



Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **12-34**

Optimal Hedging when the Underlying Asset Follows a Regime-switching Markov Process

Pascal François
Geneviève Gauthier
Frédéric Godin

Août/August 2012

François: Professor at HEC Montréal, Department of Finance and CIRPÉE Fellow
pascal.francois@hec.ca

Gauthier: Professor at HEC Montréal, Department of Management Sciences
genevieve.gauthier@hec.ca

Godin: Ph.D. student at HEC Montréal, Department of Management Sciences
frederic.godin@hec.ca

Earlier versions of this paper were presented at the IFM2 Mathematical Finance Days 2012, and CORS Annual Conference 2012. We thank the conference participants for their helpful feedback. Financial support from SSHRC (François), NSERC (Gauthier, Godin) and the Montreal Exchange (Godin) is gratefully acknowledged.

Electronic copy available at: <http://ssrn.com/abstract=2130822>

Abstract:

We develop a flexible discrete-time hedging methodology that minimizes the expected value of any desired penalty function of the hedging error within a general regime-switching framework. A numerical algorithm based on backward recursion allows for the sequential construction of an optimal hedging strategy. Numerical experiments comparing this and other methodologies show a relative expected penalty reduction ranging between 0.9% and 12.6% with respect to the best benchmark.

Keywords: Dynamic programming, hedging, risk management, regime switching

JEL Classification: G32, C61

1 Introduction and literature review

For a derivatives trading and risk management activity to be sustainable, hedging is paramount. In practice, portfolio rebalancing is performed in discrete time and the market is typically incomplete, implying that most contingent claims cannot be replicated exactly. Thus, to implement a hedging policy, the challenge is twofold: a model must be specified and hedging strategy objectives must be set.

From a modelling perspective, this article adopts a regime-switching environment. One widely studied class of regime-switching models views log-returns as a mixture of Gaussian variables. These models, introduced in finance by Hamilton (1989), have been shown to improve the statistical fit and forecasts of financial returns. They reproduce widely documented empirical properties such as heteroskedasticity, autocorrelation and fat tails. In this framework, the option pricing problem must deal with incomplete markets and requires the specification of a risk premium. Among significant contributions, Bollen (1998) presents a lattice algorithm to compute the value of European and American options. Hardy (2001) finds a closed-form formula for the price of European options. The continuous-time version of the Gaussian mixture model is studied by Mamon & Rodrigo (2005) who find an explicit value for European options by solving a partial differential equation. Elliott et al. (2005) price derivatives by means of the Esscher transform under the same continuous-time model. Buffington & Elliott (2002) derive an approximate formula for American option prices. Beyond the Gaussian mixture models, extensions address GARCH effects (Duan et al., 2002) and jumps (Lee, 2009a), for example.

Several authors study the problem of hedging an underlying asset with its futures under regime-switching frameworks. Alizadeh & Nomikos (2004) and Alizadeh et al. (2008) base their hedging strategy on minimal variance hedge ratios. Lee et al. (2006), Lee & Yoder (2007), Lee (2009a) and Lee (2009b) extend the dynamics of the underlying asset in Alizadeh & Nomikos (2004) to incorporate a time-varying correlation between the spot and futures returns, GARCH-type feedback from returns on the volatility, jumps and copulas for the dependence between futures and spot returns. Lien (2012) provides conditions under which minimal variance ratios taking into account the existence of regimes overperform their unconditional counterparts.

Option hedging under regime-switching models has recently raised interest in the literature. Rémillard & Rubenthaler (2009) identify the hedging strategy that minimizes the squared error of hedging in both discrete-time and continuous-time for European options. The implementation of this methodology is present in Rémillard et al. (2010a). Rémillard et al. (2010b) extend the hedging procedure to American options.

Another strand of literature discusses self-financing hedging policies¹ under general model assumptions. A widely known methodology is delta hedging. It consists in building a portfolio whose value variations mimic those of the hedged contingent claim when small changes in the underlying asset's value occur. In continuous-time complete markets, delta hedging is the cornerstone of any hedging strategy since it allows for perfect replication. Based on the first derivative of the option price with respect to the underlying asset price, it requires a full characterization of the risk-neutral measure. Many authors discuss the implementation of delta hedging in discrete-time and/or incomplete markets (Duan, 1995, among others). It should be stressed, however, that delta hedging is subject to model misspecification. Nevertheless, it stands as a relevant benchmark when it comes to assessing the performance of a hedging strategy.

Another approach is super-replication (e.g. El Karoui & Quenez, 1995, and Karatzas 1997). It identifies the cheapest trading strategy whose terminal wealth is at least equal to the derivative's payoff. Since the option buyer alone carries the price of the hedging risk, the initial capital required is often unacceptably large. Eberlein & Jacod (1997) show that, under many models, the initial capital required to super-replicate a call option is the price of the underlying asset itself.

An alternative to super-replication is Global Hedging Risk Minimization (GHRM), which consists in identifying trading strategies that replicate the derivative's payoff as closely as possible, or alternatively, minimize the risk associated with terminal hedging shortfalls. Xu (2006) proposes to minimize general risk measures applied to hedging errors. Several authors choose more specific risk measures: quantiles of the hedging shortfall (Föllmer & Leukert, 1999, Cvitanić & Spivak, 1999), expected hedging shortfall (Cvitanić & Karatzas, 1999), expected powers of the hedging shortfall (Pham, 2000), Tail Value-at-Risk (Sekine, 2004),

¹By contrast, local risk-minimization, which considers hedging strategies that are not self-financing, selects one that minimizes a measure of the costs related to non-initial investments in the portfolio (Schweizer, 1991).

expected squared hedging error (Schweizer, 1995, Motoczyński, 2000, Cont et al., 2007 and Rémillard & Rubenthaler, 2009) and the expectation of general loss functions (Föllmer & Leukert, 2000). Theoretical existence of optimal hedging strategies under those risk measures and their characterization are studied in a general context. However, explicit solutions exist only for some particular cases of market setups and risk measures. The implementation of the preceding methodologies in the case of incomplete markets is often not straightforward, and tractable algorithms computing the optimal strategies have yet to be identified. The presence of regimes adds an additional layer of difficulty in applying those methods.

This paper's contributions are twofold. First, on a theoretical level, we develop a discrete-time hedging methodology with the GHRM objective that minimizes the expected value of any desired penalty function of the hedging error within a general regime-switching framework (possibly including time-inhomogeneous regime shifts). This methodology is highly flexible and generalizes the quadratic hedging approach. It incorporates a large class of penalty functions encompassing usual risk measures such as Value-at-Risk and expected shortfall. The proposed framework can accommodate portfolio restrictions such as no short-selling. Portfolios can be rebalanced more frequently than the regime-switch timeframe. Second, from an implementation perspective, a numerical algorithm based on backward recursion allows for the sequential construction of an optimal hedging strategy. Numerical experiments challenge our model with existing methodologies. The relative expected penalty reduction obtained with this paper's optimal hedging approach, in comparison with the best benchmark, ranges between 0.9% and 12.6% in the different cases exposed.

This paper is organized as follows. In Section 2, the market model and the hedging problem are described. In Section 3, the hedging problem is solved. Section 4 presents a numerical scheme to compute the solution to the hedging problem. Section 5 presents the market model used for the simulations and provides numerical results. Section 6 concludes the paper.

2 Market specifications and hedging

2.1 Description of the market

Transactions take place in a discrete-time, arbitrage-free financial market. Denote by Δ_t the constant time elapsing between two consecutive observations. Two types of assets are traded. The risk-free asset is a position in the money market account with a nominal amount normalized to one monetary unit. The time- n price of the risk-free asset is

$$S_n^{(1)} = \exp(rn\Delta_t), \quad n \in \{0, 1, 2, \dots\}$$

where r is the annualized risk-free rate. The price of the risky asset, starting at $S_0^{(2)}$, evolves according to

$$S_n^{(2)} = S_0^{(2)} \exp(Y_n),$$

where Y_n is the risky asset's cumulative return over the time interval $[0, n]$. \vec{S}_n denotes the column vector $(S_n^{(1)}, S_n^{(2)})^\top$ and $\vec{S}_{0:n}$ stands for the whole price process up to time n .

The financial market is subject to various regimes that affect the dynamics of the risky asset's price. These regimes are represented by an integer-valued process $\{h_n\}_{n=0}^N$ taking values in $\mathcal{H} = \{1, 2, \dots, H\}$ where h_n is the regime prevailing during time interval $]n, n + 1]$. The joint process (Y, h) has the Markov property² with respect to the filtration $\{\mathcal{F}_n\}_{n=0}^N$ satisfying the usual conditions, where

$$\mathcal{F}_n = \sigma(\vec{S}_{0:n}, h_{0:n}) = \sigma(Y_{0:n}, h_{0:n}),$$

meaning that the distribution of (Y_{n+1}, h_{n+1}) conditional on information \mathcal{F}_n is entirely determined by Y_n and h_n .³ This assumption is consistent with Hamilton (1989) and Duan et al. (2002), among others. The transition probabilities of the regime process h are denoted by

$$P_{i,j}^{(n)}(y) = \mathbb{P}(h_{n+1} = j | h_n = i, Y_n = y) \quad i, j \in \mathcal{H}.$$

Because regimes h are not observable, a coarser filtration $\{\mathcal{G}_n\}_{n=0}^N$ modelling the information available to investors is required, that is, $\mathcal{G}_n = \sigma(Y_{0:n})$.

²A stochastic process $\{X_n\}$ has the Markov property with respect to filtration \mathcal{F} if $\forall n, x$,

$$\mathbb{P}(X_{n+1} \leq x | \mathcal{F}_n) = \mathbb{P}(X_{n+1} \leq x | X_n).$$

³Equivalently, the process (\vec{S}, h) has the Markov property with respect to filtration \mathcal{F} .

2.2 The hedging problem

A market participant (referred to as the “hedger”) wishes to replicate (or “hedge”) the payoff $\phi(S_N^{(2)})$ of a European contingent claim written on the risky asset and maturing at time N , where $\phi(\cdot)$ is some positive Borel function $\phi : [0, \infty) \rightarrow \mathbb{R}$. Alternatively, the payoff can be written as a function of the risky asset return

$$\phi(S_N^{(2)}) = \tilde{\phi}(Y_N),$$

for some function $\tilde{\phi}(\cdot)$.

To implement the replication, the hedger adopts \mathcal{G} -predictable self-financing⁴ hedging strategies $\theta = \{\vec{\theta}_n\}_{n=1}^N$ with time- n value⁵ $V_n(v_0, Y_{0:n}, \vec{\theta}_{1:n}) := \vec{\theta}_n^\top \vec{S}_n$ and initial value $V_0 := v_0 = \vec{\theta}_1^\top \vec{S}_0$. This ensures that all trading decisions are made based on up-to-date price information, regardless of the unobserved regime. Below, $\theta_n^{(k)}$ represents the number of shares of asset k held during period $]n-1, n]$ and $\vec{\theta}_n$ is the column vector $(\theta_n^{(1)}, \theta_n^{(2)})^\top$ that characterizes the hedging portfolio.

