Accepted Manuscript

A profitable modification to global quadratic hedging

Maciej Augustyniak, Frédéric Godin, Clarence Simard

PII: S0165-1889(19)30086-7
DOI: https://doi.org/10.1016/j.jedc.2019.05.008
Reference: DYNCON 3701

To appear in: *Journal of Economic Dynamics & Control*

Received date: 2 November 2018
Revised date: 14 March 2019
Accepted date: 16 May 2019

Please cite this article as: Maciej Augustyniak, Frédéric Godin, Clarence Simard, A profitable modification to global quadratic hedging, *Journal of Economic Dynamics & Control* (2019), doi: https://doi.org/10.1016/j.jedc.2019.05.008

A profitable modification to global quadratic hedging

Maciej Augustyniak[a,d], Frédéric Godin[b,d] and Clarence Simard[c]

[a] *Département de mathématiques et de statistique, Université de Montréal, PO Box 6128, Station Centre-ville, Montreal, Quebec, Canada H3C 3J7*

[b] *Department of Mathematics & Statistics, Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Quebec, Canada, H3G 1M8*

[c] *Département de mathématiques, Université du Québec à Montréal, PO Box 8888, Station Centre-ville, Montreal, Quebec, Canada H3C 3P8*

[d] *Quantact Actuarial and Financial Mathematics Laboratory, Centre de recherches mathématiques, Université de Montréal, PO Box 6128, Station Centre-ville, Montreal, Quebec, Canada H3C 3J7*

Version: May 18, 2019

Abstract

Recent research has shown that global quadratic hedging, also known as variance-optimal hedging and mean-variance hedging, can significantly reduce the risk of hedging call and put options with long-term maturities (one year or more), such as Long-Term Equity AnticiPation Securities (LEAPS). We propose a modification to global quadratic hedging that is more profitable on average to the hedger without substantially increasing his downside hedging risk, if at all. We prove mathematically that the expected terminal hedging gain of our modified strategy is greater than that of the global quadratic hedging strategy. The performance of our strategy is evaluated under simulated return paths from GARCH, regime-switching and jump-diffusion models, and under empirical S&P 500 return paths.

Keywords: risk management, variance-optimal hedging, mean-variance hedging, global risk-minimization, LEAPS

Mathematics Subject Classification (2010): 49L20, 91G20, 91G70

JEL Classification: C22, C61, G32

Maciej Augustyniak gratefully acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN-2015-05066). Frédéric Godin gratefully acknowledges financial support from NSERC (RGPIN-2017-06837) and the Fonds de recherche du Québec – Nature et technologies (FRQNT, 2017-NC-197517).

Email addresses: augusty@dms.umontreal.ca (Maciej Augustyniak), frederic.godin@concordia.ca (Frédéric Godin), simard.clarence@uqam.ca (Clarence Simard)

1 Introduction

Global quadratic hedging, also known as variance-optimal hedging and mean-variance hedging, is an alternative to delta hedging that was proposed in the seminal contribution of Schweizer (1995) to more appropriately address the hedging problem in incomplete markets. Augustyniak et al. (2017) recently showed that this hedging strategy can significantly reduce the risk of hedging European call and put options with long-term maturities (one year or more). Its use is therefore advocated for hedging Long-Term Equity AnticiPation Securities (LEAPS), which are actively traded call and put options on the Chicago Board Options Exchange (CBOE) that expire two to three years from their issuance.

The objective of global quadratic hedging is to find a self-financing trading strategy that minimizes the expected value of the squared terminal hedging error, defined as the squared shortfall between the payoff of the derivative and the terminal value of the self-financing hedging portfolio. The mathematical solution to this problem in a discrete time setting is obtained via a recursive scheme based on explicit expressions derived by Schweizer (1995). The availability of these expressions follows from the use of a quadratic penalty function in the optimization problem and considerably simplifies the strategy's implementation by way of dynamic programming (see Bertsimas et al., 2001; Černý, 2004). However, the use of a quadratic criterion also leads to an undesirable feature from a practical perspective: hedging gains are penalized in the same way as hedging losses. In other words, the strategy is aimed at replicating the derivative's payoff with minimal error, and sees discrepancies as *bad* even if they are to the advantage of the hedger. Hence, after a period of hedging gains, the strategy will put undue pressure to lose money. Although it is possible to solve the underlying optimization problem under a non-quadratic penalty function, implementing such a strategy is much more involved computationally, as illustrated by François et al. (2014). Moreover, Augustyniak et al. (2017) showed that global quadratic hedging performs well in reducing the downside risk of hedging, as measured for example by the 95% Value at Risk (VaR) of the terminal hedging error. It must also be noted that in a complete market, global quadratic hedging allows us to recover the riskless self-financing hedging strategy.

The main contribution of this article is to propose a modification to global quadratic hedging that reduces the adverse effects of using a quadratic penalty function. This strategy is more profitable on average to the hedger without substantially increasing his downside hedging risk, if at all. The main idea is to withdraw the hedging gain each time the value of the hedging portfolio is above that of the contingent claim and to *reset* the strategy. We prove mathematically that the expected terminal hedging gain of our modified strategy is greater than that of the global quadratic hedging strategy so that our strategy is more profitable on average. Moreover, the proposed modification does not come at an increased computational cost.

The paper is structured as follows. Section 2 introduces our modified global quadratic hedging strategy, discusses its properties, and explains how it can be implemented in practice. Section 3 illustrates the performance of our strategy under simulated return paths from GARCH, regime-switching and jump-diffusion models, whereas Section 4 repeats this analysis on empirical S&P 500 return data covering the period 2008–2018. Section 5 concludes. Detailed proofs of mathematical results are provided in appendix.

2 Modified global quadratic hedging

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a filtration $\{\mathcal{F}_t\}_{t=0,1,\ldots,T}$ that satisfies the usual conditions, where $\mathbb{P}$ denotes the physical (real-world, objective) probability measure. Consider a non-negative square-integrable adapted stochastic process $X := \{X_t\}_{t=0,1,\ldots,T}$ and assume that X_t is the time-t discounted value to time 0 of a tradable risky asset for $t \geq 0$. Moreover, suppose that the financial market is frictionless and arbitrage-free and assume the existence of some riskless asset whose discounted value to time 0 is 1 at all times.

Remark 1. All value processes introduced in Section 2 are expressed in discounted units to time 0. The discounting factor corresponds to the value of the riskless asset, in other words, asset prices are expressed in terms of the money market account numeraire. This set-up is standard in the quadratic hedging literature and allows for a simplified exposition.

We define the set of admissible trading strategies by

$$\Theta := \left\{ \{\theta_t\}_{t=1,\dots,T} \mid \theta_t \in \mathcal{F}_{t-1} \text{ and } \theta_t \Delta X_t \in \mathcal{L}^2(\mathbb{P}) \text{ for all } t = 1, \dots, T \right\},$$

where $\Delta X_t := X_t - X_{t-1}$ and θ_t represents the number of risky asset shares held over the time period $(t-1, t]$. Moreover, let Θ_n denote a subset of Θ such that, for $n = 0, 1, \dots, T-1$,

$$\Theta_n := \left\{ \{\theta_t\}_{t=n+1,\dots,T} \mid \theta_t \in \mathcal{F}_{t-1} \text{ and } \theta_t \Delta X_t \in \mathcal{L}^2(\mathbb{P}) \text{ for all } t = n+1, \dots, T \right\}.$$

For a trading strategy $\theta \in \Theta_t$ and $c \in \mathbb{R}$ we define, for $t = 0, 1, \dots, T-1$:

$$\begin{aligned} V_{t,t}(c, \theta) &:= c, \\ V_{t,n}(c, \theta) &:= c + \sum_{k=t+1}^{n} \theta_k \Delta X_k, \quad n = t+1, \dots, T. \end{aligned}$$

In particular, $V_{t,T}(c, \theta)$ corresponds to the time-T discounted value of a self-financing portfolio with trading strategy θ initiated at time t with a discounted capital of c. The term $\sum_{k=t+1}^{T} \theta_k \Delta X_k$ describes the process of discounted trading gains between times t and T.

2.1 Global quadratic hedging

Let H be the discounted payoff of a contingent claim, that is, H is a $\mathcal{F}_T$-measurable random variable in $\mathcal{L}^2(\mathbb{P})$. The global quadratic hedging portfolio is the pair $(V_0^*, \xi^*) \in \mathbb{R} \times \Theta$ that solves

$$\min_{(c,\theta) \in \mathbb{R} \times \Theta} \mathbb{E}\left[\{H - V_{0,T}(c, \theta)\}^2 \right]. \tag{1}$$

Since the solution (V_0^*, ξ^*) satisfies (see Schweizer, 1995, remark (3) on p. 17)

$$\mathbb{E}[V_{0,T}(V_0^*, \xi^*)] = \mathbb{E}[H], \tag{2}$$

a hedger following this strategy is not expected to make a profit as his terminal hedging gain, $V_{0,T}(V_0^*, \xi^*) - H$, is zero on average. Schweizer (1995) showed that the solution to optimization problem (1) exists if

$$\exists \zeta \in (0,1) \text{ such that } \forall t = 1, \ldots, T: \ (\mathbb{E}\left[\Delta X_t \mid \mathcal{F}_{t-1}\right])^2 \leq \zeta \mathbb{E}\left[(\Delta X_t)^2 \mid \mathcal{F}_{t-1}\right] \text{ a.s.}$$

This assumption is known as the non-degeneracy condition and we assume it to hold throughout the paper.

Moreover, let

$$V_t := V_{0,t}(V_0^*, \xi^*), \quad t = 0, 1, \ldots, T.$$

In particular, for $t = 1, \ldots, T$, we have $V_t = V_0^* + \sum_{k=1}^{t} \xi_k^* \Delta X_k$, where ξ_t^* is the element of ξ^* that represents the optimal position in the risky asset established at time $t-1$ for a global quadratic hedging portfolio initiated at time 0 with a capital of V_0^*. From Bellman's principle of optimality, ξ_t^* must satisfy $\xi_t^* = \xi_t(V_{t-1})$, where for $z \in \mathcal{F}_{t-1}$,

$$\xi_t(z) := \arg\min_{\phi} \left\{ \min_{\theta \in \Theta_t} \mathbb{E}\left[\left\{ H - \left(z + \phi \Delta X_t + \sum_{k=t+1}^{T} \theta_k \Delta X_k \right) \right\}^2 \middle| \mathcal{F}_{t-1} \right] \right\}. \tag{3}$$

Theorem 2.4 of Schweizer (1995) shows that $\xi_t(z)$ can be expressed as

$$\xi_t(z) = \varrho_t - \beta_t z, \quad t = 1, \ldots, T, \tag{4}$$

where $\varrho := \{\varrho_t\}_{t=1,\ldots,T}$ and $\beta := \{\beta_t\}_{t=1,\ldots,T}$ are predictable processes defined in Equations (19)–(22) of Appendix A. In Section 2.5, we explain how to compute the coefficients ϱ_t and β_t by way of dynamic programming.

Remark 2. The meaning of $\xi_t(z)$ in our paper differs slightly from the one of symbol $\xi_t^{(c)}$ used in Schweizer (1995, Theorem 2.4). In our paper, $\xi_t(z)$ represents the optimal global quadratic hedging position established at time $t-1$ assuming a discounted capital of z at time $t-1$, whereas in Schweizer (1995, Theorem 2.4), the symbol $\xi_t^{(c)}$ assumes a capital of c at time 0.

2.2 Modified strategy

As already alluded to in the introduction, the global quadratic hedging strategy, although convenient from a mathematical and computational standpoint, has the drawback of penalizing terminal hedging gains. In Definition 1 below, we propose a modification to this strategy that will be shown to reduce some of the adverse effects associated with the underlying quadratic penalty function at no additional computational cost. The main idea behind our modified strategy is to reset the global quadratic hedging strategy at each time t when the value of the hedging portfolio rises above an optimal capital value V_t^*, defined as

$$V_t^* := \arg\min_c \left\{ \min_{\theta \in \Theta_t} \mathbb{E}\left[\{H - V_{t,T}(c,\theta)\}^2 \mid \mathcal{F}_t \right] \right\}, \quad t = 1, \ldots, T-1,$$

and set aside the surplus.

