Empirical performance of quadratic hedging strategies applied to European call options on an equity index

Love Lindholm

Abstract

Quadratic hedging is a well developed theory for hedging contingent claims in incomplete markets by minimizing the replication error in a suitable $L^2$-norm. This theory, though, is not widely used among market practitioners and relatively few scientific papers evaluate how well quadratic hedging works on real market data. In this paper, we develop a framework for comparing hedging strategies, and use it to empirically test the performance of quadratic hedging of European call options on the Euro Stoxx 50 index modeled with an affine stochastic volatility model with and without jumps. As comparison, we use hedging in the standard Black-Scholes model. We show that quadratic hedging strategies significantly outperform hedging in the Black-Scholes model for out of the money options and options near the money of short maturity when only spot is used in the hedge. When in addition another option is used for hedging, quadratic hedging outperforms Black-Scholes hedging also for medium dated options near the money.

1 Introduction

The theory of quadratic hedging provides a framework for hedging options in incomplete markets. In a situation where not all contingent claims can be perfectly replicated, quadratic hedging aims to minimize an $L^2$-distance between the claim and a hedging portfolio. This problem was first studied by Föllmer and Sonderman [13] and has since been extensively treated in the literature. For an overview of the literature we refer to the surveys of Pham [26] and Schweizer [28] and to the more recent results due to Černý [7].

An abundant flora of stochastic models for financial assets have evolved over the past two decades. Many of these models typically lead to situations with incomplete markets, for example when only the underlying asset is used for hedging in a stochastic volatility model, or generally in models that include jumps. Among practitioners, these stochastic models are often used when handling “exotic” options (i.e., options with a payoff that depends on the trajectory of the underlying asset, rather than just the asset’s value at maturity). (See for example [25, for examples of contracts traded in the equity market.) In contrast, when it comes to “vanilla” options (i.e., standard European put- and call options), traders still often rely on the lognormal model used in the ground breaking works of Black and Scholes [5] and Merton [24]. In the sequel, we will refer to this model as the BS model. The results of El Karoui et al. [19] on the robustness of the hedging of put- and call options in the BS model might explain part of its popularity. However, the robustness results rely on assumptions that are not necessarily satisfied when trading in real financial markets, and actual hedging errors can be substantial also for vanilla products.

The motivation for this report is to investigate whether quadratic hedging in a stochastic volatility model can provide better results than the BS model when hedging standard European call options on an equity index. The purpose is promotional rather than theoretical: we try to sell the idea that a more elaborate modelling can be beneficial also for traders dealing with standard options and that quadratic hedging is a natural framework for approaching hedging in real financial markets. Also, the empirical studies of quadratic hedging on real market data seem to be relatively few in relation to the rather large number of publications on theoretical aspects of the subject. Examples of empirical tests of quadratic hedging theory can be found in Cont and Kan [10], where
quadratic hedging theory is applied to hedge credit portfolios, and in Ewald et al. [12], where
quadratic hedging in jump models are used to hedge options on crude oil. Also, it has recently
come to our knowledge that the paper by Bakshi et al. [3] contains a study that in some aspects
is similar to ours. Bakshi et al. compare the performance of hedging in the BS hedging to that of
quadratic hedging on options on S&P500 in Heston’s model 17 and Heston’s model with normally
distributed jumps in the log spot process (known as Bate’s model [4]). In this report, we work
with Heston’s model and with Heston’s model with exponentially distributed jumps. We evaluate
the effect that different methods of parameter calibration has on the actual hedging results, which
we have not encountered in other papers. We also study how hedging can be improved by making
a regression on hedging ratios on observed hedging errors in the near past, which we have not
come across in a similar context before. Our regression technique is tested on delta hedging in the
BS model as well as on quadratic hedging strategies in the stochastic volatility model with and
without jumps. We also point out the difference between the quadratic hedging problem when
written in discrete versus continuous time, and evaluate both cases to see if any improvement can
be obtained by using the discrete time formulation over the continuous time formulation.

The quadratic hedging problem is rather straightforward in the case where the assets in the
hedging portfolio are martingales. This is the framework we will use in this report. If general
semimartingales are considered, the problem gets considerably more involved (see 26, 28, 7 ). In
Section 2, we thus review the results on quadratic hedging in the martingale case, for discrete time
in Section 2.1 and continuous time in Section 2.2. In Section 3, we apply the results from Section
2 to calculate the optimal quadratic hedging ratios of options in a stochastic volatility model with
jumps when hedging is performed either with just the underlying spot, or with the underlying spot
and some other option. The model parameters used in our tests are calibrated to market data with
the methodology described in Section 4. Section 5 describes an easy method for improving delta
hedging ratios obtained from a model by means of a linear regression on the model’s past hedging
performance. The performance of the different hedging strategies are evaluated on market data
with the methodology described in Section 6 and the results are presented in Section 7. Section
8 concludes.

2 Quadratic hedging with martingales

We state and prove the quadratic hedging result in the martingale case, which is really just an
$L^2$-projection of the claim to be hedged onto an appropriate subspace. Although the result in
discrete time can be seen as a special case of the result in continuous time, we treat it separately
in Section 2.1 since it illustrates the general idea without using any of the stochastic calculus
needed in continuous time case that we review in Section 2.2.

The classic arbitrage free pricing theory in complete markets ([5, 24, 14, 15, 16]) leads to
unique prices for contingent claims. These prices can be expressed in terms of conditional expecta-
tions under a measure under which the discounted values of directly traded assets are martingales.
Quadratic hedging, in contrast, does not provide any guidance regarding the choice of measure to
use for pricing, but merely gives a method for creating a hedging portfolio once the measure has
been fixed. In the sequel, we will suppose our chosen measure is such that the discounted values of
observable price processes (without dividends) are martingales as in the complete market theory.
This is the standard choice among practitioners also in incomplete markets.

2.1 Discrete time

Let $(\Omega, \mathcal{F}_t, Q)$ be a probability triple for $t \in [0, T]$ and fix a set of discrete times
$0 = t_0 < t_1 < \ldots < t_N = T$. We will consider a process $X$ with values $X_t = (X^1_t, \ldots, X^d_t)' \in \mathbb{R}^d$ (where $'$ denotes the transpose) at $t$. We assume that the filtration $\mathcal{F}_t$ is given for all $t \in [0, T]$ so that we can think of $X_t$ as samples of some $\mathcal{F}_t$-adapted process in continuous time. This could, for example, correspond to the situation where the price of an asset continues to evolve when the financial
market on which the asset is traded is closed during the weekend or during the night.
Let \( r_t \) be the (deterministic) risk free rate of return and define the accumulated value \( B \) of a bank account between time \( \tau \) and \( T \) according to
\[
B(\tau, T) = e^{\int_{\tau}^{T} r_t dt}
\]
and set \( B_{i,j} = B(t_i, t_j) \). Suppose the process \( X \) is such that its \( t_N \)-forward value
\[
\bar{X}_i = B_{i,N} X_i
\]
is a square integrable \( \mathcal{F}_i \)-martingale, where \( \mathcal{F}_i = \mathcal{F}_{t_i} \). We want to hedge a claim \( u_N \in L^2(\Omega, \mathcal{F}_N, Q) \) by minimizing the \( L^2(\Omega, \mathcal{F}_N, Q) \)-distance between \( u_N \) and a self-financing portfolio \( P \) investing in \( X \). Set
\[
\bar{u}_i = E[u_N | \mathcal{F}_i]
\]
and define \( P \) according to
\[
P_{i+1} = P_i + (\delta_i, X_{i+1} - X_i) + (P_i - (\delta_i, X_i))(B_{i,i+1} - 1).
\]
for some value \( P_0 \in \mathbb{R} \), where \((\cdot, \cdot)\) is the \( L^2 \) inner product on \( \mathbb{R}^d \). The process \( \delta_i = (\delta_i^1, \ldots, \delta_i^d)' \) is an \( \mathcal{F}_i \)-measurable strategy such that \((\delta_i, X_i) \in L^2(\Omega, \mathcal{F}_i, Q) \) for each \( i = 0, \ldots, N \). We define the difference operator \( \Delta \) according to
\[
\Delta P_i := P_{i+1} - P_i.
\]
For the \( t_N \)-forward portfolio value \( \bar{P}_i \) we then obtain
\[
\bar{P}_i = B_{i,N} P_i
\]
\[
\Delta \bar{P}_i := (\delta_i, \bar{X}_{i+1} - \bar{X}_i) = (\delta_i, \Delta \bar{X}_i).
\]
We can write
\[
\bar{u}_N = \bar{u}_i + \sum_{j=i}^{N-1} \Delta \bar{u}_j
\]
\[
\bar{P}_N = \bar{P}_i + \sum_{j=i}^{N} \Delta \bar{P}_j = \bar{P}_i + \sum_{j=0}^{N-1} (\delta_j, \Delta \bar{X}_j)
\]
and since the martingale increments \( \Delta \bar{X}_i, \Delta \bar{u}_i \perp \mathcal{F}_i \) (where \( \perp \) denotes orthogonality in \( L^2(\Omega, \mathcal{F}_i, Q) \)), we get
\[
E[(u_N - P_N)^2] = E\left[\left((\bar{u}_0 - \bar{P}_0) + \sum_{i=0}^{N-1} \{\Delta \bar{u}_i - (\delta_i, \Delta \bar{X}_i)\}\right)^2\right]
\]
\[
= (\bar{u}_0 - \bar{P}_0)^2 + \sum_{i=0}^{N-1} E\left[\{\Delta \bar{u}_i - (\delta_i, \Delta \bar{X}_i)\}^2\right].
\]
In particular this means that the strategy \( \delta \) that minimizes (8) can be chosen independently of the portfolio values \( P_i \) and that the different \( \delta_i \) can be chosen by minimizing the components \( \{\Delta \bar{u}_j - (\delta_j, \Delta \bar{X}_j)\}^2 \) for each \( j \). We have
\[
E[(u_N - P_N)^2] = (\bar{u}_0 - \bar{P}_0)^2 + \sum_{j=0}^{N-1} E\left[\{\Delta \bar{u}_j - (\delta_j, \Delta \bar{X}_j)\}^2\right].
\]
The infimum of the norm in (9) is obtained by projecting each \( \Delta \bar{u}_j \) onto the subspace of \( L^2(\Omega, \mathcal{F}_i, Q) \) spanned by \((\delta_j, \Delta \bar{X}_j)\) for \( \delta_j^j \in L^2(\Omega, \mathcal{F}_j, Q) \), \( j = 1, \ldots, d \). However, since \( \delta_i \not\in \mathcal{F}_i \) for \( i > 0 \), only \( \delta_0 \) can be known at time \( i = 0 \). We then search for \((\bar{P}_0, \delta_0)\) satisfying
\[
(\bar{P}_0, \delta_0) = \arg\min_{(\rho, \theta) \in \mathbb{R} \times \mathbb{R}^d} \left\{ (\bar{u}_0 - \rho)^2 + E\left[\{\Delta \bar{u}_0 - (\theta, \Delta \bar{X}_0)\}^2\right]\right\}.
\]
Obviously, the optimal \( \bar{P}_0 = \bar{u}_0 \). As for \( \delta_0 \), the minimizing value is given by the projection of \( \Delta \bar{u}_0 \) onto the subspace

\[
\Theta = \{ (\theta, \bar{X}_0) : \theta \in \mathbb{R}^d \}, \tag{11}
\]

of \( L^2(\Omega, F_0, Q) \) which is finite dimensional and thus closed. The orthogonality condition for the residual of the projection gives

\[
E \left[ \Delta \bar{X}_0^k \left( \Delta \bar{u}_0 - (\delta_0, \Delta \bar{X}_0) \right) \right] = 0 \tag{12}
\]

for \( k = 1, \ldots, d \). If we thus define the components of a vector \( b \in \mathbb{R}^d \) and a matrix \( M \in \mathbb{R}^{d \times d} \) according to

\[
b_k = E \left[ \Delta \bar{X}_0^k \Delta \bar{u}_0 \right] \quad M_{kl} = E \left[ \Delta \bar{X}_0^k \Delta \bar{X}_0^l \right] \quad \tag{13}
\]

for \( k, l = 1, \ldots, d \), then (12) reads \( M \delta_0 = b \). Therefore, if \( M \) is invertible, the optimal \( \delta_0 \) is given by

\[
\delta_0 = M^{-1} b. \tag{14}
\]

Let us now assume \( \bar{X}_t, \bar{u}_t \) are sampled values at \( t_i \) of processes \( \bar{X}_t \) and \( \bar{u}_t \) respectively which are square integrable \( F_t \)-martingales in continuous time and let \( t_1 = \Delta t \). Then

\[
b_k = E \left[ \Delta \bar{X}_0^k \Delta \bar{u}_0 \right] = E \left[ \langle \bar{X}^k, \bar{u} \rangle_{\Delta t} \right] \quad M_{kl} = E \left[ \Delta \bar{X}_0^k \Delta \bar{X}_0^l \right] = E \left[ \langle \bar{X}^k, \bar{X}^l \rangle_{\Delta t} \right] \quad \tag{15}
\]

where \( \langle \cdot, \cdot \rangle \) is the quadratic variation process [27, Theorem 33]. If we divide by \( \Delta t \) in (15) then in the limit \( \Delta t \rightarrow 0 \) we get

\[
b_k^0 := \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E \left[ \langle \bar{X}^k, \bar{u} \rangle_{\Delta t} \right] = \frac{d}{d\tau} E \left[ \langle \bar{X}^k, \bar{u} \rangle_{\tau} \right] \bigg|_{\tau=0} \tag{16}
\]

\[
M_{kl}^0 := \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E \left[ \langle \bar{X}^k, \bar{X}^l \rangle_{\Delta t} \right] = \frac{d}{d\tau} E \left[ \langle \bar{X}^k, \bar{X}^l \rangle_{\tau} \right] \bigg|_{\tau=0}.
\]

In the limit when the discrete time-steps tend to zero and trading in continuous time is possible, we would therefore expect the optimal trading strategy \( \delta_0 \) at \( T = 0 \) to be obtained as

\[
\delta_0 = (M^0)^{-1} b^0, \tag{17}
\]

where \( b^0 \), \( M^0 \) are the time-derivatives at \( T = 0 \) of the quadratic variation processes \( \langle \bar{X}, \bar{u} \rangle \) and \( \langle \bar{X} \rangle \) defined componentwise in (16). We will make this argument rigorous in the section below.

### 2.2 Continuous time

We now turn to quadratic hedging with martingales in continuous time. This result can be found, for example, in the surveys of Pham [26] and Schweizer [28]. Existence of an optimal strategy in this case is usually proved by invoking the Kunita-Watanabe decomposition [21], [1]. A slight difference in our exposition is that we emphasize that only \( L^2 \)-completeness is really needed.