Definition 2.1 *The set of all \mathcal{G} -predictable self-financing hedging strategies satisfying possible additional requirements (such as no short-selling constraints⁶) is denoted by Θ . We refer to Θ as the set of admissible hedging strategies.*

Unobservable regimes and discrete-time trading make perfect replication of the European contingent claim impossible to achieve. The hedger therefore aims to best replicate the payoff $\tilde{\phi}(Y_N)$ according to a certain metric. This justifies the use of a penalty function that sanctions departure of the hedging portfolio’s terminal value V_N from $\phi(S_N^{(2)})$. Let $g(\cdot)$ be a Borel function $g : \mathbb{R} \rightarrow \mathbb{R}$ representing a penalty function. For a given amount of initial wealth v_0 , the hedger wishes to find an admissible hedging strategy solving

$$\min_{\theta \in \Theta} \mathbb{E} \left[g(\phi(S_N^{(2)}) - V_N) \right]. \quad (1)$$

The solution is referred to as the optimal hedging strategy. Admittedly, g, ϕ, θ and $S^{(2)}$ need to be well-behaved and integrable enough for this expectation to exist.

⁴ $\theta = \{\vec{\theta}_n\}_{n=1}^N$ is a self-financing hedging strategy if $\forall n \geq 1, \vec{\theta}_n^\top \vec{S}_n = \vec{\theta}_{n+1}^\top \vec{S}_n$.

⁵To ease notation, $V_n(v_0, Y_{0:n}, \vec{\theta}_{1:n})$ is denoted by V_n .

⁶Or a weaker version of it asking for V_n to be positive.

Defining the hedging problem at the terminal date does not require a pricing function for the derivatives, and in particular a characterization of the risk premium. By contrast, hedging strategies considering intermediate dates (option tracking) rely on additional assumptions about the martingale measure.

Schweizer (1995) and Rémillard & Rubenthaler (2009) work with the quadratic penalty function $g(x) = x^2$. However, this specification entails that gains and losses on the hedge are penalized equally. In practice, the hedger might be interested in treating gains and losses on the hedge differently. Among asymmetric penalty functions, Pham (2000) investigates the case $g(x) = x^p \mathbf{1}_{\{x>0\}}$ for a positive constant p , where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator variable. Another possibility is to choose $g(x) = \mathbf{1}_{\{x \geq z\}}$ where z is a constant. Such a penalty function induces the minimization of the probability that the hedging shortfall is greater than z . Föllmer & Leukert (1999) and Cvitanić & Spivak (1999) study the hedging problem in continuous time with a similar hedging goal. In this paper, we opt for a general asymmetric penalty function of the form

$$g(x) = \alpha_1 |x|^p \mathbf{1}_{\{x \leq \gamma_1\}} + \alpha_2 |x|^q \mathbf{1}_{\{x > \gamma_2\}}, \quad (2)$$

for some constants $\alpha_1, \alpha_2, \gamma_1, \gamma_2, p \geq 0$ and $q \geq 0$. This specification encompasses both symmetric and asymmetric penalties and allows different penalty weights to be put on the under- and over-replication of the terminal payoff. If $q = \alpha_1 = 0$ and $\alpha_2 = 1$, the penalty reduces to a Value-at-Risk type of measure. If $q = \alpha_2 = 1$ and $\alpha_1 = 0$, the penalty becomes an Expected shortfall type of measure. The case $p = q = 2, \alpha_1 = \alpha_2 = 1$ and $\gamma_1 = \gamma_2 = 0$ leads to the quadratic penalty.

3 Solving the hedging problem

3.1 From path-dependence to the Markov property

The tools of dynamic programming and the Bellman equation are tailor-made to solve problems of the Equation (1) type if one can invoke the Markov property for the state variables process. However, the observable process Y does not necessarily have the Markov property with respect to the filtration \mathcal{G} , because the cumulative returns depend on the regimes. Indeed, all past values of the cumulative returns path Y give information about the current

value of the unobservable regime h . This obstacle is circumvented by defining additional state variables that summarize all the relevant information of Y 's previous path. Those variables allow for the definition of a process that has the Markov property with respect to information flow \mathcal{G} .

Below, $f_{\vec{X}}(\vec{x})$ denotes the joint probability density function (pdf) of a random vector \vec{X} . In some cases, if some components of \vec{X} are discrete-type random variables, $f_{\vec{X}}(\vec{x})$ is a mixed pdf. Similarly, $f_{\vec{X}|\vec{Y}}(\vec{x}|\vec{y})$ denotes the pdf of \vec{X} conditional upon $\vec{Y} = \vec{y}$. All proofs are provided in Appendix A.

Definition 3.1 *The conditional probability $\eta_{i,n}$ of being in regime i at time n given the cumulative returns $Y_{0:n}$ is the \mathcal{G}_n -measurable function*

$$\eta_{i,n} := \mathbb{P}(h_n = i | \mathcal{G}_n) = f_{h_n | Y_{0:n}}(i | Y_{0:n}), \quad i \in \mathcal{H}.$$

As a special case, $\eta_{i,0} = \mathbb{P}(h_0 = i) = f_{h_0}(i)$. The \mathcal{G}_n -measurable vector $\vec{\eta}_n = (\eta_{1,n}, \dots, \eta_{H,n})$ denotes the set of conditional probabilities at time n .

Those η are the state variables required in the construction of a Markov process with respect to filtration \mathcal{G} . Theorem 3.1 provides a recursion formula allowing for an efficient computation of those probabilities.⁷

Theorem 3.1 *The conditional probabilities are given recursively by*

$$\eta_{i,n+1} = \frac{\sum_{j=1}^H f_{h_{n+1}, Y_{n+1} | h_n, Y_n}(i, Y_{n+1} | j, Y_n) \eta_{j,n}}{\sum_{j=1}^H \sum_{\ell=1}^H f_{h_{n+1}, Y_{n+1} | h_n, Y_n}(j, Y_{n+1} | \ell, Y_n) \eta_{\ell,n}}.$$

Moreover, if Y_{n+1} and h_{n+1} are conditionally independent upon \mathcal{F}_n , then

$$f_{h_{n+1}, Y_{n+1} | h_n, Y_{0:n}}(i, Y_{n+1} | j, Y_{0:n}) = P_{j,i}^{(n)}(Y_n) f_{Y_{n+1} | h_n, Y_{0:n}}(Y_{n+1} | j, Y_{0:n}).$$

Corollary 3.1 states that those conditional probabilities are the natural extension for the cumulative returns to retrieve the Markov property.

Corollary 3.1 $\{Y_n, \vec{\eta}_n\}_{n=0}^N$ has the Markov property with respect to \mathcal{G} .

⁷An alternative recursion formula is presented in Rémillard et al. (2010a). However, the current formula is preferred for two main reasons. First, $\eta_{i,n}$ lying in $[0, 1]$ makes it numerically more stable. Second, it benefits from a dimension reduction since $\eta_{H,n} = 1 - \sum_{j=1}^{H-1} \eta_{H,j}$.

Finally, the next corollary extends the previous one to include the hedging portfolio value. In the general case of predictable hedging strategies, this inclusion unfortunately destroys the Markov property. However, if asset reallocation is solely determined by the information about current cumulative return and portfolio value as well as the recursive conditional probabilities (as defined in Theorem 3.1), then the Markov property can be retrieved. This property is crucial, from a numerical point of view, to obtaining an implementable algorithm.

Corollary 3.2 *For any admissible hedging strategy $\theta \in \Theta$, the conditional distribution of $(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1})$ given \mathcal{G}_n is the same as if it is conditioned upon $\sigma(Y_n, \vec{\eta}_n, V_n, \vec{\theta}_{n+1})$. Moreover, if the condition that $\vec{\theta}_{n+1}$ is $\sigma(Y_n, \vec{\eta}_n, V_n)$ -measurable for any n is added, then $\{Y_n, \vec{\eta}_n, V_n\}_{n=0}^N$ has the Markov property with respect to \mathcal{G} .*

3.2 A recursive construction

In this section, an optimal hedging strategy is constructed. Let Ψ_N^* be the hedging penalty at time N ,

$$\Psi_N^* := g\left(\tilde{\phi}(Y_N) - V_N\right) \quad (3)$$

and for any $n \in \{0, 1, \dots, N-1\}$, let Ψ_n^* be the smallest possible expected hedging penalty

$$\Psi_n^* := \min_{\vec{\theta}_{n+1:N}} \mathbb{E}[\Psi_N^* | \mathcal{G}_n] \quad (4)$$

where $\vec{\theta}_{n:N} = (\vec{\theta}_n, \dots, \vec{\theta}_N)$.

Remark 3.1 *One assumes sufficient regularity in g , ϕ and the distribution of $\{Y_n\}_{n=0}^N$ such that, for all n , the minimum in (4) is attained.*

Equation (4) is stated as a minimization over $N - n$ portfolio vectors. Theorem 3.2 presents a way to optimize these portfolios one at a time.

Theorem 3.2 *For any $n \in \{0, 1, \dots, N-1\}$, the smallest expected penalty at time n may be computed using a recursive argument:*

$$\Psi_n^* = \min_{\vec{\theta}_{n+1}} \mathbb{E}[\Psi_{n+1}^* | \mathcal{G}_n]. \quad (5)$$

Furthermore, let $\vec{\theta}_{(n+2):N}^*$ denote one of the possible admissible hedging strategies that minimize the expected penalty at time $n + 1$, that is,

$$\vec{\theta}_{(n+2):N}^* = \arg \min_{\vec{\theta}_{n+2:N}} \mathbb{E}[\Psi_N^* | \mathcal{G}_{n+1}].$$

Then,

$$\vec{\theta}_{(n+1):N}^* := \left(\arg \min_{\vec{\theta}_{n+1}} \mathbb{E} [\Psi_{n+1}^* | \mathcal{G}_n], \vec{\theta}_{(n+2):N}^* \right), \quad (6)$$

is a solution to the following equation:

$$\vec{\theta}_{(n+1):N}^* = \arg \min_{\vec{\theta}_{n+1:N}} \mathbb{E}[\Psi_N^* | \mathcal{G}_n].$$

This means that the optimal admissible hedging strategy may be built up using a backward induction construction.

Equations (5) and (6) involve conditional expectations with respect to all past return realizations. Theorem 3.3 shows that it is possible to remove path-dependence and appeal only to conditional expectations with respect to the current state variables $\{Y_n, \vec{\eta}_n, V_n\}_{n=0}^N$.

Theorem 3.3 *Assume that for all n , constraints on the portfolio $\vec{\theta}_{n+1}$ depend only on the value of $(Y_n, \vec{\eta}_n, V_n)$. Then, $\forall n \leq N$, Ψ_n^* is $\sigma(Y_n, \vec{\eta}_n, V_n)$ -measurable. Moreover, there exists an optimal self-financing hedging strategy $\{\vec{\theta}_n^*\}$ that solves (1) such that $\forall n \geq 1$, $\vec{\theta}_{n+1}^*$ is $\sigma(Y_n, \vec{\eta}_n, V_n)$ -measurable. Furthermore,*

$$\vec{\theta}_{n+1}^* = \arg \min_{\vec{\theta}_{n+1} \in \sigma(Y_n, \vec{\eta}_n, V_n)} \mathbb{E} [\Psi_{n+1}^* | Y_n, \vec{\eta}_n, V_n]. \quad (7)$$

Since Ψ_n^* is $\sigma(Y_n, \vec{\eta}_n, V_n)$ -measurable, one can write $\Psi_n^* = \Psi_n(Y_n, \vec{\eta}_n, V_n)$. Finally, the next theorem combines Theorems 3.2 and 3.3 to optimize one portfolio vector at a time, searching on the space of hedging strategies for which $\{Y_n, \vec{\eta}_n, V_n\}_{n=0}^N$ has the Markov property with respect to \mathcal{G} . These two features make the algorithm numerically tractable.