Remark 3. V_t^* corresponds to the optimal capital for a global quadratic hedging strategy initiated at time t. Schweizer (1995) explains that it can be interpreted as the time-t price of the contingent claim computed under a so-called variance-optimal martingale measure. Unfortunately, this interpretation suffers from the fact that this measure is generally signed in the discrete time setting, which implies that it does not typically produce a legitimate pricing kernel. Nevertheless, as stated by Rémillard and Rubenthaler (2013), V_t^* is still the "optimal investment at period t, so that the value of the portfolio at period T is as close as possible to H, in terms of mean square error.” Therefore, in some sense, V_t^* is our best approximation at time t of the value of the contingent claim.

Definition 1 (Modified global quadratic hedging)**.** The modified global quadratic hedging portfolio starts with a capital of $Z_0 := V_0^*$ and takes a position in the risky asset of $\xi_t(Z_{t-1})$ at time $t-1$, for $t = 1, \dots, T$, where

$$Z_t := \min(V_t^*, Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t), \quad t = 1, \dots, T-1, \tag{5}$$

$$Z_T := Z_{T-1} + \xi_T(Z_{T-1})\Delta X_T. \tag{6}$$

The variable Z_t can be interpreted as the value at time t of a global quadratic hedging portfolio in which the surplus over V_t^* is withdrawn. For example, suppose that the portfolio value at time $t-1$ is Z_{t-1}, then the optimal position in the risky asset established at $t-1$ under the criterion (3) is given by $\xi_t(Z_{t-1})$. The portfolio value at time t is then obtained as $Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t$. If this value is below V_t^*, no withdrawal is made and the hedging strategy follows its course. However, if this value is above V_t^*, then the hedger withdraws the surplus and continues the hedging strategy with a capital of V_t^*. The sum of all cash flows generated by our modified global quadratic hedging strategy corresponds to

$$\underbrace{Z_T}_{\text{terminal value}} + \underbrace{\sum_{t=1}^{T-1} \max\left(Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t - V_t^*, 0\right)}_{\text{surpluses withdrawn}}.$$

Remark 4. Lemma 2 in Appendix A shows that

$$Z_T + \sum_{t=1}^{T-1} \max\left(Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t - V_t^*, 0\right) = V_0^* + \sum_{t=1}^{T} \xi_t(Z_{t-1})\Delta X_t.$$

Intuitively, this result was to be expected since our modified strategy can also be implemented in a self-financing manner, without explicitly withdrawing surpluses from the portfolio. In this case, the terminal value of the hedging portfolio would simply be $V_0^* + \sum_{t=1}^{T} \xi_t(Z_{t-1})\Delta X_t$.

2.3 Main results

Theorems 1 and 2, and Proposition 1 present mathematical results on our modified strategy. Detailed proofs are provided in Appendix A.

Theorem 1. *The expected value of the cash flows generated by our modified global quadratic hedging strategy is greater than or equal to the expected value of the contingent claim:*

$$\mathbb{E}\left[Z_T + \sum_{t=1}^{T-1} \max\left(Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t - V_t^*, 0\right)\right] \geq \mathbb{E}[H]. \tag{7}$$

Consequently, Theorem 1 indicates that $\mathbb{E}[V_0^* + \sum_{t=1}^{T} \xi_t(Z_{t-1})\Delta X_t] \geq \mathbb{E}[H]$, whereas for the classical strategy we necessarily have that $\mathbb{E}[V_0^* + \sum_{t=1}^{T} \xi_t^* \Delta X_t] = \mathbb{E}[H]$ by Equation (2). Therefore, our modification to the global quadratic hedging portfolio can only increase the expected profitability of this strategy.

Remark 5. One could interpret the result of Theorem 1 as a weak form of superhedging; the value of a superhedging portfolio is almost surely greater than or equal to the payoff of a contingent claim while the value of our modified hedging portfolio is greater than or equal to that payoff in expectation. Although superhedging excludes hedging losses, this strategy is rarely used in practice due to its prohibitively high initial cost. In contrast, our modified hedging portfolio has a cost equal to the one of the global quadratic hedging portfolio, and leads to a hedging gain on average.

Theorem 2. *Assume that:*

$$\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right] \neq 0.$$

Then, the following statements are equivalent.

1. *The inequality* (7) *in Theorem 1 is an equality.*

2. *The claim V_{T-1}^* is attainable with the global quadratic hedging portfolio (V_0^*, ξ^*) in the sense that $V_{T-1}^* = V_{T-1}$ a.s.*

3. $Z_t = V_t = V_t^*$ *a.s. for* $t = 0, 1, \dots, T-1$ *and* $Z_T = V_T$ *a.s.*

Note that Theorem 2 suggests that the inequality (7) is strict for all non-trivial cases of interest. This is because the value V_{T-1}^* of a contingent claim in an incomplete market will generally only be attainable for trivial claims such as a forward contract. Moreover, it also implies that an equality in Theorem 1 can only occur if our modified strategy falls back to the classical strategy (which will only happen if V_{T-1}^* is attainable). Therefore, our strategy will almost always deviate from the global quadratic hedging strategy and be profitable on average to the hedger.

Finally, Proposition 1 explicitly addresses two special cases where our modified strategy is equivalent to the classical strategy. The first one is that of a claim that can be replicated in a riskless manner (i.e., H is attainable) and for which a risk-minimizing strategy is therefore not needed. The second one is the setting where X is a $\mathbb{P}$-martingale, which does not represent a realistic financial environment as it implies a zero equity risk premium.

Proposition 1. *If either one of the following conditions are satisfied,*

1. *H is attainable, that is, $H = H_0 + \sum_{t=1}^{T} \xi_t^H \Delta X_t$ a.s. for some predictable process $\xi^H := \{\xi_t^H\}_{t=1,\dots,T}$ and $H_0 \in \mathbb{R}$,*

2. *X is a $\mathbb{P}$-martingale,*

then our modified strategy is equivalent to the global quadratic hedging strategy in the sense that $\xi_t(Z_{t-1}) = \xi_t^*$ *a.s. for* $t = 1, \dots, T$ *and*

$$Z_T + \sum_{t=1}^{T-1} \max\left(Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t - V_t^*, 0\right) = V_0^* + \sum_{t=1}^{T} \xi_t^* \Delta X_t \ \ a.s.$$

In particular, under either one of these conditions, the inequality (7) *in Theorem 1 is an equality.*

2.4 Discussion

The practical difference between our modified strategy and global quadratic hedging is that our strategy specifies the optimal position in the risky asset at time t as $\xi_{t+1}(Z_t)$, instead of $\xi_{t+1}(V_t)$. In other words, both methods use criterion (3) to determine the position in the risky asset, but differ in the capital assumed at t. By resetting the capital to V_t^* at each time t when the value of the hedging portfolio rises above it, our modified strategy allows the hedger to lock in realized hedging gains and thus avoids penalizing these gains in the future. The following three facts derived from Corollary 2.5 of Schweizer (1995) help to motivate our strategy:

$$\mathbb{E}\left[\{H - V_{t,T}(V_t^*, \xi^*(V_t^*))\}^2 \mid \mathcal{F}_t\right] \leq \mathbb{E}\left[\{H - V_{t,T}(c, \xi^*(c))\}^2 \mid \mathcal{F}_t\right], \quad \text{for all } c \in \mathbb{R}, \tag{8}$$

$$\mathbb{E}\left[V_{t,T}\left(V_t^* + \delta, \xi^*(V_t^* + \delta)\right) \mid \mathcal{F}_t\right] \leq \mathbb{E}\left[V_{t,T}\left(V_t^*, \xi^*(V_t^*)\right) \mid \mathcal{F}_t\right] + \delta, \quad \text{for all } \delta \geq 0, \tag{9}$$

$$\mathbb{E}\left[V_{t,T}\left(V_t^* + \delta, \xi^*(V_t^* + \delta)\right) \mid \mathcal{F}_t\right] \geq \mathbb{E}\left[V_{t,T}\left(V_t^*, \xi^*(V_t^*)\right) \mid \mathcal{F}_t\right] + \delta, \quad \text{for all } \delta \leq 0, \tag{10}$$

where $\xi^*(c)$ here refers to the global quadratic hedging strategy initiated with a capital of c at time t. Equation (8) entails that the minimal *future* expected squared hedging error is attained from a capital of V_t^* at time t. Moreover, Equation (9) implies that if the quadratic hedging strategy is implemented at time t with a capital of $V_t^* + \delta$, where $\delta \geq 0$, then the expected terminal value of the hedging portfolio is lower than if we were to withdraw the amount δ from the portfolio and resume the strategy based on a capital of V_t^*. Therefore, on the basis of the future mean and variance of the hedging error, it makes perfect sense to set aside the surplus over V_t^* (if any) and to reset the strategy to this new reference point. On the other hand, Equation (10) indicates that if the current portfolio value, say $V_t^* + \delta$, is in deficit relative to V_t^* (i.e., $\delta \leq 0$), then injecting the amount $-\delta$ into the portfolio so that the strategy can be set up with a capital of V_t^* would actually decrease the portfolio's expected terminal value. Therefore, the incentive to bring the portfolio value to V_t^* here is not clear-cut as in the previous case where $\delta \geq 0$. This is the reason why we reset the global quadratic hedging strategy only when the value of the hedging portfolio rises above V_t^*.

The price to pay for this modification is that the *overall* expected squared hedging error (i.e., from time 0) is no longer minimized. However, this allows for an advantageous tradeoff as we now explain. Indeed, as shown in Section 2.3, this specific resetting of the strategy leads to an overall positive expected hedging gain. Second, it must be noted that a higher overall expected squared hedging error does not automatically translate into a heightened downside hedging risk. To see this, we can split the expected squared hedging error in the following way:

$$\underbrace{\mathbb{E}\left[\{H - V_{0,T}(c,\theta)\}^2\right]}_{\text{MSE}} = \underbrace{\mathbb{E}\left[\{H - V_{0,T}(c,\theta)\}^2 \mathbb{1}_{\{V_{0,T}(c,\theta)\geq H\}}\right]}_{\text{MSE}^+} + \underbrace{\mathbb{E}\left[\{H - V_{0,T}(c,\theta)\}^2 \mathbb{1}_{\{V_{0,T}(c,\theta)< H\}}\right]}_{\text{MSE}^-},$$

where MSE^+ and MSE^- represent the mean squared errors associated with, respectively, terminal hedging gains and losses. Since the goal of our modification is to allow for more hedging gains, it would be incoherent to assess the performance of our strategy based on a risk measure that evaluates hedging gains as a risk, such as the MSE. As expected, our strategy does indeed increase the MSE when compared to global quadratic hedging. However, the MSE^- is a more appropriate measure of hedging risk in our context as it only penalizes the adverse events that the hedger wants to avoid. The simulation studies in Section 3 suggest that the positive average gain generated by our strategy usually comes at the cost of at most a marginal increase in the MSE^-. Intuitively, this is because our modified strategy mimics the global quadratic hedging strategy in scenarios where the value of the hedging portfolio is below that of the claim. Consequently, the increase in the MSE is almost entirely due to a potential for larger hedging gains (i.e., to an increase in the MSE^+), which is not problematic from a hedging perspective. In fact, it would be incoherent with the idea of hedging to allow for more hedging gains if this significantly increased the risk of hedging losses. Section 3 even shows that in some situations our strategy can decrease the MSE^-.

It is natural to ask at this point what optimization problem is being solved by the strategy introduced in Definition 1. Clearly, neither the global quadratic hedging portfolio nor our modified strategy are explicitly designed to minimize the MSE^{-}. The minimization of this penalty is a global hedging problem with an asymmetric criterion whose solution is computationally much more costly to implement. However, Proposition 2 indicates that one can see our strategy as minimizing a sequence of mean-variance hedging problems under a capital constraint.

