Risk Minimization in Stochastic Volatility Models: Model Risk and Empirical Performance

Rolf Poulsen†
Klaus Reiner Schenk-Hoppé‡
Christian-Oliver Ewald§

September 21, 2007

Abstract

In this paper the performance of locally risk-minimizing hedge strategies for European options in stochastic volatility models is studied from an experimental as well as from an empirical perspective. These hedge strategies are derived for a large class of diffusion-type stochastic volatility models, and they are as easy to implement as usual delta hedges. Our simulation results on model risk show that the locally risk-minimizing hedges are robust with respect to uncertainty and even misconceptions about the underlying data generating process. The empirical study indicates that locally risk-minimizing hedge strategies consistently produce lower standard deviations of profit-and-loss-ratios than delta hedges (over different time periods as well as in different markets). The more skewed the market and the more out-of-the-money the option, the higher the benefit.

Key words: Locally risk-minimizing hedge, delta hedge, stochastic volatility, model risk, empirical hedge performance.

JEL classification: C90, G13
1 Introduction

In this paper we study hedging in stochastic volatility models. These models resolve some of the problems of the standard Black-Scholes model, Derman-Dupire-type local volatility models\(^1\) and jump/Levy models\(^2\).

Stochastic volatility models are incomplete; so the seller of a contingent claim cannot eliminate all risk by trading the primary assets. The main objective is therefore to deal efficiently with this risk. Schweizer (1991) proposes the concept of locally risk-minimizing hedges which aims at minimizing the variance of the cost process of non-self-financing hedges. We derive an explicit formula for the locally risk-minimizing hedge in a general class of stochastic volatility models in a rigorous and novel fashion using results from El Karoui, Peng & Quenez (1997). The formula shows that the hedge can be decomposed as a sum of the delta hedge and a volatility/correlation-risk term. We also give an intuitive derivation illustrating the mechanics behind local risk-minimization; and we explain why this proof has a “loose end.” Applying some form of risk-minimization to stochastic volatility models is natural, so the formula appears in various guises in the previous literature ranging from Frey (1997, Prop. 6.5) and Bakshi, Cao & Chen (1997, Eq. (21)) over Ahn & Wilmott (2003, Eq. (4)) and Bouchaud & Potters (2003, Chapter 15) to Alexander & Nogueira (2007a, Eq. (9)). Still the volatility/correlation-risk term is often ignored.

The locally risk-minimizing hedge is model- and parameter-dependent. The first issue analyzed in this paper is “But by how much, really?” To this end we conduct a number of controlled experiments that test the Heston model’s locally risk-minimizing hedge against various types of model risk. We study parameter uncertainty (small effect), misspecification of the functional form of the stochastic volatility (small effect if the models are calibrated to same data), and “plain ignorance” (using vega from a constant volatility model hurts hedge performance markedly).

The empirical performance of locally risk-minimizing hedges is tested using U.S. and European stock indices and currency option markets over a period of one-and-a-half years. We find that in the stock markets (characterized by a strong negative correlation between returns on the underlying and changes in volatility, which manifests itself as a skew in implied volatilities across strikes) there is a significant improvement in hedge performance when using locally risk-minimizing hedges based on a stochastic volatility model. Compared to usual delta hedging, the standard deviation of daily profit-and-loss ratios reduced by between one-third and one-quarter for liquid options. The currency market displays a higher degree of symmetry (close to zero correlation

---

\(^1\)Matching observed option prices by functional dependence of local volatility solely on the level of the underlying and calendar time causes unreasonable time-series behavior. A further discussion can be found in Ayache, Henrotte, Nassar & Wang (2004).

\(^2\)For instance, at-the-money implied volatilities are constant through time in Merton’s jump diffusion model. This is strongly at odds with empirical behavior—as is e.g. supported by Figure 4 in this paper.
between underlying and volatility changes; smile rather than skew in implied volatilities) and there is essentially no gain from using the locally risk-minimizing strategy—as also suggested by the formula for the locally risk-minimizing hedge. However, in this market there is no loss of out-of-sample performance either. Results are robust to the use of Heston or so-called SABR as stochastic volatility model.

The paper is organized as follows. Section 2 briefly reviews local risk-minimization and introduces a class of stochastic volatility models. Section 3 studies model risk. The empirical performance is analyzed in Section 4. Section 5 concludes and suggests some ideas for future research.

2 Local Risk-Minimization and Stochastic Volatility

Hedge strategies in incomplete markets generally are not replicating self-financing portfolios. Depending on the particular restrictions on trading in the risk-free asset, hedges either do not replicate the payoff perfectly or are not self-financing. In the latter case a hedge is associated with a cost process that aggregates any additional investments. Traders describe this as “bleeding.” Consider a market with a risk-free and a risky asset with prices \( B(t) \) resp. \( S(t) \). Then the cost process associated to a trading strategy \( \varphi(t) = (\varphi^0(t), \varphi^1(t)) \) (where the components are the holdings—in number of units—of the risk-free and the risky asset) is given by

\[
\text{Cost}_\varphi(t) = V_{\varphi}(t) - \int_0^t \varphi^0(s) dB(s) - \int_0^t \varphi^1(s) dS(s)
\]

where \( V_{\varphi}(t) = \varphi^0(t) B(t) + \varphi^1(t) S(t) \) denotes the value of the trading strategy at time \( t \).

