Stochastic Volatility: Risk Minimization and Model Risk

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

January 29, 2007

Abstract

In this paper locally risk-minimizing hedge strategies for European-style contingent claims are derived and tested for a general class of stochastic volatility models. These strategies are as easy to implement as ordinary delta hedges, yet in realistic settings they produce markedly lower hedge errors. Our experimental investigations on model risk furthermore show that locally risk-minimizing hedges are robust with respect to parameter uncertainty as well as misspecifications of the stochastic volatility model.

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

JEL classification: C90, G13

1 Introduction

Our goal in this paper is to study hedging in stochastic volatility models. These models are popular as they resolve some of the problems of the standard Black-Scholes model (e.g. non-flat implied volatilities) as well as Derman-Dupire-type local volatility models (unreasonable time-series behavior caused by matching observed option prices simply by functional dependence of local volatility on the

∗School of Mathematics, University of Leeds, Leeds LS2 9JT, United Kingdom. E-mail: ewald@maths.leeds.ac.uk
†Corresponding author. Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark. E-mail: rolf@math.ku.dk
‡School of Mathematics and Leeds University Business School, University of Leeds, Leeds LS2 9JT, United Kingdom. E-mail: k.r.schenk-hoppe@leeds.ac.uk
level of the underlying and calendar time). In this class of models not every contingent claim can be replicated perfectly by a self-financing trading strategy. This incompleteness affects particularly the seller of a contingent claim because it is not possible to eliminate all risk by investing in primary assets. The main objective is therefore to deal effectively with this risk. Within a general incomplete-market semi-martingale framework Föllmer & Schweizer (1990) propose the concept risk-minimizing hedges which aims at minimizing the variance of the cost process of non-self-financing hedges. Schweizer (1991) presents a “localized” version of this approach.

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 due to El Karoui, Peng & Quenez (1997). The formula shows that the hedge can be decomposed as a sum of the delta hedge and a volatility-risk term. Since the application of risk-minimization techniques to stochastic volatility models is a quite natural avenue to explore, this formula (in various guises) has been reported elsewhere. To the best of our knowledge, it is first published in Frey (1997, Proposition 6.5) who derives the result in a different way. An alternative intuitive derivation of this formula is also provided here to illustrate the mechanics behind local risk-minimization.

These results are not well-known in the practical/applied finance literature despite considerable theoretical work, see Möller (2004) for a recent survey. This is evidenced, for instance, by the existence of several studies in which the volatility correction term is (mistakenly) ignored (see the list in Alexander & Nogueira (2005, page 14)) as well as by the fairly recent appearance of the paper Ahn & Wilmott (2003). The result is therefore derived in detail and a numerical illustration is provided to highlight the benefits of locally risk-minimizing hedging over the commonly used delta hedging in the Heston model.

A broader perspective on risk is taken by an investigation of the extent to which the performance of locally risk-minimizing hedges is sensitive to model risk. The results demonstrate the robustness of this hedging approach to parameter uncertainty and misspecification of the volatility dynamics. The latter has the caveat that a truly dynamic model must be used; the sensitivities from a static model will not do.

Section 2 briefly reviews risk-minimizing hedges as introduced by Föllmer & Schweizer (1990). Section 3 derives an explicit formula for such a hedge within a general class of stochastic volatility models. Section 4 provides numerical illustrations for the Heston model. Section 5 presents a sensitivity analysis of model risk. Section 6 concludes.
2 Cost processes and risk-minimization

In incomplete markets delta hedge strategies are usually 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 they are not self-financing. In the latter case a hedge is associated with a cost process that aggregates any additional investments. Professional delta hedgers know their hedges may “bleed,” i.e. readjusting them requires additional funds.

Our discussion is restricted to models with a money market account \( B(t) \) and a single stock \( S(t) \) whose dynamics are of the type
\[
dS(t)/S(t) = \mu(t)dt + \sum_j \sigma_{1j}(t)dW_j(t),
\]
where \( \mu(\cdot) \) and \( \sigma_{1j}(\cdot) \) are stochastic processes.

A trading strategy is a predictable process \( \varphi = (\varphi_0, \varphi_1) \), where the first and second component are the holdings (in number of units) of the risk-free asset and the stock, respectively. The cost process associated with \( \varphi \) is defined as
\[
C_\varphi(t) = V_\varphi(t) - \int_0^t \varphi_0(s)dB(s) - \int_0^t \varphi_1(s)dS(s),
\]
where \( V_\varphi(t) = \varphi^0(t)B(t) + \varphi^1(t)S(t) \) denotes the value of the trading strategy at time \( t \).