Proposition 2. *For $t = 0, 1, \ldots, T-1$, the position in the risky asset at time t prescribed by our modified global quadratic hedging strategy solves the following optimization problem:*

$$\begin{aligned} &\min_{\phi} \left\{ \min_{\substack{\theta \in \Theta_{t+1} \\ z \in \mathcal{F}_t}} \mathbb{E}\left[\left\{ H - \left(z + \phi \Delta X_{t+1} + \sum_{k=t+2}^{T} \theta_k \Delta X_k \right) \right\}^2 \,\middle|\, \mathcal{F}_t \right] \right\}, \\ &\text{subject to } z \leq Z_t \text{ a.s.}, \end{aligned} \tag{11}$$

where $Z_t := \min(V_t^*, Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t)$, defined in Equation (5), represents the value of the modified global quadratic hedging portfolio that excludes surpluses previously withdrawn.

Appendix A.5 proves that the solution to (11) is indeed given by $\phi = \xi_{t+1}(Z_t)$. Consequently, at any given time t, our modified strategy minimizes the same objective function as the global quadratic hedging portfolio (see Equation (3)), but it does so subject to $z \leq Z_t$, instead of $z = V_t$, where V_t is the value of the self-financing global quadratic hedging portfolio. The constraint $z \leq Z_t$ in optimization problem (11) is motivated by the discussion associated to Equations (8)–(10) and has for effect to relax the condition that the hedging portfolio is self-financing at all times. On one hand, it prevents the use of a capital z larger than V_t^*, and thus allows the hedger to set aside any surplus over V_t^*. On the other hand, it precludes capital injections into the hedging portfolio, which as we previously explained decrease the portfolio's expected terminal value.

The reason why our strategy is no more complicated to implement than the classical strategy is that for both approaches the solution at time $t-1$ is of the form $\xi_t(z) = \varrho_t - \beta_t z$. Since the coefficients ϱ_t and β_t are identical for both strategies and in calculating them one must also compute V_t^* (see Section 2.5), our strategy does not entail an increased computational cost.

We end this discussion by highlighting a particular property of our modified strategy: the hedging portfolio for protecting a long position in a derivative is generally not a mirror image of the one for the short position. This property results from the asymmetric treatment of hedging gains and losses that is unique to our strategy, and is therefore not shared by the classical strategy nor by delta hedging. Indeed, for a given trajectory of the risky asset, the terminal hedging errors for protecting short and long positions with either global quadratic hedging or delta hedging are of opposite sign. This implies that if a delta hedge for a short position generates a positive expected profit, then the corresponding delta hedge for the long position will necessarily lead to a loss on average. Such a symmetry does not occur with our modified strategy since Theorems 1 and 2 guarantee that the expected terminal hedging gain is positive for any derivative position.

2.5 Implementation

To implement our modified strategy, one must first solve the global quadratic hedging problem to obtain V_{t-1}^* and the coefficients ϱ_t and β_t, for $t = 1, \ldots, T$. All of these variables can be computed efficiently using dynamic programming as in Augustyniak et al. (2017) (see also Bertsimas et al., 2001; Černý, 2004; Rémillard and Rubenthaler, 2013). Algorithm 1 summarizes the resulting dynamic program.

Algorithm 1. Set $\nu_{T+1} := 1$, $V_T^* := H$, $\Delta_t := \Delta X_t$. For $t = T, \ldots, 1$, compute recursively:

$$a_t := \mathbb{E}\left[\Delta_t^2 \nu_{t+1} \mid \mathcal{F}_{t-1}\right], \tag{12}$$

$$b_t := \mathbb{E}\left[\Delta_t \nu_{t+1} \mid \mathcal{F}_{t-1}\right], \tag{13}$$

$$d_t := \mathbb{E}\left[V_t^* \Delta_t \nu_{t+1} \mid \mathcal{F}_{t-1}\right], \tag{14}$$

$$\nu_t := \mathbb{E}\left[(1 - \Delta_t b_t / a_t) \nu_{t+1} \mid \mathcal{F}_{t-1}\right], \tag{15}$$

$$V_{t-1}^* := \frac{\mathbb{E}\left[V_t^*(1-\Delta_t b_t/a_t)\nu_{t+1} \mid \mathcal{F}_{t-1}\right]}{\nu_t}. \tag{16}$$

The coefficients ϱ_t and β_t, for $t = 1, \ldots, T$, are then given by $\varrho_t = d_t/a_t$ and $\beta_t = b_t/a_t$.

Let $\{\upsilon_t\}$ be a stochastic process that influences the dynamics of $\{X_t\}$, such as a volatility or a regime process. Rémillard and Rubenthaler (2013) explain that if the payoff at time T is only a function of X_T, say $H = H(X_T)$, and if $\{(X_t, \upsilon_t)\}$ is a Markov process, then the coefficients (12)–(16) in Algorithm 1 can be expressed only in terms of $(X_{t-1}, \upsilon_{t-1})$, that is, for $t = 1, \ldots, T$, there exists functions $\tilde{a}_t$, $\tilde{b}_t$, $\tilde{d}_t$, $\tilde{\nu}_t$ and $\tilde{V}_{t-1}^*$, such that

$$\begin{aligned} &a_t = \tilde{a}_t(X_{t-1}, \upsilon_{t-1}), \quad b_t = \tilde{b}_t(X_{t-1}, \upsilon_{t-1}), \quad d_t = \tilde{d}_t(X_{t-1}, \upsilon_{t-1}), \\ &\nu_t = \tilde{\nu}_t(X_{t-1}, \upsilon_{t-1}), \quad V_{t-1}^* = \tilde{V}_{t-1}^*(X_{t-1}, \upsilon_{t-1}). \end{aligned} \tag{17}$$

The Markov property reduces the dimensionality of the dynamic program, because it is then sufficient to compute the functions in (17) on a discretization of the state space of the Markov chain $\{(X_t, \upsilon_t)\}$. At every time step, this state space is represented with an adaptive two-dimensional grid defined as

$$\mathcal{G}_t := \{X^{(t-1,i_t)} \mid i_t = 1, \ldots, I_t\} \times \{\upsilon^{(t-1,j_t)} \mid j_t = 1, \ldots, J_t\}, \quad t = 1, \ldots, T,$$

where I_t and J_t are suitably chosen integers characterizing the number of nodes on the grid. The functions in (17) are computed at every node in $\mathcal{G}_t$ by iterating backwards in time for $t = T, \ldots, 1$. Our implementation of the dynamic program follows that of Augustyniak et al. (2017). Refer to Appendix A of Augustyniak et al. (2017) for more details.

Remark 6. We implement all hedging strategies assuming that $\mathcal{F}_t$ refers to an information set containing only observed market information until time t, so that generally $\mathcal{F}_t = \sigma(X_0, X_1, \ldots, X_t)$. If the process $\{\upsilon_t\}$ is latent, then filtering techniques are used to infer this process based on the available market information.

3 Simulation study

We illustrate the performance of our modified global quadratic hedging strategy under simulated return paths from GARCH, regime-switching and jump-diffusion models estimated on a time series of daily percentage log-returns on the S&P 500 price index for the period 1986-12-31 to 2010-04-01 (5863 returns). This data set was also considered by Augustyniak et al. (2017).

In our applications, we express the time index t in trading days and suppose that one year includes 260 trading days. We let B_t denote the undiscounted value of the riskless asset at time t for $t \geq 0$, and assume that $B_t := \exp(rt)$, where r corresponds to a constant daily risk-free rate. We suppose an annualized risk-free rate of 2%, which implies $r = 0.00769\%$. Moreover, we let $S_t := B_t X_t$ denote the undiscounted value of the risky asset at time t for $t \geq 0$. The daily percentage log-return from day $t-1$ to t is then given by $y_t := 100\,(\log S_t - \log S_{t-1})$.

3.1 Description of options hedged

Our analyses assume that a short or long position in a 3-year European call option is hedged with different moneyness levels. More precisely, we consider an initial stock price of $S_0 = 100$, and the (undiscounted) payoff $B_T H = \max(S_T - K, 0)$, with strike price $K = 90$ (in-the-money, ITM), 100 (at-the-money, ATM), or 110 (out-the-money, OTM) and maturity of $T = 780$ days. We consider long-dated options because Augustyniak et al. (2017) showed that the performance of global quadratic hedging stands out at long-term maturities. Although standard exchange-traded options typically mature within one year of their issuance, call and put options expiring two to three years from the date of listing actively trade on the CBOE under the LEAPS acronym. Moreover, market-linked certificates of deposit offer embedded call option payoffs with a typical term of three to five years. Therefore, devising effective hedging strategies for long-dated derivatives is important.

3.2 GARCH model

The GARCH process that we consider models the return as follows:

$$y_t = \mu + \sigma_t \epsilon_t,$$
$$\sigma_t^2 = \omega + \alpha \sigma_{t-1}^2 \left(|\epsilon_{t-1}| - \gamma \epsilon_{t-1}\right)^2 + \beta \sigma_{t-1}^2,$$

where $\{\epsilon_t\}$ is an independent and identically distributed (i.i.d.) Gaussian innovation process with mean 0 and variance 1. To ensure positivity of the conditional variance the following parametric constraints are required: $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$. This GARCH specification is equivalent to the GJR-GARCH$(1,1)$ model introduced by Glosten et al. (1993). Table 1 reports maximum likelihood parameter estimates. Refer to Augustyniak et al. (2017) for more details on the implementation of Algorithm 1 in the context of this model. In relation to Section 2.5, note that if we let $\upsilon_t = \sigma_{t+1}$, where σ_{t+1} is $\mathcal{F}_t$-measurable, then $\{(X_t, \upsilon_t)\}$ is a Markov process.

Table 1: Maximum likelihood estimates of the GARCH model

μ	ω	α	γ	β
0.0292	0.0179	0.0545	0.591	0.911

Notes: These parameter estimates were computed on a time series of daily percentage returns on the S&P 500 price index for the period 1986-12-31 to 2010-04-01 (5863 returns).

3.3 Regime-switching model

Our regime-switching model assumes that the return distribution is Gaussian conditional on the regime of an unobserved two-state Markov chain, denoted by $\{\chi_t\}$, with homogeneous transition matrix $(p_{ij})_{i,j=1}^2$, where $p_{ij} = \Pr(\chi_t = j \mid \chi_{t-1} = i)$. The model is defined as follows:

$$y_t = \begin{cases} \mu_1 + \sigma_1 \epsilon_t, & \text{if } \chi_t = 1, \\ \mu_2 + \sigma_2 \epsilon_t, & \text{if } \chi_t = 2, \end{cases}$$

where $\{\epsilon_t\}$ is once again an i.i.d. Gaussian innovation process with mean 0 and variance 1, and (μ_1, μ_2) and (σ_1^2, σ_2^2) are, respectively, regime-dependent conditional mean and variance parameters. Table 2 gives maximum likelihood parameter estimates. Note that the states of the underlying Markov chain can be interpreted as bull and bear market regimes. If we let $\upsilon_t = \Pr(\chi_t = 1 \mid \mathcal{F}_t)$, then it can be shown that $\{(X_t, \upsilon_t)\}$ is a Markov process.

Table 2: Maximum likelihood estimates of the regime-switching model

j	μ_j	σ_j^2
1	0.070	0.551
2	-0.106	4.272

$$\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} 0.989 & 0.011 \\ 0.035 & 0.965 \end{pmatrix}$$

Notes: These parameter estimates were computed on a time series of daily percentage returns on the S&P 500 price index for the period 1986-12-31 to 2010-04-01 (5863 returns).