With the same notations as in Section 2.1, assume now that \( X \) with values \( X_t \in \mathbb{R}^d \) is such that

\[
\bar{X}_t = B(t, T) X_t \quad \tag{18}
\]

is a \( F_t \)-martingale with \( \bar{X}_T \in L^2(\Omega, F_T, Q) \). Let \( u_T \) be a random variable in \( L^2(\Omega, F_T, Q) \) and define the martingale

\[
\bar{u}_t = E[u_T | F_t]. \quad \tag{19}
\]
As before, we want to hedge \( u_T \) by minimizing \( L^2(\Omega, \mathcal{F}_T, Q) \)-distance between \( u_T \) and a self-financing portfolio \( P \) investing in \( X \) according to

\[
dP_t = (\delta_t, dX_t) + (P_t - (\delta_t, X_t)) r_t dt. \tag{20}
\]

Here we assume that \( \delta_t \in L^2(\mathcal{X}, T) \), i.e., that \( \delta_t \) with values in \( \mathbb{R}^d \) is \( \mathcal{F}_t \)-predictable with finite \( L^2(\mathcal{X}, T) \)-norm. The scalar product between two elements \( \delta, \theta \in L^2(\mathcal{X}, T) \) is defined by

\[
\langle \delta, \theta \rangle_{L^2(\mathcal{X}, T)} = E \left[ \int_0^T (\delta_t, d\langle \mathcal{X} \rangle_t, \theta_t) \right], \tag{21}
\]

where \( \langle \cdot \rangle \) denotes the bracket process [27, pp 66]. Note that for two martingales \( M, N \) with values in \( \mathbb{R}^m \) and \( \mathbb{R}^n \) such that \( M_T, N_T \in L^2(\Omega, \mathcal{F}_T, Q) \) and predictable processes \( \gamma \) with values in \( \mathbb{R}^m \), \( \phi \) with values in \( \mathbb{R}^n \), the isometry relation for the stochastic integral [27] states that

\[
E \left[ \int_0^T (\gamma_t, dM_t) \int_0^T (\phi_t, dN_t) \right] = E \left[ \int_0^T (\gamma_t, d\langle M, N \rangle_t, \phi_t) \right]. \tag{22}
\]

This gives, for the scalar products in \( L^2(\Omega, \mathcal{F}_T, Q) \) and \( L^2(\mathcal{X}, T) \),

\[
\left( \int_0^T (\delta_t, dX_t), \int_0^T (\theta_t, dX_t) \right)_{L^2(\Omega, \mathcal{F}_T, Q)} = (\delta, \theta)_{L^2(\mathcal{X}, T)} \tag{23}.
\]

The \( T \)-forward value \( \bar{P} \) of the portfolio \( P \) will evolve according to

\[
\bar{P}_t = B(T, t)P_t, \quad d\bar{P}_t = (\delta_t, d\bar{X}_t). \tag{24}
\]

and we can write

\[
\bar{u}_T = \bar{u}_0 + \int_0^T d\bar{u}_t, \tag{25}
\]

\[
\bar{P}_T = \bar{P}_0 + \int_0^T (\delta_t, d\bar{X}_t). \tag{26}
\]

We want to minimize the \( L^2(\Omega, \mathcal{F}_T, Q) \)-distance between \( u_T \) and \( P_T \), i.e.

\[
E \left[ (u_T - P_T)^2 \right] = E \left[ \left( \bar{u}_0 + \int_0^T d\bar{u}_t - \bar{P}_0 - \int_0^T (\delta_t, d\bar{X}_t) \right)^2 \right] \tag{27}
\]

\[
= (\bar{u}_0 - \bar{P}_0)^2 + \sum_{i=0}^{N-1} E \left[ \left( \int_{t_i}^{t_{i+1}} d\bar{u}_t - \int_{t_i}^{t_{i+1}} (\delta_t, d\bar{X}_t) \right)^2 \right],
\]

where the last equality follows from the orthogonality of the martingale increments. As in the discrete case, this means that the optimal hedging strategies for each of the time intervals \( [t_{i-1}, t_i] \) can be chosen independently. Also, since \( \delta_t \) is predictable, only the value \( \delta_0 \) can be known at \( t = 0 \).

If we then denote the first time by \( t_1 = \tau \) and try to find the optimal strategy \( t \in [0, \tau) \), then we search for \( (\bar{P}_0, \delta) \) satisfying

\[
(\bar{P}_0, \delta) = \arg\min_{(p, \theta) \in \mathbb{R} \times \Theta_{\tau}} \left\{ (\bar{u}_0 - p)^2 + E \left[ \left( \int_0^\tau d\bar{u}_t - \int_0^\tau (\theta_t, d\bar{X}_t) \right)^2 \right] \right\}, \tag{28}
\]

were \( \Theta_{\tau} \) is the linear subspace of \( L^2(\Omega, \mathcal{F}_T, Q) \) defined as

\[
\Theta_{\tau} = \left\{ \int_0^\tau (\theta_t, d\bar{X}_t) : \theta \in L^2(\bar{X}, \tau) \right\}. \tag{29}
\]
From the isometry relation (22) we have, for \( \theta \in L^2(\bar{X}, \tau) \),

\[
\left\| \int_0^\tau (\theta_t, d\bar{X}_t) \right\|_{L^2(\Omega, \mathcal{F}_\tau, Q)} = \|\theta\|_{L^2(\bar{X}, \tau)}
\]  

(29)
and so \( \Theta_\tau \) is closed in \( L^2(\Omega, \mathcal{F}_\tau, Q) \) from the closedness of \( L^2(\bar{X}, \tau) \). The infimum of the left-hand side of (26) is then achieved by choosing \( \bar{P}_0 = \bar{u}_0 \) and by projecting the martingale \( \Delta\bar{u}_\tau := \int_0^\tau d\bar{u}_t \) onto \( \Theta_\tau \). The orthogonality condition of the projection onto \( \Theta_\tau \) now reads (note that any martingale is orthogonal to \( \Theta_\tau \) if and only if it is orthogonal to \( \bar{X} \), for \( k = 1, \ldots, d \),

\[
E \left[ \left( \int_0^\tau d\bar{X}_t^k \right) \left( \int_0^\tau d\bar{u}_t - \int_0^\tau (\delta_t, d\bar{X}_t) \right) \right] = 0.
\]  

(30)

The orthogonality and the isometry relation (22) then gives

\[
\sum_{i=1}^d E \left[ \int_0^\tau \delta_t^i d\langle \bar{X}^k, \bar{X}^i \rangle_t \right] = E \left[ \int_0^\tau d\langle \bar{X}^k, \bar{u} \rangle_t \right].
\]  

(31)

The choice of \( \tau \) is of course arbitrary in (31), since equation (26), which says that we can chose for \( k, l = 1, \ldots, d \), we retrieve

\[
\sum_{i=1}^d \delta_t^i dE \left[ \int_0^\tau d\langle \bar{X}^k, \bar{X}^i \rangle_t \right]_{|_{\tau=0}} = \frac{d}{d\tau} E \left[ \int_0^\tau d\langle \bar{X}^k, \bar{u} \rangle_t \right]_{|_{\tau=0}}.
\]  

(32)

If we then define

\[
b^0_k := \frac{d}{d\tau} E \left[ \langle \bar{X}^k, \bar{u} \rangle_t \right]_{|_{\tau=0}}
\]  

\[
M^0_{kl} := \frac{d}{d\tau} E \left[ \langle \bar{X}^k, \bar{X}^l \rangle_t \right]_{|_{\tau=0}}
\]  

(33)

for \( k, l = 1, \ldots, d \), we retrieve

\[
\delta_t = (M^0)^{-1}b^0
\]  

(34)
as conjectured in (17).

As mentioned above, quadratic hedging in the martingale case is usually treated by making use of the Kunita-Watanabe decomposition (see [21], [1]), which gives stronger results for the projection of \( \bar{u}_\tau \) onto \( \Theta_\tau \) in (27). The Kunita-Watanabe decomposition assures that \( \bar{u}_\tau \), being an arbitrary square integrable martingale, can be decomposed according to

\[
\bar{u}_\tau = \bar{u}_0 + \int_0^\tau (\theta_t^0, d\bar{X}_t) + L^0_\tau,
\]  

(35)

where \( \theta^0 \in L^2(\bar{X}, T) \) and \( L^0 \) is a zero-mean \( \mathcal{F}_\tau \)-martingale orthogonal to \( \theta^0 \) in the sense that \( \langle \theta^0, L^0 \rangle_t = 0, \ t \in [0, T] \). In view of (35) we rewrite (27) as

\[
(P_0, \delta) = \arg \min_{(p, \theta) \in \mathbb{R} \times \Theta} \left\{ (\bar{u}_0 - p)^2 + E \left[ \left( \int_0^\tau ((\theta^0_t - \theta_t^0), d\bar{X}_t) \right)^2 \right] + E \left[ (L^0_\tau)^2 \right] \right\}.
\]  

(36)

It is obvious that (36) is minimized by setting \( p = \bar{u}_0 \) and \( \theta_t = \theta_t^0 \), so that the optimal \( \delta_t \) in (27) is equal to \( \theta_t^0 \) from the Kunita-Watanabe decomposition (35) for \( \tau \in [0, T] \).
3 Application to a stochastic volatility model with spot jumps

Let us consider a liquidly traded asset $S$ with dynamics in $(\Omega, \mathcal{F}_t, Q)$ following a stochastic volatility model with Poisson driven jumps in the log-spot process. We let

$$dS_t = (r_t - q_t)S_t dt - \left( \int_{\mathbb{R}} (e^z - 1) \nu(z) dz \right) S_t dt$$

$$+ \sqrt{y_t} S_t dW^S_t + S_t (e^{dZ_t} - 1)$$

$$dy_t = \kappa (\eta - y_t) dt + \theta \sqrt{y_t} dW^y_t,$$

where $(W^S, W^y)$ is a Brownian motion with correlation factor $\rho$, and the values $\kappa, \eta$ and $\theta$ are non-negative constants. As before $r_t$ is the risk-free rate of return and $q_t$ represents a continuous dividend yield. We let $Z$ be a compound Poisson process independent of $(W^S, W^y)$ with intensity $\lambda$ and jumps drawn from an exponential distribution on $(-\infty, 0]$ with density $\nu$ given by

$$\nu(z) = \frac{1}{\mu} \left\{ z \leq 0 \right\} e^{\mu z},$$

for some $\mu > 0$. If the jump intensity $\lambda \equiv 0$, then (37) is the model introduced in finance by Heston [17]. Bates [4] uses the same model but with normally distributed jumps. Kou [20] finds semi-analytical pricing formulas in a pure compound Poisson-process with a combination of positive and negative exponential jumps. In our numerical tests below in Section 6 we will use the model (37) both with $\lambda \equiv 0$ and with the jump part included (i.e., $\lambda$ not restricted to be zero). In the sequel, the case $\lambda \equiv 0$ will be referred to as the SV model (for stochastic volatility) and the case $\lambda \not\equiv 0$ will be referred to as the SVJ model (for stochastic volatility with jumps).

The model (37) belongs to a well known class of stochastic volatility models named affine models by Duffie et al. [11]. These models are popular in finance as the characteristic function of the log process can be calculated as the solution to a Ricatti differential equation which can sometimes (as in our case) be analytically calculated ([17], [11]). Moreover, as discovered by Carr and Madan [6], the characteristic function in turn allows for efficient numerical evaluation of prices of European put- and call options written on underlying assets with such dynamics. By “price” we here refer to the quantity

$$u(t, S_t, y_t) := B(t, T) \mathbb{E} \left[ h(S_T) \mid \mathcal{F}_t \right]$$

where, for some strike value $K$, $h(S_T) = \max(0, S_T - K)$ for a call option and $h(S_T) = \max(0, K - S_T)$ for a put option. We will use a more recent formulation due to Attari [2] to calculate these prices under the model 37. Attari expresses option prices in terms of an indefinite integral on the interval $[0, \infty)$. Transforming this integral into a definite integral on the interval $[0, 1]$ via a change of variables, allows us to avoid error from truncation of the integral. We can then calculate the definite integral to desired precision with the adaptive Gauss-Kronrod quadrature. Thus, by a change of variables and an adaptive quadrature, we are able to calculate model option prices with high precision regardless of the choice of parameters. Derivatives of $u$, when needed, are calculated by differentiating the mentioned integral under the integral sign, and a subsequent numerical evaluation of the resulting integral.

3.1 Self-financing portfolios and dividend yield

A complicating feature of the dynamics of the underlying $S$ in (37) is the dividend yield $q$ that represents a flux of cash paid to the holder of a long position in $S$. Suppose we want create a self-financing portfolio $P$ with positions in $S$ and some a contract $u$ as in (39). Let $\delta^S$ number of shares $S$ in the portfolio and $\delta^u$ the number of units held in $u$. The resulting part of the portfolio
will be invested in the bank account \( B \). With the dividend yield included, the dynamics of \( P \) in continuous time can be written
\[
dP_t = \delta^S_t dS_t + \delta^u_t du_t + (P_t - \delta^S_t S_t - \delta^u_t u_t) r_t dt + \delta^S_t S_t q_t dt,
\]
where \( u_t \) is the value of \( u \) at \( t \). Let us define
\[
Q(t, T) := e^{-\int_t^T q_s ds},
\]
\[
x_t := Q(t, T) S_t,
\]
\[
\bar{x}_t := B(t, T) x_t = B(t, T) Q(t, T) S_t,
\]
so that \( \bar{x} \), the \( T \)-forward value of \( S \), is a martingale. The \( T \)-forward portfolio value \( \bar{P}_t = B(t, T) P_t \) will evolve according to
\[
d\bar{P}_t = B(t, T) \{\delta^S_t (dS_t - S_t (r_t - q_t) dt) + \delta^u_t (du_t - u_t r_t) dt\}
\]
\[
= Q^{-1}(t, T) \delta^S_t (B(t, T) Q(t, T) S_t) + \delta^u_t (B(t, T) u_t)
\]
which can be expressed as
\[
d\bar{P}_t = \delta^S_t d\bar{x}_t + \delta^u_t d\bar{u}_t
\]
where
\[
\delta^S_t = Q^{-1}(t, T) \delta^S,
\]
\[
\bar{u}_t = B(t, T) u_t.
\]
The results for quadratic hedging in continuous time from above can then be applied by projecting onto the increments of the martingale \( (\bar{x}, \bar{u}) \) and then use the second equality in (43) to obtain the number \( \delta^S \) of shares \( S \) that should be held in the portfolio ones \( \delta^x \) is calculated.