The total cost of the hedge becomes uncertain; something that the seller might not appreciate and, therefore, seeks to reduce. One criterion, proposed by Föllmer & Schweizer (1990), is minimization of the conditional variance process of the cost process which is defined as

\[
\mathbb{E}((\text{Cost}_\varphi(T) - \text{Cost}_\varphi(t))^2 | \mathcal{F}_t).
\]


4A contingent claim \( H \) can always be hedged with a trading strategy that is not self-financing. For instance let \( \varphi^0(s) = \varphi^1(s) = 0 \) for all \( s \in [0, T) \), \( \varphi^0(T) = H \) and \( \varphi^1(T) = 0 \).
However, this dynamic optimization problem may have no solution, so Schweizer (1991) introduces a localized of the risk-minimization criterion that overcomes this existence problem. The solution is called a *locally risk-minimizing hedge*.

**A general class of stochastic volatility models.** In the remainder of this section we consider the class of stochastic volatility models of the form

\[
\begin{align*}
\frac{dS(t)}{S(t)} &= \mu dt + S(t)\gamma f(V(t)) \left[ \sqrt{1 - \rho^2} dW^1(t) + \rho dW^2(t) \right] \\
\frac{dV(t)}{V(t)} &= \beta(V(t)) dt + g(V(t)) dW^2(t)
\end{align*}
\]

(2)

with independent standard Brownian motions \( W^1(\cdot) \) and \( W^2(\cdot) \) on a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \). \( S(t) > 0 \) denotes the price of the (traded) asset and \( V(t) > 0 \) is the (non-traded) stochastic local return variance. The risk-free asset \( B \) pays the constant interest rate \( r \). These models allow for level (also known as scale) dependence (\( \gamma \neq 0 \)) and correlation between returns and variance (\( \rho \neq 0 \)). This class of stochastic volatility models contains most of those (without jump component) that are commonly used in research as well as in practice. Table 1 provides an overview.

**A three step procedure.** El Karoui et al. (1997, Proposition 1.1) introduces a technique to determine the locally risk-minimizing hedge. In the following we describe in detail a three step procedure to implement their technique for the class of models \( 5 \). The steps are: complete the market, compute the hedging strategy in the completed market, and finally project onto the original market. The method works for general contingent claims in possibly non-Markovian models, path dependent options in particular. But, of course, the practical usefulness of the method hinges on finding efficient ways of calculating prices and hedge portfolios in the completed market. For more on risk-minimizing hedges for exotic options we refer the reader to Alexander & Nogueira (2007b). For the stochastic volatility model, the steps work out like this.

**Complete.** The risk premium (process) corresponding to the minimal martingale measure is

\[
\lambda^{\text{min}}(t) = \frac{\mu - r}{S(t)^{\gamma f(V(t))}} \cdot \left( \frac{\sqrt{1 - \rho^2}}{\rho} \right).
\]

The stochastic volatility model can be completed by introducing a second, volatility-dependent traded asset; El Karoui et al. (1997, Proposition 1.1) tells us to do that in

---

\( 5 \)El Karoui et al. (1997) assume boundedness of the volatility matrix and its inverse. No commonly used stochastic volatility model satisfies this condition. However, their Proposition 1.1 can still be used as long as the Doleans-Dade exponential of the risk premium process corresponding to the minimal martingale measure is a true martingale. This property must be verified on a case-by-case basis. For the Heston model, which is employed in the empirical study, Theorem 1 in Cheridito, Filipovic & Kimmel (2007) shows that the martingale property holds if the Feller condition is satisfied under both the original measure and the minimal martingale measure.
such a way that the risk premium in the completed market is exactly $\lambda_{\text{min}}$. This is achieved by introducing as traded asset $\hat{S}$ with dynamics

$$d\hat{S}(t)/\hat{S}(t) = \left( r + \rho \frac{g(V(t))}{S(t)^\gamma f(V(t))} (\mu - r) \right) dt + g(V(t))dW^2_t.$$ 

Other choices of completing asset are possible (for instance a specific option), but there is not that much flexibility because of the constraint on market price of risk. With this choice, $\hat{S}(t)$ can be interpreted as the price of a pure volatility derivative.

**Compute.** In the completed market the hedge strategy for a European contingent claim with payoff $H(S(T))$ is found in the same way as in the Black-Scholes model. Let $X(t)$ denote the value process of a (perfect) hedge for $H$ with positions $\Delta(t)$ and $\hat{\Delta}(t)$ in the assets $S$ and $\hat{S}$, respectively. The position in the risk-free asset account is adjusted

<table>
<thead>
<tr>
<th>Author(s) &amp; year</th>
<th>Specification</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hull/White 1987</td>
<td>$f(v) = v$, $\beta(v) = 0$, $g(v) = \sigma$, $\rho = 0$, $\gamma = 0$</td>
<td>Instantaneous variance: Geometric Brownian motion. Options priced by mixing.</td>
</tr>
<tr>
<td>Wiggins 1987</td>
<td>$f(v) = e^{v/2}$, $\beta(v) = 0$, $g(v) = \sigma$, $\rho = 0$, $\gamma = 0$</td>
<td>Instantaneous volatility: Ornstein-Uhlenbeck in logarithms.</td>
</tr>
<tr>
<td>Stein/Stein 1991</td>
<td>$f(v) =</td>
<td>v</td>
</tr>
<tr>
<td>Heston 1993</td>
<td>$f(v) = \sqrt{v}$, $\beta(v) = \kappa(\theta - v)/v$, $g(v) = \sigma/\sqrt{v}$, $\rho \in [-1, 1]$, $\gamma = 0$</td>
<td>Instantaneous variance: CIR process. First model with correlation. Options priced by inversion of characteristic function.</td>
</tr>
<tr>
<td>Romano/Touzi 1997</td>
<td>$f(v) = \sqrt{v}$, $\beta$ and $g$ are free, $\rho \in [-1, 1]$, $\gamma = 0$</td>
<td>Extension of mixing to correlation.</td>
</tr>
<tr>
<td>Schöbel/Zhu 1999</td>
<td>$f(v) =</td>
<td>v</td>
</tr>
<tr>
<td>Hagan et al. 2002</td>
<td>$f(v) = v$, $\beta(v) = 0$, $g(v) = \sigma$, $\rho \in [-1, 1]$, $\gamma \in [-1, 0]$</td>
<td>Level dependence in returns. Options priced by perturbation technique. Acronym'ed as “SABR”</td>
</tr>
</tbody>
</table>