3.4 Jump-diffusion model

The jump-diffusion process that we consider is the classical one introduced by Merton (1976). It models the log-price process as follows:

$$\log S_t - \log S_0 = \left(\alpha - \sigma^2/2 - \lambda \left(e^{\mu_J + \sigma_J^2/2} - 1\right)\right) t + \sigma W_t + \sum_{j=1}^{N_t} \epsilon_j,$$

where $W := \{W_t\}$ is a Brownian motion, $N := \{N_t\}$ is a homogeneous Poisson process independent of W with intensity $\lambda > 0$, and $\{\epsilon_j\}$ is a sequence of i.i.d. Gaussian random variables with mean μ_J and variance σ_J^2 (also independent of W and N). This specification implies that the daily log-return process $\{y_t\}$ is an i.i.d. process and includes five parameters,

$(\alpha, \sigma, \lambda, \mu_J, \sigma_J)$, where α corresponds to the one-period drift of S_t, that is $\mathbb{E}[S_t \mid \mathcal{F}_{t-1}] = S_{t-1}e^{\alpha}$, and σ corresponds to the daily volatility of the process conditional on there being no jumps. Table 3 reports maximum likelihood parameter estimates.

Table 3: Maximum likelihood estimates of the Merton jump-diffusion model

260α	$\sqrt{260}\sigma$	260λ	μ_J	σ_J
0.0875	0.1036	92.42	-0.0015	0.0160

Notes: These parameter estimates were computed on a time series of daily percentage returns on the S&P 500 price index for the period 1986-12-31 to 2010-04-01 (5863 returns).

3.5 Delta hedging benchmarks

Delta hedging under a suitably chosen risk-neutral measure is implemented in our analysis to serve as a benchmark. The GARCH delta hedge is based on the definition of the delta given by Duan (1995) and follows a standard change of measure considered in the GARCH option pricing literature (e.g., Badescu et al., 2014; Christoffersen et al., 2010; Ortega, 2012). Refer to Sections 3.1 and 3.2 of Augustyniak et al. (2017) for more details on the underlying change of measure and the implementation of this strategy.

The regime-switching delta hedge is computed under a change of measure that effectively translates the conditional means in regimes 1 and 2 to $r - \sigma_1^2/2$ and $r - \sigma_2^2/2$, while leaving the conditional variances and transition probabilities unchanged. Under this risk-neutral measure, the price and delta of a call option can be obtained in closed-form. Refer to Section 5.4.2 of François et al. (2014) for more details.

The jump-diffusion delta hedge is based on the standard change of measure that modifies the drift from α to r.

3.6 Results

We simulated $N = 10{,}000$ daily stock price paths with the GARCH, regime-switching and jump-diffusion models presented in Sections 3.2, 3.3 and 3.4, respectively. The GARCH simulations were initiated with a daily stationary volatility of $\sigma_1 = 1.061\%$.[1] In simulating the regime-switching model, we assumed that the starting regime probabilities correspond to the stationary probabilities of the Markov chain. On each one of the simulated paths, and for all hedging strategies considered, we computed the terminal hedging error defined as

$$\Lambda_T := B_T \left\{ H - V_{0,T}(V_0^*, \theta) \right\}. \tag{18}$$

We assess hedging effectiveness by computing statistical measures on the $N = 10{,}000$ realizations of this hedging error; the realization of Λ_T on path i is denoted by $\Lambda_{T,i}$. We consider the root-mean-squared error (RMSE) and the RMSE$^-$, which are defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \Lambda_{T,i}^2},$$

$$\mathrm{RMSE}^- = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \Lambda_{T,i}^2 \mathbb{1}_{\{\Lambda_{T,i}>0\}}},$$

and the 95% Value at Risk (VaR), which corresponds to the 501th largest value of $\{\Lambda_{T,i}\}_{i=1}^{10000}$. Note that a negative hedging error implies a gain for the hedger.

Tables 4, 5 and 6 present our results under, respectively, GARCH, regime-switching and jump-diffusion simulated return paths. First, as expected, global quadratic hedging yields a mean of approximately zero and the smallest RMSE for the terminal hedging error. In contrast, as indicated by Theorems 1 and 2, our modified strategy generates a hedging gain on average for all cases considered. In the GARCH and jump-diffusion settings, this gain comes at the cost of at most a marginal increase in the downside risk of hedging, as measured

[1] The stationary or unconditional variance corresponds to $\mathbb{E}[\sigma_t^2] = \omega/(1 - \alpha(1+\gamma^2) - \beta)$, provided that the denominator is positive.

Table 4: Terminal hedging error under GARCH simulated return paths

	Hedging a short call position				Hedging a long call position			
	Mean	RMSE	RMSE$^-$	95% VaR	Mean	RMSE	RMSE$^-$	95% VaR
3-year ATM call option								
Global hedging	−0.01	1.73	1.41	2.71	0.01	1.73	0.99	2.31
Modified global hedging	−0.39	1.88	1.42	2.72	−0.43	2.01	1.01	2.40
Delta hedging	−0.86	2.44	1.63	3.59	0.86	2.44	1.81	3.21
3-year OTM call option								
Global hedging	−0.01	1.52	1.25	2.16	0.01	1.52	0.86	1.93
Modified global hedging	−0.33	1.64	1.27	2.17	−0.37	1.79	0.90	2.04
Delta hedging	−0.70	2.08	1.50	3.21	0.70	2.08	1.44	2.54
3-year ITM call option								
Global hedging	−0.01	1.85	1.52	2.92	0.01	1.85	1.06	2.47
Modified global hedging	−0.41	2.02	1.51	2.86	−0.47	2.14	1.06	2.54
Delta hedging	−0.94	2.65	1.67	3.64	0.94	2.65	2.06	3.77

Notes: The statistical measures (mean, RMSE, RMSE$^-$ and 95% VaR) are based on 10,000 realizations of the terminal hedging error $\Lambda_T := B_T \{H - V_{0,T}(V_0^*, \theta)\}$ computed from 10,000 780-day (3-year) return paths simulated from the GARCH model presented in Section 3.2. All hedging strategies are implemented with this model. Note that a negative hedging error implies a gain for the hedger.

by the RMSE$^-$ or the 95% VaR. For example, when protecting a short position in a 3-year ATM call option under GARCH simulated return paths, our modified strategy increases the average hedging gain by 0.38 (from 0.01 to 0.39), while raising the 95% VaR by only 0.01 (from 2.71 to 2.72). As a proportion of the downside risk of hedging, this gain is quite substantial since it is corresponds to 27% of the RMSE$^-$ and 14% of the VaR.

In the regime-switching setting, the results are strikingly in favor of our modified strategy. In fact, we are able to increase the average hedging gain and simultaneously decrease the downside hedging risk, both by a large margin. The reason for this is that the regime-switching model has a positive mean return in regime 1 (bull market regime) and a negative mean return in regime 2 (bear market regime), which allows our modified strategy to take

Table 5: Terminal hedging error under regime-switching simulated return paths

	Hedging a short call position				Hedging a long call position			
	Mean	RMSE	RMSE$^-$	95% VaR	Mean	RMSE	RMSE$^-$	95% VaR
3-year ATM call option								
Global hedging	0.01	1.13	0.84	1.71	−0.01	1.13	0.75	1.48
Modified global hedging	−1.46	2.10	0.64	0.73	−1.51	2.26	0.61	0.63
Delta hedging	−0.61	1.99	1.16	2.75	0.61	1.99	1.61	3.33
3-year OTM call option								
Global hedging	0.01	0.91	0.69	1.32	−0.01	0.91	0.59	1.13
Modified global hedging	−1.16	1.69	0.55	0.53	−1.21	1.84	0.47	0.39
Delta hedging	−0.55	1.64	0.96	2.29	0.55	1.64	1.33	2.68
3-year ITM call option								
Global hedging	0.01	1.31	0.96	2.06	−0.01	1.31	0.90	1.79
Modified global hedging	−1.67	2.42	0.72	0.89	−1.75	2.59	0.73	0.80
Delta hedging	−0.60	2.25	1.33	3.16	0.60	2.25	1.81	3.86

Notes: The statistical measures (mean, RMSE, RMSE$^-$ and 95% VaR) are based on 10,000 realizations of the terminal hedging error $\Lambda_T := B_T \{H - V_{0,T}(V_0^*, \theta)\}$ computed from 10,000 780-day (3-year) return paths simulated from the regime-switching model presented in Section 3.3. All hedging strategies are implemented with this model. Note that a negative hedging error implies a gain for the hedger.

advantage of market timing. Although the true unobserved regime is not used when solving the optimization problem in Equation (3), the hedging position in the underlying asset at time t is a function of the inferred regime probability $\Pr(\chi_t = 1 \mid \mathcal{F}_t)$. In a setting like ours where there is no model or parameter risks, this allows our strategy to take advantage of a potential bet on the direction of the market by positioning itself advantageously based on the likelihood of being in a bull or bear market regime. In contrast, the classical global quadratic hedging strategy is limited in its ability to exploit market timing because whenever the value of the hedging portfolio moves above the value of the claim, it will aim to reduce this discrepancy to more accurately replicate the derivative's payoff at maturity. In Section 4, we assess the performance of the regime-switching model on empirical data, which is a form of

Table 6: Terminal hedging error under Merton jump-diffusion simulated return paths

	Hedging a short call position				Hedging a long call position			
	Mean	RMSE	RMSE$^-$	95% VaR	Mean	RMSE	RMSE$^-$	95% VaR
3-year ATM call option								
Global hedging	0.03	0.61	0.44	0.97	−0.03	0.61	0.42	1.01
Modified global hedging	−0.07	0.63	0.43	0.97	−0.14	0.69	0.43	1.02
Delta hedging	−0.05	0.65	0.46	1.05	0.05	0.65	0.47	1.05
3-year OTM call option								
Global hedging	0.03	0.49	0.35	0.71	−0.03	0.49	0.34	0.83
Modified global hedging	−0.05	0.50	0.34	0.73	−0.12	0.56	0.35	0.87
Delta hedging	−0.06	0.52	0.36	0.83	0.06	0.52	0.38	0.83
3-year ITM call option								
Global hedging	0.03	0.71	0.52	1.14	−0.03	0.71	0.48	1.13
Modified global hedging	−0.08	0.74	0.50	1.12	−0.15	0.79	0.48	1.14
Delta hedging	−0.04	0.76	0.54	1.22	0.04	0.76	0.54	1.22

Notes: The statistical measures (mean, RMSE, RMSE$^-$ and 95% VaR) are based on 10,000 realizations of the terminal hedging error $\Lambda_T := B_T \{H - V_{0,T}(V_0^*, \theta)\}$ computed from 10,000 780-day (3-year) return paths simulated from the Merton jump-diffusion model presented in Section 3.4. All hedging strategies are implemented with this model. Note that a negative hedging error implies a gain for the hedger.

robustness check. We will see that although our modified strategy does not lead to reductions in the downside risk of hedging that are as large as those displayed in Table 5, it nevertheless still succeeds at increasing the average reward for the hedger and simultaneously diminishing his risk.

Finally, we make some remarks about the delta hedging results. We observe that when protecting a short call position under our GARCH simulations, delta hedging yields a larger mean profit than the modified strategy (see Appendix B for an explanation of why delta hedging the short call position leads to a profit on average). In particular, for the ATM option, the terminal hedging gain is larger by 0.47 on average for the delta hedging strategy (from 0.39 to 0.86). However, this is not a free lunch because delta hedging also raises the 95%

VaR by 0.87 (from 2.72 to 3.59), which corresponds to a 32% increase in the downside hedging risk. Moreover, as discussed in Section 2.4, the mean terminal hedging errors for short and long positions are necessarily of opposite sign for a delta hedging strategy. Consequently, a market agent that delta hedges the long position loses money on average, and is also exposed to a larger downside risk relative to the global quadratic hedging strategy. Our modified strategy, in contrast, increases the profitability of both short and long positions with at most a small adverse impact on the downside hedging risk.

4 Empirical study

We now illustrate the performance of our modified strategy for hedging a short or long position in a 3-year European ATM call option under empirically observed return paths of the S&P 500 price index on the period 2008–2018. More precisely, we consider 3-year (780-day) return paths starting on every trading day from 2008-01-02 to 2015-03-25 inclusively, for a total of 1,820 paths (the last path comprises returns from 2015-03-25 to 2018-04-27 inclusively). It must be noted that since these paths are based on series of overlapping returns, the hedging errors computed on these paths are not all independent. The stock price trajectory associated with each return path is assumed to start at $S_0 = 100$ and we continue to assume a constant daily risk-free rate of $r = 0.00769\%$ (2% annualized).

