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, but it is not widely used among market practitioners and relatively few papers evaluate how well it works on real market data. Here, we develop a framework for comparing hedging strategies, and use it to empirically test the performance of quadrating 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 find that the 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 [10] and has since been extensively treated in the literature. For an overview of the literature we refer to the surveys of Pham [23] and Schweizer [25] and to the more recent results due to Černý [29].

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 [22] 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 [21]. In the sequel, we will refer to this model as the BS model. The results of El Karoui et al [16] 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 Eu-
European call options on an equity index. We investigate if the more elaborate modelling can be beneficial for traders dealing with standard options. 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 [7], where quadratic hedging theory is applied to hedge credit portfolios, and in Ewald et al. [9], where quadratic hedging in jump models are used to hedge options on crude oil. Also, it came to our knowledge when finalizing the first version of this report [19] 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 [14] and Heston’s model with normally distributed jumps in the log spot process (known as Bate’s model [4]). The paper by Bakshi et al. is discussed a little further in Section 8. In this report, we work with Heston’s model and with Heston’s model with exponentially distributed jumps. 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.

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 [23], [25], [29]). 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 a straightforward 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 now 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], [21], [11], [12], [13]) leads to unique prices for contingent claims. These prices can be expressed in terms of conditional expectations 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 \(X_i = (X_{1i}, \ldots, X_{di})' \in \mathbb{R}^d\) (where `' denotes the transpose) at \(t_i\). We assume that the filtration \(\mathcal{F}_t\) is given for all \(t \in [0, T]\) so that we can think of \(X_i\) 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

\[
B(\tau, T) = e^{\int_\tau^T r_t \, dt}
\]  

(1)

and set \(B_{i,j} = B(t_i, t_j)\). Suppose the process \(X\) is such that its \(t_N\)-forward value

\[
\bar{X}_i = B_{i,N}X_i
\]  

(2)
is a square integrable \(\mathcal{F}_t\)-martingale, where \(\mathcal{F}_t = \mathcal{F}_{t_i}\). 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

\[
\bar{u}_i = E[u_N | \mathcal{F}_i]
\]  

(3)

and define \(P\) according to

\[
P_{i+1} = P_i + (\delta_i, X_{i+1} - X_i) + (P_i - (\delta_i, X_i))(B_{i,i+1} - 1).
\]  

(4)

for some value \(P_0 \in \mathbb{R}\), where \((\cdot, \cdot)\) is the \(L^2\) inner product on \(\mathbb{R}^d\). The process \(\delta_i = (\delta_{1i}, \ldots, \delta_{di})'\) is an \(\mathcal{F}_t\)-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

\[
\Delta P_i := \bar{P}_{i+1} - \bar{P}_i.
\]  

(5)

For the \(t_N\)-forward portfolio value \(\bar{P}_i\) we then obtain

\[
\bar{P}_i = B_{i,N}P_i
\]

\[
\Delta \bar{P}_i := (\delta_i, \bar{X}_{i+1} - \bar{X}_i) = (\delta_i, \Delta \bar{X}_i).
\]  

(6)

We can write

\[
\bar{u}_N = \bar{u}_i + \sum_{j=i}^{N-1} \Delta \bar{u}_j
\]

\[
\bar{P}_N = \bar{P}_i + \sum_{j=i}^{N-1} \Delta \bar{P}_j = \bar{P}_i + \sum_{j=0}^{N-1} (\delta_j, \Delta \bar{X}_j)
\]  

(7)
and since the martingale increments $\Delta \bar{X}_i, \Delta \bar{u}_i \perp \mathcal{F}_i$ (where $\perp$ denotes orthogonality in $L^2(\Omega, \mathcal{F}, Q)$), we get

$$E \left[ (u_N - P_N)^2 \right] = E \left[ (\bar{u}_0 - P_0) + \sum_{i=0}^{N-1} \{ \Delta \bar{u}_i - (\delta_i, \Delta \bar{X}_i) \}^2 \right]$$

$$= (\bar{u}_0 - P_0)^2 + \sum_{i=0}^{N-1} E \left[ \{ \Delta \bar{u}_i - (\delta_i, \Delta \bar{X}_i) \}^2 \right].$$

(8)

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

$$E \left[ (u_N - P_N)^2 \right] = (\bar{u}_0 - P_0)^2 + \sum_{j=0}^{N-1} E \left[ \{ \Delta \bar{u}_j - (\delta_j, \Delta \bar{X}_j) \}^2 \right].$$

(9)

The infimum of the norm in (9) is obtained by projecting each $\Delta \bar{u}_j$ onto the subspace of $L^2(\Omega, \mathcal{F}, Q)$ spanned by $(\delta_j, \Delta \bar{X}_j)$ for $\delta_j \in L^2(\Omega, \mathcal{F}, Q), j = 1, \ldots, d$. However, since $\delta_i \not\in \mathcal{F}_0$ for $i > 0$, only $\delta_0$ can be known at time $i = 0$. We then search for $(\bar{P}_0, \delta_0)$ satisfying

$$(\bar{P}_0, \delta_0) = \arg\min_{(p,\theta) \in \mathbb{R}^d} \left\{ (\bar{u}_0 - p)^2 + E \left[ \{ \Delta \bar{u}_0 - (\theta, \Delta \bar{X}_0) \}^2 \right] \right\}.$$

(10)

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

$$\Theta = \left\{ (\theta, \bar{X}_0) : \theta \in \mathbb{R}^d \right\},$$

(11)

of $L^2(\Omega, \mathcal{F}_0, Q)$ which is finite dimensional and thus closed. The orthogonality condition for the residual of the projection gives

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

(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

$$b_k = E \left[ \Delta \bar{X}_0^k \Delta \bar{u}_0 \right]$$

$$M_{kl} = E \left[ \Delta \bar{X}_0^k \Delta \bar{X}_0^l \right]$$

(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

$$\delta_0 = M^{-1}b.$$  

(14)

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

$$b_k = E \left[ \Delta \bar{X}_0^k \Delta \bar{u}_0 \right] = E \left[ \left\langle \bar{X}_t^k, \bar{u}_t^l \right\rangle_{\Delta t} \right]$$

$$M_{kl} = E \left[ \Delta \bar{X}_0^k \Delta \bar{X}_0^l \right] = E \left[ \left\langle \bar{X}_t^k, \bar{X}_t^l \right\rangle_{\Delta t} \right]$$