Now consider hedging in discrete time using \( S \) and \( u \) at dates \( t_i \) with \( t_0 = 0 \) and \( t_N = T \) as in Section 2.1. This amounts to creating a portfolio in which the strategy \( \delta \) is piecewise constant with value \( \delta_i \) on the interval \([t_i, t_{i+1}]\). In the absence of dividends, i.e. if \( q \equiv 0 \), the expression (4) with \( X = S \) will describe the exact evolution of such a portfolio. However, our way of modelling dividends as a continuous flux proportional to the spot \( S \) means that the total dividend amount paid in the interval \([t_i, t_{i+1}]\) can not be expressed as a deterministic function of \( S_i \) and \( S_{i+1} \). If we nonetheless make the approximation
\[
\int_{t_i}^{t_{i+1}} S_t q_t dt \approx S_{i+1} (e^{\int_{t_i}^{t_{i+1}} q_s ds} - 1) = S_{i+1} (Q_{i+1, i} - 1),
\]
where we use the notation \( S_i = S_{t_i} \) as before and also let \( Q_{i,j} = Q(t_i, t_j) \), then we can express the evolution of a self-financing portfolio \( P \) investing in \( S \) in discrete time as
\[
\Delta P_t = P_{i+1} - P_i = \delta^S_t (S_{i+1} - S_i) + \delta^u_t (u_{i+1} - u_i) + (P_i - \delta^S_i S_i - \delta^u_i u_i) (B_{i+1, i} - 1)
\]
\[
+ \delta^S_i S_{i+1} (Q_{i+1, i} - 1).
\]
For the \( T \)-forward value \( \bar{P}_t = B_{i, N} P_t \) of the portfolio, equation (46) directly leads to
\[
\Delta \bar{P}_t = \bar{P}_{i+1} - \bar{P}_i
\]
\[
= Q_{i,N}^{-1} \delta^S_t (B_{i+1, N} Q_{i+1, N} S_{i+1} - B_{i, N} Q_{i, N} S_i) + \delta^u_t (B_{i+1, N} u_{i+1} - B_{i, N} u_i)
\]
\[
= \delta^S_t (\bar{x}_{i+1} - \bar{x}_i) + \delta^u_t (\bar{u}_{i+1} - \bar{u}_i) = \delta^x_t \Delta \bar{x}_i + \delta^u_t \Delta \bar{u}_i,
\]
with \( \bar{x} \) and \( \bar{u} \) from (43). We thus see that the approximation (45) leads to a portfolio evolution (47) in discrete time in terms of increments of the martingales we use for the portfolio in continuous time. To apply the results from Section 2.1 we will use this approximation to obtain a martingale representation in discrete time for the part of the portfolio that invests in \( S \).
3.2 Optimal quadratic hedging ratios

We now consider two different contracts $u$ and $v$, both defined as in (39), but for possibly different maturities $T$ and different pay-off functions $h$. We will attempt to hedge the contract $u$ by positions in either just $S$ or in both $S$ and $u$. We will refer to the case when only the spot process $S$ is used as quadratic delta hedging and to the case where both $S$ and $u$ are used as quadratic delta-vega hedging. As above, will use $u_0$ as shorthand for the value contract $u$ at $t$ and likewise for $v$. The forward values $\bar{u}, \bar{v}$ of the contracts are as in (44), and the dividends in $S$ are handled as described above by introducing the forward value $\bar{x}$ and the hedging ratio $\delta^x$ from (44). All forward values are taken with respect to the same future date $T$ - which we choose equal to the maturity of the hedging contract $v$ - as is assumed in the expressions (6), (24) for the evolution of the forward values of the hedging portfolios.

3.2.1 Discrete time

Now consider optimal quadratic hedging of $u$ with $X = (x, u)$ at discrete times $0 = t_0 < t_1 < \ldots < t_N = T$ as in Section 2.1. To obtain the optimal hedging ratios at $t_0 = 0$ according to (13), we need to calculate $E \left[ \Delta X^b_0 \Delta X^b_0 \right]$ and $E \left[ \Delta X^b_0 \Delta u^b_0 \right]$ for $k, l = 1, 2$, i.e.,

$$
\begin{align*}
     b_1 &= E \left[ (\bar{x}_0 - \bar{x}_{t_1}) (\bar{u}_0 - \bar{u}_{t_1}) \right] = E \left[ \bar{x}_{t_1} \bar{u}_{t_1} - \bar{x}_0 \bar{u}_0 \right] \\
     b_2 &= E \left[ (\bar{u}_0 - \bar{u}_{t_1}) (\bar{v}_0 - \bar{v}_{t_1}) \right] = E \left[ \bar{u}_{t_1} \bar{v}_{t_1} - \bar{u}_0 \bar{v}_0 \right] \\
     M_{11} &= E \left[ (\bar{x}_0 - \bar{x}_{t_1})^2 \right] = E \left[ \bar{x}^2_{t_1} - \bar{x}_0^2 \right] \\
     M_{12} &= M_{21} = E \left[ (\bar{x}_0 - \bar{x}_{t_1}) (\bar{v}_0 - \bar{v}_{t_1}) \right] = E \left[ \bar{x}_{t_1} \bar{v}_{t_1} - \bar{x}_0 \bar{v}_0 \right] \\
     M_{22} &= E \left[ (\bar{v}_0 - \bar{v}_{t_1})^2 \right] = E \left[ \bar{v}^2_{t_1} - \bar{v}_0^2 \right].
\end{align*}
$$

(48)

Of course, if hedging is done with just $x$, only $b_1$ and $M_{11}$ are needed. If the joint density of the pair $(S_t, y_t)$ and the functions $u, v$ are known, the mean values in (48) are obtained by a mere integration with respect to this density. Suppose that $\varphi(t, S, y)$ is the density of $(S_t, y_t)$. Then, of course,

$$
E \left[ \bar{v}(t_1, S_{t_1}, y_{t_1}) \bar{u}(t_1, S_{t_1}, y_{t_1}) \right] = \int \bar{v}(t_1, S, y) \bar{u}(t_1, S, y) \varphi(t_1, S, y) \, dS \, dy
$$

(49)

and likewise for the other mean values. As mentioned above, the characteristic function of $(S_t, y_t)$, which corresponds to the Fourier transform of $\varphi$, is known analytically. We can therefore calculate $\varphi$ via Fourier inversion with FFT, and evaluate (49) by numerical quadrature in which values $\bar{u}, \bar{v}$ in the integral are calculated using the formulation [2] as described above. This is detailed a little further in Appendix B.

3.2.2 Continuous time

Let us now consider quadratic hedging in continuous time of $\bar{u}$. First we calculate the optimal hedge when we only use the spot process $S$ in the portfolio. According to (33), we need to calculate the time-derivative of the mean of the quadratic variation processes $(\bar{u}, \bar{x}, S)_t$ and $(\bar{x}, \bar{x})_t$. These values are readily obtained by a direct application of Itô’s lemma. We carry out the necessary calculations in Appendix A. Equations (34), (33) and (84), (86) yield the optimal number $\delta^S_t$ of shares $x_t$ to hold at $t = 0$. Changing into the optimal number $\delta^S_t$ of shares $S_t$ using (44) yields

$$
\delta^S_t = \frac{y_t \partial_x u + \frac{1}{2} \rho \theta y_t \partial_y u + \frac{1}{2} \int [u(t, S_t e^z, y_t) - u(t, S_t, y_t)] (e^z - 1) \lambda \nu(z) \, dz}{y_t + \int (e^z - 1)^2 \lambda \nu(z) \, dz}.
$$

(50)

In the special case of a deterministic variance $y_t$ (corresponding to $\theta = \kappa = 0$) and no jumps (corresponding to jump intensity $\lambda = 0$) we retrieve the classical BS hedge $\delta^S_t = \partial_S u$. For a deterministic variance but positive jump intensity $\lambda$ we get a special case of the result obtained
by Tankov, p 346) when the underlying asset follows an exponential Lévy process. In the case of the Heston model, where there are no jumps but the variance is stochastic \( (\theta > 0) \), we get

\[
\begin{align*}
\delta_t^S &= \partial_S u + \frac{1}{2\sigma^2} \rho \theta \partial_y u, \\
\end{align*}
\]

which can be seen as a BS delta corrected by a term taking into account the correlation \( \rho \) between (the Brownian motions of) the spot- and variance processes.

We now consider hedging with a portfolio investing in \( X = (x, v) \). We get, at \( t = 0 \), from (85)

\[
\begin{align*}
b_1^0 &= \frac{d}{dt} E [(\tilde{u}; \tilde{v})_t] = Q(t, T) B^2(t, T) \left[ S_t^2 y_t \partial_S u + \rho S_t \theta y_t \partial_y u \right] \\
&+ S_t \left[ u^1(t, S_t e^x, y_t) - u^1(t, S_t, y_t) \right] (e^x - 1) \lambda \nu(z) dz. \\
\end{align*}
\]

and from (85),

\[
\begin{align*}
b_2^0 &= \frac{d}{dt} E [(\tilde{u}; \tilde{v})_t] = B^2(t, T) \left[ S_t^2 y_t \partial_S u + \theta y_t \partial_y u + \rho S_t \theta y_t \partial_S v + \partial_S u \partial_y v \right] \\
&+ \lambda \int \left[ u(t, S_t e^x, y_t) - u(t, S_t, y_t) \right] \nu(z) dz. \\
\end{align*}
\]

As for the matrix \( M^0 \) we have that \( M^0_{11} = \frac{d}{dt} E [(\tilde{x}; \tilde{x})_t] \) is given by (86) and that \( M^0_{21} = M^0_{21} = \frac{d}{dt} E [(\tilde{x}; \tilde{v})_t] \) is obtained by replacing \( u \) by \( v \) and \( B_{T1} \) by \( B_{T2} \) in (52). The element \( M^0_{22} \) is given by setting \( u = v \) in (84):

\[
\begin{align*}
M^0_{22} &= \frac{d}{dt} E [(\tilde{v}; \tilde{v})_t] = B^2(t, T) \left[ S_t^2 y_t \partial_S v + \theta y_t \partial_y v \right] \\
&+ \lambda \int \left[ v(t, S_t e^x, y_t) - v(t, S_t, y_t) \right] \nu(z) dz. \\
\end{align*}
\]

In order to obtain the optimal hedging ratios in continuous time from the expressions above, we need to calculate the derivatives of prices and integrals of the type \( \int u(t, S_t e^x, y_t) e^x \nu(z) dz \). Prices and derivatives are calculated using Attari’s formulation \( 2 \), as described in the beginning of Section 3. Integrals involving the price \( u \) are evaluated numerically by an adaptive quadrature in Matlab. The calculation of the hedging ratios in continuous time is therefore easier than in the discrete case, since we do not need to bother about the density of \( (S_t, y_t) \) at future dates. It can also be noted that Kallsen and Viethauer \( 18 \) obtain semi-analytical expressions for optimal quadratic hedging ratios in affine models. Nonetheless, we have used the approach described above which is straight-forward once an efficient pricing algorithm for affine models is implemented.

4 Calibration of model parameters

We first note that the model (37) requires knowledge of the short interest rate \( r_t \) and the dividend yield \( q_t \). The interest rate is calculated via a derivation of a zero-coupon curve in Euro provided by Svenska Handelsbanken AB. The dividend yield \( q_t \) is then obtained from a differentiation of the forward prices, which are calculated from the zero-coupon values and option prices via the put-call parity applied to the quoted put- and call price pair with strike closest to the spot value.

Let \( \alpha \in \mathbb{R}^3 \) be a real valued vector containing the parameters of the model (37) except from the state variables \( S \) and \( y \), the interest rate \( r \) and the dividend yield \( q \). We can let \( \alpha = (\kappa, \eta, \theta, \rho, \lambda, \nu) \) for the SVJ model and \( \alpha = (\kappa, \eta, \theta, \rho) \) for the SV model without jumps. Suppose we observe the market at the date \( t_{i+1} \) and that market prices on call options at both dates are observable at the maturity-strike pairs \( (T^j, K^j) \), \( j = 1, \ldots, n \) (where the different \( T^j \) and \( K^j \) are not in general distinct). Denote by \( \tilde{u}_{t_{i+1}}^j \) the market mid-price (i.e., the average of the bid- and the ask price) of a call option of maturity and strike \( (T^j, K^j) \) and let \( u_{t_{i+1}}^j(y_{i+1}, \alpha) \) be the price of the corresponding
call option in the model (37) at $t_{i+1}$ with spot value $S_{i+1}$, parameters $\alpha$ as defined above and where $y_{i+1}$ is the value of the variance process at $t_{i+1}$. Let us form a vector $\text{err}_p$ of differences between model prices and market quotes with components given by

$$\text{err}_p^j := u_i^j - u_i^{j+1}(y_{i+1}, \alpha)$$

for $j = 1, \ldots, n$. A standard approach for determining $y_{i+1}$ and the parameters $\alpha$ is to apply some optimization technique to attempt to find values $(y_{i+1}, \alpha)$ that bring each component $\text{err}_p^j$ close to zero. A common choice to use a least squares optimization that attempts to solve the problem

$$\alpha_{i+1} = \arg\min_{\alpha \leq \bar{\alpha}, \bar{\alpha}} \|\text{err}_p\|^2$$

(56)

where $\|\cdot\|$ is the $L^2$-norm in $\mathbb{R}^n$ and $\alpha, \bar{\alpha}$ are vectors whose components give upper and lower bounds for the values in $\alpha$. When such a parameter optimization is performed using market data on consecutive dates, it is often observed that the values of the parameters fluctuate considerably from one day to another. One way of avoiding this is to introduce a penalty on the parameters if they deviate from their previous values. In our case, we let $\alpha_i$ be the vector of parameter values obtained at the previous date $t_i$ and define a vector $\text{err}_\alpha$ of parameter deviations according to

$$\text{err}_\alpha^j := \frac{S_{i+1}}{\max(\alpha_i^j, 0.1)} (\alpha_i^j - \alpha^j),$$

(57)

for $j = 1, \ldots, d$. The multiplier $S$ is only used as a scaling parameter. The $\alpha_i^j$ is used in the denominator to give a relative error in the parameters rather than an absolute error, and we set a lower bound on the denominator at 0.1 to avoid large error terms if a parameter value is close to zero. We can then define a penalized least squares problem according

$$\alpha_{i+1} = \arg\min_{\alpha \leq \bar{\alpha}, \bar{\alpha}} \left(\|\text{err}_p\|^2 + \beta^2 \|\text{err}_\alpha\|^2\right),$$

(58)

where $\beta$ is some positive constant and the second norm is the $L^2$-norm in $\mathbb{R}^d$. This is a Tikhonov regularization (see for example Vogel [32]) of the unpenalized problem (58). Apart from making the problem more well behaved by the adding the function $\|\text{err}_\alpha\|^2$ which is convex in $\alpha$, the penalty also incorporates “memory” into the optimization by forcing the new parameter values to be closer to the ones obtained at the previous step. Note that we do not include a penalty on the value $y_{i+1}$ of the variance process, since this is a state variable and is supposed to fluctuate as the process $(S, y)$ from (37) evolves in time. The motivation for forcing parameters to fluctuate less over time is that the parameters that give the best fit to the data observable at a given time, might not be the parameters that provide the best hedge. A good hedging results is likely to be given by a set of parameters that give a good fit to market data today as well as tomorrow. If the parameters that give an optimal fit when using just the present data fluctuate a great deal from one calibration occasion to the next, then the best guess for such a set of parameters might be given by incorporating information from several past observations. The penalty on the parameters in (58) is an indirect and simple method of introducing information from earlier observations into the calibration of a new set of parameters. A more direct way would be to let the vector of errors $\text{err}_p$ contain terms of the type (56) from previous dates $t_i, t_{i-1}, \ldots, t_{i-k}$ for some $k$, possibly with different weights on the error from different dates. This method would be more computationally costly, since model prices for a larger set of options would need to be calculated. A natural idea would therefore be to use a filtering method in which the previous parameter estimation is updated when new data becomes available. Indeed, Lindström et al. [23] use a Kalman filter method to calibrate parameters in models similar to the ones we use.