For this experiment to be fully out-of-sample, we estimate the GARCH, regime-switching and jump-diffusion models presented in Sections 3.2, 3.3 and 3.4 with returns on the S&P 500 price index covering the period from 1986-12-31 to 2007-12-31 (5296 returns). Tables 7, 8 and 9 report maximum likelihood parameter estimates. Note that the initial quadratic hedging capital associated with each empirical trajectory is model-dependent. Moreover, for GARCH and regime-switching models, the initial capital varies across paths because it is a function of υ_t, where $\upsilon_t = \sigma_{t+1}$ for the GARCH model and $\upsilon_t = \Pr(\chi_t = 1 \mid \mathcal{F}_t)$ for the regime-switching model.

Table 7: Maximum likelihood estimates of the GARCH model

μ	ω	α	γ	β
0.0312	0.0189	0.0565	0.555	0.908

Notes: These parameter estimates were computed on a time series of daily percentage returns on the S&P 500 price index for the period 1986-12-31 to 2007-12-31 (5296 returns).

Table 8: Maximum likelihood estimates of the regime-switching model

j	μ_j	σ_j^2
1	0.070	0.493
2	-0.066	2.985

$$\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} 0.989 & 0.011 \\ 0.030 & 0.970 \end{pmatrix}$$

Notes: These parameter estimates were computed on a time series of daily percentage returns on the S&P 500 price index for the period 1986-12-31 to 2007-12-31 (5296 returns).

Table 9: Maximum likelihood estimates of the Merton jump-diffusion model

260α	$\sqrt{260}\sigma$	260λ	μ_J	σ_J
0.1020	0.0894	137.67	-0.0009	0.0118

Notes: These parameter estimates were computed on a time series of daily percentage returns on the S&P 500 price index for the period 1986-12-31 to 2007-12-31 (5296 returns).

Table 10 presents the results of our empirical study. This experiment is a form of robustness check for our modified strategy, because in contrast to Section 3, it is now exposed to model risk (in the sense that return paths are not generated by the model used for hedging). We observe that in all cases considered it continues to be more profitable on average than the global quadratic hedging strategy. The gain in average profit is in fact quite substantial in most instances. For example, when protecting a long position in the call option with the GARCH model, our modified strategy increments the average hedging gain by 1.29 (from 0.22 to 1.51), which corresponds to an increase larger than the underlying risk exposure, as measured by the 95% VaR. Moreover, these gains are not associated with an elevated downside hedging risk since in all cases, our strategy actually leads to the smallest RMSE$^-$

Table 10: Terminal hedging error for a 3-year ATM call option under S&P 500 return paths

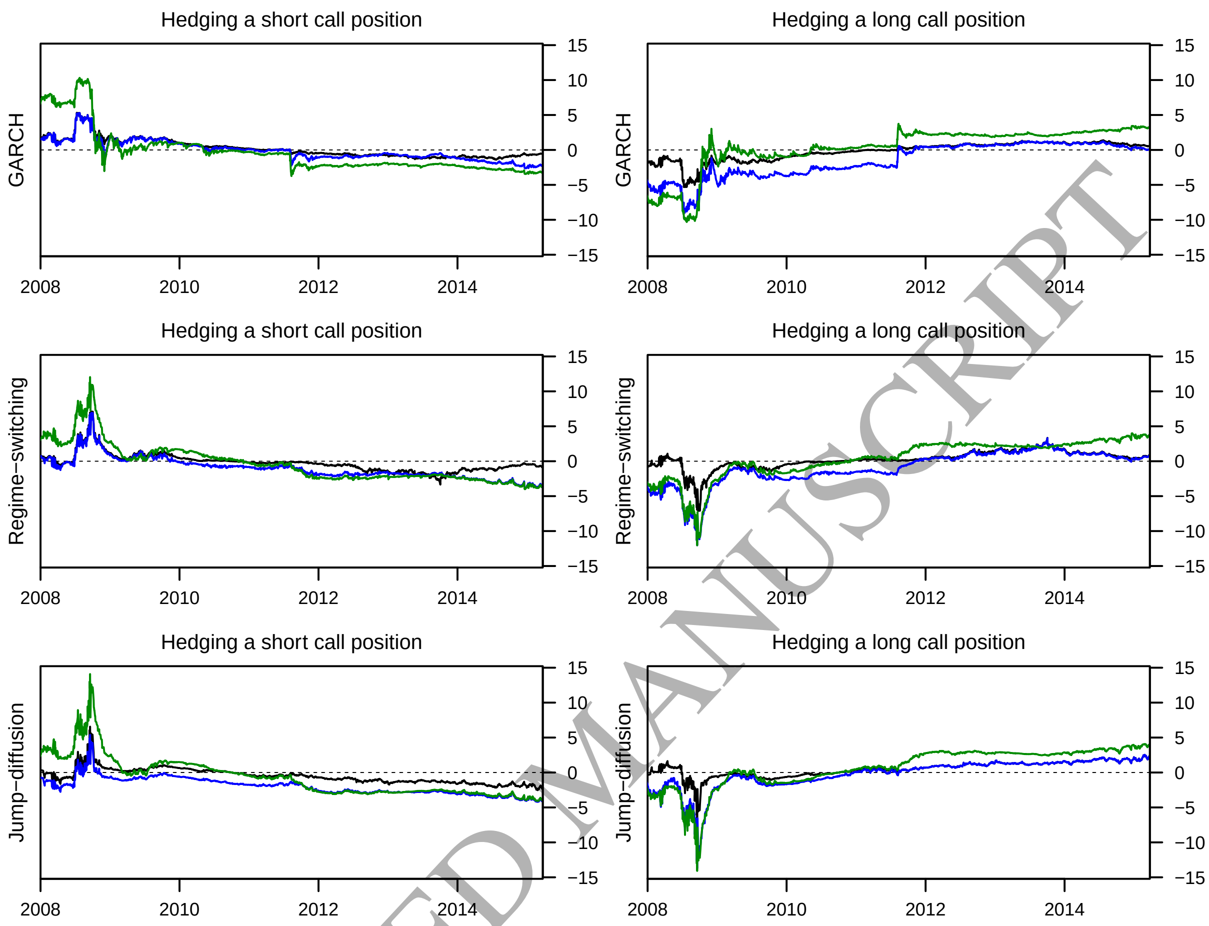
	Hedging a short call position				Hedging a long call position			
	Mean	RMSE	RMSE$^-$	95% VaR	Mean	RMSE	RMSE$^-$	95% VaR
GARCH hedging strategies based on parameters in Table 7								
Global hedging	0.22	1.34	1.20	2.25	−0.22	1.34	0.61	1.14
Modified global hedging	−0.09	1.47	1.12	2.03	−1.51	2.87	0.53	1.10
Delta hedging	−0.33	3.13	2.59	7.51	0.33	3.13	1.75	3.09
Regime-switching hedging strategies based on parameters in Table 8								
Global hedging	−0.21	1.32	0.98	2.25	0.21	1.32	0.88	1.84
Modified global hedging	−1.01	1.90	0.90	2.10	−0.98	2.63	0.81	1.75
Delta hedging	−0.35	2.78	2.12	5.52	0.35	2.78	1.80	3.32
Merton jump-diffusion hedging strategies based on parameters in Table 9								
Global hedging	−0.50	1.20	0.66	0.98	0.50	1.20	1.01	1.93
Modified global hedging	−1.95	2.32	0.41	−0.17	−0.22	2.22	0.98	1.93
Delta hedging	−0.66	2.93	2.08	5.19	0.66	2.93	2.07	3.49

Notes: The statistical measures (mean, RMSE, RMSE$^-$ and 95% VaR) are based on 1,820 values of the terminal hedging error $\Lambda_T := B_T\{H - V_{0,T}(V_0^*, \theta)\}$ computed from empirically observed 780-day (3-year) rolling return paths of the S&P 500 price index starting on every trading day from 2008-01-02 to 2015-03-25 inclusively (the last path comprises returns from 2015-03-25 to 2018-04-27 inclusively). Note that a negative hedging error implies a gain for the hedger.

and 95% VaR.

Figure 1 illustrates the evolution in time of the terminal hedging error for all cases considered. This figure allows us to gain some additional insight because the statistical measures in Table 10 are computed based on hedging error outcomes on overlapping paths. We first observe that short call positions hedged at the outset of the stock market crash of 2008–2009 generated large hedging losses. Delta hedging (green line) performed particularly poorly during this period, while global quadratic hedging (black line) was more effective at paring down losses. In scenarios leading to a terminal hedging loss (i.e., a positive hedging error), our modified strategy (blue line) performed similarly to the standard strategy, as expected following the discussion in Section 2.4. However, in other scenarios the terminal

Figure 1: Evolution of the terminal hedging error for a 3-year ATM call option under S&P 500 return paths



Notes: Each terminal hedging error $\Lambda_T := B_T \{H - V_{0,T}(V_0^*, \theta)\}$ is computed on a 780-day (3-year) return path of the S&P 500 price index. The horizontal axis specifies the path's starting date. A hedging error lying below the horizontal axis implies a gain for the hedger. The graphs on the top (bottom) row are associated with a short (long) position. The graphs on the left (right) column are related to the GARCH (regime-switching) hedging strategies. *Black line*: global hedging, *Blue line*: modified global hedging, *Green line*: delta hedging.

hedging error of our modified strategy almost always lies below that of the standard strategy. Hence, at least from an ex post perspective, our modified strategy produced an almost free lunch in comparison to the global quadratic hedging strategy.

5 Conclusion

This article proposes a modification to the global quadratic hedging strategy introduced by Schweizer (1995) that is more profitable on average to the hedger without substantially increasing his downside hedging risk, if at all. Importantly, this modification does not come at an increased computational cost. In fact, our modified strategy retains the analytical tractability of global quadratic hedging, but is able to reduce the adverse effects of using a quadratic penalty function in the hedging optimization problem by treating hedging gains and losses asymmetrically. The main idea is to reset the global quadratic hedging strategy each time a hedging gain is generated and to allow the hedger to set aside this surplus. We demonstrated mathematically that this simple mechanism can only increase the expected profitability of global quadratic hedging. Interestingly, although the terminal hedging gains for protecting short and long positions are necessarily of opposite sign for global quadratic hedging and delta hedging, our strategy can lead to positive gains for short and long positions simultaneously. Moreover, simulation and empirical studies illustrated that the positive expected profit generated by our strategy is generally not associated with an elevated hedging risk. In fact, on S&P 500 data covering the period 2008–2018, our modified strategy almost never performed worse than global quadratic hedging, and for prolonged periods of time it was able to generate a larger terminal hedging gain.

In future research, it would be interesting to investigate the performance of our strategy for path-dependent payoff structures or compare it to globally risk-minimizing strategies based on non-quadratic criteria.

References

Augustyniak, M., Godin, F., and Simard, C. (2017). Assessing the effectiveness of local and global quadratic hedging under GARCH models. *Quantitative Finance*, 17(9):1305–1318.

Badescu, A., Elliott, R. J., and Ortega, J.-P. (2014). Quadratic hedging schemes for non-Gaussian GARCH models. *Journal of Economic Dynamics & Control*, 42:13–32.

Bertsimas, D., Kogan, L., and Lo, A. W. (2001). Hedging derivative securities and incomplete markets: an ϵ-arbitrage approach. *Operations Research*, 49(3):372–397.

Carr, P. and Wu, L. (2009). Variance risk premiums. *Review of Financial Studies*, 22(3):1311–1341.

Černý, A. (2004). Dynamic programming and mean-variance hedging in discrete time. *Applied Mathematical Finance*, 11(1):1–25.