Theorem 3.4 The Bellman Equation *There exists a self-financing hedging strategy $\{\vec{\theta}_n^*\}$ that solves problem (1) and the following set of recursive equations:*

$$\forall n, \vec{\theta}_{n+1}^* = \arg \min_{\vec{\theta}_{n+1} \in \sigma(Y_n, \vec{\eta}_n, V_n)} \mathbb{E} \left[\Psi_{n+1}(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1}(\vec{\theta}_{n+1})) | Y_n, \vec{\eta}_n, V_n \right].$$

Furthermore, the minimal expected penalty can be computed as follows:

$$\Psi_N(Y_N, \vec{\eta}_N, V_N) = g(\phi(S_N^{(2)}) - V_N) = g(\tilde{\phi}(Y_N) - V_N) \quad (8)$$

$$\Psi_n(Y_n, \vec{\eta}_n, V_n) = \min_{\vec{\theta}_{n+1}} \mathbb{E} \left[\Psi_{n+1}(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1}(\vec{\theta}_{n+1})) | Y_n, \vec{\eta}_n, V_n \right], \quad n \in \{0, 1, \dots, N-1\} \quad (9)$$

Finally, $\min_{\{\vec{\theta}_n\} \in \Theta} \mathbb{E} \left[g(\phi(S_N^{(2)}) - V_N(\vec{\theta}_{1:N})) \right] = \Psi_0(Y_0, \vec{\eta}_0, V_0)$.

The proof of Theorem 3.4 is a direct consequence of Theorems 3.2 and 3.3 and the definition of Ψ_n .

4 Lattice implementation

Analytical solutions to Theorem 3.4's equations are unlikely to be found for general penalties. Therefore, numerical approximations must be considered in order to implement the algorithm. The numerical application of the hedging algorithm is discussed in this section.

4.1 Dimensionality reduction

Since $\sum_j^H \eta_{j,n} = 1$, the variable $\eta_{H,n}$ provides no additional information. Therefore, $\vec{\eta}_n = (\eta_{1,n}, \dots, \eta_{H,n})$ can be replaced with $\vec{\eta}_n := (\eta_{1,n}, \dots, \eta_{H-1,n})$ in Theorem 3.4. This reduces the dimension of the problem, which is an important numerical issue. Similarly, since for self-financing strategies $\sum_{k=1}^2 \theta_{n+1}^{(k)} S_n^{(k)} = V_n$, the optimization over $\vec{\theta}_{n+1}$ is in fact equivalent to optimizing only over $\theta_{n+1} := \theta_{n+1}^{(2)}$.

4.2 Grid values

To compute the minimal expected penalty Ψ_n and optimal portfolio position $\vec{\theta}_{n+1}$ from Theorem 3.4, one resorts to a grid whose nodes correspond to a discrete subsample of all possible values of (Y_n, η_n, V_n) . For each state variable, the largest and smallest values in the grid must be set. One can use the $[0, 1]$ bounds for $\vec{\eta}$ since it contains probabilities. Variables V_n and Y_n are unbounded. Therefore, grid bounds for V_n and Y_n are found numerically using a Monte-Carlo simulation. To this end, 10^5 sample paths of cumulative returns $Y_{0:N}$ are simulated. This yields the approximate distribution of Y_n for all n . The case of the portfolio value V_n is different since the optimal hedging strategy is not yet known. However, a proxy

$V_n^{(BS)}$ is built for V_n using the Black-Scholes delta hedging as described in Section 5.4.1. Let $Y_{n,\alpha}$, and $V_{n,\alpha}^{(BS)}$ be the α^{th} sample quantiles. Define

$$Y_{n,mid} := \frac{1}{2}(Y_{n,0.25\%} + Y_{n,99.75\%}) \text{ and } V_{n,mid}^{(BS)} := \frac{1}{2}(V_{n,0.25\%}^{(BS)} + V_{n,99.75\%}^{(BS)})$$

as the mid-points of two extreme quantiles. The largest and smallest values for the grid at time n are chosen to be

$$\begin{aligned} Y_n^{(small)} &:= (1 + \lambda_Y^{(small)})(Y_{n,0.25\%} - Y_{n,mid}) + Y_{n,mid} \\ Y_n^{(large)} &:= (1 + \lambda_Y^{(large)})(Y_{n,99.75\%} - Y_{n,mid}) + Y_{n,mid} \\ V_n^{(small)} &:= (1 + \lambda_V^{(small)})(V_{n,0.25\%}^{(BS)} - V_{n,mid}^{(BS)}) + V_{n,mid}^{(BS)} \\ V_n^{(large)} &:= (1 + \lambda_V^{(large)})(V_{n,99.75\%}^{(BS)} - V_{n,mid}^{(BS)}) + V_{n,mid}^{(BS)}. \end{aligned}$$

where $(\lambda_Y^{(small)}, \lambda_Y^{(large)}, \lambda_V^{(small)}, \lambda_V^{(large)})$ are positive stretching factors.

4.3 Algorithm solving the Bellman equation

A numerical algorithm allowing for the computation of the minimal expected penalty and the optimal portfolio position at each time step is given in this section. First, define two grids of different sizes (one finer and one coarser) containing a discrete subset of values for $(Y_n, \vec{\eta}_n, V_n)$.

4.3.1 On the coarse grid

Assume that $(Y_n, \vec{\eta}_n, V_n) = (y, \vec{\eta}, v)$. According to Theorem 3.4, the goal is to evaluate Equation (9) at each node $(y, \vec{\eta}, v)$ of the grid:

$$\Psi_n^{y, \vec{\eta}, v} = \min_{\vec{\theta}_{n+1}} \mathbb{E} \left[\Psi_{n+1} \left(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1}(\vec{\theta}_{n+1}) \right) \mid (Y_n, \vec{\eta}_n, V_n) = (y, \vec{\eta}, v) \right].$$

From Theorem 3.1, $\vec{\eta}_{n+1}$ is a function of $(Y_{n+1}, Y_n, \vec{\eta}_n)$. Seen from node $(y, \vec{\eta}, v)$, it may be denoted $\vec{\eta}_{n+1}^{y, \vec{\eta}}(Y_{n+1})$. Because the amount invested in the riskless asset is the value of the portfolio minus the investment in the risky asset, the time- $(n+1)$ value of the hedging portfolio, seen from the grid point $(Y_n, \vec{\eta}_n, V_n) = (y, \vec{\eta}, v)$, is

$$\begin{aligned} V_{n+1}(\vec{\theta}_{n+1}) &= \theta_{n+1}^{(1)} \exp(r(n+1)\Delta t) + \theta_{n+1}^{(2)} S_0^{(2)} \exp(Y_{n+1}) \\ &= \exp(r\Delta t) \left(v - \theta_{n+1}^{(2)} S_0^{(2)} \exp(y) \right) + \theta_{n+1}^{(2)} S_0^{(2)} \exp(Y_{n+1}) \\ &= V_{n+1}^{y, v}(\vec{\theta}_{n+1}, Y_{n+1}). \end{aligned}$$

Therefore, the expected penalty at time n and at grid point $(y, \vec{\eta}, v)$ satisfies

$$\begin{aligned}
\Psi_n^{y, \vec{\eta}, v} &= \min_{\vec{\theta}_{n+1}} \mathbb{E} \left[\Psi_{n+1} \left(Y_{n+1}, \vec{\eta}_{n+1}^{y, \vec{\eta}}(Y_{n+1}), V_{n+1}^{y, v}(\vec{\theta}_{n+1}, Y_{n+1}) \right) \middle| (Y_n, \vec{\eta}_n, V_n) = (y, \vec{\eta}, v) \right] \\
&= \min_{\vec{\theta}_{n+1}} \sum_{j=1}^H \eta_{j,n} \mathbb{E} \left[\Psi_{n+1} \left(Y_{n+1}, \vec{\eta}_{n+1}^{y, \vec{\eta}}(Y_{n+1}), V_{n+1}^{y, v}(\vec{\theta}_{n+1}, Y_{n+1}) \right) \middle| (h_n, Y_n, \vec{\eta}_n, V_n) = (j, y, \vec{\eta}, v) \right] \\
&\quad (\text{from Equation (20)}) \\
&= \min_{\vec{\theta}_{n+1}} \sum_{j=1}^H \eta_{j,n} \int_{-\infty}^{\infty} \Psi_{n+1}^{y, \vec{\eta}, v}(\vec{\theta}_{n+1}, z) f_{Y_{n+1}|Y_n, \vec{\eta}_n, V_n, h_n}(z|y, \vec{\eta}, v, j) dz \\
&= \min_{\vec{\theta}_{n+1}} \sum_{j=1}^H \eta_{j,n} \int_{-\infty}^{\infty} \Psi_{n+1}^{y, \vec{\eta}, v}(\vec{\theta}_{n+1}, z) f_{Y_{n+1}|Y_n, h_n}(z|y, j) dz \quad (\text{Markov property and Lemma A.1})
\end{aligned}$$

where

$$\Psi_{n+1}^{y, \vec{\eta}, v}(\vec{\theta}_{n+1}, z) = \Psi_{n+1} \left(z, \vec{\eta}_{n+1}^{y, \vec{\eta}}(z), V_{n+1}^{y, v}(\vec{\theta}_{n+1}, z) \right).$$

In general, there is no closed-form solution for this integral and it is evaluated numerically.

Therefore, the support of Y_{n+1} is partitioned in M intervals with boundaries $-\infty = z_0 < z_1 < \dots < z_{M-1} < z_M = \infty$ and $z_i^* \in [z_{i-1}, z_i]$ acts as a representative of the interval $[z_{i-1}, z_i]$.

$$\begin{aligned}
&\int_{-\infty}^{\infty} \Psi_{n+1}^{y, \vec{\eta}, v}(\vec{\theta}_{n+1}, z) f_{Y_{n+1}|Y_n, h_n}(z|y, j) dz \\
&= \sum_{i=1}^M \int_{z_{i-1}}^{z_i} \Psi_{n+1}^{y, \vec{\eta}, v}(\vec{\theta}_{n+1}, z) f_{Y_{n+1}|Y_n, h_n}(z|y, j) dz \\
&\cong \sum_{i=1}^M \Psi_{n+1}^{y, \vec{\eta}, v}(\vec{\theta}_{n+1}, z_i^*) \int_{z_{i-1}}^{z_i} f_{Y_{n+1}|Y_n, h_n}(z|y, j) dz \\
&= \sum_{i=1}^M \Psi_{n+1}^{y, \vec{\eta}, v}(\vec{\theta}_{n+1}, z_i^*) \omega_i^{y, j, n}
\end{aligned} \tag{10}$$

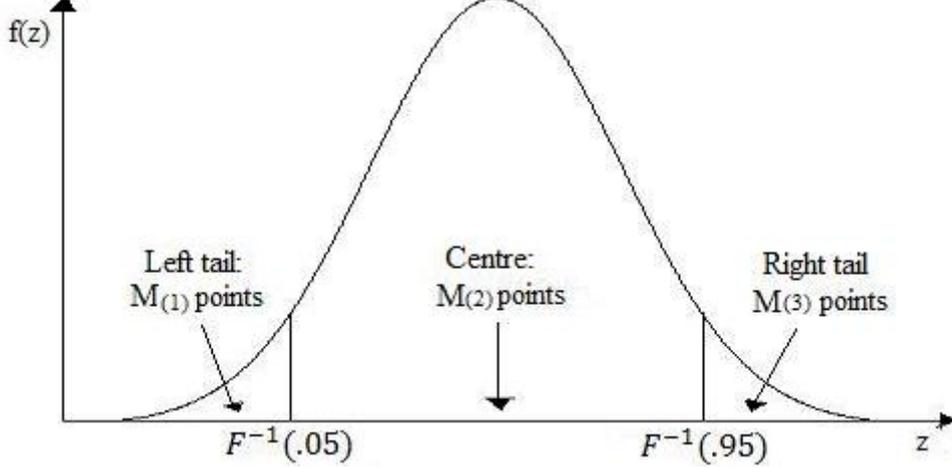
where the weights $\omega_i^{y, j, n}$ are

$$\omega_i^{y, j, n} = F_{Y_{n+1}|Y_n, h_n}(z_i|y, j) - F_{Y_{n+1}|Y_n, h_n}(z_{i-1}|y, j),$$