Table 1: Specification of stochastic volatility models for Eq. (2).
so as to make the hedge self-financing. Then, on the one hand,
\[ dX(t) = \Delta(t)dS(t) + \hat{\Delta}(t)d\hat{S}(t) + r \left[ X(t) - \Delta(t)S(t) - \hat{\Delta}(t)\hat{S}(t) \right] dt \]
\[ = (\ldots) dt + \Delta(t)S(t)^{1+\gamma}f(V(t)) \left[ \sqrt{1 - \rho^2}dW^1(t) + \rho dW^2(t) \right] \]
\[ + \hat{\Delta}(t)\hat{S}(t)g(V(t))dW^2(t). \]

On the other hand, the original market model is Markovian, which allows to write the price of the claim \( H \) as
\[ e^{-r(T-t)} \mathbb{E}^\text{min}_t(H(S(T))) = C(t, S(t), V(t)) \]
for some function \( C(\cdot, \cdot, \cdot) \). The conditional expected value \( \mathbb{E}^\text{min}_t \) is calculated under the minimal martingale measure. Since \( X(t) = C(t, S(t), V(t)) \), application of the Itô formula to find \( dC \) and comparison of the resulting diffusion terms with \( dX \) yields
\[ \Delta(t) = C_S \text{ and } \hat{\Delta}(t) = \frac{C_V}{S(t)} V(t) \]
where \( C_S \) and \( C_V \) denote partial derivatives.

**Project.** Finally, El Karoui et al. (1997, Proposition 1.1) says that the investment in asset \( S \) under the locally risk-minimizing hedge of the original market (2) is given by
\[ S(t)\varphi^1_{\text{min}}(t) = C_S S(t) + \rho \frac{V(t) g(V(t))}{S(t)^{1+\gamma} f(V(t))} C_V. \]

These findings can be summarized as follows.

**Proposition 1** Consider the stochastic volatility model (2). The locally risk-minimizing hedge of a European contingent claim with payoff \( H(S(T)) \) holds
\[ \varphi^1_{\text{min}}(t) = C_S + \rho \frac{V(t) g(V(t))}{S(t)^{1+\gamma} f(V(t))} C_V \] (3)
units of the stock, where
\[ C(t, S(t), V(t)) = e^{-r(T-t)} \mathbb{E}^\text{min}_t(H(S(T))). \]
\( \mathbb{E}^\text{min}_t \) denotes the conditional expectation with respect to the minimal martingale measure \( \mathbb{Q}^\text{min} \) under which the dynamic are given by
\[ dS(t)/S(t) = r dt + S(t)^{\gamma} f(V(t)) \left[ \sqrt{1 - \rho^2}dW^\text{min,1}_t + \rho dW^\text{min,2}_t \right] \] (4)
\[ dV(t)/V(t) = [\beta(V(t)) - \rho \frac{g(V(t))}{S(t)^{\gamma} f(V(t))} (\mu - r)] dt + g(V(t))dW^\text{min,2}_t \]
where \( dW^\text{min} = dW + \lambda^\text{min}_t dt \) defines a \( \mathbb{Q}^\text{min} \)-Brownian motion. The investment in the risk-free asset is \( C(t, S(t), V(t)) - \varphi^1_{\text{min}}(t)S(t) \).
If changes in the underlying and instantaneous variance are correlated ($\rho \neq 0$), the locally risk-minimizing hedge and the standard delta hedge do not coincide. Suppose $\rho$ is negative (as is typical in stock markets) and the payoff function $H$ convex. Then $C_V$ is positive (see e.g. Romano & Touzi (1997, Proposition 4.2)), and (3) tells us that a delta hedger invests too heavily in the stock.

The minimal martingale measure is often described loosely as “the one that changes as little as possible.” Proposition 1 highlights that when return and volatility are correlated and there is a equity risk premium ($\mu \neq r$), the minimal martingale measure does not merely change the drift rate of the stock to $r$ while leaving the volatility dynamics unaltered. In the presence of correlation a change in the stock price dynamics (when switching to the minimal martingale measure) entails a change in the volatility dynamics.

A (deceptively) simple derivation of locally risk-minimizing strategies. Intuition for the locally risk-minimizing hedge (3) (e.g. “where is something actually being minimized?”) can be gained by using direct approach from Bakshi et al. (1997, pp. 2033-4). Suppose at some point in time $t$ a trader takes a position that is (a) long one unit of the European contingent claim with payoff $H(S(T))$, which is valued at $C(t, S(t), V(t))$, and (b) short $\Delta$ units of the stock, where $\Delta$ is to be determined. Itô’s formula yields $dC = ... dt + C_SdS + C_VdV$ which implies that the change in value of the hedge over a small time-interval $[t, t + dt]$ (i.e. locally) is given by

$$dX = dC - \Delta dS = ... dt + (C_S - \Delta)dS + C_VdV.$$ 

For the conditional variance the $dt$-term does not matter, and thus

$$\text{var}_t(dX) = (C_S - \Delta)^2 \text{var}_t(dS) + C_V^2 \text{var}_t(dV) + 2(C_S - \Delta)C_V \text{cov}_t(dS, dV)$$

$$= [(C_S - \Delta)^2 S^{2(1+\gamma)} f^2(V) + C_V^2 V^2 g^2(V) + 2(C_S - \Delta)C_V S^{1+\gamma} f(V) V g(V) \rho] dt.$$ 

From the trader’s perspective a sensible choice of $\Delta$ is the one that minimizes this variance. The first-order condition

$$-2(C_S - \Delta^{\text{min}}) S^{2(1+\gamma)} f^2(V) - 2C_V S^{1+\gamma} f(V) V g(V) \rho = 0$$

yields $\Delta^{\text{min}} = \varphi^{\text{1}}_{\text{min}}$ which coincides with the above result.