(15)
where $\langle \cdot, \cdot \rangle$ is the quadratic variation process [24, Theorem 33]. If we divide by $\Delta t$ in (15) then in the limit $\Delta t \to 0$ we get

$$b_0^k := \lim_{\Delta t \to 0} \frac{1}{\Delta t} E \left[ \langle \bar{X}^k, \bar{u}^l \rangle_{\Delta t} \right] = \frac{d}{dt} E \left[ \langle \bar{X}^k, \bar{u}^l \rangle_{|r=0} \right],$$

$$M_{kl}^0 := \lim_{\Delta t \to 0} \frac{1}{\Delta t} E \left[ \langle \bar{X}^k, \bar{X}^l \rangle_{\Delta t} \right] = \frac{d}{dt} E \left[ \langle \bar{X}^k, \bar{X}^l \rangle_{|r=0} \right].$$

(16)

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

$$\delta_0 = (M^0)^{-1} b^0,$$

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 [23] and Schweizer [25]. Existence of an optimal strategy in this case is usually proved by invoking the Kunita-Watanabe decomposition [18], [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

$$\bar{X}_t = B(t,T)X_t$$

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

$$\bar{u}_t = E[u_T | \mathcal{F}_t].$$

(19)

As before, we want to hedge $u_T$ by minimizing the $L^2(\Omega, \mathcal{F}_T, Q)$-distance between $u_T$ and a self-financing portfolio $P$ investing in $\bar{X}$ according to

$$dP_t = (\delta_t, dX_t) + (P_t - (\delta_t, X_t)) r_t dt.$$  

(20)

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

$$(\delta, \theta)_{L^2(\bar{X}, T)} = E \left[ \int_0^T (\delta_t, \theta_t) \, d\bar{X}_t \right],$$

(21)

where $\langle \cdot \rangle$ denotes the bracket process [24, 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 [24] states that

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

(22)
This gives, for the scalar products in $L^2(\Omega, \mathcal{F}_T, Q)$ and $L^2(\bar{X}, T)$,
\[
\left( \int_0^T (\delta_t, d\bar{X}_t), \int_0^T (\theta_t, d\bar{X}_t) \right)_{L^2(\Omega, \mathcal{F}_T, Q)} = (\delta, \theta)_{L^2(\bar{X}, T)}.
\] (23)

The $T$-forward value $\bar{P}$ of the portfolio $P$ will evolve according to
\[
\bar{P}_t = B(t, T)P_t, \; d\bar{P}_t = (\delta_t, d\bar{X}_t),
\] (24)
and we can write
\[
\bar{u}_T = \bar{u}_0 + \int_0^T d\bar{u}_t, \; \bar{P}_T = \bar{P}_0 + \int_0^T (\delta_t, d\bar{X}_t).
\] (25)

We want to minimize the $L^2(\Omega, \mathcal{F}_T, Q)$-distance between $u_T$ and $P_T$, i.e.
\[
E \left[ (u_T - P_T)^2 \right] = E \left[ \left( \bar{u}_0 - \bar{P}_0 \right)^2 + \sum_{i=0}^{N-1} E \left[ \left( \int_{t_i}^{t_{i+1}} d\bar{u}_t - \int_{t_i}^{t_{i+1}} (\delta_t, d\bar{X}_t) \right)^2 \right] \right],
\] (26)
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
\[
(\bar{P}_0, \delta) = \text{argmin}_{(p, \theta) \in \mathbb{R} \times \Theta_\tau} \left\{ (\bar{u}_0 - p)^2 + E \left[ \left( \int_0^\tau d\bar{u}_t - \int_0^\tau (\theta_t, d\bar{X}_t) \right)^2 \right] \right\},
\] (27)
were $\Theta_\tau$ is the linear subspace of $L^2(\Omega, \mathcal{F}_\tau, Q)$ defined as
\[
\Theta_\tau = \left\{ \int_0^\tau (\theta_t, d\bar{X}_t) : \theta \in L^2(\bar{X}, \tau) \right\}.
\] (28)

From the isometry relation (22) we have, for $\theta \in L^2(\bar{X}, \tau)$,
\[
\left\| \int_0^\tau (\theta_t, d\bar{X}_t) \right\|_{L^2(\Omega, \mathcal{F}_T, Q)} = \|\theta\|_{L^2(\bar{X}, \tau)}
\] (29)
and so $\Theta_\tau$ is closed in $L^2(\Omega, \mathcal{F}_T, Q)$ 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
\[
\sum_{l=1}^{d} E \left[ \int_0^\tau \delta_l^t \, d \left( \bar{X}_k, \bar{X}_l^t \right)_t \right] = E \left[ \int_0^\tau d \left( \bar{X}_k, \bar{u}^t \right)_t \right].
\] (31)

The choice of \( \tau \) is of course arbitrary in (31), since equation (26), which says that we can chose the optimal strategy \( \delta \) independently on each time interval, will be valid for any subdivision of \([0, T]\). Since (31) is then valid for any \( \tau \in [0, T] \), we must have equality in (31) if we differentiate both sides with respect to \( \tau \). This yields
\[
\sum_{l=1}^{d} \delta_l^0 \frac{d}{d \tau} E \left[ \int_0^\tau \left( \bar{X}_k, \bar{u}^t \right)_t \right] |_{\tau=0} = \frac{d}{d \tau} E \left[ \int_0^\tau \left( \bar{X}_k, \bar{u}^t \right)_t \right] |_{\tau=0}.
\] (32)