$F_{Y_{n+1}|Y_n, h_n}$ being the cumulative distribution function of Y_{n+1} given (Y_n, h_n) . In general, the approximation (10) is good if the distances between the z_i are small and the Ψ_{n+1} function is relatively smooth. The z_i are chosen to be quantiles of the conditional distribution $F_{Y_{n+1}|Y_n, h_n}$. To better capture the impact of extreme events, particular attention is paid to the tails of the distribution. The left (right) tail is defined as the smallest (largest) 5% values of the distribution. The $M_{(1)}$ smallest z_i 's correspond to quantiles of level $k \frac{5\%}{M_{(1)}}$, $k \in \{1, 2, \dots, M_{(1)}\}$. The central part of the distribution is proxied by $M_{(2)}$ quantiles of level $k \frac{90\%}{M_{(2)}} + 5\%$, $k \in \{1, 2, \dots, M_{(2)}\}$, while the right tail is represented by $M_{(3)}$ quantiles whose

level lies in $]95\%, 100\%]$. Consequently, the weights $\omega_i^{y,j,n}$ are $\frac{5\%}{M_{(1)}}$, $\frac{90\%}{M_{(2)}}$ or $\frac{5\%}{M_{(3)}}$ depending on which part of the distribution z_i belongs to. Among possible specifications, z_i^* are chosen as quantiles whose level is the mean between the levels of z_{i-1} and z_i . This is illustrated by Figure 1.

Figure 1: Quadrature illustration



This figure illustrates the quadrature for the log-return distribution.

Because the maximization is time-consuming, especially if it must be done at all nodes of the lattice, the research area is reduced to a discrete set \mathcal{O} of values:

$$\Psi_n(y, \vec{\eta}, v) \cong \min_{\vec{\theta}_{n+1} \in \mathcal{O}} \sum_{i=1}^M \Psi_{n+1}^{y, \vec{\eta}, v}(\vec{\theta}_{n+1}, z_i^*) \omega_i^{y, \vec{\eta}, n}. \quad (11)$$

Since the backward induction on time leads to a numerical approximation $\widehat{\Psi}_{n+1}$ of Ψ_{n+1} , the latter is replaced by former in Equation (11) in applications.

Step 1: Rough estimate of optimal hedging strategy

A rough estimate of the optimal hedging strategy is

$$\hat{\theta}_{n+1}^{y, \vec{\eta}, v} = \arg \min_{\theta_{n+1} \in \mathcal{O}} \sum_{i=1}^M \widehat{\Psi}_{n+1}^{y, \vec{\eta}, v}(\theta_{n+1}, z_i^*) \omega_i^{y, \vec{\eta}, n}.$$

By construction, the z_i^* do not match the grid's discretization of next period return Y_{n+1} . For this reason, interpolation is required to evaluate each of the $\widehat{\Psi}_{n+1}^{y, \vec{\eta}, v}(\theta_{n+1}, z_i^*)$ whose arguments most likely lie between the grid nodes. This step proceeds with multivariate linear interpolation.⁸

⁸This approximation of $\widehat{\Psi}_{n+1}$ is not involved in further iterations. Therefore, while high precision is not a crucial issue at this step, computational speed is.

4.3.2 On the finer grid

Step 2 : Smoothing of the hedging strategy

From step 1, one gets an approximate portfolio position $\hat{\theta}_{n+1}^*$ for every node of the coarse grid at time n . For every value of $(y, \vec{\eta}, v)$ on the finer grid, one computes the hedging portfolio position $\hat{\vartheta}_{n+1}^{y, \vec{\eta}, v}$ using smoothing splines based on $\hat{\theta}_{n+1}^*$. $\hat{\vartheta}_{n+1}$ is now used as the final estimation of the optimal hedging portfolio position.

Step 3 : Recalculation of the value function

A finer partition of the distribution of Y_{n+1} and the corresponding weights, denoted by \tilde{z}_i^* and $\tilde{\omega}_i$, $i = 1, \dots, \tilde{M}$, serve for the approximation of the minimal expected penalty function with the new portfolio position $\hat{\vartheta}_{n+1}$. Thus, mimicking Equation (11),

$$\hat{\Psi}_n(y, \vec{\eta}, v) = \sum_{i=1}^{\tilde{M}} \hat{\Psi}_{n+1}^{y, \vec{\eta}, v} \left(\hat{\vartheta}_{n+1}^{y, \vec{\eta}, v}, \tilde{z}_i^* \right) \tilde{\omega}_i^{y, \vec{\eta}, n}.$$

The subsequent iteration of the three-step algorithm will call this new approximation for $\hat{\Psi}_n$. Thus, to minimize the accumulation of errors, the interpolation is performed with natural splines.⁹

5 Numerical results

5.1 The model

As in Hamilton (1989), the regime process is assumed to be a Markov chain, implying that the conditional distribution of h_{n+1} given \mathcal{F}_n is the same as if it were conditioned upon h_n . The model can accommodate a regime shift timeframe which is coarser than the rebalancing schedule. In that context, τ represents the number of periods between two possible regime transitions and $\{h_n\}_{n=0}^N$ becomes a time inhomogeneous Markov chain with probability transition matrix

$$P^{(n)}(y) = \begin{cases} P & \text{if } (n+1) \bmod \tau = 0 \\ I_{H \times H} & \text{otherwise,} \end{cases}$$

where $I_{H \times H}$ is the identity matrix.

⁹Natural splines in three dimensions are implemented through the *interp3* matlab function.

A basic model based on two regimes ($H = 2$) serves as benchmark to test the proposed algorithm. Conditioned on the actual regime $h_n = i$, the one-period log-return $\epsilon_{n+1} = Y_{n+1} - Y_n$ has a Gaussian distribution with mean $\mu_i \Delta_t$ and variance $\sigma_i^2 \Delta_t$.

The application of Theorem 3.4 relies on the following relations:

$$\begin{aligned} Y_{n+1} &= Y_n + \epsilon_{n+1} \\ V_{n+1} &= V_n e^{r\Delta_t} + \theta_{n+1} S_0 e^{Y_n} (e^{\epsilon_{n+1}} - e^{r\Delta_t}) \\ \eta_{1,n+1} &= \frac{\sum_{j=1}^H P_{j,1}^{(n)} \eta_{j,n} f_{\epsilon_{n+1}|h_n}(\epsilon_{n+1}|j)}{\sum_{u=1}^H \sum_{j=1}^H P_{j,u}^{(n)} \eta_{j,n} f_{\epsilon_{n+1}|h_n}(\epsilon_{n+1}|j)}, \end{aligned}$$

where $\eta_{2,n} = 1 - \eta_{1,n}$ and $f_{\epsilon_{n+1}|h_n}(\epsilon_{n+1}|j)$ is the Gaussian density function

$$f_{\epsilon_{n+1}|h_n}(\epsilon_{n+1}|j) = \frac{1}{\sqrt{2\pi\Delta_t}\sigma_j} \exp\left(-\frac{1}{2} \frac{(\epsilon_{n+1} - \mu_j \Delta_t)^2}{\sigma_j^2 \Delta_t}\right). \quad (12)$$

The conditional distribution of $\epsilon_{n+1}|(Y_n, \eta_n, V_n)$ is a mixture of two Gaussian distributions:

$$\begin{aligned} \mathbb{P}(\epsilon_{n+1} \leq x | Y_n, \eta_n, V_n) &= \mathbb{P}(\epsilon_{n+1} \leq x | h_n = 1) \eta_{1,n} + \mathbb{P}(\epsilon_{n+1} \leq x | h_n = 2) (1 - \eta_{1,n}) \\ &= \Phi\left(\frac{x - \mu_1}{\sigma_1}\right) \eta_{1,n} + \Phi\left(\frac{x - \mu_2}{\sigma_2}\right) (1 - \eta_{1,n}), \end{aligned}$$

where Φ is the standard normal cumulative distribution function.

Moreover, the following boundaries can be used for η in the algorithm of Section 4:

Proposition 5.1 *For all j, n , $\min_{i \in \mathcal{H}} P_{i,j}^{(n)} \leq \eta_{j,n+1} \leq \max_{i \in \mathcal{H}} P_{i,j}^{(n)}$.*

The proof is in Appendix B.

5.2 Estimation

Regime switches potentially occur each week and rebalancing is performed weekly ($\Delta_t = 1/52, \tau = 1$) or daily ($\Delta_t = 1/260, \tau = 5$). Maximum likelihood with the EM algorithm of Dempster et al. (1997) is applied to a sample of S&P 500 weekly log-returns from January 1, 2000 to December 31, 2010. Parameter estimates are reported in Table 1.

A p -value of 34.4% for the Cramer-Von-Mises parametric bootstrap goodness-of-fit test (see Genest & Rémillard, 2008) for the regime-switching process indicates that the model is not rejected. The first (second) regime represents an economy in expansion (recession): returns exhibit a positive (negative) mean with a low (high) volatility. The risk-free rate is set to $r = 2\%$.

Table 1: Estimated parameters of the Gaussian regime switching model

Parameter	Regime 1	Regime 2
μ_j	.0718	-.2884
σ_j	.1283	.3349
$P_{j,j}$.9736	.9091

The annualized estimated parameters of the model of Section 5 are presented. A time series from January 1, 2000 to December 31, 2010 of the weekly price of the S&P 500 is used. Estimation is performed with the EM algorithm.

5.3 Hedging strategies

The option to be hedged is a European at-the-money call option with payoff $\phi(S_N) = \max(0, S_N - E)$. The initial index value is $S_0 = 1,257.64$, which is the value of the S&P 500 on December 31, 2010. The option strike is $E = 1,257$. The maturity of the option is 12 weeks.¹⁰

The initial probability of being in regime 1 is set to $\eta_0 = 0.2318$. This value is chosen instead of the estimated value on the S&P 500 time series because it leads to the same call option price under the Black-Scholes and Hardy models (see Sections 5.4.1 and 5.4.2, respectively). Thus, both these hedging methodologies use the same initial capital, which makes the numerical results comparable. The initial hedging capital, which is the option price under those models, is $V_0 = 62.4316$.

The following penalty functions are under consideration:

$$g(x) = x^2 \quad \text{quadratic,} \quad (13)$$

$$g(x) = x^2 \mathbf{1}_{\{x>0\}} \quad \text{short quadratic,} \quad (14)$$

$$g(x) = x^2 \mathbf{1}_{\{x<0\}} \quad \text{long quadratic,} \quad (15)$$

where x represents the hedging error $\tilde{\phi}(Y_N) - V_N$. The quadratic penalty sanctions departures from the option payoff. The short (long) quadratic penalty is designed for the option seller (buyer), since it does not penalize profits; only losses are sanctioned.

