vv}} = 0.$$

3.2. Legendre transform. The nonlinear PDE (3.4) can be replaced by a *linear* PDE by applying the Legendre transform. We assume initially that H is convex in v and we consider the convex dual \hat{H} of H (with respect to the variable v). This is defined by

$$(3.5) \quad \hat{H}(t, x, z) := \sup_{v>0} \{H(t, x, v) - zv\} \quad 0 < z < \infty$$

and is convex in z . We will also use

$$(3.6) \quad g(t, x, z) := \inf \left\{ v > 0 \mid H(t, x, v) \geq zv + \hat{H}(t, x, z) \right\},$$

roughly speaking, the smallest maximizer in (3.5). Under suitable smoothness assumptions we have $g = H_v^{-1}$, the inverse of marginal utility, and $g = -\hat{H}_z$.

The Bellman PDE (3.4) for H implies the following equations for \hat{H} and g (see Appendix A.1):

$$(3.7) \quad \hat{H}_t + \mathcal{L}_x \hat{H} + \frac{1}{2} \frac{\mu^2}{\sigma^2} z^2 \hat{H}_{zz} - \mu x z \hat{H}_{xz} = 0,$$

$$(3.8) \quad g_t + \frac{\mu^2}{\sigma^2} z g_z + \frac{1}{2} \sigma^2 x^2 g_{xx} - \mu x z g_{xz} + \frac{1}{2} \frac{\mu^2}{\sigma^2} z^2 g_{zz} = 0,$$

in the domain $x > 0$, $z > 0$ and $t < T$. Notice how the partial derivatives with respect to z appear as *logarithmic derivatives*, that is, they are of the form $z^n \frac{\partial^n}{\partial z^n}$. This is the first indication that the dual variable z is related to a change-of-measure (Radon-Nikodym) process.

The terminal conditions are

$$\hat{H}(T, x, z) = \hat{U}(x, z), \quad g(T, x, z) = G(x, z),$$

where \hat{U} and G are defined by

$$\begin{aligned} \hat{U}(x, z) &:= \sup_{v>0} \{U(x, v) - zv\} \\ G(x, z) &:= \inf \left\{ v > 0 \mid U(x, v) \geq zv + \hat{U}(x, z) \right\}. \end{aligned}$$

These are illustrated in Figure 1. Both are convex in z , inherited from the convexity of U in v . It follows from (3.7) and (3.8) that \hat{H} and g are smooth away from maturity, and convex in z . In fact for $t < T$, they are strictly convex. As a consequence, $H(t, x, v)$ is convex in v , strictly convex and smooth for $t < T$.

3.3. Interpretation. Notice that (3.7) and (3.8) are both linear PDE's. This is typically the case for duals of value functions in utility maximization problems in *complete* markets. See (Karatzas and Shreve, 1998, Chapter 3) for references. In the case of a power (HARA) separable utility function of the form $U(x, v) = u(x)v^p/p$, the value function inherits the power dependence (as in the original Merton solution (Merton, 1969)) and convex duality is the same as looking for a “distortion power” solution $H(t, x, v) = \tilde{H}(t, x)^{1-p}v^p/p$ (see (Zariphopoulou, 1999)). The connection is that the dual \hat{H} also has power dependence, and in fact

$$\hat{H}(t, x, z) = \left(\frac{1-p}{p} \right) z^{-p/(1-p)} \tilde{H}(t, x).$$

In our case, the linear PDEs can be interpreted as expectations as follows. Define the process (Z_t) by the SDE

$$(3.9) \quad dZ_t = -\frac{\mu}{\sigma} Z_t dW_t.$$

Then (3.7) shows that

$$(3.10) \quad \hat{H}(t, x, z) = \mathbb{E}_{t,x,z} \{ \hat{U}(X_T, Z_T) \}.$$

In fact, depending on the initial condition, (Z_t) is proportional to the density process for the unique equivalent martingale measure (EMM) $Q \simeq P$ under which the price process (X_t) is a martingale.

By Girsanov's theorem, we have (under Q)

$$\begin{aligned}\frac{dX_t}{X_t} &= \sigma dW_t^Q \\ \frac{dZ_t}{Z_t} &= \frac{\mu^2}{\sigma^2} dt - \frac{\mu}{\sigma} dW_t^Q,\end{aligned}$$

where (W_t^Q) is a Q -Brownian motion. From (3.8), it follows that

$$(3.11) \quad g(t, x, z) = \mathbb{E}_{t,x,z}^Q \{G(X_T, Z_T)\},$$

the *no-arbitrage* price of the derivative security that pays $G(X_T, Z_T)$ on date T .

This observation is interpreted as follows: given (t, x, v) , find z (uniquely) such that the no-arbitrage price of the claim $G(X_T, Z_T)$, with $X_t = x$, $Z_t = z$, is exactly v . Then the maximum expected utility $H(t, x, v)$ is given by

$$H(t, x, v) = zv + \hat{H}(t, x, z).$$

More importantly, the optimal strategy is the hedging strategy for the claim $G(X_T, Z_T)$. To see this we first notice that, on the one hand, (3.3) implies

$$(3.12) \quad \pi^*(t, x, z) = xg_x - \frac{\mu}{\sigma^2} zg_z$$

(see Appendix A.1). On the other hand, if the price of the claim $G(X_T, Z_T)$ at time t is given by some smooth function $\tilde{g}(t, X_t, Z_t)$, with

$$(3.13) \quad \tilde{g}(T, x, z) = G(x, z),$$

then we can try to set up a riskless and self-financing portfolio containing the claim and $-\Delta_t$ units of stock, i.e.

$$\Pi_t = \tilde{g}(t, X_t, Z_t) - \Delta_t X_t.$$

By Itô's Lemma we have (under the real-world measure):

$$\begin{aligned}d\Pi_t &= \left[\tilde{g}_t + \frac{1}{2}\sigma^2 X_t^2 \tilde{g}_{xx} + \frac{1}{2}\frac{\mu^2}{\sigma^2} Z_t^2 \tilde{g}_{zz} - \mu X_t Z_t \tilde{g}_{xz} - \Delta_t \mu X_t \right] dt \\ &\quad + \left[\sigma X_t \tilde{g}_x - \frac{\mu}{\sigma} Z_t \tilde{g}_z - \Delta_t \sigma X_t \right] dW_t,\end{aligned}$$

using also the self-financing property. Thus the portfolio is instantaneously riskless if and only if

$$\Delta_t = \tilde{g}_x(t, X_t, Z_t) - \frac{\mu}{\sigma^2} \frac{Z_t}{X_t} \tilde{g}_z(t, X_t, Z_t).$$

Further, by no-arbitrage, \tilde{g} must satisfy the PDE:

$$\tilde{g}_t + \frac{1}{2}\sigma^2 x^2 \tilde{g}_{xx} - \mu x z \tilde{g}_{xz} + \frac{1}{2}\frac{\mu^2}{\sigma^2} z^2 \tilde{g}_{zz} + \frac{\mu^2}{\sigma^2} z \tilde{g}_z = 0,$$

which is the same as the PDE (3.8) satisfied by g . The terminal condition (3.13) for \tilde{g} is the same as that for g , and so, from uniqueness, $\tilde{g} \equiv g$. Thus $\pi_t^* = \Delta_t X_t$ and is exactly the hedging strategy for the claim $G(X_T, Z_T)$.

These results are given more generally, for arbitrary complete semimartingale market models in (Föllmer and Leukert, 2000) and (Cvitanic and Karatzas, 1999).

In the case of the lognormal model (3.1), it follows from the geometric structure of both (3.1) and (3.9) that

$$Z_T = cX_T^{-\mu/\sigma^2},$$

for a constant c (depending on x, z, μ, σ). Hence the optimal strategy is a hedging strategy for a *European claim* $G(X_T, Z_T) = \tilde{G}(X_T; z)$, where the initial value z is chosen so that the price of this claim is exactly the initial capital v . Thus the European structure of the hedging strategy is preserved under partial hedging. This remarkable fact is an artifact of the constant volatility assumption and will not hold in more general models.

This type of result where a modified or *reduced* terminal payoff absorbs the “departure from Black-Scholes” is seen in other contexts too, for example in superhedging of claims under portfolio constraints (Broadie et al., 1998).

3.4. Explicit Computation: Call Option. For the call option with strike price K , the payoff is

$$h(X_T) = (X_T - K)^+.$$

The modified payoff function $\tilde{G}(X_T; z)$ is given by

$$\tilde{G}(X_T; z) = \left((X_T - K)^+ - zcX_T^{-\mu/\sigma^2} \right)^+,$$

and the typical shape is shown in Figure 2.

The formula

$$g(t, x, z) = \mathbb{E}_{t,x}^Q \{ \tilde{G}(X_T; z) \}$$

gives

$$\begin{aligned} (3.14) \quad g(t, x, z) &= xN(\tilde{d}_2 + \sigma\sqrt{\tau}) - KN(\tilde{d}_2) - ze^{\frac{\mu^2}{\sigma^2}\tau}N(\tilde{d}_2 - \frac{\mu}{\sigma}\sqrt{\tau}), \\ \hat{H}(t, x, z) &= \frac{1}{2}x^2e^{(2\mu+\sigma^2)\tau}N(\tilde{d}_2 + (2\sigma + \frac{\mu}{\sigma})\sqrt{\tau}) + \frac{1}{2}K^2N(\tilde{d}_2 + \frac{\mu}{\sigma}\sqrt{\tau}) \\ &\quad + \frac{1}{2}z^2e^{\frac{\mu^2}{\sigma^2}\tau}N(\tilde{d}_2 - \frac{\mu}{\sigma}\sqrt{\tau}) - xKe^{\mu\tau}N(\tilde{d}_2 + (\sigma + \frac{\mu}{\sigma})\sqrt{\tau}) \\ (3.15) \quad &\quad - xzN(\tilde{d}_2 + \sigma\sqrt{\tau}) + KzN(\tilde{d}_2). \end{aligned}$$

where $\tau = T - t$, $N(\cdot)$ is the standard normal cumulative distribution function, and \tilde{d}_2 is the unique solution of the equation

$$xe^{-\frac{1}{2}\sigma^2\tau - \sigma\sqrt{\tau}\tilde{d}_2} - K - ze^{\frac{1}{2}\frac{\mu^2}{\sigma^2}\tau + \frac{\mu}{\sigma}\sqrt{\tau}\tilde{d}_2} = 0.$$

This is illustrated in Figure 2.

Notice that when $z = 0$,

$$\tilde{d}_2 = \frac{\log(x/K) - \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}},$$

which is the usual “ d_2 ” appearing in the Black-Scholes call option pricing formula. In this case, the investor has sufficient capital to hedge the call option perfectly, $\tilde{G}(X_T; 0) = (X_T - K)^+$, and $g(t, x, 0)$ is just the Black-Scholes call option pricing function.

Finally, the optimal hedging ratio $\Delta_t = \pi_t^*/X_t$ is given from formula (3.12) by

$$\Delta(t, x, z) = N(\tilde{d}_2 + \sigma\sqrt{\tau}) + \frac{z}{x} \frac{\mu}{\sigma^2} e^{\frac{\mu^2}{\sigma^2}\tau} N(\tilde{d}_2 - \frac{\mu\sqrt{\tau}}{\sigma}).$$

See Figure 2 for a graph.