A trading strategy is self-financing if and only if the cost process is constant in time. In this case \( C_\varphi(t) = V_\varphi(0) \) for all \( t \in [0, T] \), i.e., only an initial investment is required. The availability of a self-financing hedge \( \varphi \) for a contingent claim \( H \) permits the removal of all solvency risk. A seller of \( H \) can guarantee payment of his obligation (at the time \( H \) expires) by investing the certain initial amount \( V_\varphi(0) \) to buy \( \varphi \) because no additional investments are required afterwards. Otherwise the seller may use a non-self-financing hedge to match his obligation at expiry but that carries the risk associated to future investments into this hedge. This makes the total cost of the hedge uncertain; something that the seller does not appreciate will seeks to reduce.

The conditional variance process of the cost process is defined as
\[
R_\varphi(t) := \mathbb{E}[(C_\varphi(T) - C_\varphi(t))^2 | \mathcal{F}_t].
\]  

(1)

A trading strategy \((\psi^0, \psi^1)\) is said to be an admissible continuation of a trading strategy \((\varphi^0, \varphi^1)\) from time \( t \in [0, T] \) on, if
\[
\psi^0(s) = \varphi^0(s), \; s < t; \; \psi^1(s) = \varphi^1(s), \; s \leq t; \; \text{and} \; V_\psi(T) = V_\varphi(T) \; \mathbb{P}\text{-a.s.}
\]

Föllmer & Schweizer (1990) define the trading strategy \( \varphi \) to be risk-minimizing (\( R \)-minimizing) if, for any \( t \in [0, T] \) and for any admissible continuation \( \psi \) of \( \varphi \)

---


2A 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 \( \varphi^1(T) = 0 \).
from \( t \) on,
\[
R_\phi(t) \geq R_\varphi(t) \quad \mathbb{P}\text{-a.s. for all } t \in [0,T).
\]
When the stock price \( S \) is a \( \mathbb{P}\)-martingale this criterion guarantees that the cost process is a \( \mathbb{P}\)-martingale as well, i.e., the hedge is self-financing on average. Otherwise a risk-minimizing hedge for \( H \) may not exist, see Föllmer & Schweizer (1990). This existence problem can be overcome by employing a local version of the risk-minimization criterion. The following definition is quite technical, cf. Schweizer (1991), but an illustrative derivation is given below. Call a trading strategy \((\delta^0, \delta^1)\) a small perturbation if both \( \delta^1 \) and \( \int_0^T |\delta^1(t)S(t)b(t)| dt \) are bounded and \( \delta^0(T) = \delta^1(T) = 0 \). Given a small perturbation \((\delta^0, \delta^1)\) and a subinterval \([s, t] \subset [0, T]\), define the small perturbation \( \delta^1_{[s,t]} := (\delta^1_{[s,t],0}, \delta^1_{[s,t],1}) \) with \( \delta^1_{[s,t]}(u, \omega) := \delta^0(u, \omega) - 1_{[s,t]}(u) \) and \( \delta^1_{[s,t]}(u, \omega) := \delta^1(u, \omega) - 1_{[s,t]}(u) \). For a partition \( \tau \) of the interval \([0, T]\) and a small perturbation \( \delta \) we finally define
\[
r^\tau(t, \varphi, \delta) := \sum_{t_i \in \tau} \mathbb{E} \left[ \frac{R_{\varphi+\delta_{{[\tau_i, \tau_{i+1}]}]}(t_i) - R_{\varphi}(t_i)}{\int_{t_i}^{t_{i+1}} S(t)^2 \|\sigma(t)\| dt |\mathcal{F}_{t_i}} \right] \cdot 1_{(t_i, t_{i+1})}(t).
\]
The strategy \( \varphi \) is then called \textit{locally risk-minimizing} if, for every small perturbation \( \delta \),
\[
\liminf_{|\tau| \to 0} r^\tau(t, \varphi, \delta) \geq 0 \quad \mathbb{P}\text{-a.s. for all } t \in [0,T].
\]