¹⁰That is, $N = 60$ periods for daily rebalancing and $N = 12$ for weekly rebalancing.

The restrictions considered on the portfolio positions are that $\forall n, \theta_n \in [0, 1]$, thereby preventing short sales and excessive leverage.

5.4 Benchmarks

In order to compare the hedging model presented in this paper, benchmarks must be set. In the following, the optimal hedging strategy presented in Section 3 is referred to as "minimal expected penalty hedging" (MEPH).

The most common hedging strategy relies on delta hedging. In this case, a pricing kernel is required to compute the deltas. The first two benchmarks examine two pricing models.

5.4.1 Black-Scholes delta hedging (BSDH)

The classic Black-Scholes delta with a modified volatility determines the position held in the underlying asset:

$$\theta_{n+1}^{(\text{BSDH})} = \Phi \left(\frac{\log(S_n/E) + (r + .5\zeta^2)\Delta_t(N-n)}{\zeta\sqrt{\Delta_t(N-n)}} \right),$$

where ζ is the asymptotic stationary volatility of log-returns ϵ_n in the case $\tau = 1$:

$$\zeta = \sqrt{\left(\sum_{j=1}^H P_j^* (\sigma^{(j)^2} + \mu^{(j)^2}) \right) + \left(\sum_{j=1}^H P_j^* \mu^{(j)} \right)^2}, \quad (16)$$

P^* is the stationary distribution associated with the transition matrix P . In the case $\tau > 1$, the stationary distribution for the regimes does not exist in general because of the cyclical nature of the Markov chain transition probabilities. Nevertheless, Equation (16) is used as the presumed market volatility. The characterization of the hedging position is explicit and does not require a lattice approximation.

The initial capital used for hedging is the option price given by the Black-Scholes formula with the volatility given by (16). The Black-Scholes hedging methodology can be seen as a naive benchmark that would be applied by a hedger who ignores the presence of regimes in the market.

5.4.2 Hardy delta hedging (HDH)

In Hardy (2001)'s two-regime model, the risk-neutral dynamics of one-period log-returns ϵ_{n+1} follow a mixture of Gaussian distributions. The delta hedging strategy commands that:¹¹

$$\theta_{n+1}^{(HDH)} = \sum_{j=1}^2 \eta_{j,n} \sum_{R=0}^{N-n} \Phi \left(\frac{\log(S_n/E) + (N-n)r\Delta_t + (R\sigma_1^2/2 + (N-n-R)\sigma_2^2/2)\Delta_t}{\sqrt{(R\sigma_1^2 + (N-n-R)\sigma_2^2)}} \right) f_n^i(R),$$

where $f_n^i(R)$ is the probability, given current regime i , that the number of periods between times n and N spent in the first regime is R . Probabilities $f_n^i(R)$ can be computed recursively (see Hardy, 2001). With this benchmark, the initial capital used for hedging is the option price. The hedger acknowledges the existence of regimes, but assigns an arbitrary risk premium to price options.

5.4.3 Forecast regime quadratic hedging (FRQH)

Besides delta hedging, Rémillard & Rubenthaler (2009) propose a global hedging risk minimization approach. The hedging strategy θ minimizes the expected terminal squared error of hedging with respect to complete information \mathcal{F} . This implies perfect knowledge of the current and all past regimes. Since in practice the states h are not observable, Rémillard et al. (2010a) forecast them with the most likely regime.

Let $\bar{\Theta}$ be the set of all \mathcal{F} -predictable self-financing strategies.¹² The FRQH strategy solves

$$\min_{\theta \in \bar{\Theta}} \mathbb{E} [(\phi(S_N) - V_N)^2].$$

With this benchmark, the hedging problem is based on the terminal date. Therefore, no assumption related to the risk premium is needed, which implies in particular that this strategy works with any initial capital. However, it comes at the price of using a lattice approach to compute the strategy. The hedger acknowledges the existence of regimes. However, the hedging objective is restricted to the quadratic penalty. Furthermore, the uncertainty surrounding regime forecasts is not taken into account.

¹¹Delta hedging under this model is investigated in Augustyniak & Boudreault (2012).

¹²By contrast, the MEPH strategy is \mathcal{G} -predictable.

5.5 Lattice parameters

The grid's stretching factors are $(\lambda_Y^{(small)}, \lambda_Y^{(large)}, \lambda_V^{(small)}, \lambda_V^{(large)}) = (.6, .6, 1, 1)$. For step 1 of the algorithm in Section 4.3, $M_{(1)} = M_{(3)} = 100$ and $M_{(2)} = 200$. For step 3, $\tilde{M}_{(1)} = \tilde{M}_{(3)} = 200$ and $\tilde{M}_{(2)} = 300$. Putting more points near the tails is used to better capture the extreme events which contribute more heavily to the hedging penalty. The discrete set \mathcal{O} over which the θ_n are optimized in step 1 of the algorithm is $\mathcal{O} = \{j/99 \mid j = 0, \dots, 99\}$.

The number of grid nodes for each variable on the finer grid (step 3) is:

$$(\#Y_n, \#\eta_n, \#V_n) = \begin{cases} (200, 100, 200) & \text{if } n = N - 1 \\ (150, 100, 150) & \text{otherwise} \end{cases} \quad (17)$$

More nodes are put on the first step of the recursion as it can be computed faster because of explicit formulas.¹³ For the coarse grid in step 1, only a subset of the nodes of the finer grid in step 3 are retained. The proportion of nodes kept in the coarse grid from the finer grid across dimensions Y_n , η_n and V_n is $1/3$, $1/3$ and $1/4$.

5.6 A simulation study

The numerical efficiency of the current paper's hedging algorithm is validated by means of Monte-Carlo simulations. The MEPH and FQRH strategies are implemented through a lattice. Hedging errors $\tilde{\phi}(Y_N) - V_N$ and hedging penalties $g(\tilde{\phi}(Y_N) - V_N)$ are computed for $I = 10^6$ simulated paths of the underlying returns.

Tables 2 and 3 report estimates of the expected penalty and their standard error for each hedging methodology. Note that only the MEPH strategy is affected by the choice of penalty function. For the three benchmarks, the hedging strategy remains the same, but the calculated penalty differs.

A first observation is that the MEPH grid estimate is relatively close to the simulated expected penalty. This confirms the accuracy of the numerical implementation.

In all six cases considered, the MEPH strategy significantly reduces the expected penalty. The magnitude of the penalty dispersion is comparable across all hedging strategies.

¹³An explicit expression for Ψ_{N-1} exists for the quadratic penalty. For the short (long) quadratic penalty, an explicit expression for $\mathbb{E}[\Psi_N^*(\theta_N) | \mathcal{G}_{N-1}]$ also exists. Details are available on request.

Table 2: Estimated expected penalties (weekly rebalancing)

	MEPH	BSDH	HDH	FRQH
Quadratic Penalty				
Grid Estimate	622.57	-	-	-
Expected Penalty	622.70	662.13	670.05	664.87
Standard Error	1.010	1.082	1.052	1.148
Short Quadratic Penalty				
Grid Estimate	325.15	-	-	-
Expected Penalty	325.39	372.16	392.02	374.75
Standard Error	0.9493	1.065	1.066	1.149
Long Quadratic Penalty				
Grid Estimate	268.11	-	-	-
Expected Penalty	267.35	289.97	278.03	290.12
Standard Error	0.4365	0.5043	0.4346	0.4640

This table reports estimated expected penalties and standard errors for the optimal hedging versus the benchmarks for penalty functions (13)-(15). 10^6 paths of the stock price are simulated under the Gaussian market of Section 5 with parameters of Table 1, and the hedging algorithms are applied to each path. The models compared are the minimal expected penalty hedging (MEPH), the Black-Scholes delta hedging (BSDH), the Hardy delta hedging (HDH) and the forecast regime quadratic regime (FRQH). The grid estimate of expected penalties for the optimal hedging is taken directly from the solution of the Bellman Equation ($\hat{\Psi}_0(Y_0, \eta_0, V_0)$). The simulation parameters are found in Section 5.2, 5.3, and 5.5.

As for the quadratic penalty, the MEPH reduces the expected penalty by 6.0% in the weekly case and by 0.9% in the daily case with respect to the best benchmark, namely BSDH for weekly and FRQH for daily. The short (long) quadratic penalty is specifically designed for the call option seller (buyer). The MEPH reduces the expected penalty by 12.6% (3.8%) in the weekly case and by 4.5% (3.7%) in the daily case with respect to the best benchmark. The latter differs across penalties and rebalancing frequencies. For the weekly case, the second best strategy is HDH for the long quadratic penalty and BSDH otherwise. In the daily case, as regime forecasts are more accurate, the FRQH method performs better than the other two benchmarks.

Table 3: Estimated expected penalties (daily rebalancing)

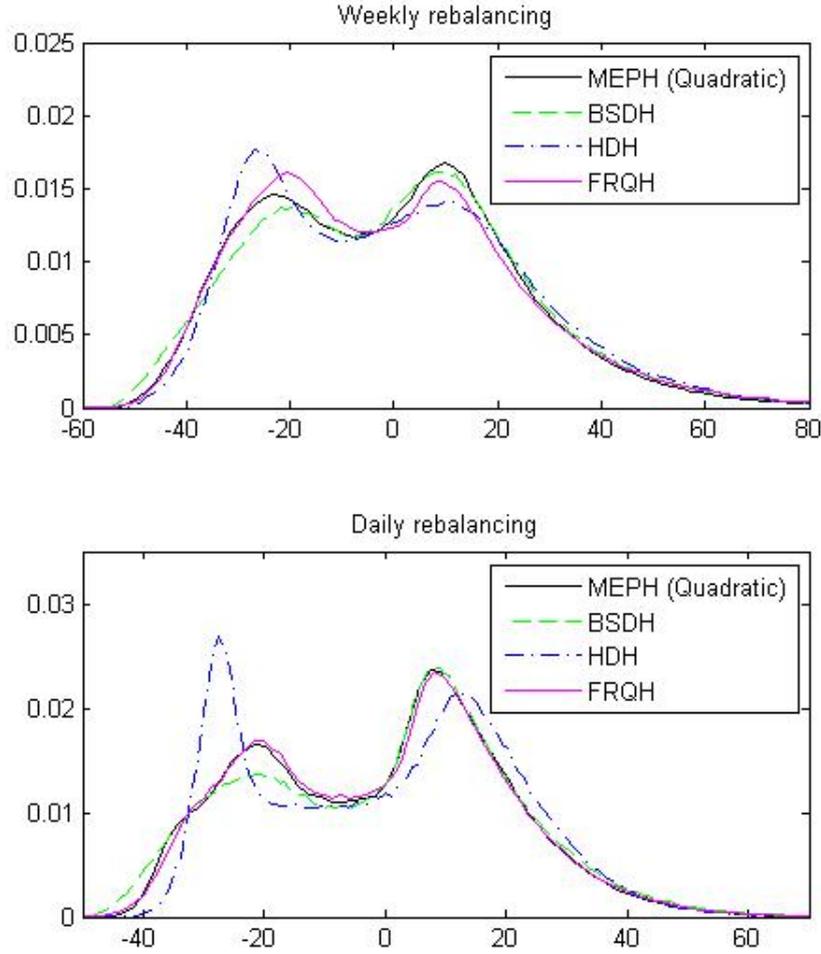
	MEPH	BSDH	HDH	FRQH
Quadratic Penalty				
Grid Estimate	417.88	-	-	-
Expected Penalty	418.77	457.16	442.40	422.57
Standard Error	0.5804	0.5605	0.4556	0.5191
Short Quadratic Penalty				
Grid Estimate	191.01	-	-	-
Expected Penalty	193.71	222.15	220.32	202.90
Standard Error	0.5321	0.4980	0.4345	0.4726
Long Quadratic Penalty				
Grid Estimate	214.87	-	-	-
Expected Penalty	211.62	235.01	222.07	219.66
Standard Error	0.3577	0.4129	0.3416	0.3678

This table reports estimated expected penalties and standard errors for the optimal hedging versus the benchmarks for penalty functions (13)-(15). The grid estimate of expected penalties for the optimal hedging is taken directly from the solution of the Bellman Equation ($\hat{\Psi}_0(Y_0, \eta_0, V_0)$). 10^6 paths of the stock price are simulated under the Gaussian market of Section 5 with parameters of Table 1, and the hedging algorithms are applied to each path. The models compared are the minimal expected penalty hedging (MEPH), the Black-Scholes delta hedging (BSDH), the Hardy delta hedging (HDH) and the forecast regime quadratic regime (FRQH). The simulation parameters are found in Section 5.2, 5.3, and 5.5.