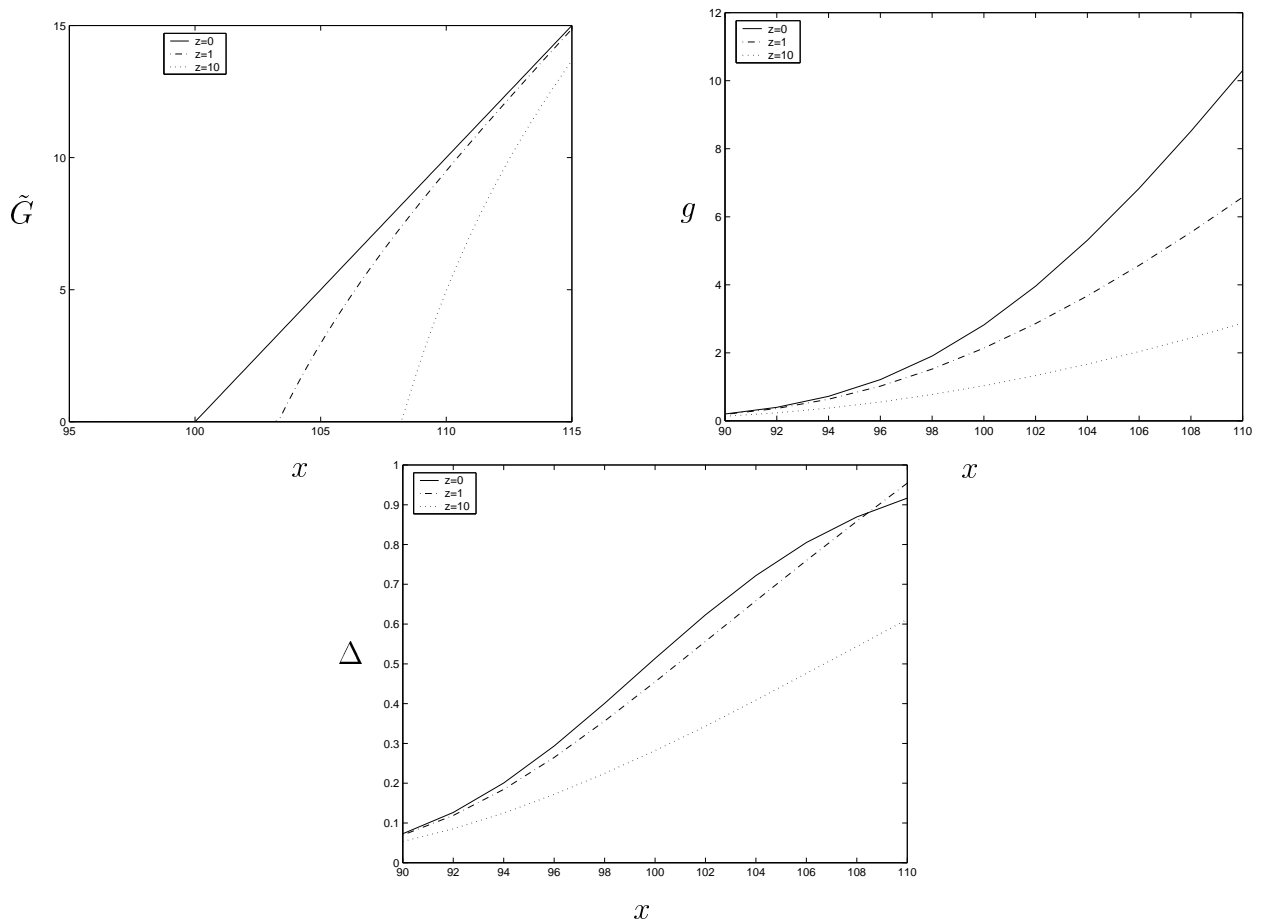


FIGURE 2. *Reduced call options.* The first picture shows the payoff functions $\tilde{G}(x; z)$, $x = X_T$ for the reduced call for $z = 0$, $z = 1$ and $z = 10$. The second picture shows the prices $g(t, x, z)$ of the three claims as a function of $x = X_t$ with $T - t = 0.5$. The third picture shows the deltas $\Delta(t, x, z)$. Notice that a large z corresponds to a cheap option and that $z = 0$ gives the standard call. The price function $g(t, x, z)$ is a decreasing function of z (for fixed t and x); this is not true for the delta $\Delta(t, x, z)$. In fact, we can have $\Delta > 1$ near the expiration date and deep in the money.

4. TIME-DEPENDENT VOLATILITY

A slight extension of the Black-Scholes model is obtained by allowing volatility to be time-dependent but deterministic: $\sigma = \sigma(t)$. This market is still complete, but the simplicity of the constant volatility calculation, in which the European structure is preserved under partial hedging, is lost, and the optimal strategy involves hedging of a path-dependent claim. We present the calculation of this case here in some detail, not because this is a useful model in practice, but because it is a natural bridge to the stochastic volatility models we want to study and, in particular, their analysis by asymptotic methods. It highlights the key features of the stock price and the volatility paths that the new claim depends upon.

While the analysis of Sections 3.1 and 3.2 carry over directly in this more general case, replacing σ by $\sigma(t)$, the interpretations will not be the same as in Section 3.3. The model

for the stock price (X_t) and the value process (V_t) is:

$$(4.1) \quad \begin{aligned} \frac{dX_t}{X_t} &= \mu dt + \sigma(t) dW_t \\ dV_t &= \pi_t \mu dt + \pi_t \sigma(t) dW_t, \end{aligned}$$

for a strategy π_t , where $\sigma(t)$ is a deterministic function of time.

The Bellman PDE for the value function $H(t, x, v)$ is

$$(4.2) \quad H_t + \mathcal{L}_x H - \frac{(\mu H_v + \sigma(t)^2 x H_{xv})^2}{2\sigma(t)^2 H_{vv}} = 0,$$

where

$$\mathcal{L}_x H = \mu x H_x + \frac{1}{2} \sigma(t)^2 x^2 H_{xx}.$$

The optimal strategy is given by

$$(4.3) \quad \pi^*(t, x, v) = -\frac{(\mu H_v + \sigma(t)^2 x H_{xv})}{\sigma(t)^2 H_{vv}}.$$

By applying the Legendre transform (3.5)- (3.6) as before, we get linear PDE's for \hat{H} and g :

$$(4.4) \quad \hat{H}_t + \mathcal{L}_x \hat{H} + \frac{1}{2} \frac{\mu^2}{\sigma(t)^2} z^2 \hat{H}_{zz} - \mu x z \hat{H}_{xz} = 0,$$

$$(4.5) \quad g_t + \frac{\mu^2}{\sigma(t)^2} z g_z + \frac{1}{2} \sigma(t)^2 x^2 g_{xx} - \mu x z g_{xz} + \frac{1}{2} \frac{\mu^2}{\sigma(t)^2} z^2 g_{zz} = 0,$$

with terminal conditions

$$\hat{H}(T, x, z) = \hat{U}(x, z), \quad g(T, x, z) = G(x, z).$$

4.1. Interpretation. The interpretation of π^* as the hedging strategy for a new claim $G(X_T, Z_T)$ still holds. Now (Z_t) satisfies

$$dZ_t = -\frac{\mu}{\sigma(t)} Z_t dW_t,$$

and so is still proportional to the density process for the unique EMM Q . However, if $\sigma(t)$ is nonconstant, then Z_T is *not* a point function of X_T , and so the claim $G(X_T, Z_T)$ is path-dependent.

In this case,

$$(4.6) \quad Z_T = z \exp \left(-\int_t^T \frac{\mu}{\sigma^2(s)} \frac{dX_s}{X_s} + \frac{1}{2} \int_t^T \frac{\mu^2}{\sigma^2(s)} ds \right),$$

so the optimal strategy is to perfectly hedge the claim which pays $G(X_T, Z_T)$, the payoff at time T depending on the terminal price X_T and the *volatility-path-weighted* return (4.6). The constant z is chosen so that the price of this claim is equal to the initial capital v .

Finally, notice from (4.4) and (4.5) that $\hat{H}(t, x, z)$ and $g(t, x, z)$ depend on the volatility path $\{\sigma(s) : t \leq s \leq T\}$ *only through the averages*

$$(4.7) \quad \bar{\sigma} := \sqrt{\frac{1}{T-t} \int_t^T \sigma(s)^2 ds}$$

$$(4.8) \quad \sigma_{\star} := \frac{1}{\sqrt{\frac{1}{T-t} \int_t^T \frac{1}{\sigma(s)^2} ds}}.$$

This can be seen, for example, by changing to logarithmic spatial variables $(\log x, \log z)$ to obtain constant (in space) coefficient PDEs and then taking the Fourier transform in both spatial variables to obtain a linear ODE which can be integrated. The dependence on the integrals in (4.7) and (4.8) survives the inverse Fourier and Legendre transforms (which act only on the spatial variables) and so the value function $H(t, x, v)$ also depends only on these averages.

However while the optimal expected utility (or equivalently the minimum expected shortfall from hedging) only depends on the volatility path through these averages, formula (4.3) shows the optimal strategy depends on the whole path because $\pi^*(t, x, v)$ depends explicitly on $\sigma(t)$. The volatility must be entirely known to use the best strategy. This observation will be important in using the asymptotic results of Section 6.

We always have the inequality $\sigma_{\star} \leq \bar{\sigma}$, and equality holds if and only if $\sigma(t)$ is a constant function. This follows from Jensen's inequality. The extent to which σ_{\star} differs from $\bar{\sigma}$ is a measure of the size of the fluctuations of volatility. We will come back to this in Section 6.

5. STOCHASTIC VOLATILITY

We are interested in more realistic market models, particularly ones in which volatility is uncertain. Motivation for stochastic volatility models and their success in describing the observed implied volatility skew from traded European option prices is described in (Fouque et al., 2000a), for example. There are a number of practical complications that arise from this generalization, in particular because volatility is not a directly observed or traded process. We cannot perfectly hedge derivatives by trading in just the underlying stock, though we may be able to by selecting another *derivative* security as part of the hedge. As explained in the introduction, we are interested in computation of less prohibitively-expensive strategies that dynamically trade only the underlying and allow possibility of shortfall.

5.1. Volatility Mean-Reversion. Mean-reversion as a much-observed characteristic of equity, FX and commodity volatility is important to model, as is the correlation between volatility and asset price shocks which directly accounts for the skew and asymmetry in empirical returns distributions. For this reason we write our canonical class of stochastic volatility models as a positive function of a simple ergodic Itô process, a mean-reverting Ornstein-Uhlenbeck process:

$$(5.1) \quad \frac{dX_t}{X_t} = \mu dt + f(Y_t) dW_t$$

$$(5.2) \quad dY_t = \alpha(m - Y_t)dt + \beta \left(\rho dW_t + \sqrt{1 - \rho^2} dB_t \right)$$

$$(5.3) \quad dV_t = \pi_t \mu dt + \pi_t f(Y_t) dW_t.$$

Here the volatility process is

$$\sigma_t = f(Y_t),$$

and (W_t) and (B_t) are independent Brownian motions with $-1 < \rho < 1$ the instantaneous correlation coefficient between asset price and volatility shocks. The factor (Y_t) is called the *volatility-driving process* and f is some positive suitably regular function whose specification is unimportant for the principal asymptotic approximation derived in Section 6.

5.2. Bellman PDE. The maximum expected utility

$$H(t, x, y, v) = \sup_{\pi} \mathbb{E}_{t,x,y,v} \{U(X_T, V_T)\}$$

is now a function of the additional state variable and is conjectured to satisfy the Bellman PDE

$$H_t + \sup_{\pi} \mathcal{L}_{x,y,v} H = 0$$

i.e.

$$(5.4) \quad H_t + \mathcal{L}_{x,y} H + \sup_{\pi} \left(\pi(\mu H_v + f(y)^2 x H_{xv} + \rho \beta f(y) H_{yv}) + \frac{1}{2} \pi^2 f(y)^2 H_{vv} \right) = 0,$$

where

$$\mathcal{L}_{x,y} H = \mu x H_x + \alpha(m - y) H_y + \frac{1}{2} f(y)^2 x^2 H_{xx} + \rho \beta f(y) x H_{xy} + \frac{1}{2} \beta^2 H_{yy},$$

the infinitesimal generator of (X_t, Y_t) .

The max in (5.4) occurs for

$$(5.5) \quad \pi^*(t, x, y, v) = - \frac{(\mu H_v + f(y)^2 x H_{xv} + \rho \beta f(y) H_{yv})}{f(y)^2 H_{vv}}$$

and so we get the fully nonlinear PDE

$$(5.6) \quad H_t + \mathcal{L}_{x,y} H - \frac{(\mu H_v + f(y)^2 x H_{xv} + \rho \beta f(y) H_{yv})^2}{2f(y)^2 H_{vv}} = 0$$

in the domain $x > 0$, $-\infty < y < \infty$, $v > 0$ and $t < T$ with terminal condition

$$H(T, x, y, v) = U(x, v).$$

Remark: It follows from a general result in (Föllmer and Leukert, 2000) based on the duality theorem in (Kramkov and Schachermayer, 1999) that H is convex in v . As in the constant volatility case, we shall work with the Legendre-transformed variables for interpretation of the optimal strategy in Section 5.4 and also computation using asymptotic approximations in Section 6.

5.3. Legendre transform. We again consider the convex dual \hat{H} of H (with respect to the variable v). This is defined by

$$\hat{H}(t, x, y, v) := \sup_{v > 0} \{H(t, x, y, v) - zv\}$$

and is convex in z (from the concavity of $H(t, x, y, v)$ in v). We will also use

$$g(t, x, y, z) := \inf \left\{ v > 0 \mid H(t, x, y, v) \geq zv + \hat{H}(t, x, y, z) \right\}.$$

The function g should be thought of the inverse of marginal utility; under suitable smoothness assumptions we have $g = H_v^{-1} = -\hat{H}_z$.

The Bellman PDE for H implies the following equations for \hat{H} and g (see Appendix A.2):

$$(5.7) \quad \hat{H}_t + \mathcal{L}_{x,y}\hat{H} + \frac{1}{2}\frac{\mu^2}{f(y)^2}z^2\hat{H}_{zz} - \mu xz\hat{H}_{xz} - \frac{\rho\beta\mu}{f(y)}z\hat{H}_{yz} - \frac{1}{2}\beta^2(1-\rho^2)\frac{\hat{H}_{yz}^2}{\hat{H}_{zz}} = 0,$$

$$(5.8) \quad g_t + \mathcal{L}_{x,y}g + \frac{1}{2}\frac{\mu^2}{f(y)^2}z^2g_{zz} - \mu xzg_{xz} - \frac{\rho\beta\mu}{f(y)}zg_{yz} + \frac{\mu^2}{f(y)^2}zg_z - \mu xg_x - \frac{\rho\beta\mu}{f(y)}g_y - \frac{1}{2}\beta^2(1-\rho^2)\left[\frac{2g_yg_{yz}}{g_z} - \frac{g_y^2}{g_z^2}g_{zz}\right] = 0,$$

with terminal conditions

$$\hat{H}(T, x, y, z) = \hat{U}(x, z), \quad g(T, x, y, z) = G(x, z).$$

Note that, in contrast to the constant volatility case, these are still nonlinear PDEs.