If we then define
\[
\theta_k^0 := \frac{d}{d \tau} E \left[ \left( \bar{X}_k, \bar{u}^\tau \right)_t \right] |_{\tau=0}
\]
\[
M_{kl}^0 := \frac{d}{d \tau} E \left[ \left( \bar{X}_k, \bar{X}_l^\tau \right)_t \right] |_{\tau=0}
\] (33)

for \( k, l = 1, \ldots, d \), we retrieve
\[
\delta_0 = (M^0)^{-1} b^0
\] (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 [18], [1]), which gives stronger results for the projection of \( \bar{u}_T \) onto \( \Theta_T \) in (27). The Kunita-Watanabe decomposition assures that \( \bar{u}_T \), being an arbitrary square integrable martingale, can be decomposed according to
\[
\bar{u}_T = \bar{u}_0 + \int_0^T \left( \theta_t^\bar{u}, d \bar{X}_t \right) + L_T^\bar{u},
\] (35)

where \( \theta^\bar{u} \in L^2(\bar{X}, T) \) and \( L^\bar{u} \) is a zero-mean \( \mathcal{F}_t \)-martingale orthogonal to \( \bar{X} \) in the sense that \( \langle \bar{X}, L^\bar{u} \rangle_t = 0, t \in [0, T] \). In view of (35) we rewrite (27) as
\[
(\bar{P}_0, \delta) = \arg \min_{(p, \theta) \in \mathbb{R} \times \Theta} \left\{ (\bar{u}_0 - p)^2 + E \left[ \left( \int_0^T ((\theta_t^\bar{u} - \theta_t^\bar{u}), d \bar{X}_t) \right)^2 \right] + E \left[ (L_T^\bar{u})^2 \right] \right\}.
\] (36)

It is obvious that (36) is minimized by setting \( p = \bar{u}_0 \) and \( \theta_t = \theta_t^\bar{u} \), so that the optimal \( \delta_t \) in (27) is equal to \( \theta_t^\bar{u} \) from the Kunita-Watanabe decomposition (35) for \( t \in [0, T] \).

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
\[
dS_t = (r_t - q_t) S_{t-} \, dt - \left( \int_{\mathbb{R}} (e^{z} - 1) \lambda \nu(z) \, dz \right) S_{t-} \, dt
\]
\[
+ \sqrt{y_t} S_{t-} \, dW^S_t + S_{t-} \left( e^{dZ_t} - 1 \right)
\] (37)
\[
dy_t = \kappa(\eta - y_t) \, dt + \theta \sqrt{y_t} \, dW^y_t,
\]
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

\[
\nu(z) = \frac{1}{\mu} 1_{\{z \leq 0\}} e^{\mu z},
\]

(38)

for some \(\mu > 0\). If the jump intensity \(\lambda \equiv 0\), then (37) is the model introduced in finance by Heston [14]. Bates [4] uses the same model but with normally distributed jumps. Kou [17] 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 \(\lambda \neq 0\) 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. [8]. These models are popular in finance since 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 ([14], [8]). 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

\[
u(t, S_t, y_t) := B(t, T) E[h(S_T) | \mathcal{F}_t]
\]

(39)

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

\[
dP_t = \delta^S_t dS_t + \delta^u_t du_t + (P_t - \delta^S_t S_t - \delta^u_t u_t) r_t dt + \delta^S_t S_t q_t dt,
\]

(40)

\[
\]
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
\[
d\bar{P}_t = B(t, T) \left\{ \delta^S_t (dS_t - S_t (r_t - q_t) dt) + \delta^u_t (du_t - u_t r_t) dt \right\}
\]
\[
= Q^{-1}(t, T) \delta^S_t (B(t, T) Q(t, T) S_t) + \delta^u_t d (B(t, T) u_t),
\]
which can be expressed as
\[
d\bar{P}_t = \delta^x_t d\bar{x}_t + \delta^u_t d\bar{u}_t,
\]
where
\[
\delta^x_t = Q^{-1}(t, T) \delta^S_t,
\]
\[
\bar{u}_t = B(t, T) u_t.
\]
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 first equality in (44) to obtain the number \( \delta^S \) of shares \( S \) that should be held in the portfolio once \( \delta^x \) 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
\[
\int_{t_i}^{t_{i+1}} q_s dt \approx S_{t_i+1} \left( e^{\int_{t_i}^{t_{i+1}} q_s dt} - 1 \right) = S_{i+1} (Q_{i+1,i} - 1),
\]
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
\[
\Delta P_{t_i} = P_{i+1} - P_i = \delta^S_{i+1} (S_{i+1} - S_i) + \delta^u_{i+1} (u_{i+1} - u_i) + (P_i - \delta^S_i S_i - \delta^u_i u_i) (B_{i+1} - 1)
\]
\[
+ \delta^S_{i+1} (Q_{i+1,i} - 1).
\]
For the \( T \)-forward value \( \bar{P}_i = B_{i,N} P_i \) of the portfolio, equation (46) directly leads to
\[
\Delta \bar{P}_i = \bar{P}_{i+1} - \bar{P}_i
\]
\[
= Q_{i,N}^{-1} \delta^S_{i+1} (B_{i+1,N} Q_{i+1,N} S_{i+1} - B_{i,N} Q_{i+1,N} S_{i+1}) + \delta^u_{i+1} (B_{i+1,N} u_{i+1} - B_{i,N} u_i)
\]
\[
+ \delta^x_{i+1} (\bar{x}_{i+1} - \bar{x}_i) + \delta^u_{i+1} (\bar{u}_{i+1} - \bar{u}_i) = \delta^x_i \Delta \bar{x}_i + \delta^u_i \Delta \bar{u}_i,
\]
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 can 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 by simultaneous positions in \( S \) and \( v \). We will refer to the case when only the spot process \( S \) is used as \textit{quadratic delta hedging} and to the case where both \( S \) and \( u \) are used as \textit{quadratic delta-vega hedging}. As above, we will use \( u_t \) 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 \( 0 = t_0 < t_1 < \ldots < t_N = T^1 \) as in Section 2.1. To obtain the optimal hedging ratios at \( t_0 = 0 \) according to (13), we need to calculate \( E[\Delta \bar{X}^k_0 \Delta \bar{X}^l_0] \) and \( E[\Delta \bar{X}^k_0 \Delta \bar{u}^l_0] \) for \( k, l = 1, 2 \), i.e.,

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

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,

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

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 of \( \bar{u} \) in continuous time. 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 \( \langle \bar{u}, \bar{x} \rangle_t \) and \( \langle \bar{x}, \bar{x} \rangle_t \). 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 (79), (81) yield the
optimal number $\delta^a_t$ of shares $x_t$ to hold at $t = 0$. Changing into the optimal number $\delta^S_t$ of shares $S_t$ using (44) yields