Hedging errors drive hedging penalties and are therefore worthy of investigation. However, descriptive statistics about hedging errors should not be the sole basis on which to judge the performance of hedging strategies. Nevertheless, analyzing those quantities sheds light on how the penalty performance is achieved. Figure 2 displays the hedging error distributions for the quadratic MEPH and the three benchmarks. All distributions exhibit bimodality with similar mode locations but different frequencies. The distribution behaviour, especially in the tails, is better described by Tables 4 and 5.

In terms of RMSE, the quadratic MEPH strategy slightly dominates all other benchmarks for both weekly and daily rebalancing. This is consistent with the quadratic objective of reducing the occurrence of large deviations of the hedging portfolio from the derivative.

Figure 2: Density plot of hedging errors for MEPH versus benchmarks



This figure illustrates empirical densities of hedging errors for the optimal hedging and the benchmarks. 10^6 paths of the stock price are simulated, and the hedging algorithms are applied to each path. The models compared are the quadratic minimal expected penalty hedging (quadratic MEPH), the Black-Scholes delta hedging (BSDH), the Hardy delta hedging (HDH) and the forecast regime quadratic regime (FRQH). The simulation parameters are found in Section 5.2, 5.3, and 5.5.

As far as Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR) are concerned, the picture is not as clear. The short quadratic MEPH used by a call seller performs slightly better than the other hedging strategies.¹⁴ For the call option buyer, MEPH is the second best behind HDH for both VaR and TVaR risk measures.

These equivocal results are due mainly to the mismatch between the penalty function and these risk measures. If a specific risk measure is the ultimate objective, the penalty function should be designed accordingly. Indeed, our methodology precisely permits to adapt the

¹⁴The 95th and 99th percentiles and TVaR are smaller than those of all benchmarks (except for the 99% TVaR of the HDH with daily rebalancing).

hedging strategy to the desired performance criterion.

To illustrate this flexibility, Figure 3 shows the effect of the penalty choice on the MEPH hedging error distribution.

Table 4: Descriptive statistics for hedging errors (weekly rebalancing)

	MEPH (Quadratic)	MEPH (Long)	MEPH (Short)	BSDH	HDH	FRQH
Mean	-0.6787	0.2152	-0.7834	0.1199	0.1695	-0.7741
Standard Deviation	24.945	31.543	25.686	25.732	25.885	25.774
RMSE	24.954	31.544	25.698	25.732	25.885	25.785
Skewness	0.5824	2.3438	0.2844	0.5479	0.6509	0.7566
Excess Kurtosis	0.6937	9.7892	0.4501	0.6640	0.4487	1.0720
99th Percentile	67.595	128.67	65.117	70.887	71.388	72.869
95th Percentile	41.431	58.540	40.272	43.762	45.292	44.219
Median	-0.9649	-3.7258	0.7948	-0.0248	-1.4455	-2.7079
5th Percentile	-36.555	-35.124	-42.559	-38.133	-34.730	-36.428
1st Percentile	-43.988	-42.834	-50.357	-46.463	-41.933	-43.614
Upper TVaR 99%	84.481	163.48	81.574	87.608	86.899	91.001
Upper TVaR 95%	57.768	98.176	55.795	60.591	61.424	62.128
Lower TVaR 5%	-41.047	-39.790	-47.352	-43.183	-39.054	-40.840
Lower TVaR 1%	-46.788	-45.883	-52.414	-49.051	-44.938	-46.575

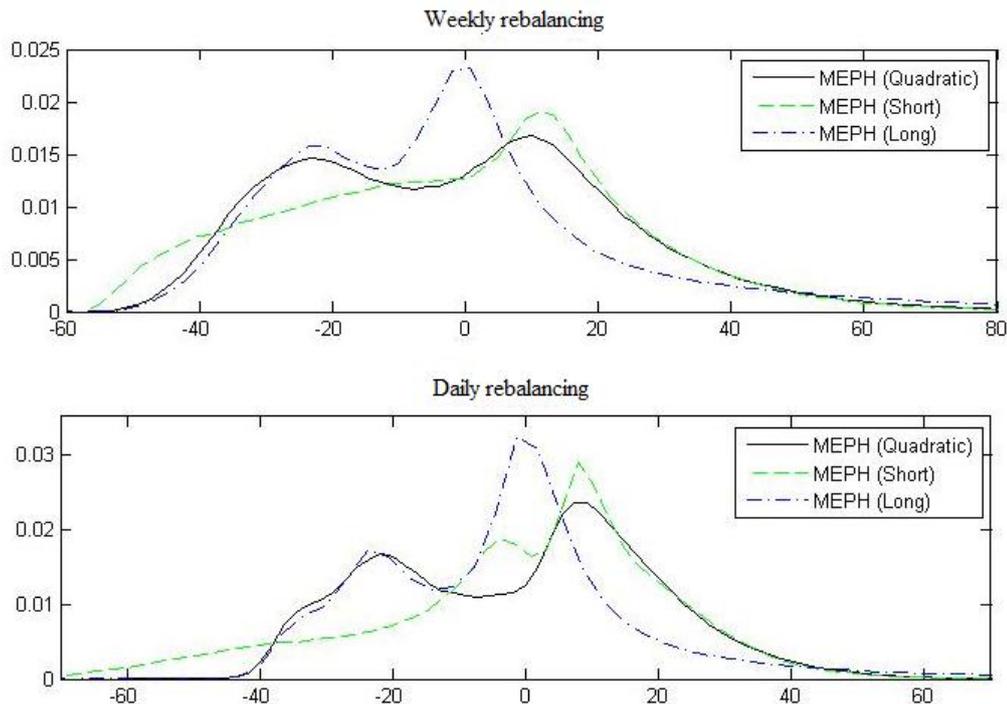
This table reports descriptive statistics for hedging errors for the optimal hedging with penalties (13)-(15) and the benchmarks. 10^6 paths of the stock price are simulated, and the hedging algorithms are applied to each path. The models compared are the minimal expected penalty hedging (MEPH), the Black-Scholes delta hedging (BSDH), the Hardy delta hedging (HDH) and the forecast regime quadratic regime (FRQH). The simulation parameters are found in Section 5.2, 5.3, and 5.5. For RMSE, extreme percentiles and TVaRs, the figure in bold characters indicates the best performing strategy.

Table 5: Descriptive statistics for hedging errors (daily rebalancing)

	MEPH (Quadratic)	MEPH (Long)	MEPH (Short)	BSDH	HDH	FRQH
Mean	-0.6067	0.7418	-0.8713	-0.0877	0.0063	-0.6306
Standard Deviation	20.455	31.951	22.800	21.381	21.033	20.547
RMSE	20.464	31.960	22.816	21.381	21.033	20.557
Skewness	0.1459	3.5405	-0.5902	0.0477	0.0769	0.1436
Excess Kurtosis	-0.0621	20.905	0.8132	-0.4963	-0.9393	-0.4742
99th Percentile	46.337	139.87	44.619	49.132	45.341	47.335
95th Percentile	31.617	50.569	31.011	33.540	32.562	32.138
Median	2.1453	-2.0499	3.3012	3.1464	2.5295	1.6114
5th Percentile	-32.851	-32.322	-45.839	-34.544	-30.521	-32.475
1st Percentile	-37.966	-37.809	-62.017	-41.267	-34.507	-38.363
Upper TVaR 99%	55.126	195.84	53.215	57.680	52.204	55.742
Upper TVaR 95%	40.850	105.89	39.585	43.142	40.423	41.529
Lower TVaR 5%	-36.063	-35.801	-55.737	-38.626	-32.949	-36.182
Lower TVaR 1%	-39.745	-39.767	-68.407	-43.477	-36.396	-41.247

This table reports descriptive statistics for hedging errors for the optimal hedging with penalties (13)-(15) and the benchmarks. 10^6 paths of the stock price are simulated, and the hedging algorithms are applied to each path. The models compared are the minimal expected penalty hedging (MEPH), the Black-Scholes delta hedging (BSDH), the Hardy delta hedging (HDH) and the forecast regime quadratic regime (FRQH). The simulation parameters are found in Section 5.2, 5.3, and 5.5. For RMSE, extreme percentiles and TVaRs, the figure in bold characters indicates the best performing strategy.

Figure 3: Density plot of hedging errors for MEPH



This figure illustrates empirical densities of hedging errors for the MEPH hedging with penalties (13)-(15). 10^6 paths of the stock price are simulated, and the hedging algorithms are applied to each path. The simulation parameters are found in Section 5.2, 5.3, and 5.5.

6 Conclusion

A flexible and tractable methodology is presented for the hedging of contingent claims in the presence of regimes. It accommodates various hedging objectives through the penalty function specification. Constraints on trading strategies, such as no short-selling, can be incorporated.

Path dependency issues are tackled by the addition of a state variable, making the hedging problem suitable for dynamic programming. The approach is implemented with the standard Gaussian two-regime model estimated from weekly S&P 500 returns. Based on the hedging of an at-the-money call option, the current methodology compares favourably with three relevant alternatives.

Since the current paper's algorithm involves lattices, the curse of dimensionality prevents the use of a large number of underlying assets and regimes. The addition of a single dimension (transaction costs, a three-regime model or stochastic interest rates) remains feasible at a substantial numerical cost.

The modeling design voluntarily avoids the identification of the pricing measure. Nevertheless, if one wishes to determine the pricing kernel (at the cost of a specification error), several extensions become feasible: option tracking, hedging American options and hedging with other derivatives.

A Appendix

Lemma A.1 is used in the proofs of Corollaries 3.1 and 3.2.