5.4. Interpretation. One can see that (5.7) is the Bellman PDE for the stochastic control problem

$$\hat{H}(t, x, y, z) = \inf_{\gamma} \mathbb{E}_{t,x,y,z} \{\hat{U}(X_T, Y_T, Z_T^{\gamma})\}$$

where (X_t, Y_t) satisfy (5.1) and (5.2), and (Z_t^{γ}) is defined through

$$(5.9) \quad dZ_t^{\gamma} = -\frac{\mu}{f(Y_t)}Z_t^{\gamma}dW_t + \gamma_t dB_t.$$

We will also see that the optimal strategy is the *perfect hedging strategy* for the claim $G(X_T, Z_T^{\gamma})$, where Z^{γ} is the density process of the “optimal” EMM. To this end we observe that (see Appendix A.2)

$$(5.10) \quad \pi^*(t, x, y, z) = xg_x + \frac{\rho\beta}{f(y)}g_y - \frac{\mu}{f(y)^2}zg_z.$$

We know that the Radon-Nikodym process of any EMM is proportional to $Z = Z^{\gamma}$ for some suitable process (γ_t) . This follows from Girsanov’s theorem and can be thought of as follows: the first Brownian motion (W_t) must be “shifted” (by a drift) so that (X_t) is a martingale or has drift zero in (5.1). However the second Brownian motion (B_t) can be shifted arbitrarily (up to regularity) by some adapted process (γ_t) .

Let us try and choose (γ_t) so that the claim $G(X_T, Z_T^{\gamma})$ can be hedged perfectly with just the underlying. Suppose first that the price of this claim at time t is given by $\tilde{g}(t, X_t, Y_t, Z_t)$, for some smooth function \tilde{g} to be determined, with

$$\tilde{g}(T, x, y, z) = G(x, z).$$

Then consider the portfolio containing the claim and $-\Delta_t$ units of stock at time t :

$$\Pi_t = \tilde{g}(t, X_t, Y_t, Z_t^{\gamma}) - \Delta_t X_t.$$

The portfolio is traded in a self-financing manner and by Itô's Lemma we have:

$$\begin{aligned}
d\Pi_t(t, X_t, Y_t, Z_t^\gamma) &= \left[\tilde{g}_t + \mathcal{L}_{x,y}\tilde{g} - \mu X_t Z_t^\gamma \tilde{g}_{xz} + \left(\gamma \beta \sqrt{1 - \rho^2} - \frac{\mu \beta \rho}{f(Y_t)} \right) \tilde{g}_{yz} \right. \\
&\quad \left. + \frac{1}{2} \left(\gamma^2 + \frac{\mu^2}{f(Y_t)^2} (Z_t^\gamma)^2 \right) \tilde{g}_{zz} - \Delta \mu X_t \right] dt \\
&\quad + \left[f(Y_t) X_t \tilde{g}_x + \rho \beta \tilde{g}_y - \frac{\mu}{f(Y_t)} Z_t^\gamma \tilde{g}_z - \Delta_t f(Y_t) X_t \right] dW_t \\
&\quad + \left[\beta \sqrt{1 - \rho^2} \tilde{g}_y + \gamma_t \tilde{g}_z \right] dB_t.
\end{aligned}$$

Thus the portfolio is instantaneously riskless if and only if we choose

$$\Delta_t = \tilde{g}_x + \frac{\beta \rho}{f(Y_t) X_t} \tilde{g}_y - \frac{\mu}{f(Y_t)^2 X_t} Z_t^\gamma \tilde{g}_z \quad \text{and} \quad \gamma_t = -\beta \sqrt{1 - \rho^2} \frac{\tilde{g}_y}{\tilde{g}_z}.$$

Normally we would not expect to be able to hedge most claims perfectly by trading in only the stock, but the extra parameter γ allows us to do this in this case. By no-arbitrage, \tilde{g} must satisfy the PDE

$$\begin{aligned}
\tilde{g}_t + \mathcal{L}_{x,y}\tilde{g} + \frac{1}{2} \frac{\mu^2}{f(y)^2} z^2 \tilde{g}_{zz} - \mu x z \tilde{g}_{xz} - \frac{\rho \beta \mu}{f(y)} z \tilde{g}_{yz} + \frac{\mu^2}{f(y)^2} z \tilde{g}_z \\
- \mu x \tilde{g}_x - \frac{\rho \beta \mu}{f(y)} \tilde{g}_y - \frac{1}{2} \beta^2 (1 - \rho^2) \left[\frac{2 \tilde{g}_y \tilde{g}_{yz}}{\tilde{g}_z} - \frac{\tilde{g}_y^2}{\tilde{g}_z^2} \tilde{g}_{zz} \right] = 0
\end{aligned}$$

But this is exactly the (transformed) Bellman PDE (5.8) satisfied by g . Since \tilde{g} and g also have the same terminal condition, we conclude that $\tilde{g} \equiv g$ by uniqueness.

Therefore, we have

$$(5.11) \quad \Delta_t = g_x + \frac{\beta \rho}{f(Y_t) X_t} g_y - \frac{\mu}{f(Y_t)^2} \frac{Z_t^\gamma}{X_t} g_z$$

$$(5.12) \quad \gamma_t = -\beta \sqrt{1 - \rho^2} \frac{g_y}{g_z}.$$

Comparing (5.11) with (5.10), we see that $\pi_t^* = X_t \Delta_t$, and so is the hedging strategy for the claim $G(X_T, Z_T^\gamma)$, where γ is chosen so that this claim can actually be hedged perfectly by trading only the stock.

6. ASYMPTOTICS FOR PARTIAL HEDGING

In this section, we study the effect of uncertain volatility on the partial hedging strategies and optimal expected shortfall. We take advantage of *fast mean-reversion* or *clustering* in market volatility that is described in (Fouque et al., 2000a) and use a singular perturbation analysis to find relatively simple hedging strategies that well approximate the optimal one. The empirical basis for this approach is described in the study of S&P 500 data in (Fouque et al., 2000c), and summarized in Section 6.1 below. The analysis in Section 5 still has not yielded a way to compute optimal strategies short of solving one of the nonlinear PDEs (5.6), (5.7) or (5.8) which have three spatial dimensions. One of the benefits of the approach described here is easing of this dimensional burden. Hedging under *uncorrelated* stochastic volatility is also studied in a nonparametric way using asymptotic methods in (Sircar, 2000; Sircar and Papanicolaou, 1999).

In the zero-order approximation derived here, two kinds of average volatilities emerge: $\bar{\sigma} := \langle \sigma^2 \rangle^{1/2}$ and $\sigma_* := \langle \sigma^{-2} \rangle^{-1/2}$, where $\langle \cdot \rangle$ denotes a particular averaging procedure described below. The first-order approximation also takes information from the observed implied volatility skew from traded *options prices* to account for the effect of volatility correlation, or asymmetric returns distributions. Other market information that is needed can be estimated from this in a relatively robust manner.

6.1. Fast Mean-Reversion of Volatility. In (Fouque et al., 2000c), we studied high-frequency S&P 500 data over the period of a year to estimate the order of the rate of mean-reversion of volatility. The major difficulty with high-frequency data is pronounced intraday phenomena associated with microscopic trading patterns as described, for example, in (Andersen and Bollerslev, 1997). In (Fouque et al., 2000c), it was shown how this 'periodic day effect' impacted the variogram and spectral methods used to analyze the data, and therefore how to account for it.

The result was, for the S&P 500 data examined, the clear presence of a fast time scale of volatility fluctuation, corresponding to a rate of mean reversion $\alpha \sim 130 - 230$ (in annualized units). The important qualitative information here is two-fold: first that the rate of mean-reversion of the unobserved volatility process is extremely difficult to estimate *precisely*, hence the large range. This was confirmed from tests on simulated data. Second, we can nonetheless be precise about the *order of magnitude* of α from this estimate: it is large.

Many empirical studies have looked at low-frequency (daily) data, with the data necessarily ranging over a period of years, and they have found a *low* rate of volatility mean-reversion. This does not contradict the empirical finding described above: analyzing data at lower frequencies over longer time periods would primarily pick up a slower time-scale of fluctuation and could not identify scales at the same order as the sampling frequency. The combined conclusions suggest that there are (at least) two important scales in volatility, and has led recently to the study of *two-factor stochastic volatility models* (Chernov et al., 2001)², where one factor is slowly mean-reverting and the other is fast mean-reverting.

If indeed the latter is a fairly accurate model of market volatility, then the traditional use of stochastic volatility models corresponds to ignoring the fast factor, on the grounds that it averages out, and concentrating on the slow factor for derivative pricing and risk management. However, when these one-factor models are calibrated from S&P 500 option prices, the estimated "v-vol" (volatility of volatility) is unreasonably large in comparison to the small rate of mean-reversion. Even adding jumps to the model does not seem to resolve this problem. See (Bakshi et al., 1997) and (Duffie et al., 2000) for details. One possible explanation is that the fast factor, modeled as having a large rate of mean-reversion and a large diffusion coefficient in order to balance the characteristic size of volatility fluctuations, shows up in option prices. In other words, if one factor could be safely ignored, it is not the fast one.

Another recent empirical study (Alizadeh et al., 2001), this time of exchange rate dynamics, finds "the evidence points strongly toward two-factor [volatility] models with one highly persistent factor and one quickly mean-reverting factor". To pull one estimate from their Table VI, they find the rate of mean-reversion of the fast volatility factor for the US Dollar-Deutsche Mark exchange rate to be $\alpha = 237.5$ (in annualized units). For the other four exchange rates they also look at, the order of magnitude of this parameter is the same (hundreds).

²We are grateful to a referee for pointing us to this reference and helpful suggestions on this topic.

In the remainder of this section, we focus on the effect of the fast volatility scale on the partial hedging problem for a European derivative with expiration on the order of a few months. In the context of a two-factor stochastic volatility model, this would correspond to ignoring the slow factor. This is reasonable if the lifetime of the derivative is on the order of the typical half-life of that factor, or less, because that factor would act approximately like a constant as far as expectations were concerned. Often, the slow factor half-life is found to be on the order of 70 – 90 days in equity indices (Engle and Patton, 2001). For partial hedging problems involving derivative with longer maturities, one would have to consider the full two-factor model and deal numerically with the resulting high-dimensional (four spatial plus time) Bellman partial differential equation.

6.2. Singular Perturbation Analysis. We introduce the scaling

$$\begin{aligned}\alpha &= 1/\varepsilon \\ \beta &= \sqrt{2}\nu/\sqrt{\varepsilon}\end{aligned}$$

where $0 < \varepsilon \ll 1$ and $\nu = \mathcal{O}(1)$ (fixed), to model fast mean-reversion (clustering) in market volatility. Recall that α measures the characteristic speed of mean-reversion of (Y_t) and ν^2 is the variance of the long-run distribution, measuring the typical size of the fluctuations of the volatility-driving process.

Then $g = g^\varepsilon$ satisfies the PDE (5.8), which we re-write with the new notation as:

$$(6.1) \quad \left(\frac{1}{\varepsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\varepsilon}} \mathcal{L}_1 + \mathcal{L}_2 \right) g^\varepsilon + \frac{\nu^2}{\varepsilon} (1 - \rho^2) \text{NL}^\varepsilon = 0,$$

where we define

$$(6.2) \quad \mathcal{L}_0 = \nu^2 \frac{\partial^2}{\partial y^2} + (m - y) \frac{\partial}{\partial y}$$

$$(6.3) \quad \mathcal{L}_1 = \sqrt{2} \rho \nu \left(f(y) x \frac{\partial^2}{\partial x \partial y} - \frac{\mu}{f(y)} z \frac{\partial^2}{\partial y \partial z} - \frac{\mu}{f(y)} \frac{\partial}{\partial y} \right)$$

$$(6.4) \quad \mathcal{L}_2 = \frac{\partial}{\partial t} + \frac{1}{2} f(y)^2 x^2 \frac{\partial^2}{\partial x^2} + \frac{\mu^2}{f(y)^2} \left(\frac{1}{2} z^2 \frac{\partial^2}{\partial z^2} + z \frac{\partial}{\partial z} \right) - \mu x z \frac{\partial^2}{\partial x \partial z},$$

and the nonlinear part is

$$\text{NL}^\varepsilon = -\frac{\partial}{\partial z} \left(\frac{(g_y^\varepsilon)^2}{g_z^\varepsilon} \right) = -\left[\frac{2g_y^\varepsilon g_{yz}^\varepsilon}{g_z^\varepsilon} - \left(\frac{g_y^\varepsilon}{g_z^\varepsilon} \right)^2 g_{zz}^\varepsilon \right].$$

Notice that \mathcal{L}_0 is the usual (scaled) OU generator and \mathcal{L}_1 takes derivatives in y and kills functions that do not depend on y .