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

(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 [27, 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

$$
\delta^S_t = \partial_S u + \frac{1}{S_t} \rho \theta \partial_y u,
$$

(51)

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 $X = (x, v)$. For $b_1^0$ from (33) at $t = 0$, we get from (80), Appendix A, that

$$
b_1^0 = \frac{d}{dt} E [\langle \bar{x}, \bar{u} \rangle] = Q(t, T) B^2(t, T) \left[ S_t^2 y_t \partial_S u + \rho S_t \theta y_t \partial_y u + S_t \int [u^1(t, S_t e^z, y_t) - u^1(t, S_t, y_t)] (e^z - 1) \lambda \nu(z)dz \right].
$$

(52)

and from (80), Appendix A, we have

$$
b_2^0 = \frac{d}{dt} E [\langle \bar{u}, \bar{v} \rangle] = B^2(t, T) \left[ S_t^2 y_t \partial_S u \partial_S v + \theta^2 y_t \partial_y u \partial_y v + \rho S_t \theta y_t (\partial_S v \partial_y u + \partial_S u \partial_y v) + \lambda \int [u(t, S_t e^z, y_t) - v(t, S_t, y_t)] [v(t, S_t e^z, y_t) - v(t, S_t, y_t)] \nu(z)dz \right].
$$

(53)

As for the matrix $M^0$ in (33), we have that $M^0_{11} = \frac{d}{dt} E [\langle \bar{x}, \bar{x} \rangle]$ is given by (81), Appendix A, and that $M^0_{12} = M^0_{21} = \frac{d}{dt} E [\langle \bar{x}, \bar{v} \rangle]$ is obtained by replacing $u$ by $v$ and $B_{T_1}$ by $B_{T_2}$ in (52). The element $M^0_{22}$ is given by setting $u = v$ in (79):

$$
M^0_{22} = \frac{d}{dt} E [\langle \bar{v}, \bar{v} \rangle] = B^2(t, T) \left[ S_t^2 y_t (\partial_S v)^2 + \theta^2 y_t (\partial_y v)^2 + 2 \rho S_t \theta y_t \partial_S v \partial_y v \right. \left. + \lambda \int [v(t, S_t e^z, y_t) - v(t, S_t, y_t)]^2 \nu(z)dz \right].
$$

(54)

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 $\int u(t, S_t e^z, y_t) e^z \nu(z)dz$. 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 Viertlauer [15] 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.

Now assume that we at a date \( t_i \) observe a vector \( \hat{u} \in \mathbb{R}^m \) with values \( \hat{u}^j, j = 1, \ldots, m \) that represent market quotes on call options with maturity and strike \((T_j, K_j)\). Market quotes are given as bid- and ask prices, so we let \( \hat{u} \) be defined as the mid prices, i.e.,

\[
\hat{u}^j = \frac{1}{2} \left( \hat{u}_{\text{bid}}^j + \hat{u}_{\text{ask}}^j \right). \tag{55}
\]

Let \( \alpha \in \mathbb{R}^d \) be a real valued vector containing the parameters of the model (37) except from the state variables \( S \), the interest rate \( r \) and the dividend yield \( q \). We can let \( \alpha = (y_0, \kappa, \eta, \theta, \rho, \lambda, \nu) \) for the SVJ model and \( \alpha = (y_0, \kappa, \eta, \theta, \rho) \) for the SV model without jumps. We let \( u_\alpha \in \mathbb{R}^m \) denote a vector whose elements \( u_\alpha^j, j = 1, \ldots, m \) correspond to the prices of the call options given by the model (37), with the same maturities and strikes as for the options in \( \hat{u} \). The prices in \( u_\alpha \) are thus given by (39), for different maturities and strikes. We will determine the parameter vector \( \alpha \) by simply minimizing the \( L^2 \)-distance between the observed prices and the model prices with respect to the parameter vector \( \alpha \). We thus let \( \| \cdot \| \) be the Euclidian norm in \( \mathbb{R}^d \), where \( d \) is the number of parameters in \( \alpha \), and solve the minimization problem

\[
\min_{\alpha \in \mathcal{A}} \| u_\alpha - \hat{u} \|^2, \tag{56}
\]

for some subspace \( \mathcal{A} \subseteq \mathbb{R}^d \). The optimization problem (56) is solved using the trust-region Newton method available in Matlab\(^1\).

In [19], we considered alternative formulations of the optimization problem used for determining the model parameters \( \alpha \). Since we wish to illustrate the effect of the use of quadratic hedging in a stochastic volatility model in the simplest setting possible, we have now omitted these formulations. In this context, it should be mentioned that calibration techniques that take into account the historical performance of the model’s hedging capability have been considered in Lindström et al. [20], where a Kalman filter is used.

5 Improving delta hedging by linear 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 [28], but there an extra parameter needs to be estimated from market data.

Suppose now that $\hat{u}_t$ 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 $u_t(S_t, \sigma) = \hat{u}_t$. 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. $\partial_{\sigma} u$, might have an impact on the optimal $\delta$ to choose 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 $\sigma$, 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 $\partial_{\sigma^2} u$. To this end, we will search for a delta hedging rule $\delta$ given by

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

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 $\delta_t^S$ from (50) and search for a strategy

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

where $u$ is evaluated using parameters calibrated as described in Section 4.

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 $\Delta \hat{u}_i^{j} = \hat{u}_i^{j+1} - \hat{u}_i^{j}$.