However, the shortcoming this derivation is its inability to tie down the price $C$ of the contingent claim. Implementation of this hedge requires taking an expectation when calculating the function $C$, but the derivation gives no indication as to which of the many martingale measures to use. It does not help to “close the model” by assuming that agents use risk-minimizing hedge strategies (and do not care about residuals) and setting the price of the claim equal to the price of this particular hedge. There is a Catch-22: the hedge depends on the pricing function which, in turn, depends on the hedge. The approach of El Karoui et al. (1997) does not have this deficiency because it derives the price as well as the hedge by considering trading in primary assets only.
3 Model Risk

In this section we analyze experimentally to what extent the performance of risk-minimizing hedge is sensitive to model risk, i.e., what happens if you get things a little wrong? Is the risk-minimizing hedge robust? As stressed by Cont (2006) this is a highly relevant practical issue. In the study of this question we use the popular Heston (1993) model as the benchmark. Four likely sources of error are considered and their effects are quantified: “Wrong” martingale measure (little effect), parameter uncertainty (detectable effect, but not nearly strong enough to outweigh the benefits), wrong Greeks (considerable negative effect) and wrong data-generating process (surprisingly small effect).

The Heston model has the dynamics (see Table 1)

\[
\begin{align*}
dS(t) &= S(t) \left( \mu dt + \sqrt{V(t)} \left[ \sqrt{1 - \rho^2} dW^1(t) + \rho dW^2(t) \right] \right) \\
dV(t) &= \kappa (\theta - V(t)) dt + \sigma \sqrt{V(t)} dW^2(t).
\end{align*}
\]

Our simulation experiments use the parameter estimates from the comprehensive study in Eraker (2004) which can be seen as consensus. Table 2 summarizes the values (annualized and in non-percentage terms) and the interpretation of these parameters. The table also includes (a) standard errors of the estimated parameters and (b) option-based parameter estimates of the pricing measure used in the market. These latter estimates, which can also be called risk-adjusted, were obtained by Eraker through a joint time-series and cross-sectional estimation of spot and option prices, and their values reflect the empirical fact that the conditional standard deviation of returns (“historical volatility”) is typically lower than the implied volatility of at-the-money options.

Proposition 1 states that the position in the stock of the locally risk-minimizing hedge is given by

\[ \varphi^1_{\min}(t) = C_S + \rho \sigma \frac{C_V}{S(t)}. \]

The option pricing formula \( C \) and the related Greeks (which are as easy—or as hard—to calculate as the option price itself) are implemented using the Lipton-Lewis reformulation of Heston’s original expression to increase computational stability, see Lipton (2002).

We report hedge errors as the standard deviation of the cost process at expiry divided by the initial option value (in percentage terms),

\[ \text{hedge error} = 100 \times \frac{\sqrt{\text{var}(\text{Cost}(T))}}{e^{-rT} \mathbb{E}_\min([S(T) - K]^+)}. \]

This error measurement tracks the hedge all the way out to expiry; it thus related to global variance in the sense of Eq. (1).
Table 2: Benchmark settings for the parameters of the Heston model. The numbers in square brackets are the standard errors of the estimates.

Hedge errors are estimated as follows: (1) simulate paths of stock prices and volatilities, (2) apply the above hedge strategy for a particular option (a forward-at-the-money call unless otherwise said) at a particular frequency (daily unless otherwise said) along each path, (3) record the path-specific terminal cost and (4) compute sample moments from many paths.

For the given parameters, a small simulation study (not reported) shows that using risk-minimizing hedges rather than usual delta hedges reduces hedge errors by a factor of 0.85 - 0.90 for the typical range of liquid options. The longer the time-to-expiry and the more out-of-the-money the option is, the larger the reduction.

Picking different martingale measures. As mentioned above, part of the result in Proposition 1 is that the $C$-function in Eq. (6) is a conditional expected value under the minimal martingale measure. But what if the hedger uses another measure? In the context of the Heston model different martingale measures can be obtained by different choices of $\kappa$ and $\theta$ in the instantaneous variance dynamics of (5). The choice of measure affects the parameters in the $C$-function, and hence the hedger’s positions. Our analysis looks at three different martingale measures: (1) the minimal martingale measure (the

---

$^6$Cheridito et al. (2007) show that this does indeed give absolutely continuous measure changes as long as the Feller conditions hold under both measures, which is true for our parameter choice in the simulation experiments.

$^7$There are numerous stories and analyses as to which martingale measure “the market picks.”
one to use if you want to minimize local variance), (2) the one obtained by just replacing \( \mu \) by \( r \) (a not uncommon misconception of the minimal martingale measure) and (3) the “market measure” as estimated from option data in Eraker (2004) (the parameters are given in Table 2).