Lemma A.1 *Let $\mathcal{I} \subseteq \mathcal{J} \subseteq \mathcal{M}$ be sigma-algebras and Z be a random variable.*

$$\text{If } \mathbb{E}[Z|\mathcal{I}] = \mathbb{E}[Z|\mathcal{M}], \text{ then } \mathbb{E}[Z|\mathcal{J}] = \mathbb{E}[Z|\mathcal{M}] = \mathbb{E}[Z|\mathcal{I}].$$

Proof of Lemma A.1

$$\begin{aligned} \mathbb{E}[Z|\mathcal{J}] &= \mathbb{E}[\mathbb{E}[Z|\mathcal{M}]|\mathcal{J}] \quad (\text{Law of iterated expectations}) \\ &= \mathbb{E}[\mathbb{E}[Z|\mathcal{I}]|\mathcal{J}] \\ &= \mathbb{E}[Z|\mathcal{I}]. \quad (\text{Law of iterated expectations}) \end{aligned}$$

QED

Proof of Theorem 3.1

$$\begin{aligned} \eta_{i,n+1} &= f_{h_{n+1}|Y_{0:n+1}}(i|Y_{0:n+1}) \\ &= \frac{f_{h_{n+1},Y_{0:n+1}}(i,Y_{0:n+1})}{f_{Y_{0:n+1}}(Y_{0:n+1})} \\ &= \frac{\sum_{j=1}^H f_{h_{n+1},h_n,Y_{0:n+1}}(i,j,Y_{0:n+1})}{\sum_{k=1}^H \sum_{\ell=1}^H f_{h_{n+1},h_n,Y_{0:n+1}}(k,\ell,Y_{0:n+1})} \\ &= \frac{\sum_{j=1}^H f_{h_{n+1},Y_{n+1}|h_n,Y_{0:n}}(i,Y_{n+1}|j,Y_{0:n}) f_{h_n|Y_{0:n}}(j|Y_{0:n}) f_{Y_{0:n}}(Y_{0:n})}{\sum_{k=1}^H \sum_{\ell=1}^H f_{h_{n+1},Y_{n+1}|h_n,Y_{0:n}}(k,Y_{n+1}|\ell,Y_{0:n}) f_{h_n|Y_{0:n}}(\ell|Y_{0:n}) f_{Y_{0:n}}(Y_{0:n})} \\ &= \frac{\sum_{j=1}^H f_{h_{n+1},Y_{n+1}|h_n,Y_{0:n}}(i,Y_{n+1}|j,Y_{0:n}) \eta_{j,n}}{\sum_{k=1}^H \sum_{\ell=1}^H f_{h_{n+1},Y_{n+1}|h_n,Y_{0:n}}(k,Y_{n+1}|\ell,Y_{0:n}) \eta_{\ell,n}}. \end{aligned}$$

The Markov property of (Y, h) and Lemma A.1 complete the proof since

$$f_{h_{n+1},Y_{n+1}|h_n,Y_{0:n}}(i,Y_{n+1}|j,Y_{0:n}) = f_{h_{n+1},Y_{n+1}|h_n,Y_n}(i,Y_{n+1}|j,Y_n).$$

QED

Proof of Corollary 3.1. Applying the Law of iterated expectations,

$$\mathbb{P}(h_n = i | \vec{\eta}_n) = \mathbb{E}(\mathbb{P}(h_n = i | \mathcal{G}_n) | \vec{\eta}_n) = \mathbb{E}(\eta_{i,n} | \vec{\eta}_n) = \eta_{i,n} = \mathbb{P}(h_n = i | \mathcal{G}_n). \quad (18)$$

By Lemma A.1, since $\sigma(\vec{\eta}_n) \subseteq \sigma(Y_n, \vec{\eta}_n) \subseteq \mathcal{G}_n$,

$$\mathbb{P}(h_n = i | \mathcal{G}_n) = \mathbb{P}(h_n = i | Y_n, \vec{\eta}_n). \quad (19)$$

Moreover, since $\{Y_n, h_n, \vec{\eta}_n\}_{n=0}^N$ has the Markov property with respect to \mathcal{F} , then for any Borel set $\mathcal{D} \subseteq \mathbb{R} \times [0, 1]^H$,

$$\begin{aligned} & \mathbb{P}[(Y_{n+1}, \vec{\eta}_{n+1}) \in \mathcal{D} | \mathcal{G}_n] \\ &= \mathbb{E}[\mathbb{P}[(Y_{n+1}, \vec{\eta}_{n+1}) \in \mathcal{D} | \mathcal{F}_n] | \mathcal{G}_n] \quad (\text{Law of iterated expectations}) \\ &= \mathbb{E}[\mathbb{P}[(Y_{n+1}, \vec{\eta}_{n+1}) \in \mathcal{D} | Y_n, h_n, \vec{\eta}_n] | \mathcal{G}_n] \quad (\text{Markov property}) \\ &= \sum_{j=1}^H \mathbb{P}[(Y_{n+1}, \vec{\eta}_{n+1}) \in \mathcal{D} | Y_n, h_n = j, \vec{\eta}_n] \mathbb{P}[h_n = j | \mathcal{G}_n] \\ &= \sum_{j=1}^H \mathbb{P}[(Y_{n+1}, \vec{\eta}_{n+1}) \in \mathcal{D} | Y_n, h_n = j, \vec{\eta}_n] \mathbb{P}[h_n = j | Y_n, \vec{\eta}_n] \quad (\text{Eq. (19)}) \\ &= \mathbb{P}[(Y_{n+1}, \vec{\eta}_{n+1}) \in \mathcal{D} | Y_n, \vec{\eta}_n] \quad (\text{Bayes' Law}). \end{aligned}$$

QED

Proof of Corollary 3.2. For any admissible strategy, because of the self-financing restriction, its time- $(n+1)$ value satisfies

$$\begin{aligned} V_{n+1} &= V_n + \vec{\theta}_{n+1}^\top (\vec{S}_{n+1} - \vec{S}_n) \\ &= V_n + \theta_{n+1}^{(1)} (\exp(r(n+1)\Delta_t) - \exp(rn\Delta_t)) + \theta_{n+1}^{(2)} S_0^{(2)} (\exp(Y_{n+1}) - \exp(Y_n)). \end{aligned}$$

Hence, V_{n+1} is $\sigma(Y_{n+1}, Y_n, V_n, \vec{\theta}_{n+1})$ -measurable. Furthermore, by Equation (18), Lemma A.1 and the fact that $\sigma(\vec{\eta}_n) \subseteq \sigma(Y_n, \vec{\eta}_n, V_n, \vec{\theta}_{n+1}) \subseteq \mathcal{G}_n$,

$$\mathbb{P}(h_n = i | \mathcal{G}_n) = \mathbb{P}(h_n = i | Y_n, \vec{\eta}_n, V_n, \vec{\theta}_{n+1}). \quad (20)$$

Therefore, for any Borel set $\mathcal{D} \subseteq \mathbb{R} \times [0, 1]^H \times \mathbb{R}$,

$$\begin{aligned}
& \mathbb{P} \left[(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1}) \in \mathcal{D} \middle| \mathcal{G}_n \right] \\
&= \mathbb{E} \left[\mathbb{P} \left[(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1}) \in \mathcal{D} \middle| \mathcal{F}_n \right] \middle| \mathcal{G}_n \right] \text{ (Law of iterated expectations)} \\
&= \mathbb{E} \left[\mathbb{P} \left[(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1}) \in \mathcal{D} \middle| Y_n, h_n, \vec{\eta}_n, V_n, \vec{\theta}_{n+1} \right] \middle| \mathcal{G}_n \right] \text{ (Corollary 3.1)} \\
&= \sum_{j=1}^H \mathbb{P} \left[(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1}) \in \mathcal{D} \middle| Y_n, h_n = j, \vec{\eta}_n, V_n, \vec{\theta}_{n+1} \right] \mathbb{P} \left[h_n = j \middle| \mathcal{G}_n \right] \\
&= \sum_{j=1}^H \mathbb{P} \left[(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1}) \in \mathcal{D} \middle| Y_n, h_n = j, \vec{\eta}_n, V_n, \vec{\theta}_{n+1} \right] \mathbb{P} \left[h_n = j \middle| Y_n, \vec{\eta}_n, V_n, \vec{\theta}_{n+1} \right] \text{ (Eq. (20))} \\
&= \mathbb{P} \left[(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1}) \in \mathcal{D} \middle| Y_n, \vec{\eta}_n, V_n, \vec{\theta}_{n+1} \right] \text{ (Bayes' Law)}.
\end{aligned}$$

QED

Proof of Theorem 3.2. In the following, all minimizations are performed over the set of admissible hedging strategies Θ . The hedging penalty Ψ_N^* depends on the initial prices \vec{S}_0 , the initial portfolio value V_0 , the cumulative returns $Y_{1:N}$, and the portfolio position $\vec{\theta}_{1:N}$, that is,

$$\Psi_N^* = g(\phi(\vec{S}_N) - V_N) = \Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:N} \right).$$

For any $n \in \{0, \dots, N-2\}$, define

$$\vec{\vartheta}_{(n+2):N} := \arg \min_{\vec{\theta}_{n+2:N}} \mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:N} \right) \middle| \mathcal{G}_{n+1} \right], \quad (21)$$

implying that

$$\begin{aligned}
\Psi_{n+1}^* &= \min_{\vec{\theta}_{n+2:N}} \mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:N} \right) \middle| \mathcal{G}_{n+1} \right] \\
&= \mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:n+1}, \vec{\vartheta}_{(n+2):N} \right) \middle| \mathcal{G}_{n+1} \right].
\end{aligned} \quad (22)$$

First direction. Therefore, for any admissible strategy $\vec{\theta}_{1:N}$,

$$\Psi_{n+1}^* = \mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:n+1}, \vec{\vartheta}_{n+2:N} \right) \middle| \mathcal{G}_{n+1} \right] \leq \mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:N} \right) \middle| \mathcal{G}_{n+1} \right].$$

Consequently, by monotonicity of the conditional expectation operator,

$$\begin{aligned}
\min_{\vec{\theta}_{n+1}} \mathbb{E} \left[\Psi_{n+1}^* \middle| \mathcal{G}_n \right] &\leq \min_{\vec{\theta}_{n+1}} \mathbb{E} \left[\mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:N} \right) \middle| \mathcal{G}_{n+1} \right] \middle| \mathcal{G}_n \right] \\
&= \min_{\vec{\theta}_{n+1}} \mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:N} \right) \middle| \mathcal{G}_n \right] \text{ (Law of iterated expectations)}.
\end{aligned}$$

Because the left-hand side of the previous inequality does not depend on $\vec{\theta}_{n+2:N}$, then

$$\min_{\vec{\theta}_{n+1}} \mathbb{E} [\Psi_{n+1}^* | \mathcal{G}_n] = \min_{\vec{\theta}_{n+1:N}} \mathbb{E} [\Psi_{n+1}^* | \mathcal{G}_n] \leq \min_{\vec{\theta}_{n+1:N}} \mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:N} \right) \middle| \mathcal{G}_n \right] = \Psi_n^*$$

where the last equality arises from Definition (4).

Second direction.

$$\begin{aligned} \Psi_n^* &= \min_{\vec{\theta}_{n+1:N}} \mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:N} \right) \middle| \mathcal{G}_n \right] \quad (\text{Definition (4)}) \\ &= \min_{\vec{\theta}_{n+1:N}} \mathbb{E} \left[\mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:N} \right) \middle| \mathcal{G}_{n+1} \right] \middle| \mathcal{G}_n \right] \quad (\text{Law of iterated expectations}) \\ &\leq \min_{\vec{\theta}_{n+1}} \mathbb{E} \left[\mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:n+1}, \vec{\vartheta}_{n+2:N} \right) \middle| \mathcal{G}_{n+1} \right] \middle| \mathcal{G}_n \right] \quad (\text{Reducing optimization domain}) \\ &= \min_{\vec{\theta}_{n+1}} \mathbb{E} [\Psi_{n+1}^* | \mathcal{G}_n] \quad (\text{Definition (22)}). \end{aligned} \tag{23}$$

Therefore, $\Psi_n^* = \min_{\vec{\theta}_{n+1}} \mathbb{E} [\Psi_{n+1}^* | \mathcal{G}_n]$, establishing Equation (5).