Christoffersen, P., Dorion, C., Jacobs, K., and Wang, Y. (2010). Volatility components, affine restrictions, and nonnormal innovations. *Journal of Business & Economic Statistics*, 28(4):483–502.

Duan, J.-C. (1995). The GARCH option pricing model. *Mathematical Finance*, 5(1):13–32.

François, P., Gauthier, G., and Godin, F. (2014). Optimal hedging when the underlying asset follows a regime-switching Markov process. *European Journal of Operational Research*, 237(1):312–322.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144.

Ortega, J.-P. (2012). GARCH options via local risk minimization. *Quantitative Finance*, 12(7):1095–1110.

Rémillard, B. and Rubenthaler, S. (2013). Optimal hedging in discrete time. *Quantitative Finance*, 13(6):819–825.

Schweizer, M. (1995). Variance-optimal hedging in discrete time. *Mathematics of Operations Research*, 20(1):1–32.

A Proofs

Before proving our main results, we introduce some notation and derive some preliminary lemmas. All equalities and inequalities are generally assumed to hold almost surely. However, in some instances where it is desirable to stress this property, we make an explicit statement to this effect.

A.1 Notation and preliminaries

The variables ϱ_t and β_t in Equation (4) are defined by Schweizer (1995) as

$$\varrho_t := \frac{\mathbb{E}\left[H\Delta X_t \prod_{j=t+1}^{T}(1-\beta_j\Delta X_j) \mid \mathcal{F}_{t-1}\right]}{\mathbb{E}\left[(\Delta X_t)^2 \prod_{j=t+1}^{T}(1-\beta_j\Delta X_j)^2 \mid \mathcal{F}_{t-1}\right]}, \quad t = 1, \ldots, T-1, \tag{19}$$

$$\varrho_T := \frac{\mathbb{E}\left[H\Delta X_T \mid \mathcal{F}_{T-1}\right]}{\mathbb{E}\left[(\Delta X_T)^2 \mid \mathcal{F}_{T-1}\right]}, \tag{20}$$

$$\beta_t := \frac{\mathbb{E}\left[\Delta X_t \prod_{j=t+1}^{T}(1-\beta_j\Delta X_j) \mid \mathcal{F}_{t-1}\right]}{\mathbb{E}\left[(\Delta X_t)^2 \prod_{j=t+1}^{T}(1-\beta_j\Delta X_j)^2 \mid \mathcal{F}_{t-1}\right]}, \quad t = 1, \ldots, T-1, \tag{21}$$

$$\beta_T := \frac{\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right]}{\mathbb{E}\left[(\Delta X_T)^2 \mid \mathcal{F}_{T-1}\right]}. \tag{22}$$

From Proposition 2.3 of Schweizer (1995), we have that, for $t = 1, \ldots T$,

$$0 \leq \mathbb{E}\left[\prod_{j=t}^{T}(1-\beta_j\Delta X_j)^2 \,\middle|\, \mathcal{F}_{t-1}\right] = \mathbb{E}\left[\prod_{j=t}^{T}(1-\beta_j\Delta X_j) \,\middle|\, \mathcal{F}_{t-1}\right] \leq 1, \tag{23}$$

or equivalently, for $k = 1, \ldots, T$,

$$0 \ \leq 1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}(1-\beta_j\Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right] \leq 1. \tag{24}$$

Lemma 1. *For each $t = 1, \ldots, T$, we have:*

$$\forall j = t, t+1, \ldots, T : \mathbb{E}\left[\Delta X_j \mid \mathcal{F}_{j-1}\right] = 0 \quad \iff \quad \mathbb{E}\left[\prod_{j=t}^{T}(1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{t-1}\right] = 1.$$

Proof. The "if" part is trivial since it implies $\beta_T = 0$ and proceeding recursively for $j = T-1, \ldots, t$, we directly obtain that $\beta_j = 0$, which leads to the result.

We prove the "only if" part, by first noting that, for $t = 1, \ldots, T-1$:

$$\begin{aligned}
&\mathbb{E}\left[\prod_{j=t}^{T}(1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{t-1}\right] \\
&= \mathbb{E}\left[(1 - \beta_t \Delta X_t)\, \mathbb{E}\left[\prod_{j=t+1}^{T}(1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_t\right] \,\middle|\, \mathcal{F}_{t-1}\right] \\
&= \mathbb{E}\left[\prod_{j=t+1}^{T}(1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{t-1}\right] - \beta_t \mathbb{E}\left[\Delta X_t \prod_{j=t+1}^{T}(1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{t-1}\right] \\
&= \underbrace{\mathbb{E}\left[\prod_{j=t+1}^{T}(1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{t-1}\right]}_{\in[0,1]} - \underbrace{\frac{\left(\mathbb{E}\left[\Delta X_t \prod_{j=t+1}^{T}(1 - \beta_j \Delta X_j) \mid \mathcal{F}_{t-1}\right]\right)^2}{\mathbb{E}\left[(\Delta X_t)^2 \prod_{j=t+1}^{T}(1 - \beta_j \Delta X_j)^2 \mid \mathcal{F}_{t-1}\right]}}_{\geq 0}.
\end{aligned} \tag{25}$$

The first term on the right-hand side of Equation (25) is contained in $[0, 1]$ by Equation (23), whereas the second term is non-negative. Therefore, if

$$\mathbb{E}\left[\prod_{j=t}^{T}(1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{t-1}\right] = 1,$$

then we must necessarily have, for $u = t+1, t+2, \ldots, T$:

$$\mathbb{E}\left[\prod_{j=u}^{T}(1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{u-1}\right] = 1. \tag{26}$$

In particular, for $u = T$, Equation (26) implies $\beta_T \mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right] = 0$, which in turn necessarily leads to $\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right] = 0$ and $\beta_T = 0$. This is simply because $\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right] \neq 0$

would imply $\beta_T \neq 0$. Proceeding recursively for $u = T-1, \dots, t$, we find that

$$\mathbb{E}\left[\prod_{j=u}^{T}(1-\beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{u-1}\right] = \mathbb{E}\left[1-\beta_u \Delta X_u \mid \mathcal{F}_{u-1}\right],$$

and hence $\beta_u \mathbb{E}\left[\Delta X_u \mid \mathcal{F}_{u-1}\right] = 0$, which once again entails $\mathbb{E}\left[\Delta X_u \mid \mathcal{F}_{u-1}\right] = 0$ and $\beta_u = 0$. The proof is therefore complete. □

Corollary 1. *If X is a $\mathbb{P}$-martingale, then the process β is identically 0, and, for $k = 1, \dots, T$,*

$$1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}(1-\beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right] = 0.$$

Proof. If X is a $\mathbb{P}$-martingale, then $\mathbb{E}\left[\Delta X_j \mid \mathcal{F}_{j-1}\right] = 0$ for all $j = 1, \dots, T$, which implies that the process β is identically 0, and the result then follows. □

Corollary 2.

$$\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right] \neq 0 \quad \Longleftrightarrow \quad \forall t = 1, \dots, T : 0 \leq \mathbb{E}\left[\prod_{j=t}^{T}(1-\beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{t-1}\right] < 1.$$

Proof. This is a direct consequence of Lemma 1 and Equation (23). □

Finally, we define:

$$C_0 := 0,$$
$$C_t := \sum_{k=1}^{t} \left(Z_{k-1} - V_{k-1}\right) \beta_k \Delta X_k, \quad t = 1, \dots, T.$$

A.2 Proof of Theorem 1

Before demonstrating Theorem 1, we prove the following lemma.

Lemma 2. *The sum of all cash flows generated by our modified global quadratic hedging strategy can be equivalently expressed as*

$$Z_T + \sum_{t=1}^{T-1} \max\left(Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t - V_t^*, 0\right) = V_0^* + \sum_{t=1}^{T} \xi_t(Z_{t-1})\Delta X_t.$$

Proof. This result follows directly from the definitions $Z_0 = V_0^*$ and $Z_T = Z_{T-1} + \xi_T(Z_{T-1})\Delta X_T$, and the fact that

$$\max\left(Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t - V_t^*, 0\right) = Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t - Z_t, \quad t = 1, \dots T-1.$$

□

We now prove Theorem 1.

Proof. Using Equation (2) and Lemma 2, we obtain

$$\begin{aligned}
&\mathbb{E}\left[H - \left(Z_T + \sum_{t=1}^{T-1} \max\left(Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t - V_t^*, 0\right)\right)\right] \\
&= \mathbb{E}\left[V_{0,T}(V_0^*, \xi^*) - \left(V_0^* + \sum_{t=1}^{T} \xi_t(Z_{t-1})\Delta X_t\right)\right] \\
&= \mathbb{E}\left[V_0^* + \sum_{t=1}^{T} \xi_t^* \Delta X_t - V_0^* - \sum_{t=1}^{T} \xi_t(Z_{t-1})\Delta X_t\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \left\{\xi_t^* - \xi_t(Z_{t-1})\right\} \Delta X_t\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \left\{\xi_t(V_{t-1}) - \xi_t(Z_{t-1})\right\} \Delta X_t\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \left\{\varrho_t - \beta_t V_{t-1} - \varrho_t + \beta_t Z_{t-1}\right\} \Delta X_t\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \left\{Z_{t-1} - V_{t-1}\right\} \beta_t \Delta X_t\right] \qquad (27) \\
&= \mathbb{E}\left[C_T\right].
\end{aligned}$$

We prove that the expectation $\mathbb{E}\left[C_T\right]$ is non-positive by way of a recursive argument. First, we show by induction that the following inequality holds for $k = 1, \ldots, T$:

$$\mathbb{E}\left[C_T \mid \mathcal{F}_{T-k}\right] \leq C_{T-k} + \left(Z_{T-k} - V_{T-k}\right)\left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}\left(1 - \beta_j \Delta X_j\right) \,\middle|\, \mathcal{F}_{T-k}\right]\right). \tag{28}$$

Observe that this inequality is trivially satisfied for $k = 1$, since

$$\begin{aligned}\mathbb{E}\left[C_T \mid \mathcal{F}_{T-1}\right] &= \mathbb{E}\left[C_{T-1} + \left(Z_{T-1} - V_{T-1}\right)\beta_T \Delta X_T \mid \mathcal{F}_{T-1}\right] \\ &= C_{T-1} + \left(Z_{T-1} - V_{T-1}\right)\mathbb{E}\left[\beta_T \Delta X_T \mid \mathcal{F}_{T-1}\right].\end{aligned} \tag{29}$$

Moreover, since we have, for $t = 1, \ldots, T-1$,

$$\begin{aligned}Z_t &= \min(V_t^*, Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t) \leq Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t, \\ V_t &= V_{t-1} + \xi_t(V_{t-1})\Delta X_t,\end{aligned}$$

then, for $k = 1, \ldots, T-1$,

$$\begin{aligned}Z_{T-k} - V_{T-k} &\leq Z_{T-k-1} - V_{T-k-1} + \{\xi_{T-k}(Z_{T-k-1}) - \xi_{T-k}(V_{T-k-1})\}\Delta X_{T-k} \\ &= Z_{T-k-1} - V_{T-k-1} + \{\varrho_{T-k} - \beta_{T-k}Z_{T-k-1} - \varrho_{T-k} + \beta_{T-k}V_{T-k-1}\}\Delta X_{T-k} \\ &= \left(Z_{T-k-1} - V_{T-k-1}\right)\left(1 - \beta_{T-k}\Delta X_{T-k}\right).\end{aligned} \tag{30}$$

Combining inequalities (24) and (30), we can conclude that, for $k = 1, \ldots, T-1$,

$$\begin{aligned}&\left(Z_{T-k} - V_{T-k}\right)\left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}\left(1 - \beta_j \Delta X_j\right) \,\middle|\, \mathcal{F}_{T-k}\right]\right) \\ &\quad \leq \left(Z_{T-k-1} - V_{T-k-1}\right)\left(1 - \beta_{T-k}\Delta X_{T-k}\right)\left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}\left(1 - \beta_j \Delta X_j\right) \,\middle|\, \mathcal{F}_{T-k}\right]\right).\end{aligned} \tag{31}$$