The approach is now to think of the actual market (and our hedging problem) as embedded in a family of similar problems parametrized by (small) values of ε . For $\varepsilon = 0$, volatility is mean-reverting “infinitely fast” and can be replaced by some average as far as expectations are concerned. However two different averages are needed for different facets of the optimal strategy. This principal approximation may be sufficient for many purposes. It can be improved by perturbation or expansion around $\varepsilon = 0$, but the cost is greater reliance on model specification, as we describe.

We also point out that the asymptotic analysis presented here differs substantially from the problems described in (Fouque et al., 2000a) on derivative pricing problems in the following regard:

- the partial differential equations in the previous work were linear whereas (6.1) is *nonlinear*;
- the analysis here is on a dual equation arrived at after the Legendre transform has isolated the nonlinearity due to market incompleteness;
- the nonlinearity is *strong* in that it appears as an order $1/\varepsilon$ term and so plays a role immediately in the construction of the expansion.

6.2.1. *Expansion.* We look for an expansion

$$g^\varepsilon(t, x, y, z) = g^{(0)}(t, x, y, z) + \sqrt{\varepsilon}g^{(1)}(t, x, y, z) + \varepsilon g^{(2)}(t, x, y, z) + \dots$$

for small ε .

6.2.2. *Term of Order $1/\varepsilon$.* Inserting the expansion and comparing terms of order $1/\varepsilon$ gives

$$(6.5) \quad \nu^2 g_{yy}^{(0)} + (m - y)g_y^{(0)} - \nu^2(1 - \rho^2) \frac{\partial}{\partial z} \left(\frac{(g_y^{(0)})^2}{g_z^{(0)}} \right) = 0,$$

This implies that $g^{(0)}$ does not depend on y , assuming that $g^{(0)}$ is smooth and of controlled growth. To see this ³, we define

$$k(y, z) := - \int_0^z g^{(0)}(y, \zeta) d\zeta;$$

this would be the leading term in the asymptotic expansion of \hat{H} . Clearly $k_z = -g^{(0)}$ and so (6.5) leads to the following PDE for k :

$$(6.6) \quad \nu^2 k_{yy} + (m - y)k_y - \nu^2(1 - \rho^2) \frac{(k_{yz})^2}{k_{zz}} = 0.$$

The pricing function $g(t, x, z)$ is strictly decreasing in z for $t < T$; thus the same must be true for the leading term $g^{(0)}$, and so $k_{zz} = -g_z > 0$ everywhere. Using this, and fixing z we obtain the following ordinary differential inequality:

$$\nu^2 k_{yy} + (m - y)k_y \geq 0.$$

By integrating we get

$$\begin{aligned} k_y(y, z) &\geq k_y(m, z) e^{\frac{(y-m)^2}{2\nu^2}} \quad \text{for } y \geq m \\ k_y(y, z) &\leq k_y(m, z) e^{\frac{(y-m)^2}{2\nu^2}} \quad \text{for } y \leq m \end{aligned}$$

We conclude that $k_y(m, z) = 0$, because otherwise k_y would grow too fast as $y \rightarrow \pm\infty$. This implies that any solution k to (6.6) is independent of y . To see this, write $l(z) := k(m, z)$ and define $\tilde{k}(y, z) := l(z)$. Then \tilde{k} is clearly a solution to (6.6) and satisfies $\tilde{k}_y(m, z) = k_y(m, z) = 0$, $\tilde{k}(m, z) = k(y, z) = l(z)$, so by uniqueness $k = \tilde{k}$. Thus neither k nor $g^{(0)} = -k_z$ depend on y .

³We are grateful to J. Rauch for suggesting the following argument.

6.2.3. *Term of Order $1/\sqrt{\varepsilon}$.* At the order $1/\sqrt{\varepsilon}$,

$$\mathcal{L}_1 g^{(0)} + \mathcal{L}_0 g^{(1)} = 0,$$

which implies $g^{(1)}$ also does not depend on y because $\mathcal{L}_1 g^{(0)} = 0$ and \mathcal{L}_0 has null space spanned by constants. This is a general property of generators of “nice” ergodic processes like the OU.

Since both $g^{(0)}$ and $g^{(1)}$ do not depend on y , the nonlinear term is effectively

$$\text{NL}^\varepsilon = \mathcal{O}(\varepsilon^2)$$

and only contributes to the asymptotics when we compare order ε and higher. We will go as far as order $\sqrt{\varepsilon}$ so we are dealing essentially with linear asymptotics (except for the very first equation).

6.2.4. *Zeroth-order Term.* At order 1, we have

$$\mathcal{L}_0 g^{(2)} + \mathcal{L}_1 g^{(1)} + \mathcal{L}_2 g^{(0)} = 0.$$

The middle term is zero because $g^{(1)}$ does not depend on y . We have a Poisson equation (in y) for $g^{(2)}$. The solvability condition is that $\mathcal{L}_2 g^{(0)}$ must be centered with respect to the invariant distribution of the OU process (Y_t) (equivalently, orthogonal to the null space of the adjoint of \mathcal{L}_0 , the Fredholm alternative). Therefore

$$(6.7) \quad \langle \mathcal{L}_2 g^{(0)} \rangle = \langle \mathcal{L}_2 \rangle g^{(0)} = 0,$$

where $\langle \cdot \rangle$ denotes the averaging

$$\langle \Psi \rangle = \frac{1}{\sqrt{2\pi\nu^2}} \int_{-\infty}^{\infty} \Psi(y) e^{-\frac{(m-y)^2}{2\nu^2}} dy,$$

that is, the average with respect to the $\mathcal{N}(m, \nu^2)$ distribution, the invariant or long-run distribution of the OU process (Y_t).

The averaged operator is

$$\langle \mathcal{L}_2 \rangle = \frac{\partial}{\partial t} + \frac{1}{2} \bar{\sigma}^2 x^2 \frac{\partial^2}{\partial x^2} + \frac{\mu^2}{\sigma_\star^2} \left(\frac{1}{2} z^2 \frac{\partial^2}{\partial z^2} + z \frac{\partial}{\partial z} \right) - \mu x z \frac{\partial^2}{\partial x \partial z},$$

where we define

$$(6.8) \quad \bar{\sigma}^2 = \langle f^2 \rangle$$

$$(6.9) \quad \frac{1}{\sigma_\star^2} = \left\langle \frac{1}{f^2} \right\rangle.$$

The terminal condition is

$$g^{(0)}(T, x, z) = G(x, z).$$

The problem for $g^{(0)}(t, x, z)$ is similar to the constant volatility problem (3.8), with two important differences:

- (1) The zeroth-order approximation $g^{(0)}$ depends not just on the usual long-run average historical volatility $\bar{\sigma}$, but also on the *harmonically-averaged volatility* σ_\star defined by (6.9). Thus the asymptotic approximation of the optimal strategy will depend on estimating this unusual volatility too. By Jensen’s inequality, $\sigma_\star \leq \bar{\sigma}$ and equality holds if and only if volatility is constant a.s. This statistic also arose in the asymptotic analysis of the Merton problem under stochastic volatility in (Fouque et al., 2000a, Chapter 10).

- (2) The “homogenized” operator $\langle \mathcal{L}_2 \rangle$ is nondegenerate even though \mathcal{L}_2 is. As a result, $g^{(0)}(t, x, z)$ is the expectation of a functional of a *two-dimensional* Brownian motion, unlike the expectation in (3.11). In other words, the zero-order asymptotic approximation is not simply the complete market problem with constant averaged volatility.

The consequences of this are discussed within the call option example in Section 6.3.

6.2.5. *Zero-order Strategy.* The optimal zero-order strategy is given by

$$(6.10) \quad \pi^* = \left(x \frac{\partial}{\partial x} - \frac{\mu}{f(y)^2} z \frac{\partial}{\partial z} \right) g^{(0)}.$$

Notice that this *does* depend on tracking volatility $f(y)$ even though the corrected minimum expected loss does not (to zero-order).

6.2.6. *Interpretation and Estimation of σ_* .* One possible way to estimate σ_* is to use the Taylor expansion

$$\begin{aligned} \frac{1}{\sigma_t^2} &= \frac{1}{\bar{\sigma}^2 + (\sigma_t^2 - \bar{\sigma}^2)} \\ &\approx \frac{1}{\bar{\sigma}^2} \left[1 - \left(\frac{\sigma_t^2 - \bar{\sigma}^2}{\bar{\sigma}^2} \right) + \left(\frac{\sigma_t^2 - \bar{\sigma}^2}{\bar{\sigma}^2} \right)^2 \right], \end{aligned}$$

so that

$$\left\langle \frac{1}{\sigma_t^2} \right\rangle \approx \frac{\langle \sigma_t^4 \rangle}{\bar{\sigma}^6}.$$

The long-run volatility $\bar{\sigma}$ and the fourth-moment $\langle \sigma_t^4 \rangle$ can be estimated stably from the second and fourth-moments of high-frequency historical returns (see (Fouque et al., 2000b), for example). There is no need to specify a volatility model $f(Y_t)$.

This rough estimator also shows that

$$(6.11) \quad \frac{\bar{\sigma}^2}{\sigma_*^2} \approx \frac{\langle \sigma_t^4 \rangle}{\bar{\sigma}^4},$$

and so $\bar{\sigma}/\sigma_*$ is a measure of excess kurtosis.

6.3. Explicit Computation: Call Option. As in Section 3.4, we consider partial hedging of a European call option with strike price K . To compute the zeroth order approximation $g^{(0)}(t, x, z)$ to the dual variable $g(t, x, y, z)$, we start with the probabilistic representation of the PDE problem (6.7).

Let $(B_t^{(1)}, B_t^{(2)})$ be a two-dimensional Brownian motion on some probability space and define Itô processes $(\hat{X}_s, \hat{Z}_s)_{t \leq s \leq T}$ by

$$\begin{aligned} \frac{d\hat{X}_s}{\hat{X}_s} &= \bar{\sigma} dB_s^{(1)} \\ \frac{d\hat{Z}_s}{\hat{Z}_s} &= \frac{\mu^2}{\sigma_*^2} ds - \frac{\mu}{\sigma_*} \left(\hat{\rho} dB_s^{(1)} + \sqrt{1 - \hat{\rho}^2} dB_s^{(2)} \right), \end{aligned}$$

where

$$(6.12) \quad \hat{\rho} := \frac{\sigma_*}{\bar{\sigma}},$$

If we define starting values

$$\hat{X}_t = x \quad \hat{Z}_t = z,$$

then

$$g^{(0)}(t, x, z) = \mathbb{E}_{t,x,z} \left\{ ((\hat{X}_T - K)^+ - \hat{Z}_T)^+ \right\}.$$

Note, we can think of (\hat{X}_t) as a stock price in a “shadow” market with constant volatility $\bar{\sigma}$, and (\hat{Z}_t) as something reminiscent of a Radon-Nikodym process. The parameter $\hat{\rho}$, which satisfies $0 < \hat{\rho} \leq 1$ by Jensen’s inequality, can be thought of as a correlation coefficient and is a measure of how much volatility is fluctuating in the real market. In particular, $\hat{\rho} = 1$ if and only if volatility is constant a.s. In the expOU model for example, the approximation (6.11) turns out to be exact and $\hat{\rho} = e^{-2\nu^2}$, where $\nu^2 = \beta^2/2\alpha$ is a measure of the *size* of volatility fluctuations.

The expectation can be simplified to reveal $g^{(0)}$ as an average of prices of “reduced” calls of the form (3.14) with $\sigma = \bar{\sigma}$ and varying z . To see this, define a random variable S by

$$S = \frac{1}{\sqrt{\tau}}(B_T^{(2)} - B_t^{(2)}),$$

and write

$$a(S) = e^{\frac{\mu}{\sigma^*} \sqrt{1-\hat{\rho}^2} \sqrt{\tau} S + \frac{1}{2} \frac{\mu^2}{\sigma^*} (1-\hat{\rho}^2) \tau}.$$

Then $S \sim \mathcal{N}(0, 1)$ and we may alternatively write

$$\hat{Z}_T = \hat{Z}_T^{(S)}, \quad \text{in distribution,}$$

where $\hat{Z}_s^{(S)}$ satisfies the one-dimensional SDE

$$\frac{d\hat{Z}_s^{(S)}}{\hat{Z}_s^{(S)}} = \frac{\mu^2}{\bar{\sigma}^2} ds - \frac{\mu}{\bar{\sigma}} dB_s^{(1)},$$

with the random independent initial condition

$$\hat{Z}_t^{(S)} = za(S).$$

Thus, comparing with the analysis in Section 3.4, we see that

$$g^{(0)}(t, x, z) = \mathbb{E} \{ g(t, x, za(S); \bar{\sigma}) \},$$

where $g(t, x, z; \bar{\sigma})$ is given by (3.14) and S is drawn from a standard $\mathcal{N}(0, 1)$ distribution.