3 Local risk-minimization in stochastic volatility models

In this paper we consider a class of stochastic volatility models which encompasses most of those commonly used in research as well as in practice. The class is given by models of the form
\[
\begin{align*}
\frac{dS_t}{S_t} &= \mu dt + S_t^\gamma f(V_t) \left[ \sqrt{1-\rho^2} dW^1_t + \rho dW^2_t \right] \\
\frac{dV_t}{V_t} &= \beta(V_t) dt + g(V_t) dW^2_t
\end{align*}
\]
\tag{2}
with independent standard Brownian motions \( W^1 \) and \( W^2 \). \( S(t) > 0 \) denotes the price of the (traded) stock and \( V(t) > 0 \) is the (untraded) stochastic local return variance. These models allow for level dependence \((\gamma \neq 0)\) and correlation between returns and variance \((\rho \neq 0)\). Table 4 collects several (named) models that fit into the context of \( \text{(2)} \).

The model \( \text{(2)} \) is formally similarity to the framework considered in El Karoui et al. (1997); their model (1.2) is obtained by setting
\[
\begin{align*}
P_t = \begin{pmatrix} S(t) \\ V(t) \end{pmatrix}, b_t = \begin{pmatrix} \mu \\ \beta(V(t)) \end{pmatrix}, \sigma_t = \begin{pmatrix} S(t)^\gamma \sqrt{1-\rho^2} f(V(t)) & S(t)^\gamma \rho f(V(t)) \\ 0 & g(V(t)) \end{pmatrix}
\end{align*}
\]
Table 1: Specification of stochastic volatility models for Equation (2).

<table>
<thead>
<tr>
<th>Authors &amp; year</th>
<th>Specification</th>
<th>Remarks</th>
</tr>
</thead>
</table>
| Hull-White     | \( f(v) = v, \)  
\( \beta(v) = 0, \)  
\( g(v) = \sigma, \)  
\( \rho = 0, \gamma = 0 \) | Local variance: Geometric Brownian motion. Options priced by mixing. |
| Wiggins        | \( f(v) = e^{v/2}, \)  
\( \beta(v) = \kappa(\theta - v)/v, \)  
\( g(v) = \sigma, \)  
\( \rho = 0, \gamma = 0 \) | Local volatility: Ornstein-Uhlenbeck in logarithms. |
| Stein-Stein    | \( f(v) = |v|, \)  
\( \beta(v) = \kappa(\theta - v)/v, \)  
\( g(v) = \sigma/v, \)  
\( \rho = 0, \gamma = 0 \) | Local volatility: Reflected Ornstein-Uhlenbeck. |
| Heston         | \( f(v) = \sqrt{v}, \)  
\( \beta(v) = \kappa(\theta - v)/v, \)  
\( g(v) = \sigma/\sqrt{v}, \)  
\( \rho \in [-1, 1], \gamma = 0 \) | Local variance: CIR process. First model with correlation. Options priced by inversion of characteristic fct. |
| Romano-Touzi   | \( f(v) = \sqrt{v}, \)  
\( \beta \) and \( g \) are free,  
\( \rho \in [-1, 1], \gamma = 0 \) | Extension of mixing to correlation. |
| SABR           | \( f(v) = v \)  
\( \beta(v) = 0, \)  
\( g(v) = \sigma, \)  
\( \rho \in [-1, 1], \gamma \in [-1, 0] \) | Level dependence in volatility. Options priced perturbation tech. |

Proposition 1.1 in El Karoui et al. (1997) states that the locally risk-minimizing hedge can be computed in three steps:

Completion. The original market is completed by introducing a second stock in such a way that the risk premium in the extended (and now complete) market is the one corresponding to the minimal martingale measure, \( \mathbb{Q}_{\min} \);

Hedging. Compute the self-financing hedge strategy in the completed market;

---

*El Karoui et al. (1997) assume boundedness of the volatility matrix and its inverse. No commonly used stochastic volatility model satisfies this. 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 used in later, Theorem 1 in Cheridito, Filipovic & Kimmel (forthcoming) shows that the martingale property holds if the Feller condition is satisfied under both \( \mathbb{P} \) and \( \mathbb{Q}_{\min} \).*

---

5
**Projection.** The hedge strategy in the extended market is projected onto the original market (taking into account the geometry of the market extension).