Assuming that the inequality in (28) holds for k, we can write, for $k = 1, \ldots, T-1$,

$$\mathbb{E}\left[C_T \mid \mathcal{F}_{T-k-1}\right] = \mathbb{E}\left[\mathbb{E}\left[C_T \mid \mathcal{F}_{T-k}\right] \mid \mathcal{F}_{T-k-1}\right] \tag{32}$$

$$\leq \mathbb{E}\left[C_{T-k} + (Z_{T-k} - V_{T-k})\left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}(1-\beta_j \Delta X_j)\,\middle|\,\mathcal{F}_{T-k}\right]\right)\,\middle|\,\mathcal{F}_{T-k-1}\right] \tag{33}$$

$$\leq \mathbb{E}\Bigg[C_{T-k} + (Z_{T-k-1} - V_{T-k-1})(1-\beta_{T-k}\Delta X_{T-k}) \times \left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}(1-\beta_j \Delta X_j)\,\middle|\,\mathcal{F}_{T-k}\right]\right)\,\Bigg|\,\mathcal{F}_{T-k-1}\Bigg] \tag{34}$$

$$= \mathbb{E}\Bigg[C_{T-k} + (Z_{T-k-1} - V_{T-k-1}) \times \left(1 - \beta_{T-k}\Delta X_{T-k} - \mathbb{E}\left[\prod_{j=T-k}^{T}(1-\beta_j \Delta X_j)\,\middle|\,\mathcal{F}_{T-k}\right]\right)\,\Bigg|\,\mathcal{F}_{T-k-1}\Bigg] \tag{35}$$

$$= \mathbb{E}\Bigg[C_{T-k-1} + (Z_{T-k-1} - V_{T-k-1})\beta_{T-k}\Delta X_{T-k} + (Z_{T-k-1} - V_{T-k-1}) \times \left(1 - \beta_{T-k}\Delta X_{T-k} - \mathbb{E}\left[\prod_{j=T-k}^{T}(1-\beta_j \Delta X_j)\,\middle|\,\mathcal{F}_{T-k}\right]\right)\,\Bigg|\,\mathcal{F}_{T-k-1}\Bigg] \tag{36}$$

$$= C_{T-k-1} + (Z_{T-k-1} - V_{T-k-1})\left(1 - \mathbb{E}\left[\prod_{j=T-k}^{T}(1-\beta_j \Delta X_j)\,\middle|\,\mathcal{F}_{T-k-1}\right]\right), \tag{37}$$

which completes the proof by induction of inequality (28). Finally, taking $k = T$ in inequality (28), we obtain

$$\mathbb{E}\left[C_T\right] \leq C_0 + (Z_0 - V_0)\left(1 - \mathbb{E}\left[\prod_{j=1}^{T}(1-\beta_j \Delta X_j)\right]\right) = 0,$$

because by definition $C_0 = 0$, $Z_0 = V_0^*$ and $V_0 = V_0^*$, which completes the proof of Theorem 1.

□

A.3 Proof of Theorem 2

Before demonstrating Theorem 2, we prove the following lemmas.

Lemma 3. *The inequality* (7) *in Theorem 1 is an equality if and only if for all $k = 1, \ldots, T$:*

$$\mathbb{E}\left[C_T \mid \mathcal{F}_{T-k}\right] = C_{T-k} + (Z_{T-k} - V_{T-k})\left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}(1-\beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right]\right) \quad a.s. \quad (38)$$

Proof. The "if" part is trivial since if we take $k = T$, we obtain $\mathbb{E}\left[C_T\right] = 0$, which directly implies an equality in the result of Theorem 1.

To prove the "only if" part, we proceed by contraposition, that is, we demonstrate that if Equation (38) does not hold, then $\mathbb{E}\left[C_T\right] < 0$. First, note that by Equation (28), we have, for $k = 1, \ldots, T$,

$$\mathbb{E}\left[C_T \mid \mathcal{F}_{T-k}\right] \leq C_{T-k} + (Z_{T-k} - V_{T-k})\left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}(1-\beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right]\right) \quad \text{a.s.}$$

Hence, if Equation (38) does not hold, then for some k we must have

$$\mathbb{E}\left[C_T \mid \mathcal{F}_{T-k}\right] < C_{T-k} + (Z_{T-k} - V_{T-k})\left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}(1-\beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right]\right),$$

on some subset of Ω with probability greater than zero. By Equations (32)–(37), this in turn implies that on some subset of Ω with probability greater than zero we have

$$\begin{aligned}
\mathbb{E}\left[C_T \mid \mathcal{F}_{T-k-1}\right] &= \mathbb{E}\left[\mathbb{E}\left[C_T \mid \mathcal{F}_{T-k}\right] \mid \mathcal{F}_{T-k-1}\right] \\
&< \mathbb{E}\left[C_{T-k} + (Z_{T-k} - V_{T-k})\left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T}(1-\beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right]\right) \,\middle|\, \mathcal{F}_{T-k-1}\right] \\
&\leq C_{T-k-1} + (Z_{T-k-1} - V_{T-k-1})\left(1 - \mathbb{E}\left[\prod_{j=T-k}^{T}(1-\beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k-1}\right]\right),
\end{aligned}$$

which recursively leads to

$$\mathbb{E}\left[C_T \mid \mathcal{F}_{T-h}\right] < C_{T-h} + (Z_{T-h} - V_{T-h})\left(1 - \mathbb{E}\left[\prod_{j=T-h+1}^{T}(1-\beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-h}\right]\right),$$

for all $h = k, k+1, \dots, T$. Taking $h = T$, we obtain $\mathbb{E}[C_T] < 0$, which completes the proof by contraposition. $\square$

Lemma 4. *The inequality* (7) *in Theorem 1 is an equality if and only if for all* $k = 1, \dots, T-1$*:*

$$(Z_{T-k} - V_{T-k}) \left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right]\right)$$
$$= (Z_{T-k-1} - V_{T-k-1})(1 - \beta_{T-k}\Delta X_{T-k}) \left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right]\right) \quad a.s. \tag{39}$$

Proof. To prove the "if" part, we show that Equation (38) in Lemma 3 is satisfied for all $k = 1, \dots, T$. From Equation (29), we know that it is satisfied for $k = 1$. Proceeding by induction, we assume that Equation (38) holds for k. Based on Equations (32)–(37) we can then write,

$$\begin{aligned}
&\mathbb{E}[C_T \mid \mathcal{F}_{T-k-1}] = \mathbb{E}[\mathbb{E}[C_T \mid \mathcal{F}_{T-k}] \mid \mathcal{F}_{T-k-1}] \\
&\quad = \mathbb{E}\left[C_{T-k} + (Z_{T-k} - V_{T-k}) \left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right]\right) \,\middle|\, \mathcal{F}_{T-k-1}\right] \\
&\quad = \mathbb{E}\Bigg[C_{T-k} + (Z_{T-k-1} - V_{T-k-1})(1 - \beta_{T-k}\Delta X_{T-k}) \\
&\qquad\qquad \times \left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right]\right) \,\Bigg|\, \mathcal{F}_{T-k-1}\Bigg] \\
&\quad = C_{T-k-1} + (Z_{T-k-1} - V_{T-k-1}) \left(1 - \mathbb{E}\left[\prod_{j=T-k}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k-1}\right]\right),
\end{aligned}$$

which implies that Equation (38) is also satisfied for $k+1$, and hence for all $k = 1, \dots, T$.

We also rely on the equivalence identified in Lemma 3 to demonstrate the "only if" part. Assume that Equation (38) holds for all $k = 1, \dots, T$. Then, based on Equations (32)–(37), we can establish that necessarily, for $k = 1, \dots, T-1$,

$$\begin{aligned}
&\mathbb{E}\left[(Z_{T-k} - V_{T-k}) \left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k} \right] \right) \,\middle|\, \mathcal{F}_{T-k-1} \right] \\
&\quad = \mathbb{E}\Bigg[(Z_{T-k-1} - V_{T-k-1}) (1 - \beta_{T-k} \Delta X_{T-k}) \\
&\qquad\qquad \times \left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k} \right] \right) \,\Bigg|\, \mathcal{F}_{T-k-1} \Bigg]. \qquad (40)
\end{aligned}$$

Since Equation (31) implies that, for $k = 1, \dots, T-1$,

$$\begin{aligned}
&(Z_{T-k} - V_{T-k}) \left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k} \right] \right) \\
&\quad \le (Z_{T-k-1} - V_{T-k-1}) (1 - \beta_{T-k} \Delta X_{T-k}) \left(1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k} \right] \right),
\end{aligned}$$

then Equation (40) can only be satisfied if both sides in this inequality are equal almost surely. Hence, we can conclude that Equation (39) must hold for $k = 1, \dots, T-1$. $\square$

Lemma 5. *Assume that:*

$$\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1} \right] \neq 0.$$

Then, the inequality (7) in Theorem 1 is an equality if and only if $Z_t = V_t$ a.s. for $t = 0, 1, \dots, T$.

Proof. The "if" part is trivial since if $Z_t = V_t$ a.s., for $t = 0, 1, \dots, T$, then, from Equation (27), we obtain

$$\mathbb{E}\left[C_T \right] = \mathbb{E}\left[\sum_{t=1}^{T} \left\{ Z_{t-1} - V_{t-1} \right\} \beta_t \Delta X_t \right] = 0,$$

which directly implies an equality in the result of Theorem 1. The condition $\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1} \right] \neq 0$ is in fact not needed for the "if" part.

We rely on the equivalence identified in Lemma 4 to prove the "only if" part. Therefore, assume that Equation (39) in Lemma 4 holds for $k = 1, \ldots, T-1$. Since by Corollary 2 the condition $\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right] \neq 0$ implies

$$1 - \mathbb{E}\left[\prod_{j=T-k+1}^{T} (1 - \beta_j \Delta X_j) \,\middle|\, \mathcal{F}_{T-k}\right] > 0,$$

then Equation (39) can be equivalently written as

$$\forall k = 1, \ldots, T-1 : Z_{T-k} - V_{T-k} = \left(Z_{T-k-1} - V_{T-k-1}\right)\left(1 - \beta_{T-k}\Delta X_{T-k}\right) \text{ a.s.}$$

Because $Z_0 = V_0$ by definition, this equation recursively implies $Z_t = V_t$ a.s. for $t = 1, \ldots, T-1$. Moreover, since $Z_T = Z_{T-1} + \xi_T(Z_{T-1})\Delta X_T$ and $V_T = V_{T-1} + \xi_T(V_{T-1})\Delta X_T$ by definition, we obtain $Z_T = V_T$ a.s., which completes the proof. □

Lemma 6. *Assume that:*

$$\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right] \neq 0.$$

Then, the inequality (7) *in Theorem 1 is an equality if and only if V^*_{T-1} is attainable with the global quadratic hedging portfolio (V^*_0, ξ^*) in the sense that $V^*_{T-1} = V_{T-1}$ a.s.*

Proof. To prove the "only if" part, we rely on Lemma 5, which implies that if the inequality (7) in Theorem 1 is an equality, then we must necessarily have $Z_t = V_t$ a.s. for $t = 0, 1, \ldots, T$. First, note that the statement $Z_t = V_t$ a.s. for $t = 0, 1, \ldots, T$ is equivalent to $V^*_t \geq V_t$ a.s. for $t = 0, 1, \ldots, T-1$. This equivalence follows directly from the definition of Z_t in Equations (5) and (6). Now, consider the optimization problem,

$$\min_{(c,\phi)} \mathbb{E}\left[\left\{H - V_{T-1,T}(c,\phi)\right\}^2 \,\middle|\, \mathcal{F}_{T-1}\right], \tag{41}$$

which, by definition, has the solution $\left(V^*_{T-1}, \xi_T(V^*_{T-1})\right)$. Since this solution is an optimal global quadratic hedging portfolio, it must satisfy, similarly to Equation (2),

$$\mathbb{E}\left[V_{T-1}^* + \xi_T(V_{T-1}^*)\Delta X_T \mid \mathcal{F}_{T-1}\right] = \mathbb{E}\left[H \mid \mathcal{F}_{T-1}\right],$$

and hence also

$$\mathbb{E}\left[V_{T-1}^* + \xi_T(V_{T-1}^*)\Delta X_T\right] = \mathbb{E}\left[H\right].$$