<table>
<thead>
<tr>
<th>Martingale measure; ( Q )</th>
<th>Minimal</th>
<th>Misconceived minimal</th>
<th>Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Q )-parameters ( \theta )</td>
<td>0.229(^2)</td>
<td>0.220(^2)</td>
<td>0.289(^2)</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>4.75</td>
<td>4.75</td>
<td>2.75</td>
</tr>
<tr>
<td>Hedge error</td>
<td>19.7</td>
<td>19.7</td>
<td>20.3</td>
</tr>
</tbody>
</table>

Table 3: Hedge errors under different conceptions about the Heston model’s martingale measures. Hedge portfolios are adjusted daily and the target option is a 1-year forward-at-the-money call.

The results of this exercise are summarized in Table 3. We see that the choice of measure has little effect. That is comforting because the minimal martingale measure dynamics depend on \( \mu \) (the expected stock return), which is notoriously hard to estimate. And it means that loose end in the deceptive derivation of the risk-minimizing strategy in the previous section is not of major practical importance. Note also that the risk-minimizing strategy based on local considerations does have stable (or even: close-to-optimal) global behavior.

Parameter estimation risk. Rather than a question of picking the right or wrong measure, the previous analysis can be seen as an investigation of (a very particular form of) parameter uncertainty. This analysis can be extended to the effects of the uncertainty that is entailed in using estimated parameters. Eraker (2004) reports standard errors\(^8\) of his estimates, see Table 2. With these at hand, we quantify the estimation risk by the following experiment. Suppose that the true parameters (\( \mu, \theta, \kappa, \sigma, \rho \)) are given by the estimates in Table 2 but that the hedger (along a path) uses parameters drawn from the (asymptotically normal) distribution of the estimator. Repeat this simulation over many paths (each time drawing a new hedge parameter, but keeping it fixed along the path).

Table 4 compares the performance of the “random-parameter hedger” to that of someone who uses the true parameter. Results are reported in terms of the relative increase in the hedge errors. As one would expect, the hedge quality deteriorates when the true parameter value is not known. Indeed, the more frequently you hedge, the bigger the effect. (With infrequent hedging, the differences “drown.”) The main message from Table 4 is that the adverse effects (in the range of 0-4% in relative terms) of

---

\(^{8}\)Only standard errors, not correlations, are reported. We treat things as independent, which should give conservative estimates of the effects.
Hedge frequency

<table>
<thead>
<tr>
<th>Expiry</th>
<th>Moneyness</th>
<th>Hedge frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>monthly</td>
</tr>
<tr>
<td>3M</td>
<td>At-the-money</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>10% Out-of-the-money</td>
<td>1.1%</td>
</tr>
<tr>
<td>1Y</td>
<td>At-the-money</td>
<td>1.2%</td>
</tr>
<tr>
<td></td>
<td>10% Out-of-the-money</td>
<td>2.8%</td>
</tr>
</tbody>
</table>

Table 4: Effects of Heston-parameter uncertainty on locally risk-minimizing hedges. The table shows the relative increase in the hedge errors when the hedger uses parameters drawn from the distribution of Eraker’s estimator rather than the true parameter.

Parameter estimation risk in a stochastic volatility model are small compared to what is gained from using risk-minimizing hedges (in the range of 10%-15%). This parameter uncertainty analysis applies to a hedger who bases his inference solely on observations of the underlying. In practice, hedgers are likely to calibrate at least some of the parameters to option prices observed in the market. That situation raises different questions and is analyzed in the end of this section and at length in the empirical section.

Using Black-Scholes’ Greeks. The partial derivatives $C_S$ and $C_V$ in Eq. (3) could be computed within a Black-Scholes model. This might indeed be tempting for a trader who generally has these functions readily available; using implied volatilities even adds a flavor of achieving “consistency with market data.” It therefore of interest to study the performance of a hedger who does just that.

---

9Note that $C_V$ is the derivative with respect to the variance. As the usual vega is a derivative with respect to the standard deviation, the chain rule is needed.
Figure 1 shows the hedge errors for different values of the correlation parameter \( \rho \). For \( \rho = 0 \) using Black-Scholes’s Greeks (or Greek, as only the delta enters the formula) does little harm as both hedges show almost identical performance. But, as \( |\rho| \) increases, the quality of the Black-Scholes hedge deteriorates. Indeed it quickly reaches the point where any benefit from risk-minimizing behavior is wiped out. The performance of the locally risk-minimizing hedge based on the correct Heston Greeks in contrast improves as the absolute value of the correlation becomes higher. The results in Figure 1 show the importance of using a genuine stochastic volatility model. One cannot just use sensitivities from a static model to hedge successfully in the dynamic model. We like to think of this as a finance analogy of the Lucas critique.

**SABR as the data-generating process.** What if the Heston model is not the correct stochastic volatility model (after all alternatives abound)? Hagan, Kumar, Lesniewski & Woodward (2002) suggest the SABR-acronym’ed model given by

\[
\begin{align*}
    dS(t) &= a(t)S^b(t) dW_1(t) \\
    da(t) &= \nu a(t) dW_2(t)
\end{align*}
\]

where \( W_1 \) and \( W_2 \) have correlation \( \rho \), and we put \( r = \mu = 0 \) for simplicity. This model is qualitatively different from the Heston model in several ways. Instantaneous variance \( a(t)^2 \) is log-normal (rather than non-central \( \chi^2 \)) and does not mean-revert (which is likely to have effects at long horizons). The main difference, though, is the possibility of models where the return variance depends on the level of the underlying; this happens when \( b \neq 1 \). If \( b = 1 \), the model is called scale-invariant; Alexander & Nogueira (2007b) explain why and investigate consequences. Having \( b < 1 \) creates a negative relation (sometimes referred to as leverage) between returns and their volatility and thus (elements of) a skew in implied volatilities. In fact, as Figure 2 shows, negative correlation (and scale-invariance) can give the same (to the naked eye) option prices as scale-dependence (and zero correlation). Dependence of volatility on the (absolute) level of the process is quite reasonable when modeling quantities that are thought of as exhibiting “more stationarity” than stock prices (such as interest rates or commodity prices).