**Completion.** Let \( r \) be the interest rate for the risk-free money account. Denote by \( \lambda_t^\top = (\lambda_1^t, \lambda_2^t) \) an arbitrary risk premium. Eq. (1.13) in El Karoui et al. (1997) implies that the risk premium corresponding to the minimal martingale measure is given by

\[
\lambda_t^{\text{min}} = (\sigma_1^t)^\top \left[ \sigma_1^t (\sigma_1^t)^\top \right]^{-1} \sigma_1^t \lambda_t
\]

\[
= \left( \frac{\sqrt{1 - \rho^2}}{\rho} \right) \frac{1}{f(V(t))} \cdot f(V(t)) \left( \sqrt{1 - \rho^2}, \rho \right) \left( \begin{array}{c} \lambda_1^t \\ \lambda_2^t \end{array} \right)
\]

\[
= \frac{\mu - r}{f(V(t))} \cdot \left( \frac{\sqrt{1 - \rho^2}}{\rho} \right)
\]

The equality under the bracelets holds by the definition of the risk premium. The completed market is therefore given by

\[
dS_1(t)/S_1(t) = \mu dt + S(t) \gamma f(V(t)) \left[ \sqrt{1 - \rho^2} dW_1^{\text{min},1} + \rho dW_1^{\text{min},2} \right]
\]

\[
dS_2(t)/S_2(t) = \left( r + \rho g(V(t)) \right) dt + g(V(t)) dW_2^{\text{min},2}
\]

where \( dW^{\text{min}} = dW + \lambda_{\text{min}}^{\top} dt \) defines a \( \mathbb{Q}_{\text{min}} \)-Brownian motion.

**Hedging.** 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 \( \Delta_1(t) \) and \( \Delta_2(t) \) in the assets \( S_1 \) and \( S_2 \), respectively. The position in the money market account is adjusted so as to make the hedge self-financing. Then, on the one hand,

\[
\begin{align*}
\frac{dX(t)}{X(t)} &= \Delta_1(t)\frac{dS_1(t)}{S_1(t)} + \Delta_2(t)\frac{dS_2(t)}{S_2(t)} + r \left[ X(t) - \Delta_1(t)S_1(t) - \Delta_2(t)S_2(t) \right] dt \\
&= (\ldots) dt + \Delta_1(t)S_1(t)\gamma f(V(t)) \left[ \sqrt{1 - \rho^2} dW_1^{\text{min},1} + \rho dW_1^{\text{min},2} \right] \\
&\quad + \Delta_2(t)S_2(t)g(V(t)) dW_2^{\text{min},2}.
\end{align*}
\]

On the other hand, the original market model is Markovian, which allows to write the price of the claim \( H \) as

\[
e^{-r(T-t)} \mathbb{E}_{\mathbb{Q}_{\text{min}}}^\text{min} H(S(T)) = C(t, S_1(t), V(t))
\]

for some function \( C(\cdot, \cdot, \cdot) \). The conditional expected value is calculated under the minimal martingale measure. Since \( X(t) = C(t, S_1(t), V(t)) \), application of
the Itô formula to find \(dC\) and comparison of the resulting diffusion terms with \(dX\) yields

\[
\Delta_1(t) = C_S \quad \text{and} \quad \Delta_2(t) = \frac{C_V}{S_2(t)} V(t),
\]

where \(C_S\) and \(C_V\) denote partial derivatives.

**Projection.** Proposition 1.1 in El Karoui et al. (1997) implies that the investment in the stock \(S(t) = S_1(t)\) under the locally risk-minimizing hedge of the original market \([2]\) is given by

\[
S(t) \varphi_{\text{min}}^1(t) = \left[\sigma_t^1 (\sigma_t^1) \right]^{-1} \sigma_t^1 \sigma_t \left( \frac{C_S S(t)}{C_V V(t)} \right) = C_S S(t) + \rho \frac{V(t) g(V(t))}{S(t)^{1+\gamma} f(V(t))} C_V.
\]

The following proposition summarizes these findings.