We can also use (6.3) to compute derivatives of $g^{(0)}$. For instance, we get

$$\begin{aligned} g_x^{(0)}(t, x, z) &= \mathbb{E} \{ g_x(t, x, za(S); \bar{\sigma}) \} \\ g_z^{(0)}(t, x, z) &= \mathbb{E} \{ a(S) g_z(t, x, za(S); \bar{\sigma}) \}. \end{aligned}$$

Similar results also hold for $\hat{H}^{(0)}(t, x, z)$, the zero order approximation to $\hat{H}(t, x, y, z)$.

6.4. First Correction. The zeroth-order approximation may be sufficient and the most practical. It depends only on estimation from data of $\bar{\sigma}$ and σ_* which can be done in a model-independent manner (no need to specify a function f). We can compute the correction $g^{(1)}$ in order to improve the approximation to the optimal partial hedging strategy. Of course, this will depend on greater specification of a stochastic volatility model (so far we have only had to estimate $\bar{\sigma}$ and σ_*). We shall discuss robustness to model specification in Section 6.5.

A standard multiple-scales argument (see Appendix B) shows that $g^{(1)}$ satisfies

$$(6.13) \quad \langle \mathcal{L}_2 \rangle g^{(1)} = \langle \mathcal{L}_1 \mathcal{L}_0^{-1} (\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle) \rangle g^{(0)},$$

with zero terminal condition. It is more convenient to multiply by $\sqrt{\varepsilon}$ and write the equation for

$$\widetilde{g^{(1)}} = \sqrt{\varepsilon} g^{(1)},$$

namely

$$\langle \mathcal{L}_2 \rangle \widetilde{g^{(1)}} = \mathcal{A} g^{(0)}.$$

It remains to compute the operator

$$\mathcal{A} = \sqrt{\varepsilon} \langle \mathcal{L}_1 \mathcal{L}_0^{-1} (\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle) \rangle.$$

We have that

$$(6.14) \quad \mathcal{L}_2 - \langle \mathcal{L}_2 \rangle = \frac{1}{2} (f(y)^2 - \bar{\sigma}^2) x^2 \frac{\partial^2}{\partial x^2} + \mu^2 \left(\frac{1}{f(y)^2} - \frac{1}{\sigma_*^2} \right) \left(\frac{1}{2} z^2 \frac{\partial^2}{\partial z^2} + z \frac{\partial}{\partial z} \right),$$

so that

$$\mathcal{A} = \frac{\rho\nu}{\sqrt{2\alpha}} \left\langle \left(f(y)x \frac{\partial^2}{\partial x \partial y} - \frac{\mu}{f(y)} \left(z \frac{\partial^2}{\partial y \partial z} + \frac{\partial}{\partial y} \right) \right) \left(\phi x^2 \frac{\partial^2}{\partial x^2} + \mu^2 \psi \left(z^2 \frac{\partial^2}{\partial z^2} + 2z \frac{\partial}{\partial z} \right) \right) \right\rangle,$$

where ϕ and ψ satisfy the Poisson equations

$$(6.15) \quad \mathcal{L}_0 \phi = f(y)^2 - \bar{\sigma}^2$$

$$(6.16) \quad \mathcal{L}_0 \psi = \frac{1}{f(y)^2} - \frac{1}{\sigma_*^2}.$$

This becomes

$$\begin{aligned} \mathcal{A} = & (2C_1 - \mu C_3) x^2 \frac{\partial^2}{\partial x^2} + C_1 x^3 \frac{\partial}{\partial x^3} + C_2 \mu^2 \left(x z^2 \frac{\partial^3}{\partial x \partial z^2} + 2x z \frac{\partial^2}{\partial x \partial z} \right) \\ & - \mu C_3 z x^2 \frac{\partial^3}{\partial x^2 \partial z} - C_4 \mu^3 \left(1 + z \frac{\partial}{\partial z} \right) \left(z^2 \frac{\partial^2}{\partial z^2} + 2z \frac{\partial}{\partial z} \right), \end{aligned}$$

where

$$(6.17) \quad \begin{aligned} C_1 &= \frac{\rho\nu}{\sqrt{2\alpha}} \langle f \phi' \rangle \\ C_2 &= \frac{\rho\nu}{\sqrt{2\alpha}} \langle f \psi' \rangle \\ C_3 &= \frac{\rho\nu}{\sqrt{2\alpha}} \left\langle \frac{\phi'}{f} \right\rangle \\ C_4 &= \frac{\rho\nu}{\sqrt{2\alpha}} \left\langle \frac{\psi'}{f} \right\rangle. \end{aligned}$$

The solution is

$$\begin{aligned}
\widetilde{g}^{(1)} &= -(T-t)\mathcal{A}g^{(0)} \\
&= -(T-t) \left[(2C_1 - \mu C_3)x^2 \frac{\partial^2}{\partial x^2} + C_1 x^3 \frac{\partial}{\partial x^3} + C_2 \mu^2 \left(xz^2 \frac{\partial^3}{\partial x \partial z^2} + 2xz \frac{\partial^2}{\partial x \partial z} \right) \right. \\
(6.18) \quad &\left. - \mu C_3 z x^2 \frac{\partial^3}{\partial x^2 \partial z} - C_4 \mu^3 \left(z^3 \frac{\partial^3}{\partial z^3} + 5z^2 \frac{\partial^2}{\partial z^2} + 4z \frac{\partial}{\partial z} \right) \right] g^{(0)}.
\end{aligned}$$

6.4.1. *Corrected Strategy.* Given that $g^{(0)}$ and $g^{(1)}$ do not depend on y , the corrected optimal strategy is given by

$$\begin{aligned}
(6.19) \quad \pi^* &= \left(x \frac{\partial}{\partial x} - \frac{\mu}{f(y)^2} z \frac{\partial}{\partial z} \right) (g^{(0)} + \widetilde{g}^{(1)}) \\
&\quad - \frac{\sqrt{2} \rho \nu}{\sqrt{\alpha} f(y)} \left(\frac{1}{2} \phi'(y) x^2 g_{xx}^{(0)} + \mu^2 \psi'(y) \left(\frac{1}{2} z^2 g_{zz}^{(0)} + z g_z^{(0)} \right) \right).
\end{aligned}$$

This is derived in Appendix B.

6.5. **Estimation of Parameters and Robustness.** The approximation to the optimal strategy in Section 6.4.1 depends on the average volatilities

$$\bar{\sigma}, \sigma_*$$

and the group market parameters

$$C_1, C_2, C_3, C_4.$$

Given a fully specified and estimated model of volatility $(\alpha, \nu, m, \rho, f(\cdot))$, these can all be computed using the formulas (6.8)- (6.17). The function $\phi'(y)$ and $\psi'(y)$ which are also needed in (6.19) are found as a by-product of these calculations. However, it is well known that estimates of the correlation ρ and the mean-reversion rate α are extremely unstable because volatility is not a directly observed process. See, for example (Fouque et al., 2000c) for a discussion of this issue. The advantage of the asymptotic approximation is that what is needed about these two parameters can be obtained more stably from the *implied volatility skew from liquidly traded option prices*.

First, notice that ρ, ν and α appear as $\rho\nu/\sqrt{\alpha}$ in the formulas (6.17), so we only need (to this level of approximation) this grouping. Further, the parameter C_1 is exactly the parameter V_3 that arises from the asymptotic approximation of derivative prices, as discussed in (Fouque et al., 2000a). It is estimated from fitting the implied volatility surface I to an affine function of the log-moneyness-to-maturity ratio (LMMR)

$$I = a \frac{\log(K/x)}{T-t} + b.$$

From the estimated slope a , $V_3 = -a\bar{\sigma}^3$.

Having obtained $C_1 = V_3$ from the skew, it remains to find the other C 's. This will depend on further model specification, in particular choosing a function $f(\cdot)$.

If we define

$$\Phi(y) = \frac{1}{\sqrt{2\pi\nu^2}} e^{-\frac{(m-y)^2}{2\nu^2}},$$

the density of the invariant distribution $\mathcal{N}(m, \nu^2)$ of the OU process (Y_t) , then from (6.15), it follows that

$$\phi'(y) = \frac{1}{\Phi(y)} \int_{-\infty}^y (f(z)^2 - \bar{\sigma}^2) \Phi(z) dz,$$

so that, after integration-by-parts,

$$\langle f \phi' \rangle = -\langle F(f^2 - \bar{\sigma}^2) \rangle,$$

where $F(y)$ is an antiderivative of $f(y)$. Similarly, using also (6.16),

$$(6.20) \quad \langle f \psi' \rangle = -\langle F\left(\frac{1}{f^2} - \frac{1}{\sigma_\star^2}\right) \rangle,$$

and

$$(6.21) \quad \left\langle \frac{\phi'}{f} \right\rangle = -\langle \tilde{F}(f^2 - \bar{\sigma}^2) \rangle,$$

$$(6.22) \quad \left\langle \frac{\psi'}{f} \right\rangle = -\langle \tilde{F}\left(\frac{1}{f^2} - \frac{1}{\sigma_\star^2}\right) \rangle,$$

where $\tilde{F}(y)$ is an antiderivative of $1/f(y)$.

For the *expOU model*, $f(y) = e^y$

$$(6.23) \quad \begin{aligned} \langle f \phi' \rangle &= e^{3m + \frac{5\nu^2}{2}} - e^{3m + \frac{9\nu^2}{2}} \\ \left\langle \frac{\phi'}{f} \right\rangle &= e^{m + \frac{\nu^2}{2}} - e^{m + \frac{5\nu^2}{2}} \\ \langle f \psi' \rangle &= e^{-m + \frac{5\nu^2}{2}} - e^{-m + \frac{\nu^2}{2}} \\ \left\langle \frac{\psi'}{f} \right\rangle &= e^{-3m + \frac{5\nu^2}{2}} - e^{-3m + \frac{9\nu^2}{2}}, \end{aligned}$$

and

$$(6.24) \quad \bar{\sigma}^2 = e^{2m + 2\nu^2}$$

$$(6.25) \quad \sigma_\star^2 = e^{2m - 2\nu^2}.$$

The estimation procedure is then as follows:

- (1) From historical returns data, estimate $\bar{\sigma}$ and σ_\star as discussed in Section 6.2.6. These estimates are typically very stable and do not use any correlation structure.
- (2) Compute m and ν from (6.24) and (6.25).
- (3) From the skew, estimate $V_3 = C_1$.
- (4) Compute C_2 , C_3 and C_4 using (6.17) and the formulas (6.23).

We shall illustrate the effectiveness and computational simplicity of the asymptotic formulas from simulations in Section 7.

Remark: The expOU model violates conditions for applying Girsanov's theorem. Therefore we shall always assume a cutoff version in the application and in simulations, where the cutoffs are sufficiently above and below not to affect any of the calculations to the order of the asymptotic approximations.

7. NUMERICAL RESULTS AND SIMULATIONS

In this section we present numerical results that illustrate the results described above. In particular, we demonstrate the performance of the zero- and first-order hedging strategies.

First we show how these two strategies behave along a simulated stock path in a specific stochastic volatility model.

Then we compare the two loss functions associated to the zero- and first-order strategies, respectively.

7.1. Strategies along a volatility path. We examined the behavior of the zero-order and first-order hedging strategies along a typical stock path, and compared with the Föllmer-Leukert hedging strategy, pretending that volatility is constant.

For this we used the explicit model

$$\begin{aligned}\frac{dX_t}{X_t} &= \mu dt + e^{Y_t} dW_t \\ dY_t &= \alpha(m - Y_t)dt + \beta\rho dW_t + \beta\sqrt{1 - \rho^2} dB_t \\ dV_t &= \pi_t \frac{dX_t}{X_t},\end{aligned}$$

where $\mu = 0.2$, $K = 100$, $\alpha = 200$, $\rho = -0.2$, $m = -2.3651$ and $\nu = 0.25$ for simulation. This gives $\bar{\sigma} = 0.1$ and $\sigma_* = 0.0882$.

To compute the three strategies along the path we did as follows.

- (1) For the Föllmer-Leukert strategy we pretended that volatility is constant, equal to $\bar{\sigma}$. Given t , X_t and V_t we solved numerically the equation $g(t, X_t, z; \bar{\sigma}) = V_t$ for z , where g satisfies (3.14). The hedging ratio $\Delta = \pi/X$ is given by

$$\Delta_t = g_x(t, X_t, z) - \frac{\mu}{\bar{\sigma}^2} \frac{z}{X_t} g_z(t, X_t, z)$$

and the value of the portfolio was updated using $dV_t = \Delta_t dX_t$. (Note the z changes with time too although we do not denote its dependence here).

- (2) For the zero-order strategy we did as follows. Given t , X_t and V_t we solved the equation $g^{(0)}(t, X_t, z) = V_t$ for z , with $g^{(0)}$ from (6.3). The hedging ratio was then chosen as

$$\Delta_t = g_x^{(0)}(t, X_t, z) - \frac{\mu}{e^{2Y_t}} \frac{z}{X_t} g_z^{(0)}(t, X_t, z)$$

and the value of the portfolio updated using $dV_t = \Delta_t dX_t$.