For notational convenience, let us put

$$
D_i := S_{i+1} Q_{i+1, i} - S_i B_{i,i+1}.
$$
Suppose that $\delta^j_i$ in (60) is on the form (57) and let us denote the model function that corresponds to the contract $\hat{u}^j_i$ by $u^j_i$. If the portfolio increment (60) were to perfectly match the contract price increment (59) with the strategy (57), we would have

$$
\left(\alpha_1 \partial_S u^j_{i-1} + \alpha_2 \partial_u u^j_{i-1} + \alpha_3 \partial_S \sigma u^j_{i-1}\right) D_i = u^j_i - u^j_{i-1}\{1 + r_{i-1}(t_i - t_{i-1})\}.
$$

(62)

where $u^j_{i-1}\{1 + r_{i-1}(t_i - t_{i-1})\}$ represents the forward price of $u$ from time $t_{i-1}$ to $t_i$. Denote the right-hand side in (62) by $\beta^j_i := u^j_i - u^j_{i-1}\{1 + r_{i-1}(t_i - t_{i-1})\}$

(63)

and let $\beta$ be a column vector given by $\beta = (\beta^1_1, \ldots, \beta^n_1, \ldots, \beta^1_m, \ldots, \beta^n_m)'$. Likewise, we use the notation

$$
\Lambda^1_i = \partial_S u^j_{i-1} D_i
$$

$$
\Lambda^2_i = \partial_u u^j_{i-1} D_i
$$

$$
\Lambda^3_i = \partial_S \sigma u^j_{i-1} D_i
$$

(64)

and form the column vectors $\Lambda^k = \left(\Lambda_{1}^{k,1}, \ldots, \Lambda_{1}^{k,n}, \ldots, \Lambda_{m}^{k,1}, \ldots, \Lambda_{m}^{k,n}\right)'$ for $k = 1, 2, 3$. If our portfolio strategy were to perfectly match all contract increments $\Delta u^j_i$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, we would now have

$$\Lambda \alpha = \beta,
$$

(65)

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 (65) holds. What we will do is to find the $\alpha$ that satisfies (65) in a least squares’ sense:

$$\alpha = \arg\min_{a \in \mathbb{R}^3} \|\Lambda a - \beta\|,
$$

(66)

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 (57) in the BS case and from (58) 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 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 recorded 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

---

2Courtesy of Svenska Handelsbanken AB.

3A small number of the trading days were missing from our data.
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 denote by \( t_{0}, t_{1}, \ldots, t_{m} \) the times at which we have sampled the market’s option prices. We will consider a series of hedging portfolios that only last for one time-step, from a time \( t_{i-1} \) to \( t_{i} \) for \( i = 1, \ldots, m \). For each consecutive pair \( \{t_{i-1}, t_{i}\} \) 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 \( I_{l} \), \( l = 1, \ldots, 6 \), of the maturity-strike pairs among the quoted options in the market. Each subset \( I_{l} \) 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

\[
I_{l} = \{(T, K) \text{ such that } T_{\min} < T < T_{\max}, k_{\min} < K/S_{t_{k-1}} < k_{\max} \text{ and a call option } C(T, K) \text{ is quoted at } t_{i-1} \text{ and } t_{i}\}. \tag{67}
\]

We will test our strategies on observed call contract increments of the type (59) 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 1, where we give the total number of increments used for each subset and the total number of options summed over all hedging times. For each of the two time intervals, we will call the category of options in Table 1 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 1, is to try to replicate a portfolio of total capital \( c \) and with the same amount of money invested in each of the contracts \( \hat{u}_{i-1}(T, K) \) quoted at \( t_{i} \) for \( (T, K) \in I_{l} \). The values \( \hat{u}_{i-1} \) 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

\[
\omega_{i-1}(T, K) = \frac{c}{\#I_{l}} \hat{u}_{i-1}(T, K), \tag{68}
\]

number of contracts \( \hat{u}_{i-1}(T, K) \) for each \( (T, K) \in I_{l} \), where \#\( I_{l} \) is the total number of options in \( I_{l} \). At a given observation time \( t_{i-1} \), we construct a hedging portfolio \( P_{i-1}(T, K) \) for each
Table 1: Total number of increment dates (and the total number of options used) in the different subcategories considered.

<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 ≤ 0.25</td>
<td>1283 (24911)</td>
<td>1624 (73190)</td>
<td>1091 (19294)</td>
</tr>
<tr>
<td>Medium dated</td>
<td>0.25 &lt; T ≤ 2</td>
<td>1150 (29815)</td>
<td>1624 (158840)</td>
<td>1238 (41422)</td>
</tr>
</tbody>
</table>

Expiry Moneyness

| 0.7 ≤ \( \frac{K}{S_0} \) < 0.9 | 0.9 ≤ \( \frac{K}{S_0} \) ≤ 1.1 | 1.1 < \( \frac{K}{S_0} \) ≤ 1.3 |

Table 1: Total number of increment dates (and the total number of options used) in the different subcategories considered.

\( (T, K) \in \mathcal{I} \), aiming to replicate one unit of the option \( \hat{u}_{t-1}(T, K) \) and thus with an initial value at \( t_{i-1} \) that equals the option price: \( P_{t-1}(T, K) = \hat{u}_{t-1}(T, K) \). Our total replicating portfolio, which we will denote by \( \hat{P} \), will be the weighted sum of the portfolios \( P_{t-1}(T, K) \) according to

\[
\hat{P}(T, K) := \sum_{(T,K) \in \mathcal{I}_k} \omega_{t-1}(T, K) P_{t-1}(T, K).
\]

The evolution of each sub-portfolio \( P_{t-1}(T, K) \) 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 \( \bar{T} \) that is closest to 180 days and, for this maturity, the strike \( \bar{K} \) 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

\[
\text{err}_i = \hat{P}_i - \hat{P}_{t-1} - \sum_{(T,K) \in \mathcal{I}_k} \omega_{t-1}(T, K) (\hat{u}_i(T, K) - \hat{u}_{t-1}(T, K)),
\]

where \( \omega_{t-1}(T, K) \) are given by (68). 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 \( \omega_{t-1}(T, K) \) units of each contract with maturity-strike \( (T, K) \in \mathcal{I}_k \) 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 \( t_i \).

6.1 Comparing hedging performance

Given the hedging error series \( \text{err}_i \) from (70) 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 à 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

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

(71)

where $n$ is the length of the series (see Table 1) and

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

(72)

with $\text{err}_j$ from (70). 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 [26] 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 [26]. Let $a$ and $b$ denote different data series, both with 9 data points. Table 2 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 1) the Siegel-Tukey test uses an approximately normal statistics $z$ defined as

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

(73)
Table 2: Example from [26] 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>a</td>
<td>a</td>
<td>b</td>
<td>a</td>
<td>a</td>
<td>b</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 3: 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>

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 \( \pm 1 \) in the numerator is chosen to maximize the module of \( z \).