The vector of errors (55) only takes into account how well the model prices fit observed market prices for a particular choice of parameters $\alpha$. However, we want to use the model for hedging purposes, the parameters that give the best fit to quoted prices might not be the parameters
\[
\begin{array}{c|cc}
\text{Parameter value} & \beta = 0 & \beta = 5 \\
\gamma = 0 & 1A & 1B \\
\gamma = 0.75 & 2A & 2B \\
\end{array}
\]

Table 1: Parameter values in the four calibration cases (1A, 1B, 2A and 2B) considered.

that will give the best hedge. One way include hedging performance in the calibration is to let
the objective function include the hedging error that a given choice of parameters would have
given between the last rehedging time \( t_i \) and the present time \( t_{i+1} \). Let \( \Delta P^j_i \) be the increment
of a portfolio evolving according to (46) with \( P^j_i = \hat{u}^j_i \) and a delta \( \delta^j \) calculated according to
(50) for \( u^j_i(y_i, \alpha) \), thus using the parameters \( \alpha \) and the value of the variance process \( y \), obtained
from calibration at the previous date \( t_i \). For each observable market quote, we can then define an
observed hedging error as a function of \((y_{i+1}, \alpha)\) according to
\[
\text{err}^j_h = \Delta \hat{u}^j_i - \Delta P^j_i,
\]
where
\[
\Delta \hat{u}^j_i := \hat{u}^j_{i+1} - \hat{u}^j_i.
\]

We can then define a least squares problem that takes the observed hedging error into account
according to
\[
\alpha_{i+1} = \arg\min_{\frac{\gamma}{2} \leq \alpha \leq \frac{\bar{\alpha}}{2}} (1 - \gamma)^2 \| \text{err}_p \|^2 + \gamma^2 \| \text{err}_h \|^2 + \beta^2 \| \text{err}_\alpha \|^2).
\]

where \( \gamma \in [0, 1] \). Again, filtering can be used, which is done by Lindström and Guo [22] who apply
Kalman filtering techniques to solve the parameter calibration problem with the hedging error \( \text{err}_h \) present.

We will compare results from four different calibration cases, with and without the parameter
penalty \( \text{err}_\alpha \) and with and without the hedging error term \( \text{err}_h \). The cases we consider are defined
by the parameter choices for \( \beta \) and \( \gamma \) given in Table 1. These parameter values are subjectively
chosen, i.e., we have not used any mathematical criteria when choosing the values. The value \( \beta = 5 \)
has been chosen to be large enough to give a clear reduction in the fluctuations of parameter values
over time, but small enough to still allow for changing values. The value \( \gamma = 0.75 \) has been chosen
so that the error \( \gamma^2 \| \text{err}_h \|^2 \) is larger than the error \( (1 - \gamma)^2 \| \text{err}_p \|^2 \) without making the latter
insignificant. Empirically, we obtain an average value over our roughly 1600 calibration dates (see
Table 2 below) of around 2.5 for the ratio
\[
\frac{\gamma^2 \| \text{err}_h \|^2}{(1 - \gamma)^2 \| \text{err}_p \|^2}
\]
for the case 1B in both the SV model without jumps and the SVJ model, and an average of around
3 for the case 2B.

The optimization problems (56), (58) and (61) are solved using the trust-region Newton method
available in Matlab\(^1\) in the lsqnonlin (non-linear least squares) function. (The references given in
Matlab’s user manual for the trust-region algorithm are [9], [8].)

5 Improving delta hedging by simple regression

A comment often made in relation to the use of the BS model as a tool for calculating hedging
ratios, is that although the model has its flaws, traders compensate for the shortcomings by using
their experience and intuition for the market’s behaviour. In this section, we suggest our own way of

improving a hedging strategy by making a simple regression on the strategy’s recent performance. Our motivation is that since we intend to use the BS model as our “benchmark” when we test the hedging strategies from the sections above on real data, we also want to test our strategies when “formalized intuition” (in the form of linear regression) is used in an attempt to improve them. This will give us an idea of whether or not the BS model for calculating hedging ratios can be improved without entirely abandoning the model. We will apply the same technique to test if the delta hedging strategies stemming from the quadratic hedging theory can be improved.

What we find appealing about the methodology we propose below, is that we only make use of information inherent in the model itself; no other information than the implicit BS volatility needs to be calculated from market data. Equivalently, when we apply the same technique to the SVJ and the SV model, we only make use of the models’ internal parameters. Another example of a formalized way of improving the BS delta hedge can be found in Vähämaa [31], but there an extra parameter needs to be estimated from market data.

Suppose now that $\hat{u}$ is the market’s mid-price for a European call option at time $t$ and let $\sigma$ be its corresponding BS volatility. Denote by $u_t(S_t, \sigma)$ the function that gives the option price in the BS model, so that $u_t(S_t, \sigma) = \hat{u}$. A major issue with the BS framework is that the volatility parameter $\sigma$ that yields the correct price $u$ at time $t$ will not remain constant, but will change when the spot price $S$ moves, which indicates that not only the spot derivative $\partial_S u$ but also the change of this derivative with the volatility $\sigma$, i.e. $\partial_{S\sigma} u$, might have an impact on the optimal $\delta$ to chose in (40) or (46). Also, in the quadratic hedging case the “volatility derivative” $\partial_{\sigma} u$, enters the expression for the optimal $\delta$ in (50). The process $\gamma$, however, represents the instantaneous variance of the model (37) and in a BS framework, practitioners tend to work with the derivative $\partial_{\sigma} u$, where $\sigma$ represents volatility rather than variance. We might therefore also want to consider if the derivative $\partial_{\sigma} u$ in the BS model could improve the standard delta hedging strategy. The simple idea we exploit here is thus to search for a $\delta$ that takes into account not only the spot derivative $\partial_S u$, but also the derivatives $\partial_{\sigma} u$ and $\partial_{S\sigma} u$. To this end, we will search for a delta hedging rule $\delta$ given by

$$
\delta_t = \alpha_1 \partial_S u + \alpha_2 \partial_{\sigma} u + \alpha_3 \partial_{S\sigma} u,
$$

where $\alpha_i$, $i = 1, 2, 3$, are constants that we will attempt to determine at each instant $t$. When we work with the stochastic volatility model (37) we will try to improve the delta hedging by correcting the value $\delta^S_t$ from (50) and search for a strategy

$$
\delta_t = \alpha_1 \delta^S_t + \alpha_2 \partial_{\sqrt{\sigma}} u + \alpha_3 \partial_{\sqrt{\sigma^2}} u,
$$

where $u$ is evaluated using parameters calibrated as described at the end of Section 3.2.2.

Assume now that we are at a time $t$ and that we have observations of option prices at a number of times $t_0 < t_1 < \ldots < t_{m-1} < t_m = t$. Suppose that for each $i \in \{1, \ldots, m\}$ a number $n_i$ of contracts where observable in the market at time $t_{i-1}$ as well as at $t_i$ and let $\hat{u}_i^1, \ldots, \hat{u}_i^{n_i}$ be their quoted prices at $t_i$. We form the differences $\Delta \hat{u}_{i-1}$ from (60) for $j = 1, \ldots, n_i$. For each contract $\hat{u}_i^j$ we form a replicating portfolio $P_i^j$ holding $\delta_i^j$ positions in $S$ at time $t_i$. The portfolio increment will evolve according to (46) - with $\delta^S_i$ replaced by $\delta_i^j$ and $\delta^u \equiv 0$ - and can be rewritten as

$$
\Delta P_i^j = \delta_i^j \left( S_{i+1} Q_{i+1,i} - S_i B_{i,i+1} \right) + P_i^j (B_{i,i+1} - 1)
$$

(65)

For notational convenience, let us put

$$
D_i := S_{i+1} Q_{i+1,i} - S_i B_{i,i+1}.
$$

(66)

Suppose that $\delta_i^j$ in (65) is on the form (63) and let us denote the model function that corresponds to the contract $\hat{u}_i^j$ by $u_i^j$. If a portfolio increment on the form (65) for a portfolio value $P_{i-1} = \hat{u}_{i-1}^j$ were to perfectly match the market’s contract price increment (60) with the strategy (63), we would have

$$
\left( \alpha_1 \partial_S u_{i-1}^j + \alpha_2 \partial_{\sigma} u_{i-1}^j + \alpha_3 \partial_{S\sigma} u_{i-1}^j \right) D_i = \hat{u}_i^j - \hat{u}_{i-1}^j (B_{i,i+1} - 1).
$$

(67)
Denote the right-hand side in (67) by
\[ \beta^j_i := \hat{u}^j_i - \hat{u}^{j-1}_i (B_{i,i+1} - 1) \] (68)
and let \( \beta \) be a column vector given by \( \beta = (\beta^1_1, \ldots, \beta^m_1, \ldots, \beta^1_n, \ldots, \beta^m_n)' \). Likewise, we use the notation
\[
\Lambda_{k,i}^j = \partial S u_{j,i}^{i-1} D_i \\
\Lambda_{k,ij} = \partial \sigma S u_{j,i}^{i-1} D_i \\
\Lambda_{k,ij} = \partial S \sigma u_{j,i}^{i-1} D_i
\] (69)
and form the column vectors \( \Lambda_k = (\Lambda_{k,1}^1, \ldots, \Lambda_{k,m}^m)' \) for \( k = 1, 2, 3 \). If our portfolio strategy were to perfectly match all contract increments \( \Delta C_{j,i}^i \), \( i = 1, \ldots, m; j = 1, \ldots, n \), we would now have
\[ \Lambda \alpha = \beta, \] (70)
where \( \alpha = (\alpha_1, \alpha_2, \alpha_3)' \) is a column vector and \( \Lambda \) is a matrix with columns given by the vectors \( \Lambda^k, k = 1, 2, 3 \) according to \( \Lambda = (\Lambda^1, \Lambda^2, \Lambda^3) \). Of course, there will not in general be a vector \( \alpha \) such that the equality (70) holds. What we will do is to find the \( \alpha \) that satisfies (70) in a least squares’ sense:
\[ \alpha = \arg\min_{\alpha \in \mathbb{R}^3} \| \Lambda \alpha - \beta \|, \] (71)
where \( \| \cdot \| \) is the \( L^2 \) norm in \( \mathbb{R}^d \) with \( d \) being the number of elements in \( \beta \). At each hedging time \( t \) we make this linear regression to find the \( \alpha \) that would have been optimal using the past 20 hedging occasions that we have available. That is, we use 20 time-steps, so \( m = 20 \). We then apply the hedging strategy (46) between time \( t_m \) and a following time \( t_{m+1} \) with \( \delta \) calculated from (63) in the BS case and from (64) in the stochastic volatility case. We have not tested for other values than \( m = 20 \) and we therefore do not know if some other choice of \( m \) would yield better or worse results for the data we have used.

6 Tests on market data on European call options

6.1 Constructing test portfolios

We use market data on call options written the Euro Stoxx 50 Index (the European equity index with the most liquid option market). Our data consists of market quotes registered three times daily at (approximately) 10.07, 13.07 pm and 17.07 CET on the days open for trading in the period from January 3rd 2011 to April 3rd 2013. This gives us a data set of options sampled at 1645 distinct times. We need the first 21 observations to calculate the first values in our “regression strategies” from Section 5 above, which leaves us with 1624 remaining observation dates. We only use options of maturity up to a maximum of two years, since we did not consistently have quotes for longer maturities during the whole period. We have quotes of options until the open trading day preceding the day of expiration, meaning that our shortest maturities will be around 19 hours from our observations at around 17.07 CET the day before expiry until expiry at 12.00 CET on the last day.

In order to test our hedging strategies, we first calibrate the parameters of the model (37) to market data as described above in Section 4. As mentioned, we use the model both with and without jumps, i.e., we calibrate the pure SV model (setting \( \lambda \equiv 0 \)) and we calibrate the SVJ (allowing \( \lambda > 0 \)). Once the model’s parameters are calibrated, we can calculate the hedging strategies from Section 3.2.1 in the discrete time case and 3.2.2 for the continuous. We will refer

2 Courtesy of Svenska Handelsbanken AB.
3 A small number of the trading days were missing from our data.
I. six different subsets that are quoted both at the market. Each subset \(K/S\) selected according to “moneyness”, \(i\), for \(t\). We will consider a series hedging portfolios that only last for one time-step, from a time and some other option as delta-vega hedging.

We will test our strategies on observed call contract increments of the type \((60)\) for our different subsets of market options. If a subset at a given time contains less than 10 quoted contracts, that particular option set will not be considered. We summarize the six different choices of intervals \(k\) of sampling times we select the call options that are quoted both at \(t_{i-1}\) and \(t_i\). Among those quotes, we select for each date pair \(\{t_{i-1}, t_i\}\) six different subsets \(I^j_{l}\) of maturity-strike pairs among the quoted options in the market. Each subset \(I^j_{l}\) will be defined by a range of maturities and strikes, where the strikes are selected according to “moneyness”, \(K/S_{t_{k-1}}\) where \(K\) is the market strike and \(S\) the spot price. We chose a moneyness interval \((k_{min}, k_{max})\) and a maturity interval \((T_{min}, T_{max})\) and let

\[
I^j_{l} = \{ (T, K) \text{ such that } T_{min} < T < T_{max}, k_{min} < K/S_{t_{k-1}} < k_{max} \text{ and a call option } C(T, K) \text{ is quoted at } t_{i-1} \text{ and } t_i \}. \tag{72}
\]

We will test our strategies on observed call contract increments of the type \((60)\) for our different subsets of market options. If a subset at a given time contains less than 10 quoted contracts, that particular option set will not be considered. We summarize the six different choices of intervals \((k_{min}, k_{max})\) and \((T_{min}, T_{max})\) we use in Table 2, where we give the total number of increments used for each subset and the total number of options summed over all hedging times. We have chosen “moneyness” intervals \((k_{min}, k_{max})\) in units of one standard deviation of a log-normal diffusion process with constant volatility 20% observed at time \(T = 0.5\) and \(T = 2\) respectively. For \(T = 0.5\), the standard deviation is approximately 10% of the initial value and at \(T = 2\) it is approximately 20%. In the sequel, we will refer to the options with maturities up to 0.5 years as short dated options and those of maturities above 0.5 years and up to 2 years as medium dated options. For each of the two time intervals, we will call the category of options in Table 2 with lowest moneyness as ITM (“in the money”) options, whereas those in the middle section (containing moneyness level 1) will be called ATM (“at the money”) options (although this term is usually not used for such a large span of strikes) and those with higher moneyness OTM (“out of the money”) options.