- (3) The computation for the first-order strategy is very similar. Given t , X_t and V_t we solved the equation $(g^{(0)} + \widetilde{g^{(1)}})(t, X_t, z) = V_t$ for z , with $g^{(0)}$ as above and $\widetilde{g^{(1)}}$ from (6.18). The hedging ratio was then chosen as

$$\Delta_t = (g^{(0)} + \widetilde{g^{(1)}})_x(t, X_t, z) - \frac{\mu}{e^{2Y_t}} \frac{z}{X_t} (g^{(0)} + \widetilde{g^{(1)}})_z(t, X_t, z)$$

and the value of the portfolio updated using $dV_t = \Delta_t dX_t$. For simplicity of implementation, we omitted the last grouped term in the formula (6.19) for the first-order approximation to π^* . As discussed below, we still observe an improvement over the other strategies without this term.

The portfolios were re-hedged 200 times over the life of the option.

Figures 3, 4 and 5 illustrate the relative performances along a typical simulated path when the initial capital was 60% of the Black-Scholes price of the claim computed with the volatility $\bar{\sigma}$ and the option started at the money. Clearly on such a path, the zeroth-order strategy has outperformed the constant volatility strategy and the first-order has come closer to the terminal payoff than either of them. An interesting observation from Figure 4 is that the hedging deltas are often bigger than one when the option is in the money. This can be thought of as the investor's aversion to high losses (measured by the shortfall measure of risk) driving him to capitalize on the rising stock.

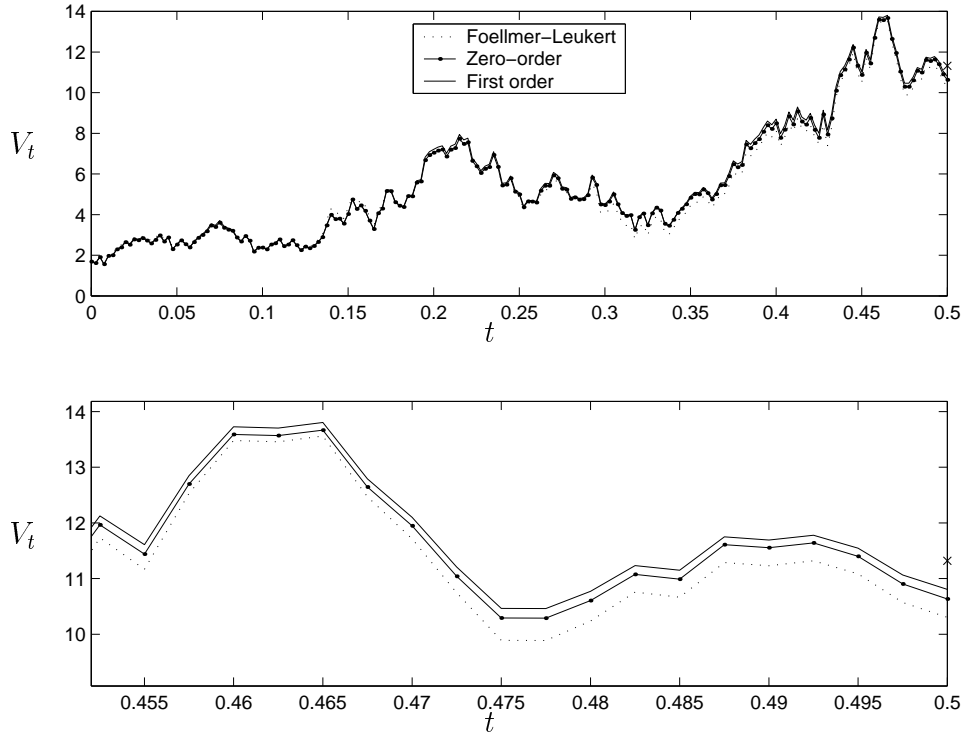


FIGURE 3. *Portfolio values along a path. The cross at the right side of the picture illustrates the value of the option $(X_T - K)^+$ at expiration. The full stock price and volatility paths are shown in Figure 5. The initial capital was 60% of the Black-Scholes price of the claim computed with the volatility $\bar{\sigma}$. The bottom graph magnifies how the values of the three hedging portfolios differ near expiration.*

Figures 6, 7 and 8 show the differences in a more dramatic situation, when the investor starts with a low initial hedging premium (20% of the Black-Scholes price of the option). In the path shown, the differences in performance of the first-order strategy over the zeroth-order and even the zeroth-order above the constant volatility strategy are quite significant.

7.2. Loss functions. It is also of interest to study the expected losses under the zero- and first-order strategies. The principal ($\varepsilon = 0$) approximation $\hat{H}^{(0)}$ to the dual value function \hat{H} satisfies the PDE

$$(7.1) \quad \hat{H}_t^{(0)} + \mu x \hat{H}_x^{(0)} + \frac{1}{2} \bar{\sigma}^2 x^2 \hat{H}_{xx}^{(0)} + \frac{1}{2} \frac{\mu^2}{\sigma_*^2} z^2 \hat{H}_{zz}^{(0)} - \mu x z \hat{H}_{xz}^{(0)} = 0, \quad \hat{H}^{(0)}(T, x, z) = \hat{U}(x, z).$$

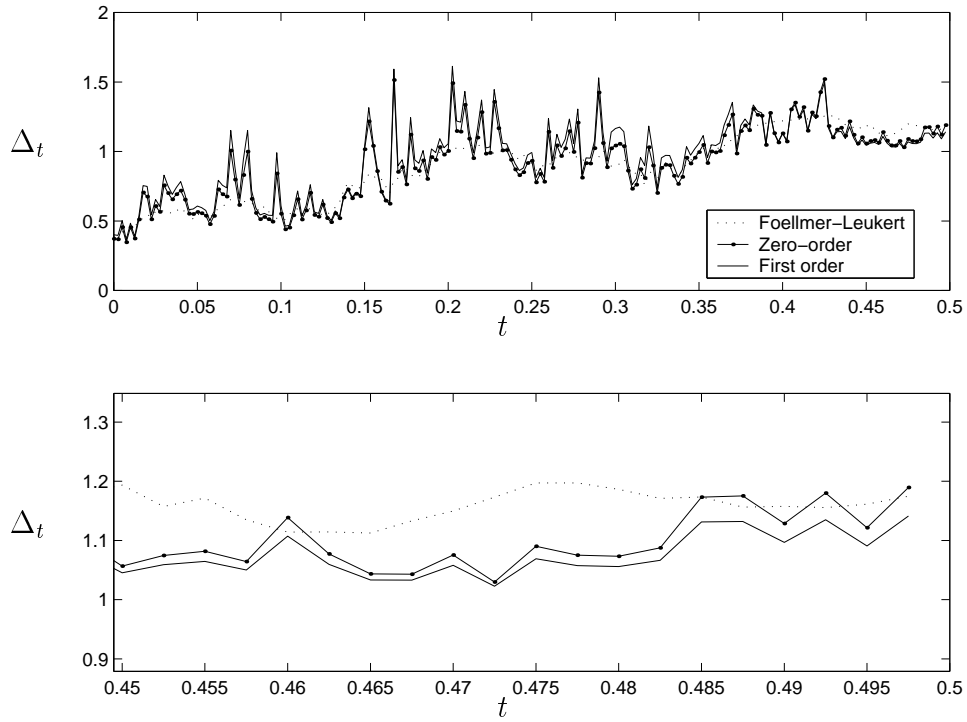


FIGURE 4. Strategies along the path shown in Figure 5 measured by the number of stocks in the portfolios corresponding to Figure 3. The bottom graph magnifies how the deltas of the three hedging portfolios differ near expiration.

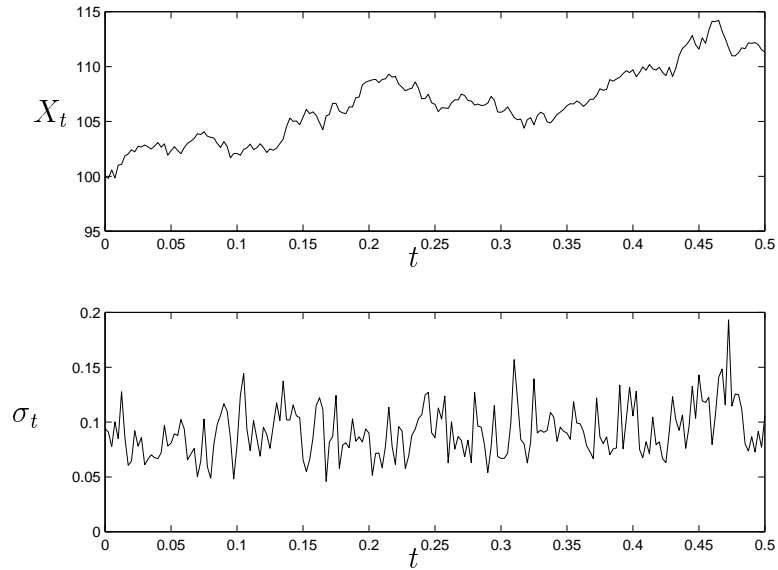


FIGURE 5. Stock price and volatility paths for the simulation whose results are shown in Figures 3 and 4. The starting stock price is at the money of the option being hedged ($X_0 = K = 100$.)

This follows from the analogous calculation of Section 6.2 starting with the PDE (5.7) for \hat{H} . The PDE (7.1) for the approximation is linear, and we have a semi-explicit formula for

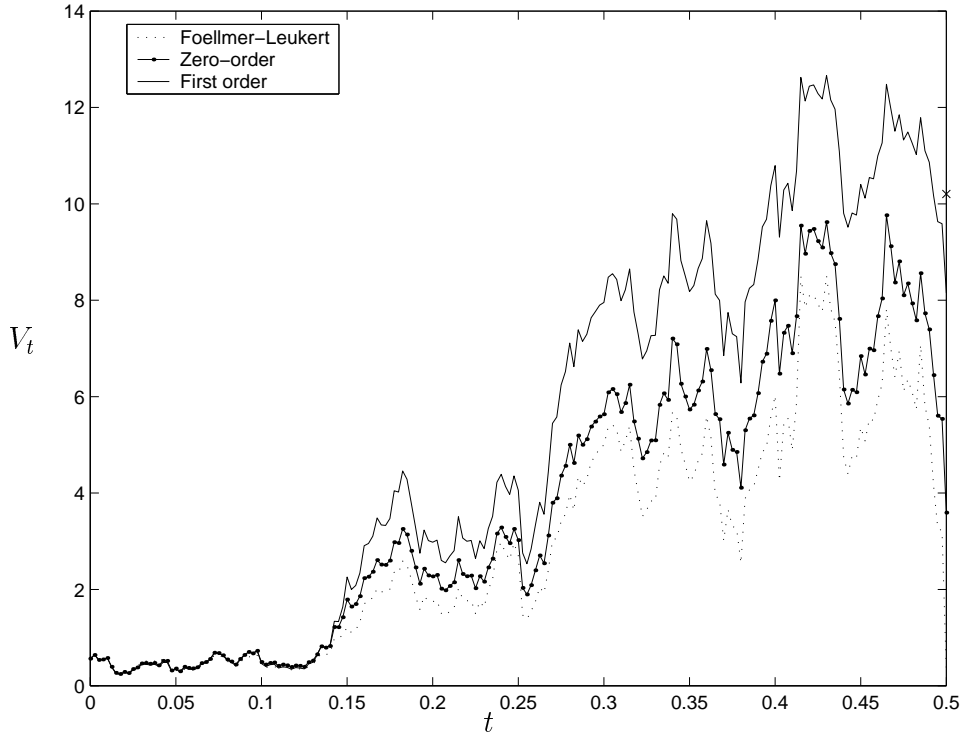


FIGURE 6. *Portfolio values along a path. The cross at the right side of the picture illustrates the value of the option $(X_T - K)^+$ at expiration. The full stock price and volatility paths are shown in Figure 8. The initial capital was 20% of the Black-Scholes price of the claim computed with the volatility $\bar{\sigma}$.*

its solution, analogous to computing $g^{(0)}$ in Section 6.3. From that formula we can easily compute $\hat{H}^{(0)}$ and its derivatives to very high order of accuracy, and that allows us to find the principal approximation $H^{(0)}(t, x, v)$ to the value function $H(t, x, y, v)$ using the formula

$$(7.2) \quad H^{(0)}(t, x, v) = zv + \hat{H}^{(0)}(t, x, z), \quad v = g^{(0)}(t, x, z).$$

Namely, given (t, x) we compute $\hat{H}^{(0)}(t, x, z)$ and $g^{(0)}(t, x, z)$ for a grid on the z -axis. Then we pick z such that $g^{(0)}(t, x, z) = v$, and we use (7.2) with this z to compute $H^{(0)}(t, x, v)$.