Rather than calculating \( z \) from (73) directly from the series \( \text{err}_i \) from (70), we use the error series \( \hat{\text{err}}_i \) from (72) 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 (71), the choice of using the adjusted series \( \hat{\text{err}}_i \) rather than \( \text{err}_i \) 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

\[
H_0 \quad \text{The two series have the same variability.}
\]

\[
H_a \quad \text{Series } a \text{ has lower variability than series } b.
\]

\[
H_b \quad \text{Series } b \text{ has lower variability than series } a.
\]

If \( R_1 \) in (73) 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 3 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.

In [19], we compared results for Heston’s SV model using the quadratic hedging strategy in discrete time from Section 3.2.1 and the strategy defined in continuous time from Section (3.2.2). We found no evidence to support that the two strategies show any differences in hedging quality when option data separated by a time of 7 days was used. This indicates that the error in the models themselves is larger than the difference between calculating the optimal hedging ratios for continuous versus discrete time. We do not display the results from the discrete time strategy here, and the results discussed below are all obtained using the continuous time hedging ratios from Section (3.2.2).
The sample standard deviation of the hedging error 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 4. 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 roughly 1.5 to 2 times higher in the BS model than in the stochastic models. The stochastic models perform better for medium dated options OTM options, in both delta and delta-vega hedging.

Figures 1(a) to 1(d) yield 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. It is clear from the pictures that the errors are smaller for the stochastic model. Figures 1(e) to 1(h) show the time series and histograms of errors for delta-vega hedging of short dated ATM options in the BS and the SVJ model. Again, it is clear that the stochastic model shows a lower error. The pictures thus tell the same story as the standard deviations in Table 4.

Table 5 display the values of the statistic $z$ from (73) for comparing the hedging results from Table 4 of the BS model and the stochastic models. The $z$ statistic mostly gives the same picture as the standard deviation. For short dated options, quadratic hedging in the stochastic models is preferable. For medium dated options, the stochastic models perform better for OTM options. The stochastic models also perform better for vega hedging of medium dated ATM options. For delta hedging of medium dated ATM options, the $z$ statistic favours the BS model, although the standard deviation of the errors are of similar size for the different models, as can be seen in Table 4.

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. Table 6 shows the error standard deviations for delta hedging with regression in the different models. It is apparent that the values of the different models are close for all of the displayed subcategories of options. This picture is supported by the $z$ statistic displayed in Table 7, which compares delta hedging with regression between the BS model and the two stochastic models. In none of the cases do we find any significant advantage for either the BS model or the stochastic models.

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 model through more or less formalized procedures, we also developed a simple regression technique to improve delta hedging performance in both the BS model 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 in the money options and for at the money 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.

The type of regression technique used in the delta hedging case is less evident to apply
when we hedge with other contracts, since it would require different regression calculations de-
pending 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.

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 the results of Bakshi et al. and ours. Reasons for this are that
we use larger moneyness intervals in categorizing our options, that options of higher and lower
strikes are traded today than in that period, and also that 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 options are 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]. 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 general 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>
<th>ITM 0.7 ≤ ( \frac{K}{S_0} &lt; 0.9 )</th>
<th>ATM 0.9 ≤ ( \frac{K}{S_0} &lt; 1.1 )</th>
<th>OTM 1.1 ≤ ( \frac{K}{S_0} ≤ 1.3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short dated</td>
<td>Delta</td>
<td>BS</td>
<td>0.659%</td>
<td>6.33%</td>
<td>20.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>0.658%</td>
<td>4.52%</td>
<td>10.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>0.949%</td>
<td>4.84%</td>
<td>10.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>BS</td>
<td>0.552%</td>
<td>4.99%</td>
<td>17.4%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>0.548%</td>
<td>3.78%</td>
<td>9.40%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>0.462%</td>
<td>3.25%</td>
<td>9.06%</td>
<td></td>
</tr>
<tr>
<td>Medium dated</td>
<td>Delta</td>
<td>BS</td>
<td>0.360%</td>
<td>1.48%</td>
<td>7.00%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SV</td>
<td>0.400%</td>
<td>1.56%</td>
<td>5.28%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>0.456%</td>
<td>1.60%</td>
<td>4.98%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>BS</td>
<td>0.242%</td>
<td>0.788%</td>
<td>5.99%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>0.255%</td>
<td>0.503%</td>
<td>3.47%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>0.229%</td>
<td>0.421%</td>
<td>3.19%</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Sample standard deviation (in percentage of invested capital) for the error series \( \text{err}_i \) from (70) obtained from hedging in the BS model and quadratic hedging in the stochastic model (37) with jumps (SVJ) and without jumps (SV, \( \lambda \equiv 0 \)).
Figure 1: Time series for $\text{err}_t$ from (70) and their corresponding histograms hedging errors in the BS model and quadratic hedging in the stochastic volatility models. Figures 1(a) to 1(d) show results for delta hedging short dated OTM options in the BS model and quadratic delta hedging in the SV model (model (37) with $\lambda = 0$). Figures 1(e) to 1(h) show results for delta-vega hedging of short dated ATM options in the BS model and quadratic delta-vega hedging in the SVJ model (37).
Expiry Hedge Model Moneyness

<table>
<thead>
<tr>
<th>Expiry</th>
<th>Hedge</th>
<th>Model</th>
<th>ITM (0.7 \leq \frac{K}{S_0} &lt; 0.9)</th>
<th>ATM (0.9 \leq \frac{K}{S_0} \leq 1.1)</th>
<th>OTM (1.1 &lt; \frac{K}{S_0} \leq 1.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Short dated</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0 &lt; T \leq 0.25)</td>
<td>Delta</td>
<td>SV</td>
<td>0.358</td>
<td>-7.16</td>
<td>-16.3</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>13.3</td>
<td>-6.20</td>
<td>-16.6</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>-0.134</td>
<td>-5.65</td>
<td>-15.2</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>-3.13</td>
<td>-9.49</td>
<td>-15.5</td>
</tr>
<tr>
<td><strong>Medium dated</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.25 &lt; T \leq 2)</td>
<td>Delta</td>
<td>SV</td>
<td>3.37</td>
<td>6.96</td>
<td>-5.73</td>
</tr>
<tr>
<td></td>
<td>Delta</td>
<td>SVJ</td>
<td>8.39</td>
<td>7.63</td>
<td>-8.03</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SV</td>
<td>0.587</td>
<td>-11.9</td>
<td>-13.4</td>
</tr>
<tr>
<td></td>
<td>Delta-vega</td>
<td>SVJ</td>
<td>-1.87</td>
<td>-17.2</td>
<td>-15.2</td>
</tr>
</tbody>
</table>