**Proposition 1** Consider the stochastic volatility model \([2]\). The locally risk-minimizing hedge of a European contingent claim with payoff \(H(S(T))\) holds

\[
\varphi_{\text{min}}^1(t) = C_S + \rho \frac{V(t) g(V(t))}{S(t)^{1+\gamma} f(V(t))} C_V
\]

units of the stock, where

\[
C(t, S(t), V(t)) = e^{-r(T-t) E_t^{\text{min}} H(S(t))}.
\]

\(E_t^{\text{min}}\) denotes the conditional expectation with respect to the minimal martingale measure under which the dynamic are given by

\[
dS(t)/S(t) = r dt + S(t)^{\gamma f(V(t))} \left[ \sqrt{1 - \rho^2 dW_t^{\text{min},1}} + \rho dW_t^{\text{min},2} \right]
\]

\[
dV(t)/V(t) = \left[ \beta(V(t)) - \rho g(V(t)) f(V(t)) (\mu - r) \right] dt + g(V(t)) dW_t^{\text{min},2}.
\]

The investment in the money market is by \(C(t, S(t), V(t)) - \varphi_{\text{min}}^1(t) S(t)\).

The locally risk-minimizing hedge and the standard delta hedge do not coincide in general. The difference in stock holdings is given by

\[
\varphi_{\text{min}}^1(t) - C_S = \rho \frac{V(t) g(V(t))}{S(t)^{1+\gamma} f(V(t))} C_V
\]

see \([3]\). If stock returns and variance are uncorrelated (\(\rho = 0\)), the hedges are identical. However, the typical case is that the correlation coefficient \(\rho\) is negative and the pay-off function convex and, hence (e.g. from Romano & Touzi (1997, Proposition 4.2)) \(C_V\) is positive. Proposition \([4]\) therefore implies that an ordinary delta hedger invests too much money 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 ($\rho \neq 0$) and there is a 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.

Proposition 1 is formulated for the hedge of a European claim. The three step (complete-compute-project) procedure indeed works for general claims, but its usefulness hinges on the possibility of finding the replicating strategy in the completed model. The strategy can be obtained for pure volatility derivatives, barrier options and Asian options. Implementation of the locally risk-minimizing hedge requires the computation of the partial derivatives (“Greeks”) $C_S$ and $C_V$.

When analytical expressions for these are not available, Monte Carlo estimates can be used instead, e.g., by employing the Malliavin calculus techniques from Ewald & Zhang (2006).

A deceptively simple derivation of locally risk-minimizing strategies. We give a brief interlude to highlight a short-cut for deriving (3). Suppose at some point in time $t$ one 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 $dC = \ldots dt + C_S dS + C_V dV$ which implies that the change in value of the hedge over a small time-interval $[t, t + dt]$ is given by

$$dX = dC - \Delta dS = \ldots dt + (C_S - \Delta) dS + C_V dV.$$

For the local conditional variance the $dt$-term does not matter, and one has

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

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

From the hedger’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 $\Delta_{\text{min}} = \varphi_{\text{min}}^1$ which coincides with the above result.

This derivation confirms that one has indeed a minimum local variance hedge. However the calculation is too short to hide that it is an attempt of deception. Ignoring the variances from the drift term in the derivation can be seen as local (rather than global) risk-minimization; it could be justified formally. The same holds for the hands-on approach to discretization.

The main shortcoming of the above 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. Föllmer and Schweizer’s approach does not have this deficiency because it derives the price as well as the hedge by considering trading in primary assets only.

4 Hedge performance in the Heston model

The stochastic volatility model by Heston (1993),

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

(6)

is a benchmark and large number of studies have fit the model to data. Parameter estimates, however, vary across markets, time periods and (more discomfortingly) econometric methods. This paper uses the parameter estimates obtained by Eraker (2004). 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 which 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. Both of these measures are of importance for the analysis in Section 5.

The Heston model is obtained from (2) by setting $\gamma = 0$, $f(v) = \sqrt{v}$, $\beta(v) = \kappa(\theta - v)/V$ and $g(v) = \sigma/\sqrt{v}$. The Feller condition $2\kappa\theta > \sigma^2$ guarantees strict positivity of the process $V$.