For the first-order strategy, we instead use

$$(7.3) \quad (H^{(0)} + \widetilde{H}^{(1)})(t, x, v) = zv + (\hat{H}^{(0)} + \widetilde{H}^{(1)})(t, x, z), \quad v = (g^{(0)} + \widetilde{g}^{(1)})(t, x, z).$$

Notice that value of z corresponding to a given triple (t, x, v) will depend on whether we use the zero- or first-order strategy. If we would use the full strategy (no asymptotics), then we would get another value of z , depending on the quadruple (t, x, y, v) , but all these values are close for small ε .

In the computations illustrated below we used $\bar{\sigma} = 0.1$, $\sigma_* = 0.0874$, $\mu = 0.2$, $K = 100$, $T - t = 0.5$. We also converted the two utilities to *losses*, using the formula

$$\text{loss} = \sqrt{\max \text{ utility} - \text{utility}}.$$

The three graphs show the two losses (Figures 9 and 10) and their difference (Figure 11) as a function of the initial stock price x and the fraction ζ of the initial capital to the Black-Scholes price of the call using constant volatility $\bar{\sigma}$, i.e. $\zeta = v/C_{BS}(t, x; \bar{\sigma}, K)$. We

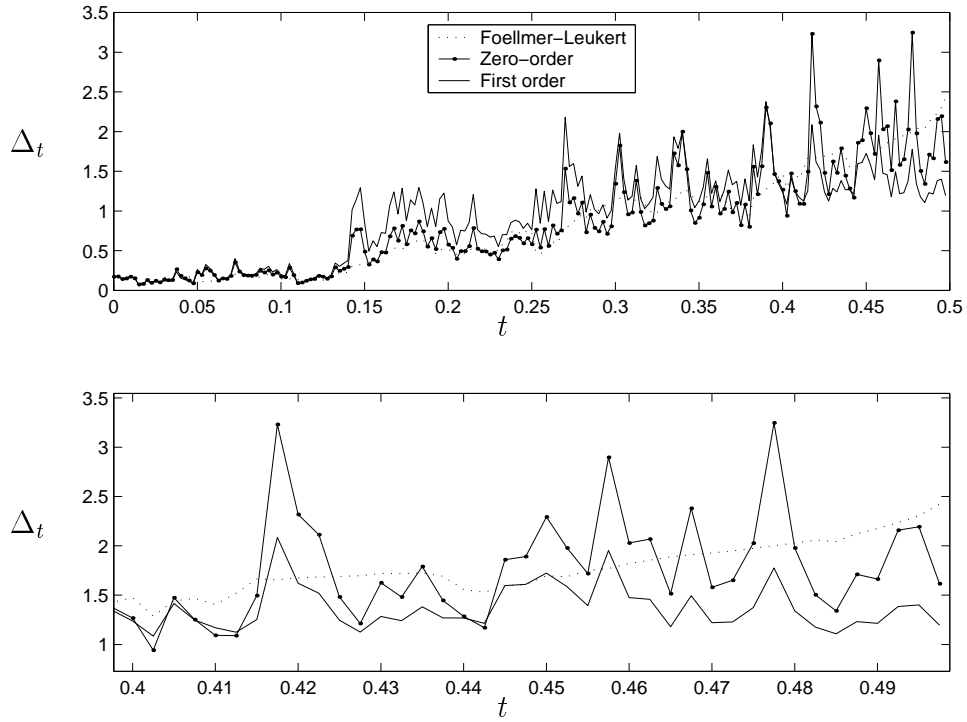


FIGURE 7. Strategies along the path shown in Figure 8 measured by the number of stocks in the portfolios corresponding to Figure 6. The bottom graph magnifies how the deltas of the three hedging portfolios differ near expiration.

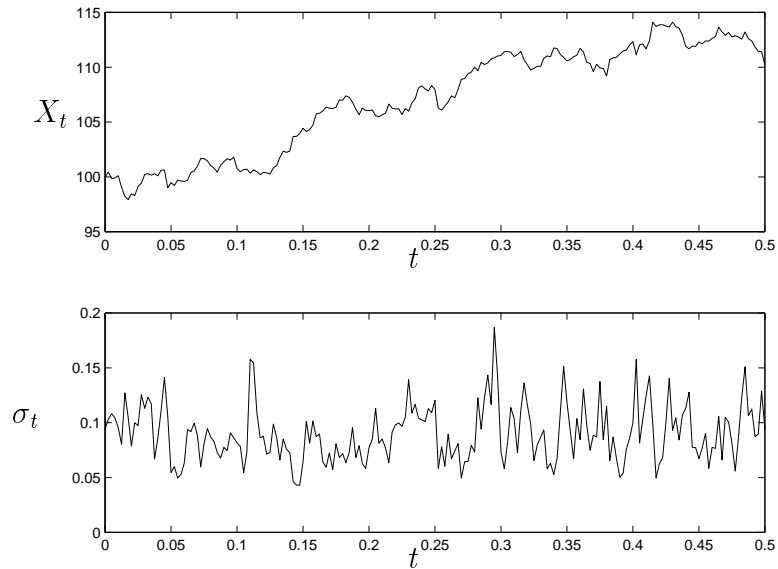


FIGURE 8. Stock price and volatility paths for the simulation whose results are shown in Figures 6 and 7. The starting stock price is at the money of the option being hedged ($X_0 = K = 100$.)

see that the expected loss is smaller when using the first-order strategy. Also notice that both losses appear close to zero for $\zeta = 1$. This is because $\zeta = 1$ corresponds to $z = \mathcal{O}(\varepsilon)$ and from the asymptotic formulas the loss is $\mathcal{O}(\varepsilon)$ in this case. As a consequence, if the

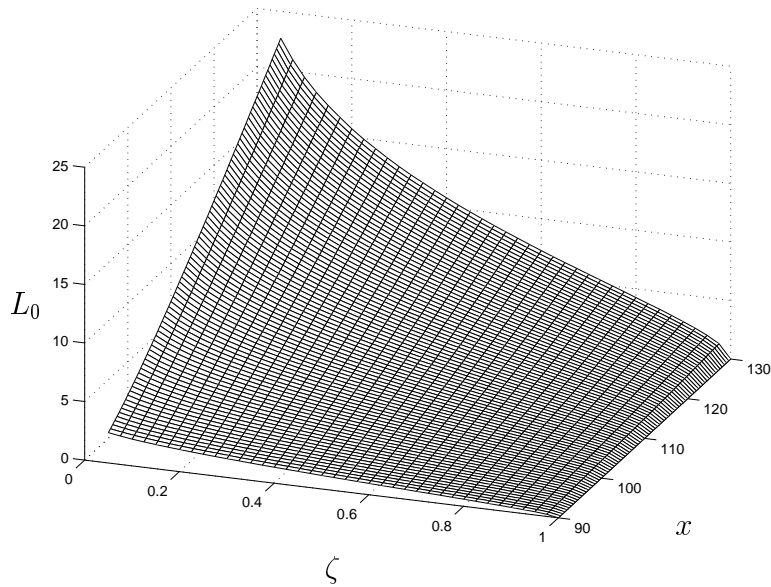


FIGURE 9. *Zero-order loss.*

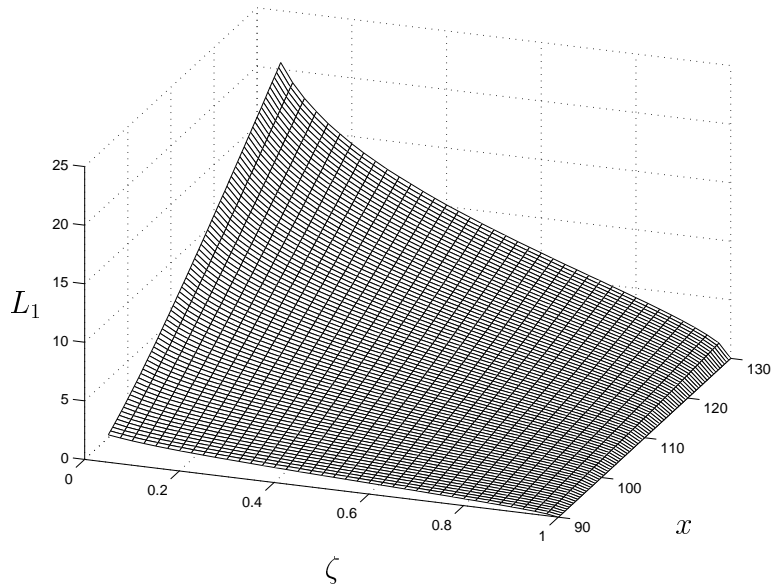


FIGURE 10. *First-order loss.*

volatility is fast mean-reverting, then given initial capital $v = C_{BS}(t, x; \bar{\sigma}, K)$ we have a partial hedging strategy that gives rise to very small losses on average. This should be contrasted with the fact that we may need much more capital to *superhedge* the call option, namely $C_{BS}(t, x; \sigma_{\max}, K)$, where σ_{\max} is the largest possible value of the volatility.

8. CONCLUSIONS

In this article, we have discussed the problem of hedging derivative risk with the underlying security in an incomplete stochastic volatility market when the investor pays a reasonable insurance premium and wants to minimize a measure of his or her expected shortfall. It

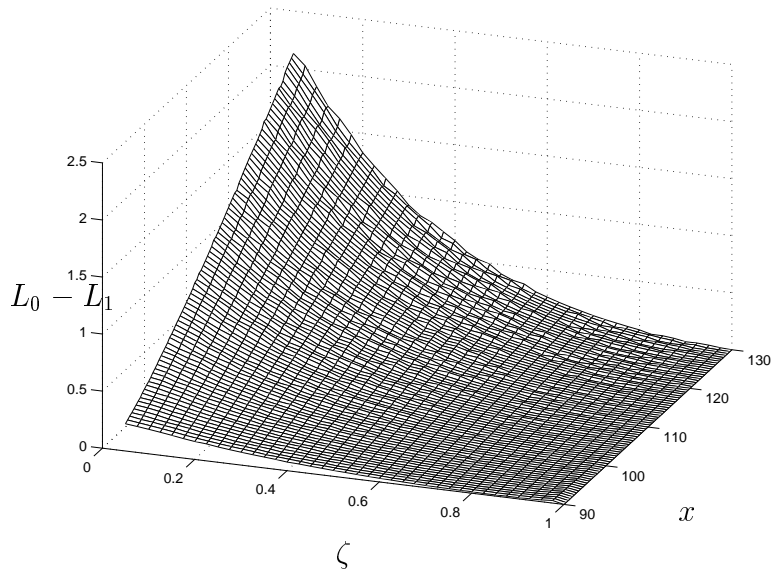


FIGURE 11. *Difference between the two losses*

follows from convex duality that the optimal strategy is to perfectly hedge a cheaper claim whose payoff depends on the terminal values of the stock price process *and* one of the change-of-measure processes (Z_t^γ). The analysis reveals that the “correct” change-of-measure process is chosen so that the new claim can indeed be perfectly hedged. However, optimal strategies depend on solution of a high-dimensional nonlinear Bellman equation and in general will be sensitive to volatility model specification.

To compute robust approximations to the optimal partial hedging strategy, we exploit volatility clustering. An asymptotic analysis shows

- The long-run mean volatility $\bar{\sigma}$ and the harmonically-averaged volatility σ_\star are important statistics for this problem, independent of a specific volatility model.
- If more model details are specified, the approximation can be improved. The difficulty of estimating the correlation ρ and the rate of mean-reversion α of the volatility-driving process is removed by using information from the slope of the implied volatility skew.

In future work, we will investigate estimators for the unusual volatility σ_\star and the sensitivity of the strategies given here to model misspecification. Another interesting problem we have not considered is the effect of filtering the volatility level $f(Y_t)$ from data.

The asymptotic method has many applications as a computational tool to related stochastic control problems in finance. It can also be thought of as a device for *robust control* in which stochastic volatility models uncertainty about what volatility the hedger will face. The zeroth-order strategy is then a robust modification of the constant volatility optimal strategy.

APPENDIX A. CALCULATIONS

A.1. Constant volatility.