**Table 5:** The statistic \(z\) from (73) for comparing the error series \(err_i\) from (70) obtained from hedging in the BS model and quadratic hedging in the stochastic model (37) with jumps (SVJ) and without jumps (SV, \(\lambda \equiv 0\)). 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.
Table 6: Sample standard deviation (in percentage of invested capital) for the error series err$_i$ from (70) obtained from delta hedging with regression as in Section 5 in the BS model as well as in the stochastic model (37) with jumps (SVJ) and without jumps (SV, $\lambda \equiv 0$).

<table>
<thead>
<tr>
<th>Expiry</th>
<th>Hedge Model</th>
<th>Moneyness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 &lt; $T \leq 0.25$</td>
<td>Delta BS</td>
<td>0.658%</td>
</tr>
<tr>
<td></td>
<td>Delta SV</td>
<td>0.659%</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ</td>
<td>0.681%</td>
</tr>
<tr>
<td>0.6 ≤ $\frac{K}{S_0} &lt; 0.8$</td>
<td>Delta BS</td>
<td>0.373%</td>
</tr>
<tr>
<td></td>
<td>Delta SV</td>
<td>0.376%</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ</td>
<td>0.381%</td>
</tr>
</tbody>
</table>

Table 7: The statistic $z$ from (73) 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. High values favour the BS model and low values the stochastic model. At a confidence level of 99.5%, none of the cases yields of $|z| > 2.807$ that would refute the hypothesis that the two compared models have equal variability.

<table>
<thead>
<tr>
<th>Expiry</th>
<th>Hedge Model</th>
<th>Moneyness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 &lt; $T \leq 0.25$</td>
<td>Delta SV</td>
<td>-0.267</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ</td>
<td>1.70</td>
</tr>
<tr>
<td>0.6 ≤ $\frac{K}{S_0} &lt; 0.8$</td>
<td>Delta SV</td>
<td>0.0426</td>
</tr>
<tr>
<td></td>
<td>Delta SVJ</td>
<td>0.461</td>
</tr>
</tbody>
</table>
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

\[
\begin{align*}
\text{d}U(t, S_t, y_t) &= \partial_t U(t, S_t, y_t) \text{d}t + \partial_S U(t, S_t, y_t) \text{d}S_t + \partial_y U(t, S_t, y_t) \text{d}y_t \\
&\quad + \frac{1}{2} \partial_{SS} U(t, S_t, y_t) \text{d} \langle S^c \rangle_t + \frac{1}{2} \partial_{yy} U(t, S_t, y_t) \text{d} \langle y \rangle_t + \partial_S U(t, S_t, y_t) \text{d} \langle S^c, y \rangle_t \\
&\quad + U(t, S_t, e^{dZ_t}, y_t) - U(t, S_t, y_t) \\
&\quad + \partial_S U(t, S_{t-}, y_t) S_t - (e^{dZ_t} - 1),
\end{align*}
\tag{74}
\]

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

\[
U(t, S_t, y_t) := \bar{u}(t, S_t, y_t) \bar{v}(t, S_t, y_t)
\tag{75}
\]

then (74) and a development of the partial derivatives of \( U \) gives

\[
\begin{align*}
\text{d}U(t, S_t, y_t) &= (\bar{v} \partial_t \bar{u} + \bar{u} \partial_t \bar{v}) \text{d}t + (\bar{v} \partial_S \bar{u} + \bar{u} \partial_S \bar{v}) \text{d}S_t + (\bar{v} \partial_y \bar{u} + \bar{u} \partial_y \bar{v}) \text{d}y_t \\
&\quad + \frac{1}{2} (\bar{v} \partial_{SS} \bar{u} + 2 \partial_S \bar{u} \partial_S \bar{v} + \bar{u} \partial_{SS} \bar{v}) \text{d} \langle S^c \rangle_t + \frac{1}{2} (\bar{v} \partial_{yy} \bar{u} + 2 \partial_y \bar{u} \partial_y \bar{v} + \bar{u} \partial_{yy} \bar{v}) \text{d} \langle y \rangle_t \\
&\quad + (\partial_S \bar{v} \partial_y \bar{u} + \bar{v} \partial_S \bar{u} + \bar{u} \partial_S \bar{v} + \partial_S \bar{v} \partial_y \bar{v}) \text{d} \langle S^c, y \rangle_t \\
&\quad + \left[ \bar{u}(t-, S_{t-}, e^{dZ_t}, y_t) \bar{v}(t-, S_{t-}, e^{dZ_t}, y_t) - \bar{u}(t-, S_{t-}, y_t) \bar{v}(t-, S_{t-}, y_t) \right] \\
&\quad + \left[ \partial_S \bar{v} \bar{u} + \bar{u} \partial_S \bar{v} \right] S_t - (e^{dZ_t} - 1).
\end{align*}
\tag{76}
\]

Now,

\[
\text{d} \langle \bar{u}, \bar{v} \rangle_t = \text{d} \langle \bar{u}_t, \bar{v}_t \rangle - \bar{u}_t \text{d} \bar{v}_t - \bar{v}_t \text{d} \bar{u}_t,
\tag{77}
\]

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 \( \text{d} \bar{u} \) and \( \text{d} \bar{v} \) according to (74) and subtracting \( \text{d} \bar{u} \text{d} \bar{v} \) from (76),

\[
\begin{align*}
\text{d} \langle \bar{u}, \bar{v} \rangle_t &= \partial_S \bar{u} \partial_S \bar{v} S_t^2 \text{d} \langle S^c \rangle_t + \partial_y \bar{u} \partial_y \bar{v} \text{d} \langle y \rangle_t \\
&\quad + \left[ \bar{u}(t, S_t, e^{dZ_t}, y_t) - \bar{u}(t, S_t, y_t) \right] \left[ \bar{v}(t, S_t, e^{dZ_t}, y_t) - \bar{v}(t, S_t, y_t) \right].
\end{align*}
\tag{78}
\]