The methodology we use, for each subcategory of options \(l = 1, \ldots, 6\) represented in Table 2, is to try to replicate a portfolio of total capital \(c\) and with the same amount of money invested in each of the contracts \(\hat{u}_{t_{i-1}}(T, K)\) quoted at \(t_{i-1}\) for \((T, K) \in I^j_{l}\). The values \(\hat{u}_{t_{i-1}}\) here denotes market mid-prices, the arithmetic average of the bid- and ask prices quoted in the market. To invest the same amount of money in each option, we create a portfolio with

\[
\omega_{t_{i-1}}(T, K) = \frac{c}{(\#I^j_{l}) \hat{u}_{t_{i-1}}(T, K)}, \tag{73}
\]

number of contracts \(\hat{u}_{t_{i-1}}(T, K)\) for each \((T, K) \in I\), where \((\#I^j_{l})\) is the total number of options in \(I^j_{l}\). At a given observation time \(t_{i-1}\), we construct a hedging portfolio \(P_{t_{i-1}}(T, K)\) for each

<table>
<thead>
<tr>
<th>Expiry</th>
<th>Moneyness</th>
<th>ITM</th>
<th>ATM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short dated</td>
<td>0 &lt; (T \leq 0.5)</td>
<td>1285 (25658)</td>
<td>1624 (73277)</td>
<td>1091 (19294)</td>
</tr>
<tr>
<td>Medium dated</td>
<td>0.5 &lt; (T \leq 2)</td>
<td>1150 (29817)</td>
<td>1624 (158840)</td>
<td>1238 (41422)</td>
</tr>
</tbody>
</table>

Table 2: Total number of increment dates (and the total number of options used) in the different sub-categories considered.
\((T, K) \in I\), aiming to replicate one unit of the option \(\hat{u}_{i-1}(T, K)\) and thus with an initial value at \(t_{i-1}\) that equals the option price: \(P_{i-1}(T, K) = \hat{u}_{i-1}(T, K)\). Our total replicating portfolio, which we will denote by \(\hat{P}\), will be the weighted sum of the portfolios \(P_{i-1}(T, K)\) according to

\[
\hat{P}_i(T, K) = \sum_{(T,K)\in I_i^k} \omega_{i-1}(T, K) P_{i-1}(T, K).
\]

(74)

The evolution of each sub-portfolio \(P_{i-1}(T, K)\) is given by (4), with hedging ratios \(\delta\) calculated in discrete- and continuous time respectively with the methodology described in Section 3.2 using parameters for the model (37) obtained as described in Section 4. The parameters used are the ones calibrated at time \(t_{i-1}\). We consider hedging with either just the underlying index \(S\) or with both \(S\) and an option. For the latter case, we will at each \(t_{i-1}\) choose the quoted maturity \(T\) that is closest to 180 days and, for this maturity, the strike \(K\) that is closest to the spot price \(S_0\). We then create a hedging portfolio with positions in \((S, C(T, K))\). At the next observation time \(t_i\), we get the new values of the quoted options and the index spot price. The evolution of the bank account \(B\) from \(t_{i-1}\) to \(t_i\) is deduced from the zero-coupon curve at \(t_{i-1}\). The hedging error between the evolution of the contracts we invest in and our replicating portfolio \(\hat{P}\) can then be written as

\[
\text{err}_i = \hat{P}_i - \hat{P}_{i-1} - \sum_{(T,K)\in I_i^k} \omega_{i-1}(T, K) (\hat{u}_i(T, K) - \hat{u}_{i-1}(T, K)),
\]

(75)

where \(\omega_{i-1}(T, K)\) are given by (73). We will analyse the series \(\text{err}_i\) obtained from our different hedging methodologies to test the validity of each strategy. The errors \(\text{err}_i\) correspond to the profit or loss obtained if \(\omega_{i-1}(T, K)\) units of each contract with maturity-strike \((T, K)\) \(\in I_k^i\) were sold at the mid-price at \(t_{i-1}\) and repurchased at mid-price at \(t_i\), while investing in the portfolio \(\hat{P}\) at \(t_{i-1}\) and liquidating the portfolio at \(t_i\).

A few comments are in place regarding our approach. First, by considering that we buy and sell at the market’s mid price (rather than buying at ask price and selling at bid price) we neglect the effect of the bid-ask price spread. Taking the spread into account would most likely affect the different models in a similar fashion by introducing a small loss at each trading occasion, and we do not believe that it would change our conclusions regarding the performance of the models. It can also be noted that an important part of a trader’s job is to decide when to hedge, so as to avoid to frequent rehedging which leads to losses of the price spread. Therefore, a proper consideration of the impact of the spread ought to include a strategy for the times at which hedging occur, which is not within the scope of this study. We look at incremental hedging of new portfolios of options at each time step, rather than following a fix portfolio of options over time. We use this method partly because it gives us a larger number of observations in our time series. If we follow fix portfolios over, say, 3 months, our 27 months of data would give only 9 independent observations in time. Also, by looking at options that are quoted in the market at two consecutive times that are close in time, we increase the number of observed quotes. (The quoted options that are observable at both of two dates separated by 3 months will be fewer than those observable at two times that lie a few hours or days apart). The tests we perform differ somewhat from how hedging would typically be performed in a financial institute. An obvious aspect is the constitution of the portfolio. As we described above, we will test portfolios consisting only of short positions of call options with maturity-strike pairs in different intervals. As a contrast, a market maker might try to take both long and short positions that conveniently hedge each other and earn money by being on the right side of the bid-ask spread. A trader in vanilla options will continuously monitor her positions and re-hedge her book when considered necessary. As a contrast, we have data at certain fixed instants in time and consider portfolios that are created and liquidated at those exact times. If the market moves significantly between two of our sample times, we might get a larger error than would have been the case with an intermediate adjustment of the positions. Futures on equity indices are the instruments used in practice for “delta hedging” (i.e., hedging with the underlying asset) of equity index options. We do not have data on future prices on the indices and have instead considered the spot value of each index and constructed portfolios as if the spot could be directly traded.

Our model (37) treats dividends via the continuous yield \(q\). Although convenient for analytic
purposes, this dividend model is obviously false since stocks pay dividends at discrete dates. Also, we calculate a dividend yield at each sampling time using an implicit forward price deduced from put-call parity. Noise in the option prices in our data can therefore result in fluctuations in our calculated dividend yield that is not necessarily reflected by the market’s view on future dividend payments. A financial institute trading in put- and call options would use a more realistic dividend model when calculating hedging ratios and also typically use other sources of market data (such as index future prices and stock analysts’ views on future dividends) to set the values on the dividends in the models. Since the dividend assumptions have an impact on the hedging ratios, our simplified treatment of dividends might have a detrimental effect on hedging, especially near dates when dividends are paid. In spite of these differences between our tests and real life hedging, we believe that our methodology is appropriate for comparing the differences in hedging performance between strategies emanating from different volatility models.

6.2 Comparing hedging performance
Given the hedging error series $\text{err}_i$ from (75) obtained using our different strategies, we want to decide which strategy is most “successful”. We are trying to achieve an error as close to zero as possible and since we do not expect our hedging strategies to be gaining (remember that we buy and sell at mid-price in our tests) or losing, we have an equivalent view on a positive and a negative error.

We have little a priori knowledge about the time series $\text{err}_i$. The values $\text{err}_i$ are obtained using different model parameters and spot price at each time $t_i$, we do not know the distribution of the increments of prices quoted in the market and the hedging is done over different time increments. In particular, this means that we have no reason to believe that the variables $\text{err}_i$ (for a given hedging strategy) be equally distributed. To test our results, we therefore want to use methods that use as little assumptions as possible about the data.

We will calculate two values for each strategy that we test. First we will calculate a sample standard deviation $\hat{\sigma}$ according to

$$\hat{\sigma} = \sqrt{\frac{1}{n - 1} \sum_{i=1}^{n} \text{err}_i^2},$$

(76)

where $n$ is the length of the series (see Table 2) and

$$\text{err}_i = \text{err}_i - \frac{1}{n} \sum_{j=1}^{n} \text{err}_j.$$

(77)

with $\text{err}_j$ from (75). Subtracting the mean value from the samples as in a sample variance calculation means that a series with a mean not equal to zero is treated in the same way as a series with an expected zero error. However, the mean of our series deviate little from zero, so the correction for the mean does not affect the values to a great extent. The sample variance $\hat{\sigma}^2$ does not obviously have an interpretation as an estimator of the variance of a random variable (with mean zero) in this case since the values $\text{err}_i$ can not be assumed to be sampled from the same distribution. None the less, $\hat{\sigma}$ is an easily interpreted measure of a sample’s deviation from a mean around zero.

Second, we use a test from non-parametric statistics to evaluate if the difference in variability away from zero between any pair of two hedging error series can be considered large enough to discard the hypothesis that the “variability” of the series is equal. In the Siegel-Tukey rank sum test 29 a set of combined data from two different data series is ordered by giving low ranks to data points having either high or low values while keeping track of which set each point originates from. The sum of the ranks of the data from one of the sets is then calculated and used to obtain a statistic based on the idea that if the two data series were equally spread around a common mean value, then the rank sum of both series should tend to have the same mean. To be concrete, we cite the example used in [29]. Let $a$ and $b$ denote different data series, both with 9 data points.
Table 3: Example from [29] of ranking procedure for the Siegel-Tukey test.

<table>
<thead>
<tr>
<th>Value</th>
<th>0</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>8</th>
<th>10</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>19</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
<td>a</td>
<td>b</td>
<td>a</td>
<td>b</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>Rank</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>9</td>
<td>12</td>
<td>13</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>15</td>
<td>14</td>
<td>11</td>
<td>10</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4: Two sided confidence intervals for a variable $z$ with standard normal distribution.

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>95.00%</th>
<th>99.00%</th>
<th>99.50%</th>
<th>99.90%</th>
<th>99.99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interval</td>
<td>$</td>
<td>z</td>
<td>\leq 1.960$</td>
<td>$</td>
<td>z</td>
</tr>
</tbody>
</table>

Table 3 shows the two series combined, their respective values and the rank attributed to each data point. As the table illustrates, the lowest value is given rank 1, the two highest values rank 2 and 3 respectively, after which rank 4 and 5 is given to the points that have the second and third lowest values. The ranking proceeds in this staggering manner for all points in the combined set.

For fairly large series of data (as in our case, where the shortest series has 1091 points as shown in Table 2) the Siegel-Tukey test uses an approximately normal statistics $z$ defined as

$$z = \frac{2R_1 - n_1(n_1 + n_2 + 1)}{\sqrt{n_1(n_1 + n_2 + 1)(n_2/3)}}, \quad (78)$$

where $R_1$ is the sum of the ranks attributed to the values in one of the series, $n_1$ the number of points in this series and $n_2$ the number of points in the other series. The sign of the addition ±1 in the numerator is chosen to maximize the module of $z$.

Rather than calculating $z$ from (78) directly from the series $err_i$ from (75), we use the error series $\hat{err}_i$ from (77) adjusted by their sample mean values in accordance with the way the sample standard deviation is calculated. As was the case for the sample standard deviation (76), the choice of using the adjusted series $\hat{err}_i$ rather than $err_i$ does not change our conclusions, since the sample mean values of our series are close to zero.

For two series $a$ and $b$ with the same approximate mean, the null hypothesis $H_0$ and the alternative hypothesis $H_a$ and $H_b$ tested for by the statistic $z$ are

- $H_0$ The two series have the same variability.
- $H_a$ Series $a$ has lower variability than series $b$.
- $H_b$ Series $b$ has lower variability than series $a$.

If $R_1$ in (78) is the rank sum for series $a$, we will reject $H_0$ in favour of $H_a$ for high values of $z$ and in favour of $H_b$ for low values of $z$. The statistic $z$ will be close to normal and Table 4 gives two sided confidence intervals for $H_0$ - based on the cumulative normal distribution - at different levels.

7 Results

We now present the results for the hedging errors in the different models, for delta hedging (using only the underlying spot) and for delta-vega hedging (using spot and some quoted option). As mentioned, for delta-vega hedging we use the quoted maturity closest to 180 days and for this maturity the observable strike closest to the spot price.

We first comment on the hedging strategies in discrete time from Section 3.2.1. As can be seen from equations (33) and (34) for the continuous time case and equations (13) to (17) for discrete time, the quadratic hedging ratios in discrete time can be seen as approximations of the ones in continuous time if the derivatives in (33) are replaced by the finite difference approximations. It
should therefore be no surprise if the results from using the hedging ratios calculated from the
result from continuous time are very similar to those calculated using the result from discrete time.
We have successfully implemented the method from Section 3.2.1 to calculate the discrete hedging
ratios in the Heston case (i.e., \( \lambda = 0 \) in the model (37)). The case including jumps is more difficult
to implement, mostly due to the larger support of the joint density of the spot and variance (\( \varphi \) in
(49)) in this case. An alternative method to calculate the hedge ratios would be to use Monte-Carlo
simulations to calculate the variances in (48) - something we have not implemented.

We have compared the discrete time strategy to the continuous time strategy for the data
described above, but with most of the observation dates excluded to create a series of option data
with a smallest time increment of 5 \( \times \) 24 hours, which (since the markets are closed on weekends)
in practice gives us a time series where almost every time-increment is exactly 7 \( \times \) 24 hours. We
use this data with weekly observations to construct our series \( \text{err}_i \) from (75). The time series are,
of course, much shorter than the full series, with a total of 117 points in the series \( \text{err}_i \). In order
not to get too short series, we have not (as in Table 2) in this case excluded observation dates
at which we have few options in certain category. The model parameters for this comparison has
not been calibrated with one of the cases displayed in Table 1. Instead, the values \( \gamma = 0 \) and
\( \beta = 1 \) are used in (61), and Matlab’s line-search routine in the fmincon function with the option
“active-set” is used rather than the Newton based method with lsqnonlin as stated in Section
4. On this data with time steps of 7 days, we do not observe any significant difference between
the result with the strategy from discrete- and continuous time. Although we do not display the
hedging ratios calculated for individual options in the discrete- and continuous case, they usually
differ very little. Consequently, the sample standard deviations (76) and the test statistic \( z \) from
(78) give no reason to believe there is a difference in performance between the two cases. Table 5
displays the sample standard deviations from delta- and delta-vega hedging in the two cases and
Table 6 shows the values of the corresponding values for \( z \).