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

$$
\varphi_{\min}^1(t) = C_S + \rho \frac{C_V}{S(t)}.
$$

The option pricing formula $C$ and the related Greeks are implemented using the Lipton-Lewis reformulation of Heston’s original expression to increase computational stability, see Lipton (2002).
Table 2: Benchmark settings for the parameters of the Heston model. The numbers in square brackets are the standard errors of the estimates.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Text reference/interpretation</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>risk-free rate</td>
<td>0.04</td>
</tr>
<tr>
<td>$\mu$</td>
<td>expected stock return</td>
<td>0.10 [0.022]</td>
</tr>
<tr>
<td>$\theta$</td>
<td>typical local variance</td>
<td>0.0483 [0.0012] ($\approx 0.220^2$)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>speed of mean reversion</td>
<td>4.75 [1.8]</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>volatility of volatility</td>
<td>0.550 [0.018]</td>
</tr>
<tr>
<td>$\rho$</td>
<td>correlation</td>
<td>-0.569 [0.014]</td>
</tr>
<tr>
<td>$S(0)$</td>
<td>initial stock price</td>
<td>100</td>
</tr>
<tr>
<td>$V(0)$</td>
<td>initial local variance</td>
<td>$\theta$</td>
</tr>
<tr>
<td>$T$</td>
<td>expiry date of option</td>
<td>varies; often 1 year</td>
</tr>
<tr>
<td>$K$</td>
<td>strike of (call) option</td>
<td>varies; often forward at-the-money: $S(0)e^{rT}$</td>
</tr>
<tr>
<td>$\theta_{\text{option}}$</td>
<td>typical at-the-money implied volatility (squared)</td>
<td>0.0834 ($\approx 0.289^2$)</td>
</tr>
<tr>
<td>$\kappa_{\text{option}}$</td>
<td>option market implied speed of mean reversion</td>
<td>2.75</td>
</tr>
</tbody>
</table>

Hedge errors are reported as the standard deviation\(^7\) of the cost process at expiry divided by the initial option value (in percentage terms),

\[
\text{hedge error} = 100 \times \frac{\sqrt{\text{var}^P(\text{cost}(T; n))}}{e^{-rT} \mathbb{E}^\mu(\min(|S(T) - K|^+, T))}.
\]

Here “cost” denotes the cost-process and $n$ indicates dependence on the adjustment frequency of hedges which is $n$-times in a year at equidistant points in time. Hedge errors are estimated as follows: (1) simulate paths of stock prices and volatilities, (2) apply the above hedge strategy along each path (with prescribed frequency), (3) record the path-specific terminal cost and (4) compute sample moments from many paths.

The typical dependence of the hedge error on the hedging frequency is depicted in Figure 1. The locally risk-minimizing hedge gives a smaller error than the

---

\(^7\)The expected (discounted) profit is not a particularly relevant measure for the quality of a hedge because this amount is very close to being an initial investment. If $\mu = r$ it exactly coincides with the initial contribution. Put differently: you hedge to reduce risk, not to make money. Since the distribution of hedge errors is quite close to being normal (at least, it is almost symmetric), the standard deviation tells the same story as one-sided risk-measures such as the value-at-risk of expected shortfall. Also, for less-than-perfect hedges it is important to calculate the standard deviation under $P$, not under some martingale measure; the catch-phrase explanation being that “hedging is done in the real world,” see Nalholm & Poulsen (2006, page 56) for further discussion.
ordinary delta hedge throughout the range of hedging frequencies. Both hedge errors can initially be reduced by increasing the hedge frequency but the errors do not tend to 0. Hedging more frequently than twice a week does not lead to a further improvement in the hedge error. The error of the locally risk-minimizing hedge, however, flattens out at a distinctively lower level than that of the delta hedge. These findings are in perfect agreement with the model: the delta hedge is inferior and market incompleteness excludes a perfect hedge of the call-option using only the stock. In the Black-Scholes model daily hedging of a similar option gives an error of about 5%.

Information on typical hedge errors (for daily adjusted hedges) across moneyness and expiry date are provided in Table 3. The results show that the hedge of out-of-the-money options is much worse than for those at-the-money; this effect is particularly strong for short-dated options. The superiority of locally risk-minimizing strategies is confirmed: compared to the delta hedge they reduce the hedge error by a factor in the range of 0.85 to 0.90. This finding is robust across

---

*Absolute standard deviations for out-of-the-money options are smaller than the reported values which are measured relative to the initial value of the option. The latter are deemed to be more relevant here.*
strikes and expiry dates. Note that this 10-15% reduction of the hedge error with respect to the delta hedge is virtually free; a locally risk-minimizing hedge is in no way harder to use than an ordinary delta hedge.

5 Model risk

This section analyzes to what extent the performance of risk-minimizing hedge in the Heston model is sensitive to model risk, i.e., what happens if you get things a little bit wrong? Are the nice results from the previous section robust or are they on the “razor’s edge”? As stressed by Cont (2006) this is a highly relevant practical issue; not at least in light of the variety of parameter estimates for the Heston model. 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 of local risk-minimization),
- wrong Greeks (considerable negative effect) and
- wrong data-generating process (surprisingly small effect).