Moreover, we also have from Equation (2) that

$$\mathbb{E}\left[V_{T-1} + \xi_T(V_{T-1})\Delta X_T\right] = \mathbb{E}\left[H\right],$$

from which we deduce the following equality

$$\mathbb{E}\left[V_{T-1}^* + \xi_T(V_{T-1}^*)\Delta X_T\right] = \mathbb{E}\left[V_{T-1} + \xi_T(V_{T-1})\Delta X_T\right].$$

Therefore, we can write

$$\begin{aligned}
&\mathbb{E}\left[V_{T-1}^* + \xi_T(V_{T-1}^*)\Delta X_T - V_{T-1} - \xi_T(V_{T-1})\Delta X_T\right] \\
&\quad = \mathbb{E}\left[\left(V_{T-1}^* - V_{T-1}\right)\left(1 - \beta_T \Delta X_T\right)\right] \\
&\quad = \mathbb{E}\big[\underbrace{\left(V_{T-1}^* - V_{T-1}\right)}_{\geq 0} \underbrace{\mathbb{E}\left[(1 - \beta_T \Delta X_T) \mid \mathcal{F}_{T-1}\right]}_{>0}\big] = 0,
\end{aligned}$$

which from the assumptions $V_{T-1}^* \geq V_{T-1}$ a.s. and $\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right] \neq 0$ allows us to conclude that $V_{T-1}^* = V_{T-1}$ and $\xi_T(V_{T-1}^*) = \xi_T^*$ a.s. In other words, the claim V_{T-1}^* is attainable with the global quadratic hedging portfolio (V_0^*, ξ^*).

For the "if" part, we assume that $V_{T-1}^* = V_{T-1}$ a.s. (the condition $\mathbb{E}\left[\Delta X_T \mid \mathcal{F}_{T-1}\right] \neq 0$ is not needed in this part). Consequently, (V_{T-1}, ξ_T^*) solves the optimization problem in Equation (41), which implies that for all $(c, \theta) \in \mathbb{R} \times \Theta_t$ and $t = 0, 1, \ldots, T-1$, we have

$$\mathbb{E}\left[\{H - V_{t,T}(c,\theta)\}^2 \mid \mathcal{F}_{T-1}\right] \geq \mathbb{E}\left[\{H - V_{T-1,T}(V_{T-1}, \xi_T^*)\}^2 \mid \mathcal{F}_{T-1}\right],$$

and hence also

$$\mathbb{E}\left[\{H - V_{t,T}(c,\theta)\}^2 \mid \mathcal{F}_t\right] \geq \mathbb{E}\left[\{H - V_{T-1,T}(V_{T-1}, \xi_T^*)\}^2 \mid \mathcal{F}_t\right]. \tag{42}$$

Since the choices $c = V_t$ and $\theta = \{\xi_k^*\}_{k=t+1,\ldots,T}$ lead to

$$V_{t,T}(c,\theta) = V_t + \sum_{k=t+1}^{T-1} \xi_k^* \Delta X_k + \xi_T^* \Delta X_T = V_{T-1} + \xi_T^* \Delta X_T = V_{T-1,T}(V_{T-1}, \xi_T^*),$$

the lower bound can actually be attained in Equation (42), which implies that $c = V_t$ and $\theta = \{\xi_k^*\}_{k=t+1,\ldots,T}$ must solve the following optimization problem:

$$\min_{(c,\theta)} \mathbb{E}\left[\{H - V_{t,T}(c,\theta)\}^2 \mid \mathcal{F}_t\right].$$

Since the value of c that solves this problem is equal to V_t^* by definition, we can hence conclude that $V_t^* = V_t$ a.s. for $t = 0, 1, \ldots, T-1$. This then implies $Z_t = V_t$ a.s. for $t = 0, 1, \ldots, T$, which, after invoking Lemma 5, completes the proof. □

We now prove Theorem 2.

Proof. Theorem 2 is essentially a restatement of Lemmas 5 and 6. For completeness, we add a remark concerning statement 3 in the theorem. From the proof of Lemma 6, we get that $V_{T-1}^* = V_{T-1}$ a.s. implies $V_t^* = V_t$ a.s. for $t = 1, \ldots, T-1$, which allows us to strengthen Lemma 5 in the sense that $Z_t = V_t = V_t^*$ a.s. for $t = 0, 1, \ldots, T-1$. □

A.4 Proof of Proposition 1

We prove Proposition 1.

Proof. If H is attainable, then there exists a predictable process ξ^H such that $H = H_0 + \sum_{t=1}^{T} \xi_t^H \Delta X_t$. The optimal global quadratic hedging portfolio for this contingent claim is then obviously given by the pair $(V_0^* = H_0, \xi^* = \xi^H)$, as shown in Section 4.2 of Schweizer

(1995). Moreover, for $t = 1, \dots, T-1$, we have that $V_t^* = H_0 + \sum_{k=1}^{t} \xi_t^H \Delta X_t$ because the expectation,

$$\begin{aligned}
\mathbb{E}\left[\{H - V_{t,T}(c,\theta)\}^2 \mid \mathcal{F}_t\right] &= \mathbb{E}\left[\left(H_0 + \sum_{k=1}^{T} \xi_k^H \Delta X_k - c - \sum_{k=t+1}^{T} \theta_k \Delta X_k\right)^2 \,\middle|\, \mathcal{F}_t\right] \\
&= \mathbb{E}\left[\left(H_0 + \sum_{k=1}^{t} \xi_k^H \Delta X_k - c + \sum_{k=t+1}^{T} \left(\xi_k^H - \theta_k\right) \Delta X_k\right)^2 \,\middle|\, \mathcal{F}_t\right],
\end{aligned}$$

equals zero, and is hence minimized over all $(c,\theta) \in \mathbb{R} \times \Theta_t$ if we take $c = H_0 + \sum_{k=1}^{t} \xi_t^H \Delta X_t$ and $\theta = \{\xi_k^H\}_{k=t+1,\dots,T}$. Consequently, since $Z_0 = V_0^* = H_0$ and $\xi_1(Z_0) = \xi_1^* = \xi_1^H$, then one finds that $Z_1 = \min(V_1^*, Z_0 + \xi_1(Z_0)\Delta X_1) = H_0 + \xi_1^H \Delta X_1 = V_1$, which in turns recursively implies $\xi_t(Z_{t-1}) = \xi_t^* = \xi_t^H$ and $Z_t = H_0 + \sum_{k=1}^{t} \xi_t^H \Delta X_t = V_t$, for $t = 2, \dots, T$. Hence, our modified strategy is in this case identical to the global quadratic hedging strategy.

Furthermore, if X is a $\mathbb{P}$-martingale, then the process β is identically 0, which implies by Equation (4) that

$$\xi_t(z) = \varrho_t = \frac{\mathbb{E}\left[H \Delta X_t \mid \mathcal{F}_{t-1}\right]}{\mathbb{E}\left[(\Delta X_t)^2 \mid \mathcal{F}_{t-1}\right]}, \quad t = 1, \dots, T,$$

independently of z. Hence, the optimal global quadratic hedging position does not depend on the current (or past) value of the hedging portfolio in the martingale setting (see also Schweizer, 1995, Section 4.1). This leads to $\xi_t(Z_{t-1}) = \xi_t(V_{t-1}) = \xi_t^*$ a.s., for $t = 1, \dots, T$, and hence by Lemma 2 we obtain

$$Z_T + \sum_{t=1}^{T-1} \max\left(Z_{t-1} + \xi_t(Z_{t-1})\Delta X_t - V_t^*, 0\right) = V_0^* + \sum_{t=1}^{T} \xi_t^* \Delta X_t \text{ a.s.}$$

□

A.5 Proof of Proposition 2

We prove Proposition 2.

Proof. From Equation (3), we know that the solution to (11) is given by $\phi = \xi_{t+1}(z)$ for some $z \in \mathcal{F}_t$ that satisfies the constraint $z \leq Z_t$. Thus, to conclude that $\phi = \xi_{t+1}(Z_t)$, it remains to show that the objective function in (11) is minimized at $z = Z_t$.

For each fixed z, we have that,

$$\begin{aligned}
&\min_{\phi}\left\{\min_{\theta\in\Theta_{t+1}} \mathbb{E}\left[\left\{H - \left(z + \phi\Delta X_{t+1} + \sum_{k=t+2}^{T}\theta_k\Delta X_k\right)\right\}^2 \,\middle|\, \mathcal{F}_t\right]\right\} \\
&= \mathbb{E}\left[\left\{H - V_{t,T}(z, \xi^*(z))\right\}^2 \,\middle|\, \mathcal{F}_t\right] \\
&\geq \mathbb{E}\left[\left\{H - V_{t,T}(V_t^*, \xi^*(V_t^*))\right\}^2 \,\middle|\, \mathcal{F}_t\right],
\end{aligned}$$

where $\xi^*(z)$ here refers to the global quadratic hedging strategy initiated with a capital of z at time t.

Let

$$\tilde{\mathcal{Z}}^{(t)} := \prod_{j=t+1}^{T}(1 - \beta_j\Delta X_j), \quad t = 0, 1, \ldots, T-1,$$

and note that from Equation (23), we have that,

$$0 \leq \mathbb{E}\left[\left(\tilde{\mathcal{Z}}^{(t)}\right)^2 \,\middle|\, \mathcal{F}_t\right] = \mathbb{E}\left[\tilde{\mathcal{Z}}^{(t)} \,\middle|\, \mathcal{F}_t\right] \leq 1.$$

Corollary 2.5 of Schweizer (1995) allows us to deduce that

$$\begin{aligned}
&\mathbb{E}\left[\left\{H - V_{t,T}(z, \xi^*(z))\right\}^2 \,\middle|\, \mathcal{F}_t\right] - \mathbb{E}\left[\left\{H - V_{t,T}(V_t^*, \xi^*(V_t^*))\right\}^2 \,\middle|\, \mathcal{F}_t\right] \\
&= \mathbb{E}\left[\tilde{\mathcal{Z}}^{(t)} \,\middle|\, \mathcal{F}_t\right](V_t^* - z)^2.
\end{aligned}$$

Consequently, the objective function in (11) is minimized at the value of z that satisfies the constraint $z \leq Z_t$ and is closest to V_t^*. Since by the definition of Z_t in Equation (5), we have that $Z_t \leq V_t^*$, the optimal value of z is thus $z = Z_t$ as required. $\square$

B Why delta hedging the short call position in the GARCH setting leads to a profit on average

The reason why delta hedging a short call (or put) position in our GARCH model leads to a profit on average can be explained as follows. First, we stress that although our initial capital assumption of V_0^* obviously impacts the profit distribution of the delta hedging strategy, it is not the main cause of the positive expected gain. In fact, Table 2 in Augustyniak et al. (2017) shows that V_0^* is almost identical to the risk-neutral expectation of the discounted call option payoff in our estimated GARCH model, and this expectation would be the natural choice for the initial capital in the delta hedging strategy.

The main reason for the positive expected profit is that, under our model assumptions, the expected future volatility under the risk-neutral measure can be shown to be larger than under the physical measure $\mathbb{P}$. This discrepancy entails that our estimated models give rise to an implicit negative variance risk premium, which is in line with empirical research (see Carr and Wu, 2009). A negative sign on the variance risk premium indicates that option buyers are willing to pay a premium (i.e., accept a negative average excess return) to protect themselves against upward movements in stock market volatility. Therefore, whenever a market agent sells an option, he receives extra compensation for taking volatility risk. Since delta hedging this short position only hedges away movements in the stock and not volatility, the agent is rewarded for taking volatility risk by earning a positive average excess return.