The sample standard deviation for the full series (which typically has three daily observations)
for delta hedging and delta-vega hedging in the BS model, Heston’s SV model and our SVJ model
with jumps (37) are given in Table 7. A first observation is that in terms of standard deviation,
the BS model performs poorly compared to the stochastic models for short dated options, both
for pure delta hedging and for delta-vega hedging. For ATM and OTM short dated options, the
error standard deviation is about 1.5 to 2 times higher in the BS model than in the stochastic
models.

In Figure 1 we display the error time series and the corresponding histograms for delta hedging
of short dated options in the BS model and the quadratic delta hedging in the SV model as well
as for delta-vega hedging in the BS model and quadratic delta-vega hedging in the SVJ model.
We look at OTM options and for the delta hedge and ATM options for the delta-vega hedge. The
smaller error obtained with the stochastic models is apparent. If we return to Table 7, we see that
for medium dated options, the BS model behaves well for delta hedging of ITM and ATM options,
but still gives poor results for OTM options. For delta-vega hedging of the medium dated options,
the BS model has the worst results for ATM and OTM options but compares to the performance
of the stochastic models for ITM options. It can be noted that the standard deviation is around
15 to 30 times larger for OTM options than for ITM options, and roughly 2 to 3 times larger for
OTM options than for ATM options. Quadratic hedging with the stochastic models thus give a
much better control of the contracts that are most “dangerous” in terms of their hedging errors
(measured relative to the invested capital). In Table 8 we display the values of the \( z \)-statistic (78)
when hedging in the BS model is compared to the stochastic models. We have chosen a confidence
level of 99.5\%, thus - according to Table 4 - rejecting the null hypothesis that any two tested
series have the same variability if \( |z| > 2.807 \). The \( z \)-statistic mostly tells us the same thing as
the standard deviation: The stochastic models with quadratic hedging significantly outperforms
the BS model for hedging of OTM options in all cases. When it comes to ITM options, the \( z \-
statistic does not give us reason to chose one over the other. This category of call options also
are the “easiest” to hedge which is translated into a low standard deviation. For ATM options, all
the stochastic models outperform the BS model for delta and delta-vega hedging of short dated
options, and likewise for delta-vega hedging of medium dated options. The only result yielded by
the $z$-statistic that seems a little surprising given the sample standard deviations is that the null hypothesis is rejected in favour of the hypothesis that BS hedging has lower variability than each of the stochastic models for delta hedging of medium dated ATM options. Nonetheless, the BS model does not yield the lowest standard deviation for this subcategory of options. According to Table 7, the SV model with the calibration 2B yields a slightly lower standard deviation for the error than the BS model. A pictorial explanation of the situation is given in Figure 2(a) to 2(d) where we display the time series for the error time series and the corresponding histograms for delta hedging of medium dated ATM options in the BS model and in the SV model calibrated as in case 2B. We see that the histogram corresponding to the BS model is more leptokurtic than the histogram resulting from the SV model. The $z$-statistic and the sample standard deviation therefore do not measure quite the same thing in this case. Often, risk managers in financial institutes want to avoid strategies leptokurtic distribution of returns because of the high associated risk of large losses, which would favour the stochastic models over the BS model in this case. Interestingly enough, the difference between the distributions of the errors from delta hedging in the BS model and the stochastic models disappears when the regression technique from Section 5 is applied. In Figure 2(e) to 2(h) we display the error series and histograms when regression is applied for the BS model and the SV model for the same case as just discussed. It is visually apparent that the two series and their resulting histograms are very similar. This idea is supported by the standard deviations obtained from delta hedging with regression for all subcategories of options in each of our tested models, which are displayed in Table 9. There are no large differences between any of the models or calibration cases. The $z$-statistic for comparing delta hedging with regression in the BS model to the quadratic hedging with regression in the stochastic models with different calibration displayed in Table 10 supports this idea: in none of the cases can we refute the null hypothesis that the error from hedging in the BS model has the same variability as the error obtained with the stochastic models. The regression technique thus seems to neutralize the differences between the models as far as delta hedging is concerned. We also want to test if the regression actually improves the delta hedging. We therefore calculate the $z$-statistic for comparing delta hedging results from each model (with the different calibration setups) with regression and without regression. The values are displayed in Table 11, where high values of $z$ are in favour of hedging with regression. We see that the $z$-statistic is in favour of hedging with the regression technique for almost all cases with options of medium maturity. For short maturities, the stochastic models are not in general improved. The BS model is improved for ATM and OTM options of short and medium maturities. These observations are in line with the results for the standard deviation with and without regression in Table 7 and 9.

If we return to Table 7 with the standard deviations for hedging errors in the different models, it is apparent that the errors obtained with the different calibration setups from Table 1 are quite similar. This could potentially be because the calibrations find the same parameters, so that we actually use the same parameters in each case. But that the parameters are indeed different is illustrated in Figure 3 which displays the values of the state variable $y_0$ and the diffusion parameters ($\kappa, \eta, \theta, \rho$) of the SVJ model (37) obtained at different dates for each calibration setup, and in Figure 4 which shows the corresponding plots for the jump parameters ($\lambda, \mu$). The evolution of $y_0$ over time is remarkably similar with all calibration setups. The model parameters, however, display a different behaviour in each case, with the values obtained in the un-penalized cases (1A and 2A in Table 1) showing an erratic behaviour over time. The hedging ratios in the different cases have thus been calculated from different model parameters.

Even though the standard deviations in Table 1 are rather similar for the different calibration setups, we can observe that the SVJ model gives the lower standard deviations for delta-vega hedging, and the best results are obtained with the unconstrained calibration case 1A. In Table (12), we display the values for the $z$-statistic when comparing hedging in the SV model versus the SVJ model. These values support the idea that the SVJ jump model performs better for delta-vega hedging than the purely diffusive SV model. All the $z$-values for comparing the delta-vega hedging have a negative sign - supporting the SVJ model - although the values only in a few cases are low enough to refute the null hypothesis at our confidence level of 99.5%. For pure delta hedging, the results are mixed, but the SV model without jumps seem to behave slightly better for
medium dated options. As for the different calibrations setups, we show the z-statistic for pairwise comparing hedging using parameters from different calibrations for the SVJ model in Table 13. The clearest observation here is that the calibrations 2A and 2B involving the observed delta hedging error actually do yield better results for delta hedging. However, this is not the case for delta-vega hedging, where these two calibrations yield worse results than the calibration cases 1A and 1B using only information about market prices. Overall, there seem to be a certain robustness in the models regarding the calibration. Hedging performance with the models does not vary to a great extent when different parameters are used.

8 Conclusion

We have evaluated quadratic hedging strategies in a stochastic volatility setting with and without jumps in the spot process. To this end, we have calibrated model parameters to market data and performed hedging on day-to-day portfolios of market quoted option prices on a major equity index. The quadratic hedging strategies were developed both in a discrete time setting and in continuous time. As a comparison, we have used standard hedging strategies calculated with the BS model. The motivation for this study was to test if better hedging results could be obtained by changing from the BS model - still widely used for this type of financial products - to affine stochastic volatility models. Since it is not unusual among practitioners to try to improve the performance of hedging ratios provided by the BS through more or less formalized procedures, we also developed a simple regression technique to improve delta hedging performance in both BS and in the stochastic volatility case.

Our results on European call options written on the Euro Stoxx 50 equity index indicate that when only using the underlying equity index to hedge (i.e., pure delta hedging) there is a substantial performance gain to be made in using quadratic hedging in the stochastic volatility model (with or without jumps) when it comes to those options that have the highest errors (measured in sample standard deviation of the hedging error relative to the amount invested in the options). These options are out of the money options of all expiries and at the money options of shorter expiries. For ITM options and for ATM options of medium termed expiries, the hedging errors are much smaller and the difference in sample standard variation is small between the models.

When we apply our regression technique to improve delta hedging, the difference between the models mostly vanish. The performance of the BS hedge is substantially improved, whereas smaller improvements can be seen for the stochastic models with the quadratic hedging. Our tests do not indicate that there is any significant difference between the models in this case. However, it should be noticed that the results from delta hedging with the regression techniques are only marginally better than the results from the unaltered quadratic delta hedging techniques with the stochastic models (at least in terms of the hedging error sample standard deviation). This type of regression technique is less evident to apply when we hedge with other contracts, since it would require different regression calculations depending on the contracts chosen for hedging. Since real life hedging typically involves hedging of the “vega-risk” (or volatility risk) of a trading agent’s aggregated positions, we argue that the better performance for the quadratic hedging for delta-vega hedging is a strong argument for financial institutes to use the stochastic models and quadratic hedging theory to estimate the exposures they have to this type of risk. When it comes to hedging with a portfolio consisting both of the underlying index and another option (which we call delta-vega hedging) our results indicate a clear advantage for the quadratic hedging in stochastic models over the BS approach. The sample standard deviation of hedging errors of ATM and OTM options in the BS model is roughly 1.5 to 2 times higher than the corresponding values obtained with the stochastic models.

We have also tested different setups for calibrating the parameters in the stochastic models. Our tests imply that the hedging performance in these models posses a certain robustness when it comes to choosing the parameters. Our tested cases do not display any remarkable differences in terms of hedging performance. Nonetheless, pure delta hedging is somewhat improved when incorporating information about realized hedging errors into the calibration procedure.
The paper [3] Bakshi et al. contains a study of hedging errors of European options written on the S&P 500-index in the BS model, Heston’s SV model and an SVJ model with normally distributed jumps. This study was conducted on option data from 1988 to 1991. It is difficult to make an exact comparison of these results and ours. In part because we use larger moneyness intervals in categorizing our options, because options of higher and lower strikes are traded today than in that period, and also because they use a different error measure than we do. Also, when they hedge with both the underlying index and other options, it is not clear what option is used. When including another option in the hedge, we have used the maturity closest to 180 days and for that maturity, the quoted option whose strike is closest to the spot price. This means that the hedging instrument can be an option with a strike and maturity that significantly differs from the maturities and strikes in the category of options we hedge. Our results seem to be more favourable for delta-vega hedging than the results in [3]. This might possibly be because they hedge with options that are closer in strike and maturity to the options they aim to replicate.

Our results are also more favourable for the stochastic models when it comes to hedging of short dated ITM and OTM options, and our results for delta-vega hedging are more positive for the SVJ model with jumps when compared to the SV model without jumps. This might be because we use a different jump specification. It should be emphasized, of course, that since we perform the studies at different time periods for different data, there is no guarantee that the conclusions regarding different models’ behaviour should be the same. Nonetheless, the common conclusion is that hedging can be improved by replacing the BS model by a stochastic volatility model.

Stochastic volatility models are often used in financial institutes to price and hedge more complex products than the standard European call options used in this study. The positive results for the quadratic hedging theory when applied to these standard European call options is an indication that financial institutes might be able to achieve markedly lower hedging errors for such “vanilla” products if the simple BS modelling typically used for these contracts were to be replaced by more elaborate models.
<table>
<thead>
<tr>
<th>Expiry</th>
<th>Hedge</th>
<th>Model</th>
<th>Moneyness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ITM</td>
<td>ATM</td>
<td>OTM</td>
</tr>
<tr>
<td>Short dated</td>
<td>0.7 ≤ (\frac{K}{S_0}) &lt; 0.9</td>
<td>0.9 ≤ (\frac{K}{S_0}) ≤ 1.1</td>
<td>1.1 ≤ (\frac{K}{S_0}) ≤ 1.1</td>
</tr>
<tr>
<td>0 &lt; T ≤ 0.5</td>
<td>Delta</td>
<td>SV</td>
<td>2.14%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SVD</td>
<td>2.16%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>1.62%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SVD</td>
<td>1.52%</td>
</tr>
<tr>
<td></td>
<td>ITM</td>
<td>ATM</td>
<td>OTM</td>
</tr>
<tr>
<td></td>
<td>0.6 ≤ (\frac{K}{S_0}) &lt; 0.8</td>
<td>0.8 ≤ (\frac{K}{S_0}) ≤ 1.2</td>
<td>1.2 ≤ (\frac{K}{S_0}) ≤ 1.4</td>
</tr>
<tr>
<td>Medium dated</td>
<td>0.5 &lt; T ≤ 2</td>
<td>Delta</td>
<td>1.22%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SV</td>
<td>1.26%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SVD</td>
<td>0.65%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SVD</td>
<td>0.65%</td>
</tr>
</tbody>
</table>

**Table 5:** Sample standard deviation (in percentage of invested capital) for delta hedging in the SV model ((37) with \(\lambda = 0\)) using the continuous time hedge (here denoted SV) and the discrete time hedge (denoted SVD) with time increments of 7 days.

<table>
<thead>
<tr>
<th>Expiry</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Moneyness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ITM</td>
<td>ATM</td>
<td>OTM</td>
</tr>
<tr>
<td>Short dated</td>
<td>0.7 ≤ (\frac{K}{S_0}) &lt; 0.9</td>
<td>0.9 ≤ (\frac{K}{S_0}) ≤ 1.1</td>
<td>1.1 ≤ (\frac{K}{S_0}) ≤ 1.1</td>
</tr>
<tr>
<td>0 &lt; T ≤ 0.5</td>
<td>SV</td>
<td>SVD</td>
<td>0.464</td>
</tr>
<tr>
<td></td>
<td>SV</td>
<td>SVD</td>
<td>-0.0850</td>
</tr>
<tr>
<td>Medium dated</td>
<td>0.5 &lt; T ≤ 2</td>
<td>SV</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>SV</td>
<td>SVD</td>
<td>0.30</td>
</tr>
</tbody>
</table>

**Table 6:** The statistic \(z\) from (78) for comparing delta hedging with time increments of 7 days using the continuous time hedge (SV) and the discrete time hedge (SVD) in the SV model ((37) with \(\lambda = 0\)). In none of the cases is \(|z| < 2.807\), which is the value at which we would refute the hypothesis that the compared series of hedging errors have the same variability at a confidence level of 99.5%.
### Table 7: Sample standard deviation (in percentage of invested capital) for the error series $err_i$ obtained from hedging in the BS model and quadratic hedging in the stochastic model (37) with jumps (SVJ) and without jumps (SV, $\lambda = 0$). The calibration cases refer to Table 1. The highest value in each subcategory of options (maturity interval and moneyness interval) for delta- and delta-vega hedging is written in *italic* and the lowest value in *bold*.