With parameters as given in Figure 2 let us assume the SABR model is the true data-generating process, and look at hedgers who use a Heston model. More specifically, the investigation of the performance of the locally risk-minimizing hedge and the delta hedge in the (wrong) Heston model is carried out as follows: (1) simulate stock prices and volatilities from the SABR model, (2) for each path implement the Heston-based locally risk-minimizing strategy (using the initially calibrated parameters and the simulated Heston-sense local variance along each path) as well as a delta hedge and (3) implement the SABR model’s delta hedge (which, because of zero correlation of the Brownian motions, coincides with the locally risk-minimizing hedge) using the pricing formula given in Hagan et al. (2002).
Figure 2: Correlation and scale-dependence can both explain a skew. The picture shows 1-year implied volatilities in the Heston model (circles) and in the SABR model (solid line). Parameters for the Heston model are as specified in Table 2 (except for $r = \mu = 0$), and the SABR settings are $a(0) = 1.92$, $\nu = 0.2$, $b = 0$, and $\rho = 0$, with the two latter being fixed and the two former then calibrated.

<table>
<thead>
<tr>
<th>Hedge method</th>
<th>SABR RiskMin</th>
<th>Heston RiskMin</th>
<th>Heston Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hedge error</td>
<td>13.3</td>
<td>13.9</td>
<td>18.4</td>
</tr>
</tbody>
</table>

Table 5: Hedge error under a misspecified data-generating process (SABR). The target option is a 1-year forward-at-the-money call.

Table 5 presents the results. The locally risk-minimizing hedge from the (incorrect) Heston model is almost as good as the one from the (correct) SABR model. It is important, however, to use the calibrated Heston model’s (spurious) correlation; using a Heston-based ordinary delta hedge increases the error by about one third. In essence the result of this controlled experiment can be interpreted as follows: if the stochastic volatility model is reasonably calibrated, it does not matter much which particular one is used. As explained below, the same is true for real market data.

4 Empirical Performance

In this section we present an empirical test of the performance of locally risk-minimizing hedge strategies. In light of the previous section the two main questions are: Are there any gains from risk minimization based on a stochastic volatility model? Does it matter which parametric model one uses? (The answers are: “Yes, sometimes, and at least nothing’s lost” and “No, apparently not.”)
Data. We collected\footnote{Time series of foreign exchange rate options can be found at British Bankers’ Association home page \url{http://www.bba.org.uk/bba/jsp/polopoly.jsp?d=129&a=799}. The data on stock index options were kindly provided by two investment banks, and can be found on the corresponding author’s home page \url{http://www.math.ku.dk/~rolf}} times series for spot and option prices for three different markets: The U.S. S&P 500 index, the European EUROSTOXX 50 index and the USD/EURO exchange rate. The data, which cover the period from early 2004 to mid 2005, are summarized in Figure\textsuperscript{3}. The markets display both differences and similarities. The varying implied volatilities of at-the-money options are very hard to explain without a stochastic volatility model. In both the U.S. and the European stock markets there is a strong negative correlation between returns and (at-the-money implied) volatility, and generally the European market is a bit more “wild” (higher volatility, higher volatility of volatility and a more pronounced skew in implied volatilities across strikes). In the exchange rate market there is very low correlation between changes in the underlying and in the implied volatility. Moreover implied volatilities across strikes display a smile, rather than a skew.

Experimental design. At a point in time $t$ the hedger sets up a position that holds $\phi(t)$ (prescribed by some model and strategy) units of the underlying $S$ and $b(t) = \text{option}(t) - S(t)\phi(t)$ in a money market account, where $\text{option}(t)$ denotes the observed market price of the option to be hedged. This position then requires an investment of $\text{option}(t)$. As common among traders, we measure the quality (or riskiness) of a hedge portfolio by its daily (i.e. $dt = 1/250$) profit-and-loss(P&L)-ratio

$$P\&L_{t+dt} = \frac{\phi(t)S(t+dt) + e^{rdt}b(t) - \text{option}(t+dt)}{\text{option}(t)}$$

This reflects the profit or loss (or alternatively viewed: the instantaneous cost) relative to the size of the transaction of someone who has sold an option and is trying to cover the position by the underlying and the money market. We always consider hedgers who sell at- or out-of-the-money options.

Calibration. We consider hedges based on the Heston and SABR models. At a given date $t$ we assume that the hedger calibrates his model parameters to that date’s observed option prices in the following way:

**Heston** The stationary mean $\theta$ and speed of mean-reversion $\kappa$ are estimated from a first-order auto-regression of the implied 3-month at-the-money (squared) volatility. On any given day $V_t$, $\sigma$ and $\rho$ are then chosen to minimize the sum of squared differences between observed and model-based implied volatilities for 3-month options.

**SABR** We fix $b$ at some value (typically 1) and chose $a_t$, $\eta$ and $\rho$ to minimize differences between the model’s and the market’s implied volatilities. This is done separately for each expiry.
Both U.S. and European interest rates are assumed to be 4% and differences to $\mathbb{P}$-drift rates are ignored.