Furthermore, define $\vec{\theta}_{n+1}^*$ for any $n \in \{0, \dots, N-1\}$ as a solution of

$$\vec{\theta}_{n+1}^* := \arg \min_{\vec{\theta}_{n+1}} \mathbb{E} [\Psi_{n+1}^* | \mathcal{G}_n]. \tag{24}$$

Then,

$$\begin{aligned} \Psi_n^* &= \min_{\vec{\theta}_{n+1}} \mathbb{E} [\Psi_{n+1}^* | \mathcal{G}_n] \quad (\text{Equation (5)}) \\ &= \mathbb{E} [\Psi_{n+1}^* (\vec{\theta}_{n+1}^*) | \mathcal{G}_n] \quad (\text{Equation (24)}) \\ &= \mathbb{E} \left[\mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:n}, \vec{\theta}_{n+1}^*, \vec{\vartheta}_{(n+2):N} \right) \middle| \mathcal{G}_{n+1} \right] \middle| \mathcal{G}_n \right] \quad (\text{Equation (22)}) \\ &= \mathbb{E} \left[\Psi_N^* \left(\vec{S}_0, V_0, Y_{1:N}, \vec{\theta}_{1:n}, \vec{\theta}_{n+1}^*, \vec{\vartheta}_{(n+2):N} \right) \middle| \mathcal{G}_n \right] \quad (\text{Law of iterated expectations}). \end{aligned}$$

Therefore, by Equation (4),

$$\vec{\vartheta}_{(n+1):N} := \left(\arg \min_{\vec{\theta}_{(n+1)}} \mathbb{E} [\Psi_{n+1}^* | \mathcal{G}_n], \vec{\vartheta}_{(n+2):N} \right)$$

is a solution (possibly not the only one) to the following equation:

$$\vec{\vartheta}_{(n+1):N} := \arg \min_{\vec{\theta}_{n+1:N}} \mathbb{E} [\Psi_N^* | \mathcal{G}_n]. \tag{25}$$

Hence, if Equation (24) is satisfied $\forall n \in \{0, \dots, N-1\}$, a recursive argument implies that $\theta^* := (\vec{\theta}_1^*, \dots, \vec{\theta}_N^*)$ solves Problem (1).

QED

Proof of Theorem 3.3.

The proof hinges on a backward induction over time. Clearly, $\Psi_N^* = g\left(\tilde{\phi}(Y_N) - V_N\right)$ is $\sigma(Y_N, \vec{\eta}_N, V_N)$ -measurable. Assume that Ψ_{n+1}^* is $\sigma(Y_{n+1}, \vec{\eta}_{n+1}, V_{n+1})$ -measurable. From Corollary 3.2, there is a Borel-measurable function φ such that

$$\begin{aligned}\Psi_n^* &= \min_{\vec{\theta}_{n+1} \in \mathcal{G}_n} \mathbb{E}[\Psi_{n+1}^* | \mathcal{G}_n] \quad (\text{Equation (5)}) \\ &= \min_{\vec{\theta}_{n+1} \in \mathcal{G}_n} \mathbb{E}[\Psi_{n+1}^* | Y_n, \vec{\eta}_n, V_n, \vec{\theta}_{n+1}] \quad (\text{Corollary 3.2}) \\ &= \min_{\vec{\theta}_{n+1} \in \mathcal{G}_n} \varphi\left(Y_n, \vec{\eta}_n, V_n, \vec{\theta}_{n+1}\right).\end{aligned}$$

Therefore, a necessary and sufficient condition for $\vec{\theta}_{n+1}^* = \underset{\vec{\theta}_{n+1}}{\operatorname{argmin}} \mathbb{E}[\Psi_{n+1}^* | \mathcal{G}_n]$ is to minimize $\xi(\cdot) := \varphi(Y_n, \vec{\eta}_n, V_n, \cdot)$ which only depends on $(Y_n, \vec{\eta}_n, V_n)$. Consequently, there exists $\vec{\theta}_{n+1}^*$ which is $\sigma(Y_n, \vec{\eta}_n, V_n)$ -measurable. Hence, $\Psi_n^* = \varphi\left(Y_n, \vec{\eta}_n, V_n, \vec{\theta}_{n+1}^*\right)$ is also $\sigma(Y_n, \vec{\eta}_n, V_n)$ -measurable and

$$\vec{\theta}_{n+1}^* = \underset{\vec{\theta}_{n+1} \in \sigma(Y_n, \vec{\eta}_n, V_n)}{\operatorname{argmin}} \mathbb{E}[\Psi_{n+1}^* | \mathcal{G}_n] = \underset{\vec{\theta}_{n+1} \in \sigma(Y_n, \vec{\eta}_n, V_n)}{\operatorname{argmin}} \mathbb{E}[\Psi_{n+1}^* | Y_n, \vec{\eta}_n, V_n],$$

that is, the set of admissible hedging strategies Θ may be restricted to keep only strategies that also satisfy that θ_{n+1} is $\sigma(Y_n, \vec{\eta}_n, V_n)$ -measurable.

QED

B Appendix

The following proof applies to the simple market of Section 5.

Proof of Proposition 5.1.

$$\begin{aligned}\eta_{j,n+1} &= \mathbb{P}(h_{n+1} = j | \mathcal{G}_n) \\ &= \mathbb{E}[\mathbb{P}(h_{n+1} = j | h_n) | \mathcal{G}_n] \\ &= \sum_{i=1}^H \mathbb{P}(h_{n+1} = j | h_n = i) \mathbb{P}(h_n = i | \mathcal{G}_n) \\ &= \sum_{i=1}^H P_{i,j}^{(n)} \eta_{i,n} \geq \sum_{i=1}^H \min_{u \in \mathcal{H}} P_{u,j}^{(n)} \eta_{i,n} = \min_{u \in \mathcal{H}} P_{u,j}^{(n)}\end{aligned}$$

The case $\eta_{j,n+1} \leq \max_{u \in \mathcal{H}} P_{u,j}$ is similar.

QED

References

- [1] Alizadeh, A., Nomikos, N. (2004). A Markov regime switching approach for hedging stock indices. *Journal of Futures Markets*. **(24)**. pp. 649-674
- [2] Alizadeh, A.H., Nomikos, N.K., Pouliasis, P.K. (2008). A Markov regime switching approach for hedging energy commodities. *Journal of Banking and Finance*. **(32)**. pp. 1970-1983
- [3] Augustyniak, M., Boudreault, M. (2012). An out-of-sample analysis of investment guarantees for equity-linked products: lessons from the financial crisis of the late 2000s. Working paper, Université de Montréal.
- [4] Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*. **(81)**. pp. 637-654
- [5] Bollen, N.P.B. (1998). Valuing options in regime-switching models. *Journal of Derivatives*. **(6)**. pp. 38-49
- [6] Buffington, J., Elliott, R.J. (2002). American options with regime switching. *International Journal of Theoretical and Applied Finance*. **(5)**. pp. 497-514
- [7] Cont, R., Tankov, P., Voltchkova, E. (2007). Hedging with Options in Models with Jumps. *Stochastic Analysis and Applications*. **(2)**. pp. 197-217
- [8] Cvitanić, J., Spivak, G. (1999). Maximizing the probability of a perfect hedge. *The Annals of Applied Probability*. **(9)**. pp. 1303-1328
- [9] Cvitanić, J., Karatzas, I. (1999). On dynamic measures of risk. *Finance and Stochastics*. **(3)**. pp. 451-482
- [10] Dempster, A.P., Laird, N.M., Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*. **(39)**. pp. 1-38
- [11] Duan, J.C. (1995). The GARCH Option Pricing Model. *Mathematical Finance*. **(5)**. pp. 13-32
- [12] Duan, J.C., Popova, I. and Ritchken, P. (2002). Option pricing under regime switching. *Quantitative Finance*. **(2)**. pp. 116-132
- [13] Eberlein, E., Jacod, J. (1997). On the range of option prices. *Finance and Stochastics*. **(1)**. pp. 131-140
- [14] El Karoui, N., Quenez, M.-C. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal on Control and Optimization*. **(33)**. pp. 29-66

- [15] Elliott, R.J., Chan, L.L., Siu, T.K., (2005). Option pricing and Esscher transform under regime switching. *Annals of Finance*. **(1)**. pp. 423-432
- [16] Föllmer, H., Leukert, P. (2000). Efficient hedging: cost versus shortfall risk. *Finance and Stochastics*. **(4)**. pp. 117-146
- [17] Föllmer, H., Leukert, P. (1999). Quantile hedging. *Finance and Stochastics*. **(3)**. pp. 251-273
- [18] Genest, C., Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*. **(44)**. pp. 1096-1127
- [19] Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*. **(57)**. pp. 357-384
- [20] Hardy, M. (2001). A regime switching model of long term stock returns. *North American Actuarial Journal*. **(5)**. pp. 41-53
- [21] Karatzas, I. (1997). Lectures in mathematical finance. *Providence: American Mathematical Society*
- [22] Lee, H.T., Yoder, J.K., Mittelhammer, R.C., McCluskey, J.J. (2006). A random coefficient autoregressive Markov regime switching model for dynamic futures hedging. *Journal of Futures Markets*. **(26)**. pp. 103-129
- [23] Lee, H.T., Yoder, J.K. (2007). Optimal hedging with a regime-switching time-varying correlation GARCH model. *Journal of Futures Markets*. **(27)**. pp. 495-516
- [24] Lee, H.T. (2009a). Optimal futures hedging under jump switching dynamics. *Journal of Empirical Finance*. **(16)**. pp. 446-456
- [25] Lee, H.T. (2009b). A copula-based Markov regime switching GARCH model for optimal futures hedging. *Journal of Futures Markets*. **(29)**. pp. 946-972
- [26] Lien, D. (2012). A note on the performance of regime switching hedge strategy. *The Journal of Futures Markets*. **(32)**. pp. 389-396
- [27] Mamon, R.S., Rodrigo, M.R. (2005). Explicit solutions to European options in a regime-switching economy. *Operations Research Letters*. **(33)**. pp. 581-586
- [28] Motoczyński, M. (2000). Multidimensional variance-optimal hedging in discrete-time model - a general approach. *Mathematical Finance*. **(10)**. pp. 243-257
- [29] Pham, H. (2000). Dynamic L^p -hedging in discrete time under cone constraints. *SIAM Journal on Control and Optimization*. **(38)**. pp. 665-682

- [30] Rémillard, B., Hocquard, A., and Papageorgiou, N. (2010a). Option pricing and dynamic discrete time hedging for regime-switching geometric random walks models. Working paper series, HEC Montréal.
- [31] Remillard, B., Langlois, H., Hocquard, A., Papageorgiou, N. (2010b). Optimal hedging of American options in discrete time. Working paper series, HEC Montréal.
- [32] Rémillard, B., Rubenthaler, S. (2009). Optimal hedging in discrete and continuous time. Technical Report G-2009-77, Gerad.
- [33] Schweizer, M. (1991). Option hedging for semimartingales. *Stochastic Processes and their Applications*. (**37**). pp. 339-363
- [34] Schweizer, M. (1995). Variance-optimal hedging in discrete time. *Mathematics of Operation Research*. (**20**). pp. 1-32
- [35] Sekine, J. (2004). Dynamic minimization of worst conditional expectation of shortfall. *Mathematical Finance*. (**14**). pp. 605-618
- [36] Xu, M. (2006). Risk measure pricing and hedging in incomplete markets. *Annals of Finance*. (**2**). pp. 51-71