<table>
<thead>
<tr>
<th>Expiry</th>
<th>Hedge</th>
<th>Model</th>
<th>Calib.</th>
<th>Moneyness</th>
<th>ITM</th>
<th>ATM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.7 ≤ $\frac{K}{S_0}$ &lt; 0.9</td>
<td>0.9 ≤ $\frac{K}{S_0}$ ≤ 1.1</td>
<td>1.1 &lt; $\frac{K}{S_0}$ ≤ 1.2</td>
</tr>
<tr>
<td>Short dated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 &lt; $T$ ≤ 0.5</td>
<td>Delta</td>
<td>BS</td>
<td>-</td>
<td></td>
<td>0.712%</td>
<td>6.31%</td>
<td>20.7%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>1A</td>
<td></td>
<td>0.657%</td>
<td>4.52%</td>
<td>10.8%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>1A</td>
<td></td>
<td>0.948%</td>
<td>4.84%</td>
<td>10.7%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>1B</td>
<td></td>
<td>0.652%</td>
<td>4.52%</td>
<td>10.8%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>1B</td>
<td></td>
<td>0.875%</td>
<td>4.74%</td>
<td>10.7%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>2A</td>
<td></td>
<td>0.657%</td>
<td>4.53%</td>
<td>10.9%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>2A</td>
<td></td>
<td>0.655%</td>
<td>4.59%</td>
<td>11.6%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>2B</td>
<td></td>
<td>0.656%</td>
<td>4.53%</td>
<td><strong>10.6%</strong></td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>2B</td>
<td></td>
<td>0.652%</td>
<td><strong>4.42%</strong></td>
<td><strong>10.6%</strong></td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>BS</td>
<td>-</td>
<td></td>
<td>0.616%</td>
<td>4.97%</td>
<td>17.4%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>1A</td>
<td></td>
<td>0.548%</td>
<td>3.78%</td>
<td>9.40%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>1A</td>
<td></td>
<td><strong>0.463%</strong></td>
<td><strong>3.24%</strong></td>
<td>9.06%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>1B</td>
<td></td>
<td>0.545%</td>
<td>3.83%</td>
<td>9.52%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>1B</td>
<td></td>
<td>0.495%</td>
<td>3.33%</td>
<td><strong>8.90%</strong></td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>2A</td>
<td></td>
<td>0.548%</td>
<td>3.64%</td>
<td>9.33%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>2A</td>
<td></td>
<td>0.492%</td>
<td>3.37%</td>
<td>9.10%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>2B</td>
<td></td>
<td>0.558%</td>
<td>3.96%</td>
<td>9.48%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>2B</td>
<td></td>
<td>0.548%</td>
<td>3.71%</td>
<td>9.10%</td>
</tr>
<tr>
<td>Medium dated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5 &lt; $T$ ≤ 2</td>
<td>Delta</td>
<td>BS</td>
<td>-</td>
<td></td>
<td><strong>0.360%</strong></td>
<td>1.48%</td>
<td>7.00%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>1A</td>
<td></td>
<td>0.400%</td>
<td>1.56%</td>
<td>5.28%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>1A</td>
<td></td>
<td>0.456%</td>
<td>1.60%</td>
<td><strong>4.98%</strong></td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>1B</td>
<td></td>
<td>0.401%</td>
<td>1.56%</td>
<td>5.30%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>1B</td>
<td></td>
<td>0.446%</td>
<td>1.72%</td>
<td>5.59%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>2A</td>
<td></td>
<td>0.392%</td>
<td>1.52%</td>
<td>5.21%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>2A</td>
<td></td>
<td>0.396%</td>
<td>1.52%</td>
<td>5.32%</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>2B</td>
<td></td>
<td>0.387%</td>
<td><strong>1.45%</strong></td>
<td><strong>4.98%</strong></td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>2B</td>
<td></td>
<td>0.390%</td>
<td>1.46%</td>
<td>5.05%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>BS</td>
<td>-</td>
<td></td>
<td>0.242%</td>
<td>0.788%</td>
<td>5.99%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>1A</td>
<td></td>
<td>0.255%</td>
<td>0.503%</td>
<td>3.47%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>1A</td>
<td></td>
<td><strong>0.229%</strong></td>
<td><strong>0.421%</strong></td>
<td><strong>3.19%</strong></td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>1B</td>
<td></td>
<td>0.255%</td>
<td>0.507%</td>
<td>3.48%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>1B</td>
<td></td>
<td>0.241%</td>
<td>0.478%</td>
<td>3.39%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>2A</td>
<td></td>
<td>0.255%</td>
<td>0.514%</td>
<td>3.53%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>2A</td>
<td></td>
<td>0.249%</td>
<td>0.489%</td>
<td>3.36%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>2B</td>
<td></td>
<td>0.255%</td>
<td>0.504%</td>
<td>3.51%</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>2B</td>
<td></td>
<td>0.254%</td>
<td>0.493%</td>
<td>3.46%</td>
</tr>
</tbody>
</table>
Table 8: The statistic $z$ from (78) for comparing the error series error from (75) obtained from hedging in the BS model and quadratic hedging in the stochastic model (37) with jumps (SVJ) and without jumps (SV, $\lambda \equiv 0$). The calibration of the stochastic models correspond to the cases in Table 1. High values favour the BS model and low values the stochastic model. Values of $|z|$ > 2.807 that refute the hypothesis that the compared cases yield the same variability of the hedging error at a confidence level of 99.5% are written in italic if the BS model is more favourable and in bold if the stochastic model is favoured.
Figure 1: Time series for $e_{it}$ from (75) and their corresponding histograms. Figures 1(a) to 1(d) show results for short dated OTM options (Table 1) for delta hedging in the BS model and quadratic delta hedging in the SV model (model (37) with $\lambda = 0$). Figure 1(e) to 1(h) show results for short dated ATM options (Table 1) for delta-vega hedging in the BS model and quadratic delta-vega hedging in the SVJ model (37).

Figure 2: Time series for $e_{it}$ from (75) for delta hedging of medium dated OTM options their corresponding histograms. Figures 2(a) to 2(d) show results for the BS model and in quadratic delta hedging in the SV model (model (37) with $\lambda = 0$) calibrated according to case 2B (Table 1). Figures 2(e) to 2(h) show results for the same two models when the regression technique from Section 5 is applied.
Table 9: Comparing delta hedging with regression in different models. All results refer to cases where regression as in Section 5 is used. Sample standard deviation (in percentage of invested capital) for the error series $err_i$ from (75) obtained from delta hedging with regression in the BS model as well as in the stochastic model (37) with jumps (SVJ) and without jumps (SV, $\lambda \equiv 0$). The calibration cases refers to Table 1. The highest value in each subcategory of options (maturity interval and moneyness interval) for delta- and delta-vega hedging is written in *italic* and the lowest value in *bold*.
<table>
<thead>
<tr>
<th>Expiry</th>
<th>Hedge Model Calib.</th>
<th>Moneyness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 &lt; T ≤ 0.5</td>
<td>Delta SV 1A</td>
<td>-1.33</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 1A</td>
<td>0.541</td>
</tr>
<tr>
<td></td>
<td>Delta SV 1B</td>
<td>-1.37</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 1B</td>
<td>-0.0253</td>
</tr>
<tr>
<td></td>
<td>Delta SV 2A</td>
<td>-1.36</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 2A</td>
<td>-0.703</td>
</tr>
<tr>
<td></td>
<td>Delta SV 2B</td>
<td>-1.34</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 2B</td>
<td>-1.32</td>
</tr>
<tr>
<td>0.6 ≤ K_{SO} &lt; 0.8</td>
<td>Delta SV 1A</td>
<td>0.0363</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 1A</td>
<td>0.456</td>
</tr>
<tr>
<td></td>
<td>Delta SV 1B</td>
<td>0.0995</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 1B</td>
<td>0.216</td>
</tr>
<tr>
<td></td>
<td>Delta SV 2A</td>
<td>0.0499</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 2A</td>
<td>-0.266</td>
</tr>
<tr>
<td></td>
<td>Delta SV 2B</td>
<td>-0.00581</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 2B</td>
<td>-0.0643</td>
</tr>
<tr>
<td>0.8 ≤ K_{SO} ≤ 1.2</td>
<td>Delta SV 1A</td>
<td>0.0499</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 1A</td>
<td>0.456</td>
</tr>
<tr>
<td></td>
<td>Delta SV 1B</td>
<td>0.0995</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 1B</td>
<td>0.216</td>
</tr>
<tr>
<td></td>
<td>Delta SV 2A</td>
<td>0.0499</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 2A</td>
<td>-0.266</td>
</tr>
<tr>
<td></td>
<td>Delta SV 2B</td>
<td>-0.00581</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 2B</td>
<td>-0.0643</td>
</tr>
<tr>
<td>1.2 ≤ K_{SO} ≤ 1.4</td>
<td>Delta SV 1A</td>
<td>0.0363</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 1A</td>
<td>0.456</td>
</tr>
<tr>
<td></td>
<td>Delta SV 1B</td>
<td>0.0995</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 1B</td>
<td>0.216</td>
</tr>
<tr>
<td></td>
<td>Delta SV 2A</td>
<td>0.0499</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 2A</td>
<td>-0.266</td>
</tr>
<tr>
<td></td>
<td>Delta SV 2B</td>
<td>-0.00581</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ 2B</td>
<td>-0.0643</td>
</tr>
</tbody>
</table>

Table 10: Comparing delta hedging with regression in different models. All results refer to cases where regression as in Section 5 is used. The statistic $z$ from (78) for comparing the error series $err_i$ from delta hedging when regression as in Section 5 is applied to the BS model and to the stochastic models calibrated in the cases in Table 1. High values favour the BS model and low values the stochastic model. At a confidence level of 99.5%, none of the cases yields $|z| > 2.807$ that would refute the hypothesis that the two compared models have equal variability.
Table 11: Comparing delta hedging with and without regression in different models. The statistic $z$ from (78) for comparing delta hedging in each of the models with and without regression as in Section 5 The calibration of the stochastic models correspond to the cases in Table 1 High values favour hedging with regression. Values of $|z| > 2.807$ that refute the hypothesis that the hedging error in the compared cases yield the same variability at a confidence level of 99.5% are written in italic if the regression technique is more favourable and in **bold** if hedging without regression is favoured.
Figure 3: Time-series for the state variable $y_0$ and the diffusion parameters $\kappa, \eta, \theta, \rho$ obtained at each calibration occasion for the SVJ model (37) for the four different calibration setups in Table 1.
Figure 4: Time-series for the jump parameters $\lambda, \mu$ obtained at each calibration occasion for the SVJ model (37) for the four different calibration setups in Table 1.
<table>
<thead>
<tr>
<th>Expiry</th>
<th>Hedge</th>
<th>Calib.</th>
<th>Moneyness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ITM</td>
<td>ATM</td>
<td>OTM</td>
</tr>
<tr>
<td></td>
<td>$0.7 \leq \frac{S}{K} &lt; 0.9$</td>
<td>$0.9 \leq \frac{S}{K} \leq 1.1$</td>
<td>$1.1 &lt; \frac{S}{K} \leq 1.2$</td>
</tr>
<tr>
<td>Short dated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0 &lt; T \leq 0.5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delta</td>
<td>1A</td>
<td>12.9</td>
<td>1.12</td>
</tr>
<tr>
<td>Delta</td>
<td>1B</td>
<td>8.46</td>
<td>0.376</td>
</tr>
<tr>
<td>Delta</td>
<td>2A</td>
<td>0.487</td>
<td>-0.24</td>
</tr>
<tr>
<td>Delta</td>
<td>2B</td>
<td>-0.059</td>
<td>-0.643</td>
</tr>
<tr>
<td>Delta-vega</td>
<td>1A</td>
<td>-2.97</td>
<td>-4.04</td>
</tr>
<tr>
<td>Delta-vega</td>
<td>1B</td>
<td>-1.84</td>
<td>-4.14</td>
</tr>
<tr>
<td>Delta-vega</td>
<td>2A</td>
<td>-0.734</td>
<td>-1.7</td>
</tr>
<tr>
<td>Delta-vega</td>
<td>2B</td>
<td>-0.162</td>
<td>-1.27</td>
</tr>
<tr>
<td>Medium dated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.5 &lt; T \leq 2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delta</td>
<td>1A</td>
<td>5.12</td>
<td>0.83</td>
</tr>
<tr>
<td>Delta</td>
<td>1B</td>
<td>3.69</td>
<td>3.12</td>
</tr>
<tr>
<td>Delta</td>
<td>2A</td>
<td>0.334</td>
<td>0.0276</td>
</tr>
<tr>
<td>Delta</td>
<td>2B</td>
<td>0.103</td>
<td>0.161</td>
</tr>
<tr>
<td>Delta-vega</td>
<td>1A</td>
<td>-2.43</td>
<td>-5.91</td>
</tr>
<tr>
<td>Delta-vega</td>
<td>1B</td>
<td>-1.09</td>
<td>-2.71</td>
</tr>
<tr>
<td>Delta-vega</td>
<td>2A</td>
<td>-0.225</td>
<td>-1.23</td>
</tr>
<tr>
<td>Delta-vega</td>
<td>2B</td>
<td>-0.186</td>
<td>-0.389</td>
</tr>
</tbody>
</table>

**Table 12:** The statistic $z$ from (78) for comparing the errors obtained from hedging in the SV model ((37) with $\lambda \equiv 0$) versus the SVJ ((37)) model for the different calibration setups from Table 1. High values indicates a lower variability for the hedging errors from SV model whereas low values favour the SVJ model. Values of $|z| > 2.807$ that refute the hypothesis that the hedging error in the compared cases yield the same variability at a confidence level of 99.5% are written in *italic* if the SV model is favoured and in **bold** if the SVJ model is favoured.
\[
\begin{array}{cccccc}
\text{Expiry} & \text{Hedge} & \text{Calib. 1} & \text{Calib. 2} & \text{Moneyness} \\
\hline
\end{array}
\]

| Short dated $0 < T \leq 0.5$ | Delta 1A 1B | -4.28 | -0.958 | 0.0135 |
| Delta 1A 2A | -12.4 | -1.16 | 1.55 |
| Delta 1A 2B | -13.2 | -2.14 | 0.297 |
| Delta 1B 2A | -8.01 | -0.191 | 1.54 |
| Delta 1B 2B | -8.66 | -1.21 | 0.306 |
| Delta 2A 2B | -0.595 | -0.932 | -1.26 |

| Delta-vega 1A 1B | 1.26 | 0.348 | -1.21 |
| Delta-vega 1A 2A | 2.11 | 1.62 | -0.668 |
| Delta-vega 1A 2B | 2.92 | 3.47 | -0.807 |
| Delta-vega 1B 2A | 0.913 | 1.33 | 0.548 |
| Delta-vega 1B 2B | 1.71 | 3.13 | 0.413 |
| Delta-vega 2A 2B | 0.782 | 1.9 | -0.103 |

| Medium dated $0.5 < T \leq 2$ | Delta 1A 1B | -1.36 | 2.42 | 3.26 |
| Delta 1A 2A | -5.33 | -2.69 | 1.5 |
| Delta 1A 2B | -5.48 | -2.75 | 1.58 |
| Delta 1B 2A | -3.96 | -5.06 | -1.79 |
| Delta 1B 2B | -4.1 | -5.16 | -1.75 |
| Delta 2A 2B | -0.0938 | -0.0456 | 0.0567 |

| Delta-vega 1A 1B | 1.37 | 3.14 | 1.18 |
| Delta-vega 1A 2A | 2.18 | 4.53 | 1.05 |
| Delta-vega 1A 2B | 2.26 | 5.09 | 1.23 |
| Delta-vega 1B 2A | 0.852 | 1.38 | -0.172 |
| Delta-vega 1B 2B | 0.96 | 1.94 | 0.0232 |
| Delta-vega 2A 2B | 0.0966 | 0.547 | 0.194 |