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

\[
\frac{\text{d}}{\text{d}t} E \left[ \langle \bar{u}, \bar{v} \rangle_t \right] = \partial_S \bar{u} \partial_S \bar{v} S_t^2 \text{d} \langle S^c \rangle_t + \partial_y \bar{u} \partial_y \bar{v} \text{d} \langle y \rangle_t \\
&\quad + \left[ \bar{u}(t, S_t, e^{z}, y_t) - \bar{u}(t, S_t, y_t) \right] \left[ \bar{v}(t, S_t, e^{z}, y_t) - \bar{v}(t, S_t, y_t) \right] \nu(z) \text{d}z.
\tag{79}
\]

For the specific case \( \bar{v}(t, S_t, y_t) = \bar{x}_t = B(t, T^1)QS_t \) from (41) with \( Q = e^{-\int_0^t \text{d}s} \) we have

\[
\frac{\text{d}}{\text{d}t} E \left[ \langle \bar{u}, \bar{x} \rangle_t \right] = QB(t, T^1) \left[ \partial_S \bar{u} S_t^2 y_t + \partial_y \bar{u} \rho S_t \theta y_t \\
&\quad + \left[ \bar{u}(t, S_t, e^{z}, y_t) - \bar{u}(t, S_t, y_t) \right] S_t(e^{z} - 1) \nu(z) \text{d}z \right].
\tag{80}
\]
and if \( \bar{v}(t, S_t, y_t) = \bar{u}(t, S_t, y_t) = \bar{x}_t \)

\[
\frac{d}{dt} E[(\bar{x}, \bar{x})_t] = Q^{i\pi} B^2 T^3 \left[ S^2_t y_t + \lambda \int S^2_t (e^x - 1)^2 \nu(z) dz \right].
\]  

(81)

**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 \( x_k = (k - 1 - \frac{M}{2}) \frac{B}{\lambda} \), for \( k = 1, \ldots, M \), \( y_l = (l - 1 - \frac{N}{2}) \frac{Q}{\lambda} \), for \( l = 1, \ldots, N \). To this end, we will perform the integration in the frequency space over a number of points \( u_a = (a - 1 - \frac{M}{2}) \frac{A}{\lambda} \) for \( a = 1, \ldots, M \) and where \( AB = 2\pi m \), and likewise for \( v_c = (c - 1 - \frac{N}{2}) \frac{C}{\lambda} \) for \( c = 1, \ldots, N \) with \( CD = 2\pi N \).

We denote the density 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

\[
\nu(x_a, y_c) \approx \frac{1}{4\pi^2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{-i(ux_a + vy_c)} \Phi(u, v) du dv \\
= \frac{1}{4\pi^2} \sum_{j=1}^{M} \sum_{k=1}^{N} \int_{(j-1-\frac{M}{2})\frac{\pi}{2}}^{(j-1+\frac{M}{2})\frac{\pi}{2}} \int_{(k-1-\frac{N}{2})\frac{\pi}{2}}^{(k-1+\frac{N}{2})\frac{\pi}{2}} e^{-i(ux_a + vy_c)} \Phi(u, v) du dv \\
\approx \frac{1}{4\pi^2} \frac{A}{M} \frac{C}{N} \sum_{j=1}^{M} \chi(j, M) e^{-i(j-1-\frac{M}{2})\frac{\pi}{2}+i(a-1-\frac{M}{2})\frac{\pi}{2}} \left\{ \sum_{k=1}^{N} \chi(k, N) \Phi \left( \left( j - 1 - \frac{M}{2} \right) \frac{A}{M}, \left( k - 1 - \frac{N}{2} \right) \frac{C}{N} \right) e^{-i(k-1-\frac{N}{2})\frac{\pi}{2}+i(c-1-\frac{N}{2})\frac{\pi}{2}} \right\} \\
= \frac{1}{4\pi^2} \frac{A}{M} \frac{C}{N} e^{i\pi(a-1-\frac{M}{2})\frac{\pi}{2}} e^{i\pi(c-1-\frac{N}{2})\frac{\pi}{2}} \sum_{j=1}^{M} \chi(j, M) e^{i\pi(j-1)(a-1-\frac{M}{2})\frac{\pi}{2}} \left\{ \sum_{k=1}^{N} \chi(k, N) \Phi \left( \left( j - 1 - \frac{M}{2} \right) \frac{A}{M}, \left( k - 1 - \frac{N}{2} \right) \frac{C}{N} \right) e^{i\pi(k-1)(c-1-\frac{N}{2})\frac{\pi}{2}} \right\}.
\]

(82)

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 $\{K^i_j, 1 \leq j \leq N^i\}$. Let $u^{BS}(T, K)$ be the BS price function (where the dependence 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 $(T^i, K^i_j)$ we choose $\sigma = \sigma_{BS}(T^i, K^i_j)$ exactly replicate the quoted prices $\hat{u}(T^i, K^i_j)$. The standard BS delta hedge consists in buying

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

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 (83).

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\(^4\) strategy then consist in a portfolio with initial value $P_0$ from (84) which invests in a number $\delta^S$ of shares $S$ and $\delta^C$ contracts $\hat{u}(\hat{T}, \hat{K})$ such that

$$\delta^C \partial_{\sigma_{BS}(\hat{T}, \hat{K})} u^{BS}(\hat{T}, \hat{K}) = \sum_{i=1}^{M} \sum_{j=1}^{N^i} \omega^i_j \partial_{\sigma_{BS}(T^i, K^i_j)} u^{BS}(T^i, K^i_j)$$

and

$$\delta^S \partial_{\sigma_{BS}(\hat{T}, \hat{K})} u^{BS}(\hat{T}, \hat{K}) = \sum_{i=1}^{M} \sum_{j=1}^{N^i} \omega^i_j \delta^i_j,$$

where the $\delta^i_j$ are again given by (83). 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 (85) and the delta-vega strategy (86) to the results from the quadratic hedging strategies in the affine model described in Section 3.

\(^4\)The 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.