**Results.** Averages and standard deviations of the profit-and-loss-ratio across markets, models, moneyness and option expiries are given in Table 6. The main message is that the risk-minimizing hedges offer a clear benefit; in markets where there is correlation between changes in the underlying and the instantaneous variance (skew in implied volatilities), the standard deviations of the profit-and-loss-ratios are reduced by between one-third and one-quarter. Differently put, based on recent, large, and varied data-sets we give a strongly affirmative answer to the questions-in-title posed by authors from Nandi (1998) to Doran, Peterson & Tarrant (2007).

If correlation is close to zero (as in the exchange rate market), there is (as it should be) no gain from the suggested risk-minimization. But nothing is lost either: one could fear that a complicated model and frequent re-calibrations would lead to “over-fitting the data” and, thus, deteriorate the out-of-sample (as is this study) hedge performance.

The table also reveals that long-term options are (somewhat) easier to hedge than short-terms ones and that risk-minimization is comparatively most effective for out-of-the-money options; both these results match the simulation studies. Moreover, the Heston and SABR models perform (slightly) better for long- resp. short-term options. This supports (though only weakly) the view held by some\(^\text{12}\) that local variance is “more log-normal than non-central $\chi^2$,” and it indicates that you should take mean-reversion into account for long expiries. But the big picture is that the magnitude of the standard deviations of the profit-and-loss ratios is quite stable across markets, moneyness, models and expiry. The scale-dependence parameter $b$ seems to matter little; you may as well use a scale-invariant model. This lends empirical support to the suggestions in Ayache et al. (2004) and the methodology in Cont & da Fonseca (2002).

Average $P&L$'s are small, but not zero; this reflects a volatility risk-premium. In our sample there are (almost always) positive expected profits to sellers of at- or out-of-the-money options, who hedge their positions with the underlying. Traders will see this as reward for “short volatility (or gamma) exposure”, while closer inspection, see Branger & Schlag (2007), reveals that one should be careful about what to read into the signs of average hedge errors.

Advocates of Levy-type models, e.g. Carr, Geman, Madan & Yor (2007), say that “stock markets are all jumps.” Our results show that there is a (return, volatility)-correlation component that can be treated (i.e. better hedged or risk-managed) by a diffusive stochastic volatility model. The skew can be tamed, so to say.

\[^{11}\]1-month EUROSTOXX 50 options were not available, but other than that the empty or non-existent cells are intentional; nothing interesting or surprising happens here.