Table 13: The statistic $z$ from (78) for comparing the errors obtained from hedging in the SVJ model in for the different calibration setups from Table 1. High values indicate a lower variability for the hedging errors from the calibration in the left column (Calib. 1) whereas low values favour the case in the right column (Calib. 2). Values of $|z| > 2.807$ that refute the hypothesis that the hedging error in the compared cases yield the same variability at a confidence level of 99.5% are written in italic if Calib. 1 is favoured and in bold if Calib. 2 is favoured.
References


A Quadratic variation between contracts

Suppose \( \bar{u}(t, S, y) \in C^{1,2,2} \) and likewise for \( \bar{v} \). Let \((S_t, y_t)\) be given by (37). For a general function \( U \in C^{1,2,2} \), Itô’s lemma gives

\[
dU(t, S_t, y_t) = \partial U(t, S, y)dt + \partial_S U(t, S_t, y)dS_t + \partial_y U(t, S_t, y)dy_t
\]

\[
+ \frac{1}{2} \partial_{SS} U(t, S, y) d\langle S \rangle_t + \frac{1}{2} \partial_{yy} U(t, S, y) d\langle y \rangle_t
\]

\[
+ U(t, S_t, e^{\Delta Z_t}, y) - U(t, S_t, y) + \partial_S U(t, S_t, y) S_t e^{\Delta Z_t} - 1,
\]

(79)

where \( S \) denotes the continuous part of \( S \) and \( S_{t-} \) the left limit at \( t \). If we now let

\[
U(t, S_t, y_t) := \bar{u}(t, S_t, y_t) \tilde{v}(t, S_t, y_t)
\]

then (79) and a development of the partial derivatives of \( U \) gives

\[
dU(t, S_t, y_t) = \left( \bar{v} \partial_{y} \bar{u} + \bar{u} \partial_{y} \bar{v} \right) dt + \left( \bar{v} \partial_{y} \bar{y} + \bar{u} \partial_{y} \bar{y} \right) d\langle y \rangle_t + \left( \bar{v} \partial_{y} \bar{y} + \bar{u} \partial_{y} \bar{y} \right) d\langle y \rangle_t
\]

\[
+ \frac{1}{2} \left( \bar{v} \partial_{y} \bar{y} + \bar{u} \partial_{y} \bar{y} \right) d\langle y \rangle_t
\]

\[
+ \left[ \bar{u}(t, S_t, e^{\Delta Z_t}, y_t) \tilde{v}(t, S_t, e^{\Delta Z_t}, y_t) - \bar{u}(t, S_t, y_t) \tilde{v}(t, S_t, y_t) \right]
\]

\[
+ \left[ \partial_S \bar{v} \bar{u} + \bar{u} \partial_S \bar{v} \right] S_t e^{\Delta Z_t} - 1.
\]

(80)

Now,

\[
d \langle \bar{u}, \bar{v} \rangle_t = d(\bar{u} \bar{v}_t) - \bar{u} \bar{v}_t dt - \bar{v} \bar{u}_t dt,
\]

(82)

where \( \bar{u}_t \) is shorthand for \( \bar{u}(t, S_t, y_t) \) and likewise for \( \bar{v}_t \). We therefore obtain, by developing \( \bar{d} \bar{u} \bar{v} \) and \( \bar{d} \bar{v} \bar{u} \) according to (79) and subtracting \( \bar{d} \bar{u} \bar{v} \) and \( \bar{d} \bar{v} \bar{u} \) from (81),

\[
d \langle \bar{u}, \bar{v} \rangle_t = \partial_S \bar{u} \partial_S \bar{v} S_t^2 dt + \partial_y \bar{u} \partial_y \bar{v} \theta^2 dt + \left( \partial_S \bar{v} \partial_y \bar{u} + \partial_S \bar{u} \partial_y \bar{v} \right) \rho S_t \bar{y} dt
\]

\[
+ \left[ \bar{u}(t, S_t, e^{\Delta Z_t}, y_t) \tilde{v}(t, S_t, e^{\Delta Z_t}, y_t) - \bar{u}(t, S_t, y_t) \tilde{v}(t, S_t, y_t) \right].
\]

(83)

We then get, at \( t = 0 \),

\[
\frac{d}{dt} \mathbb{E} [\langle \bar{u}, \bar{v} \rangle_t] = \partial_S \bar{u} \partial_S \bar{v} S_t^2 \mathbb{E} \bar{y}^2 + \partial_y \bar{u} \partial_y \bar{v} \theta^2 \mathbb{E} \bar{y}^2 + \left( \partial_S \bar{v} \partial_y \bar{u} + \partial_S \bar{u} \partial_y \bar{v} \right) \rho S_0 \mathbb{E} \bar{y}_{0 t}
\]

\[
+ \lambda \int [\bar{u}(t, S_t e^{\bar{z}}, y_t) - \bar{u}(t, S_t, y_t)] [\tilde{v}(t, S_t e^{\bar{z}}, y_t) - \tilde{v}(t, S_t, y_t)] \nu(z) dz.
\]

(84)

For the specific case \( \bar{v}(t, S_t, y_t) = \bar{x}_t = B(t, T)QS_t \) from (41) with \( Q = e^{-\int_t^1 q_s ds} \) we have

\[
\frac{d}{dt} \mathbb{E} [\langle \bar{u}, \bar{x} \rangle_t] = Q B(t, T) \left[ \partial_S \bar{u} S_t^2 \mathbb{E} \bar{y} + \partial_y \bar{u} \rho S_t \mathbb{E} \bar{y}_{tt} \right]
\]

\[
+ \lambda \int [\bar{u}(t, S_t e^{\bar{z}}, y_t) - \bar{u}(t, S_t, y_t)] S_t (e^{\bar{z}} - 1) \nu(z) dz.
\]

(85)

and if \( \bar{v}(t, S_t, y_t) = \bar{u}(t, S_t, y_t) = \bar{x}_t \)

\[
\frac{d}{dt} \mathbb{E} [\langle \bar{x}, \bar{x} \rangle_t] = Q^2 \frac{B_Z^2}{B_t} S_t^2 \mathbb{E} \bar{y} + \lambda \int S_t (e^{\bar{z}} - 1)^2 \nu(z) dz.
\]

(86)
B Density from characteristic functions and FFT

The densities of the log-spot distributions are not known in analytical form, but their characteristic functions are, so we can perform an inverse Fourier transform to obtain the densities. To do this efficiently, we will exploit the fast (discrete) Fourier transform (FFT). Concretely, we will calculate approximate values of the densities at values \( x_k = (k - 1 - \frac{M}{2}) \frac{a}{N} \), for \( k = 1, \ldots, M \), \( y_l = (l - 1 - \frac{N}{2}) \frac{a}{N} \), for \( l = 1, \ldots, N \). To this end, we will perform the integration in the frequency space over a number of points \( u_a = (a - 1 - \frac{M}{2}) \frac{a}{N} \) for \( a = 1, \ldots, M \) and where \( AB = 2\pi m \), and likewise for \( v_c = (c - 1 - \frac{N}{2}) \frac{a}{N} \) for \( c = 1, \ldots, N \) with \( CD = 2\pi N \). We denote the density function by \( \nu \), the characteristic function by \( \Phi \) and with the notation \( \chi(j, M) = 1 - \frac{1}{2} \left( 1_{\{j=1\}} + 1_{\{j=M\}} \right) \) use the trapezoid quadrature rule to obtain

\[
\nu(x_a, y_c) \approx \frac{1}{4\pi^2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-i(ux_a + vy_c)} \Phi(u, v) du dv
\]

\[
= \frac{1}{4\pi^2} \sum_{j=1}^{M} \sum_{k=1}^{N} \int_{(j-1-\frac{N}{2})\frac{a}{N}}^{(j-\frac{N}{2})\frac{a}{N}} \int_{(k-1-\frac{N}{2})\frac{a}{N}}^{(k-\frac{N}{2})\frac{a}{N}} e^{-i(ux_a + vy_c)} \Phi(u, v) du dv
\]

\[
\approx \frac{1}{4\pi^2} \frac{A}{M} \frac{C}{N} \sum_{j=1}^{M} \chi(j, M) e^{-i(j-1-\frac{N}{2})\frac{a}{N}(a-1-\frac{N}{2})\frac{a}{N}} \left\{ \right. \\
\sum_{k=1}^{N} \chi(k, N) \Phi\left( j - 1 - \frac{M}{2}, k - 1 - \frac{N}{2} \right) e^{-i(k-1-\frac{N}{2})\frac{a}{N}(c-1-\frac{N}{2})\frac{a}{N}} \left. \right\}
\]

\[
(87)
\]

\[
= \frac{1}{4\pi^2} \frac{A}{M} \frac{C}{N} e^{i\pi(a-1-\frac{N}{2})} e^{i\pi(c-1-\frac{N}{2})} \sum_{j=1}^{M} \chi(j, M) e^{i\pi(j-1)} e^{-i(j-1)(a-1)\frac{2\pi}{N}} \left\{ \right. \\
\sum_{k=1}^{N} \chi(k, N) \Phi\left( j - 1 - \frac{M}{2}, k - 1 - \frac{N}{2} \right) e^{i\pi(k-1)} e^{-i(k-1)(c-1)\frac{2\pi}{N}} \left. \right\}.
\]

The sums in the last expression are readily evaluated using Matlab’s implementation of FFT.

C Hedging in the Black-Scholes model

The BS model can be obtained as a degenerate version of the dynamics (37) by letting \( \lambda = \theta = \kappa = 0 \), which gives a spot process with constant volatility \( y_0 \). As is customary, let us denote this constant volatility by \( \sigma \). Denote by \( F(0, T) \) T-forward price of \( S \) and let, as before, \( B \) denote the value of a bank account so that \( B^{-1}(0, T) \) is the value of a zero coupon bond with maturity \( T \). From arbitrage arguments, it can be deduced that the price \( C(T, K) \) of a European call option on \( S \) of maturity \( T \) and strike \( K \) is bounded above by \( F(0, T) \) and bounded below by \( B^{-1}(0, T)(F(0, T) - K) \). Since the price of a call option in the BS model is strictly increasing in \( \sigma \), for any call option price \( C(T, K) \) within the arbitrage free bounds there exists a unique \( \sigma_{BS}(T, K) \) with a corresponding BS price exactly equal to \( C(T, K) \).

Now suppose market option quotes are available for a set of maturities \( \{T^i, 1 \leq i \leq M\} \) and - for each \( T^i \) - the strikes \( \{K^i, 1 \leq j \leq N_i\} \). Let \( u^{BS}(T^i, K^i) \) be the BS price function (where the dependance on the spot price \( S_0 \), the volatility \( \sigma \), the interest rate and dividend yield are omitted from the notation) an suppose for each pair \( (T^i, K^i) \) we choose \( \sigma = \sigma_{BS}(T^i, K^i) \) exactly replicate the quoted prices \( \hat{u}(T^i, K^i) \). The standard BS delta hedge consists in buying

\[
\delta^i_j = \partial_{S_0} u^{BS}(T^i, K^i)
\]

(88)
shares $S$ to hedge one unit of the option $\hat{u}(T^i, K^i_j)$. Suppose we want to create a portfolio $P$ to replicate the behaviour of $\omega^i_j$ number of options with strike $(T^i, K^i_j)$ for $(1 \leq i \leq M, 1 \leq j \leq N^i)$. In a delta-strategy, we let the initial portfolio value $P_0$ be given by

$$P_0 = \sum_{i=1}^{M} \sum_{j=1}^{N^i} \omega^i_j \hat{u}(T^i, K^i_j)$$

and buy a total number of shares

$$\delta^S = \sum_{i=1}^{M} \sum_{j=1}^{N^i} \omega^i_j \delta^i_j$$

where the $\delta^i_j$ come from (88).

In a BS model any option can be perfectly replicated by continuous time trading in only the underlying asset $S$. Nonetheless, even when using the BS model, market practitioners often hedge their positions by counter positions in other options as a means of taking into account the model’s incorrect assumption of a constant volatility. Suppose that we choose one option of maturity $\hat{T}$ and strike $\hat{K}$ to hedge the portfolio of other options above. A BS delta-vega strategy then consist in a portfolio with initial value $P_0$ from (89) which invests in a number $\delta^S$ of shares $S$ and $\delta^C$ contracts $\hat{u}(\hat{T}, \hat{K})$ such that

$$\delta^C \frac{\partial \hat{u}}{\partial \sigma}(\hat{T}, \hat{K}) \hat{u}(\hat{T}, \hat{K}) = \sum_{i=1}^{M} \sum_{j=1}^{N^i} \omega^i_j \frac{\partial \sigma}{\partial \sigma}(T^i, K^i_j) \hat{u}(T^i, K^i_j)$$

$$\delta^S + \delta^C \frac{\partial S}{\partial \sigma}(\hat{T}, \hat{K}) = \sum_{i=1}^{M} \sum_{j=1}^{N^i} \omega^i_j \delta^i_j,$$

where the $\delta^i_j$ are again given by (88). Once again, the BS model does not in its theoretical form justify a hedging strategy as the one above. Market practitioners use it as a way of trying to deal with the model’s imperfections. When we test our strategies on option market data below in Section (6), we will compare the outcome of the BS delta-strategy (90) and the delta-vega strategy (91) to the results from the quadratic hedging strategies in the affine model described in Section 3.

---

4The derivative of the price with respect to the volatility is usually referred to as the option’s vega and the spot derivative is known as the delta.