Picking the wrong martingale measure. In the context of the Heston model different martingale measures are obtained by different choices of $\kappa$ and $\theta$ in the local volatility dynamics of (6). Cheridito et al. (forthcoming) 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 choice of parameters, see Table 2). The choice of measure affects the conditional expectation on which the risk-minimizing strategy is based.

Our analysis focuses on three different martingale measures: (1) the minimal

<table>
<thead>
<tr>
<th>Forward (call) moneyness</th>
<th>Hedge type</th>
<th>Expiry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1/12</td>
</tr>
<tr>
<td>At-the-money</td>
<td>Delta</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>RiskMin.</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Reduction factor</td>
<td>0.90</td>
</tr>
<tr>
<td>10% Out-of-the-money</td>
<td>Delta</td>
<td>229</td>
</tr>
<tr>
<td></td>
<td>RiskMin</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td>Reduction factor</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 3: Hedge error for locally risk-minimizing hedges and delta hedges across different moneyness and expiry dates.
martingale measure (the correct one when you want to minimize the 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</td>
<td>( \theta )</td>
<td>0.229(^2)</td>
<td>0.220(^2)</td>
</tr>
<tr>
<td></td>
<td>( \kappa )</td>
<td>4.75</td>
<td>4.75</td>
</tr>
<tr>
<td>Hedge error</td>
<td>19.7</td>
<td>19.7</td>
<td>20.3</td>
</tr>
</tbody>
</table>

Table 4: Hedge error under different misspecifications of the volatility model.

The results of this exercise are summarized in Table 4. 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.

**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 of his estimates, see our 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 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>
<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 5: Effects of parameter uncertainty on locally risk-minimizing hedges. The table shows the relative increases in hedge error when the hedger uses parameters drawn from the distribution of Eraker’s estimator rather than the true parameter.

Table 5 compares the performance of the “random-parameter hedger” to that of someone who uses the true parameter. Results are reported in terms of the

---

11 Only standard errors, not correlations, are reported. We treat things as independent, which should give conservative estimates of the effects.
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 5 is that the adverse effects (in the range of 0-4% in relative terms) of 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%).

Using Black-Scholes’ Greeks. The partial derivatives $C_S$ and $C_V$ in equation (8) 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.”

![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](image)

---

12 One could fear that, in reality, a complex model could prove unstable because its parameters cannot be accurately assessed.

13 Note 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.
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 2 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.) But what if the Heston model is not the correct stochastic volatility model (after all alternatives abound)? This question leads to our last investigation.

**SABR as the data-generating process.** Let us put \( r = \mu = 0 \) for simplicity. Hagan, Kumar, Lesniewski & Woodward (2002) suggest the so-called SABR model, a version of which is given by

\[
\frac{dS(t)}{S(t)} = V(t)S^\gamma(t)dW_1^t, \\
\frac{dV(t)}{V(t)} = \nu dW_2^t.
\]

With appropriate parameter settings the SABR model can generate option prices (for a specific expiry) that are similar to those in the Heston model, see Figure 3. Yet the model is quite different: the skew in implied volatilities is captured by a level effect, not through negative correlation. This 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). Furthermore local variance is a log-normal process and shows no mean-reversion.

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).

Table 6 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 this result can be interpreted as follows: if the

---

14 If \(|\rho| = 1\), one is in a “one Brownian motion case” and perfect hedging is possible.
15 SABR is Anglo-acronymization of “stochastic \(\alpha-\beta-\rho\);” parameter names used by some authors.
16 Some people argue that log-normal processes fit local variance data better than the square root process in the Heston model, see, e.g., Paul Wilmott’s view on this issue: http://www.wilmott.com/images/246/VolForecastingOpTradingCM.wmv.
stochastic volatility model is reasonably calibrated, it does not matter much which particular one is used.

6 Conclusion

In this paper we calculated locally risk-minimizing hedge strategies for a general class of stochastic volatility models. In realistic settings they distinctively improve ordinary delta hedges at no extra cost. We further presented experimental evidence on the importance of model risk (or lack thereof.) 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. Moreover the gains from using locally risk-minimizing hedges are of a larger magnitude than the losses incurred from parameter estimation risk.

References