A.1.1. Convex duals.

$$\begin{aligned}
 H(t, x, v) &= v\hat{g}(t, x, v) + \hat{H}(t, x, \hat{g}(t, x, v)) & H_v(t, x, g(t, x, z)) &= z \\
 \hat{H}(t, x, z) &= -zg(t, x, z) + H(t, x, g(t, x, z)) & \hat{H}_z(t, x, \hat{g}(t, x, v)) &= -v \\
 & & H_v(t, x, v) &= \hat{g}(t, x, v) \\
 & & \hat{H}_z(t, x, z) &= -g(t, x, z)
 \end{aligned}$$

A.1.2. Equations for \hat{H} .

$$\begin{aligned}
 H_v &= z & H_{vv} &= -\frac{1}{\hat{H}_{zz}} \\
 H_t &= \hat{H}_t & H_{xv} &= -\frac{\hat{H}_{xz}}{\hat{H}_{zz}} \\
 H_x &= \hat{H}_x & H_{xx} &= \frac{\hat{H}_{zz}\hat{H}_{xx} - \hat{H}_{xz}^2}{\hat{H}_{zz}}
 \end{aligned}$$

This implies:

$$\begin{aligned}
 H_t + \mathcal{L}_x H - \frac{(\mu H_v + \sigma^2 x H_{xv})^2}{2\sigma^2 H_{vv}} &= \hat{H}_t + \mathcal{L}_x \hat{H} - \frac{1}{2}\sigma^2 x^2 \frac{\hat{H}_{xz}^2}{\hat{H}_{zz}} - \frac{(\mu z - \sigma^2 x \frac{\hat{H}_{xz}}{\hat{H}_{zz}})^2}{-2\sigma^2 / \hat{H}_{zz}} \\
 &= \hat{H}_t + \mathcal{L}_x \hat{H} + \frac{1}{2}\frac{\mu^2}{\sigma^2} z^2 \hat{H}_{zz} - \mu x z \hat{H}_{xz} \\
 &= 0.
 \end{aligned}$$

A.1.3. *Equations for g .* The PDE for g is most easily obtained by using $g = -\hat{H}_z$. Namely, by differentiating (3.7) with respect to z we get

$$\begin{aligned}
 0 &= \hat{H}_{tz} + \mathcal{L}_x \hat{H}_z + \frac{1}{2}\frac{\mu^2}{\sigma^2} \frac{\partial}{\partial z} \left(z^2 \hat{H}_{zz} - \mu x z \hat{H}_{xz} \right) \\
 &= - \left(g_t + \mathcal{L}_x g + \frac{1}{2}\frac{\mu^2}{\sigma^2} \frac{\partial}{\partial z} \left(z^2 g_z - \mu x z g_x \right) \right),
 \end{aligned}$$

which leads to (3.8).

A.1.4. *Optimal controls.* We have

$$\begin{aligned}
 \pi^* &= -\frac{(\mu H_v + \sigma^2 x H_{xv})}{\sigma^2 H_{vv}} \\
 &= \frac{\mu}{\sigma^2} z \hat{H}_{zz} - x \hat{H}_{xz} \\
 &= x g_x - \frac{\mu}{\sigma^2} z g_z.
 \end{aligned}$$

A.2. Stochastic volatility.

A.2.1. *Convex duals.*

$$\begin{aligned}
H(t, x, y, v) &= v\hat{g}(t, x, y, v) + \hat{H}(t, x, y, \hat{g}(t, x, y, v)) & H_v(t, x, y, g(t, x, z)) &= z \\
\hat{H}(t, x, y, z) &= -zg(t, x, y, z) + H(t, x, y, g(t, x, y, z)) & \hat{H}_z(t, x, y, \hat{g}(t, x, y, v)) &= -v \\
& & H_v(t, x, y, v) &= \hat{g}(t, x, y, v) \\
& & \hat{H}_z(t, x, y, z) &= -g(t, x, y, z)
\end{aligned}$$

A.2.2. *Equations for \hat{H} .*

$$\begin{aligned}
H_v &= z & H_{vv} &= -\frac{1}{\hat{H}_{zz}} & H_{xx} &= \frac{\hat{H}_{zz}\hat{H}_{xx} - \hat{H}_{xz}^2}{\hat{H}_{zz}} \\
H_t &= \hat{H}_t & H_{xv} &= -\frac{\hat{H}_{xz}}{\hat{H}_{zz}} & H_{yy} &= \frac{\hat{H}_{zz}\hat{H}_{yy} - \hat{H}_{yz}^2}{\hat{H}_{zz}} \\
H_x &= \hat{H}_x & H_{yv} &= -\frac{\hat{H}_{yz}}{\hat{H}_{zz}} & H_{xy} &= \frac{\hat{H}_{zz}\hat{H}_{xy} - \hat{H}_{xz}\hat{H}_{yz}}{\hat{H}_{zz}} \\
H_y &= \hat{H}_y & & & &
\end{aligned}$$

This implies:

$$\begin{aligned}
& H_t + \mathcal{L}_{x,y}H - \frac{(\mu H_v + f^2 x H_{xv} + \rho\beta f H_{yv})^2}{2f^2 H_{vv}} \\
&= \hat{H}_t + \mathcal{L}_{x,y}\hat{H} - \frac{1}{2}f^2 x^2 \frac{\hat{H}_{xz}^2}{\hat{H}_{zz}} - \rho\beta f x \frac{\hat{H}_{xz}\hat{H}_{yz}}{\hat{H}_{zz}} - \frac{1}{2}\beta^2 \frac{\hat{H}_{yz}^2}{\hat{H}_{zz}} - \frac{(\mu z - f^2 x \frac{\hat{H}_{xz}}{\hat{H}_{zz}} - \rho\beta f \frac{\hat{H}_{yz}}{\hat{H}_{zz}})^2}{-2f^2/\hat{H}_{zz}} \\
&= \hat{H}_t + \mathcal{L}_{x,y}\hat{H} + \frac{1}{2}\frac{\mu^2}{f^2}z^2 \hat{H}_{zz} - \mu x z \hat{H}_{xz} - \frac{\rho\beta\mu}{f}z \hat{H}_{yz} - \frac{1}{2}(1 - \rho^2)\beta^2 \frac{\hat{H}_{yz}^2}{\hat{H}_{zz}} = 0.
\end{aligned}$$

A.2.3. *Equations for g .* As in the case of constant volatility, the PDE for g is most easily obtained by using $g = -\hat{H}_z$. Namely, by differentiating (5.7) with respect to z we get

$$\begin{aligned}
0 &= \hat{H}_{tz} + \mathcal{L}_{x,y}\hat{H}_z + \frac{1}{2}\frac{\mu^2}{f^2}\frac{\partial}{\partial z}\left(z^2 \hat{H}_{zz} - \mu x z \hat{H}_{xz} - \frac{\rho\beta\mu}{f}\hat{H}_{yz} - \frac{1}{2}(1 - \rho^2)\beta^2 \frac{\hat{H}_{yz}^2}{\hat{H}_{zz}}\right) \\
&= -\left(g_t + \mathcal{L}_{x,y}g + \frac{1}{2}\frac{\mu^2}{f^2}\frac{\partial}{\partial z}\left(z^2 g_z - \mu x z g_x - \frac{\rho\beta\mu}{f}g_y - \frac{1}{2}(1 - \rho^2)\beta^2 \frac{g_y^2}{g_z}\right)\right),
\end{aligned}$$

which leads to (5.8).

A.2.4. *Optimal controls.* We have

$$\begin{aligned}
\gamma^* &= -\beta\sqrt{1 - \rho^2}\frac{\hat{H}_{yz}}{\hat{H}_{zz}} \\
&= \beta\sqrt{1 - \rho^2}H_{yv} \\
&= -\beta\sqrt{1 - \rho^2}\frac{g_y}{g_z}
\end{aligned}$$

and

$$\begin{aligned}
\pi^* &= -\frac{(\mu H_v + f^2 x H_{xv} + \rho\beta f H_{yv})}{f^2 H_{vv}} \\
&= \frac{\mu}{f^2} z \hat{H}_{zz} - x \hat{H}_{xz} - \frac{\rho\beta}{f} \hat{H}_{yz} \\
&= x g_x + \frac{\rho\beta}{f} g_y - \frac{\mu}{f^2} z g_z.
\end{aligned}$$

APPENDIX B. EQUATION FOR THE FIRST-ORDER TERM

We present the argument for the equation (6.13) for the first-order term $g^{(1)}$ in the approximation to the dual function g . It is the same mathematical argument seen in pricing problems (Fouque et al., 2000a), but the differential operators \mathcal{L}_1 and \mathcal{L}_2 are different in the present context.

The key features are

- \mathcal{L}_0 is the generator of an ergodic Markov process with a unique invariant distribution;
- \mathcal{L}_1 takes derivatives in y and so kills any function that does not depend on y ;
- The zeroth-order approximation $g^{(0)}(t, x, z)$ and the first-order term $g^{(1)}(t, x, z)$ do not depend on y .

Comparing terms of $\mathcal{O}(\sqrt{\varepsilon})$ in the PDE for g , we find

$$(B.1) \quad \mathcal{L}_0 g^{(3)} = -(\mathcal{L}_1 g^{(2)} + \mathcal{L}_2 g^{(1)}),$$

which we look at as a Poisson equation for $g^{(3)}(t, x, y, z)$. Just as the Fredholm solvability condition for $g^{(2)}$ determined the equation for $g^{(0)}$, the solvability for (B.1) will give us the equation for $g^{(1)}(t, x, v)$. Substituting for $g^{(2)}(t, x, y, z)$ with

$$(B.2) \quad g^{(2)} = -\mathcal{L}_0^{-1} (\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle) g^{(0)} + c(t, x, z),$$

for some function c not depending on y , this condition is

$$\langle \mathcal{L}_2 g^{(1)} - \mathcal{L}_1 \mathcal{L}_0^{-1} (\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle) g^{(0)} \rangle = 0,$$

where

$$\langle \mathcal{L}_2 g^{(1)} \rangle = \langle \mathcal{L}_2 \rangle g^{(1)}$$

since $g^{(1)}$ does not depend on y . Hence we obtain equation (6.13).

To obtain the expression for the optimal strategy π^* up to order $\sqrt{\varepsilon}$, we insert the expansion for g into (5.10), which gives

$$\pi^* = \left(x \frac{\partial}{\partial x} - \frac{\mu}{f(y)^2} z \frac{\partial}{\partial z} \right) (g^{(0)} + \widetilde{g^{(1)}}) + \frac{\sqrt{2} \rho \nu}{\sqrt{\alpha} f(y)} g_y^{(2)},$$

omitting terms of higher order. From (B.2) and (6.14), we compute

$$g_y^{(2)} = \frac{1}{2} \phi'(y) x^2 g_{xx}^{(0)} + \mu^2 \psi'(y) \left(\frac{1}{2} z^2 g_{zz}^{(0)} + z g_z^{(0)} \right),$$

which leads to (6.19).

REFERENCES

- Alizadeh, S., Brandt, M., and Diebold, F. (2001). Range-based estimation of stochastic volatility models. *Journal of Finance*. To appear.
- Andersen, T. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *J. Empirical Finance*, 4:115–158.
- Avellaneda, M., Levy, A., and Paras, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 1(2):73–88.
- Bakshi, G., Cao, C., and Chen, Z. (1997). Empirical performance of alternative option pricing models. *J. Fin.*, 52(5):2003–2049.
- Broadie, M., Cvitanić, J., and Soner, H. M. (1998). Optimal replication of contingent claims under portfolio constraints. *Review of Financial Studies*, 11:59–79.
- Chernov, M., Gallant, R., Ghysels, E., and Tauchen, G. (2001). Alternative models for stock price dynamics. Preprint, Columbia University.
- Cvitanić, J. (2000). Minimizing expected loss of hedging in incomplete and constrained markets. *SIAM J. Control & Optim.*, 38(4):1050–1066.
- Cvitanić, J. and Karatzas, I. (1999). On dynamic measures of risk. *Finance and Stochastics*, 3(4):451–482.
- Kramkov, D. and Schachermayer, W. (1999). The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annals of Applied Probability*, 9(3):904–950.
- Duffie, D., Pan, J., and Singleton, K. (2000). Transform analysis and option pricing for affine jump-diffusions. *Econometrica*, 68:1343–77.
- Engle, R. and Patton, A. (2001). What good is a volatility model? *Quantitative Finance*, 1:237–45.
- Föllmer, H. and Leukert, P. (2000). Efficient hedging: Cost versus shortfall risk. *Finance and Stochastics*, 4(2):117–146.
- Fleming, W. and Soner, H. M. (1993). *Controlled Markov Processes*. Springer-Verlag.
- Fouque, J.-P., Papanicolaou, G., and Sircar, K. (2000a). *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press.
- Fouque, J.-P., Papanicolaou, G., and Sircar, K. (2000b). Mean-Reverting Stochastic Volatility. *Intl. J. Theor. Appl. Fin.*, 3(1):101–42.
- Fouque, J.-P., Papanicolaou, G., Sircar, K., and Solna, K. (2000c). Mean-Reversion of S&P500 volatility. *Preprint*.
- Frey, R. (1996). Derivative asset analysis in models with level-dependent and stochastic volatility. *CWI Quarterly*, 10(1):1–34.
- Karatzas, I. and Shreve, S. (1998). *Methods of Mathematical Finance*. Springer-Verlag.
- Karoui, N. E. and Quenez, M. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control and Optimization*, 33(1):29–66.
- Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: the continuous-time case. *Rev. Econom. Statist.*, 51:247–257.
- Sircar, K. (2000). Hedging under stochastic volatility. In Avellaneda, M., editor, *Quantitative Analysis in Financial Markets : Collected Papers of the New York University Mathematical Finance Seminar*, volume 2. World Scientific Publishing. To appear.
- Sircar, K. and Papanicolaou, G. (1999). Stochastic volatility, smile and asymptotics. *Applied Mathematical Finance*, 6(2):107–145.

Zariphopoulou, T. (1999). Optimal investment and consumption models with nonlinear stock dynamics. *Math. Meth. Oper. Res.*, 50:271–296.