Figure 3: The data used in the study. Each row corresponds to a different market: S&P500 (top), EUROSTOXX 50 (middle) and USD/EURO exchange rate (bottom). The columns correspond to: spot prices over the observation period (early 2004 to mid 2005) (left), time series of implied volatilities of 3-month at-the-money options (middle), and time-series averages of implied volatilities across strikes for 1-month (red), 3-month (black) and 1-year (green) expiries (right). The scales are identical within the middle resp. right column.
<table>
<thead>
<tr>
<th></th>
<th>SPX</th>
<th>EUX</th>
<th>FX</th>
<th>SPX</th>
<th>EUX</th>
<th>FX</th>
<th>SPX</th>
<th>EUX</th>
<th>FX</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>strike/forward</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market</td>
<td>0.94</td>
<td></td>
<td></td>
<td>1.07</td>
<td></td>
<td></td>
<td>1.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Expiry 1M</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston Delta</td>
<td>-2.9</td>
<td>0.9</td>
<td>3.9</td>
<td>-2.7</td>
<td>1.0</td>
<td>4.3</td>
<td>-1.5</td>
<td>0.1</td>
<td>4.9</td>
</tr>
<tr>
<td>Heston RiskMin</td>
<td>-2.7</td>
<td>1.0</td>
<td>4.3</td>
<td>-1.5</td>
<td>0.1</td>
<td>4.9</td>
<td>0.1</td>
<td>4.9</td>
<td>0.1</td>
</tr>
<tr>
<td>SABR Delta ($b = 1$)</td>
<td>-1.4</td>
<td>0.5</td>
<td>0.0</td>
<td>0.9</td>
<td>0.5</td>
<td>0.0</td>
<td>4.9</td>
<td>0.1</td>
<td>4.9</td>
</tr>
<tr>
<td>SABR RiskMin ($b = 1$)</td>
<td>-1.5</td>
<td>0.1</td>
<td>4.9</td>
<td>0.1</td>
<td>4.9</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston Delta</td>
<td>48</td>
<td>7.9</td>
<td>44</td>
<td>37</td>
<td>5.4</td>
<td>32</td>
<td>28</td>
<td>4.9</td>
<td>31</td>
</tr>
<tr>
<td>Heston RiskMin</td>
<td>32</td>
<td>5.0</td>
<td>2.5</td>
<td>28</td>
<td>4.9</td>
<td>2.5</td>
<td>28</td>
<td>4.9</td>
<td>2.5</td>
</tr>
<tr>
<td>SABR Delta ($b = 1$)</td>
<td>32</td>
<td>5.0</td>
<td>2.5</td>
<td>28</td>
<td>4.9</td>
<td>2.5</td>
<td>28</td>
<td>4.9</td>
<td>2.5</td>
</tr>
<tr>
<td>SABR RiskMin ($b = 1$)</td>
<td>28</td>
<td>4.9</td>
<td>2.5</td>
<td>28</td>
<td>4.9</td>
<td>2.5</td>
<td>28</td>
<td>4.9</td>
<td>2.5</td>
</tr>
<tr>
<td><strong>Expiry 3M</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston Delta</td>
<td>0.5</td>
<td>0.7</td>
<td>0.4</td>
<td>0.4</td>
<td>1.7</td>
<td>1.3</td>
<td>0.5</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>Heston RiskMin</td>
<td>0.5</td>
<td>0.6</td>
<td>0.4</td>
<td>0.3</td>
<td>1.7</td>
<td>1.0</td>
<td>0.5</td>
<td>0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>SABR Delta ($b = 1$)</td>
<td>0.6</td>
<td>0.1</td>
<td>0.3</td>
<td>0.0</td>
<td>1.2</td>
<td>0.2</td>
<td>0.6</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>SABR RiskMin ($b = 1$)</td>
<td>0.6</td>
<td>0.1</td>
<td>0.3</td>
<td>0.0</td>
<td>1.2</td>
<td>0.2</td>
<td>0.6</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>Std. deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston Delta</td>
<td>8.2</td>
<td>8.3</td>
<td>3.5</td>
<td>4.2</td>
<td>12</td>
<td>11</td>
<td>8.3</td>
<td>8.3</td>
<td>3.5</td>
</tr>
<tr>
<td>Heston RiskMin</td>
<td>5.8</td>
<td>5.1</td>
<td>2.4</td>
<td>2.8</td>
<td>8.3</td>
<td>9.0</td>
<td>5.8</td>
<td>5.1</td>
<td>2.4</td>
</tr>
<tr>
<td>SABR Delta ($b = 1$)</td>
<td>8.3</td>
<td>3.2</td>
<td>3.5</td>
<td>0.6</td>
<td>12</td>
<td>3.0</td>
<td>8.3</td>
<td>3.2</td>
<td>3.5</td>
</tr>
<tr>
<td>SABR RiskMin ($b = 1$)</td>
<td>5.2</td>
<td>4.8</td>
<td>2.2</td>
<td>0.6</td>
<td>7.6</td>
<td>2.9</td>
<td>5.2</td>
<td>4.8</td>
<td>2.2</td>
</tr>
<tr>
<td>SABR RiskMin ($b = 1/2$)</td>
<td>5.2</td>
<td>2.2</td>
<td>7.6</td>
<td>5.2</td>
<td>2.2</td>
<td>7.6</td>
<td>5.2</td>
<td>2.2</td>
<td>7.6</td>
</tr>
<tr>
<td><strong>Expiry 1Y</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston Delta</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Heston RiskMin</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>SABR Delta ($b = 1$)</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>SABR RiskMin ($b = 1$)</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Std. deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heston Delta</td>
<td>1.6</td>
<td>2.0</td>
<td>1.0</td>
<td>1.3</td>
<td>1.8</td>
<td>2.3</td>
<td>1.6</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Heston RiskMin</td>
<td>1.1</td>
<td>1.3</td>
<td>0.7</td>
<td>1.0</td>
<td>1.2</td>
<td>2.1</td>
<td>1.1</td>
<td>1.3</td>
<td>0.7</td>
</tr>
<tr>
<td>SABR Delta ($b = 1$)</td>
<td>3.0</td>
<td>1.5</td>
<td>2.0</td>
<td>0.8</td>
<td>4.4</td>
<td>1.6</td>
<td>3.0</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>SABR RiskMin ($b = 1$)</td>
<td>2.0</td>
<td>1.6</td>
<td>1.4</td>
<td>0.8</td>
<td>3.0</td>
<td>1.5</td>
<td>2.0</td>
<td>1.6</td>
<td>1.4</td>
</tr>
<tr>
<td>SABR RiskMin ($b = 1/2$)</td>
<td>2.0</td>
<td>1.4</td>
<td>3.0</td>
<td>2.0</td>
<td>1.4</td>
<td>3.0</td>
<td>2.0</td>
<td>1.4</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Table 6: Averages and standard deviations over the period early 2004 to mid 2005 for daily P&L-ratios across different models and hedge strategies for options on the S&P 500 index (SPX), the EUROSTOXX 50 index (EUX), and the USD/EURO exchange rate (FX).
5 Conclusion and Future Research

In this paper we calculated locally risk-minimizing hedge strategies for a general class of stochastic volatility models. Our empirical tests (across markets, time and option-types) showed that the risk-minimizing hedges offer what an economist would call a Pareto-improvement over usual delta hedges; they are as easy and reliable to implement, you are never worse off, and sometimes better off by quite some margin. We presented experimental and empirical evidence on the importance of model risk (or the lack of it). Our findings reveal that when volatility is stochastic, it is important to model it as such; short-cuts will not do. However, as long as the modeling is done sensibly, the exact model seems to matter little for hedging plain vanilla options.

An interesting topic for future research is the application of the locally risk-minimizing strategies to exotic options. For example, the form of the risk-minimizing hedge in Proposition 1 carries over verbatim to barrier options; we just don’t know of any truly closed-form expressions for the relevant conditional expectations and partial derivatives. Traders view barrier options as “skew products,” and as we have seen local risk-minimization deals well with skews. But on the other hand there could be a stronger element of model risk, because as pointed out by among others Ayache et al. (2004) models that produce similar prices of plain vanilla options (by calibration) may give markedly different prices of barrier (and other exotic) options.

In practice barrier options are primarily used in exchange rate markets. We saw that return/volatility-correlations were low in these markets and, therefore, the risk-minimizing hedges from the stochastic volatility models offered no benefits. However, as investigated in Carr & Wu (2007), there are day-to-day variations that can be captured by models that introduce a third stochastic factor controlling the asymmetry in implied volatilities: the stochastic skew. While Carr & Wu (2007) use general Levy-processes, we think that it is both possible and sufficient to use diffusive models when considering exchange rates markets, and the effectiveness of different hedging techniques based on such models is the focus of ongoing research.

References


