Maturity Risk for Variational Hedge with Composed Risk Measure

João Ramos Jungblut[a*] Marcelo Brutti Righi[a]

joaoramosjungblut@hotmail.com marcelo.righi@ufrgs.br

[a]Federal University of Rio Grande do Sul, Washington Luiz, 855, Porto Alegre, Brazil, zip 90010-460

Abstract

We study the maturity risk of contingent claims based on convex risk measures combined with utility functions, allowing the solution of an optimal hedging problem equivalent to the dual of a decision under variational preferences. Our analysis starts from the construction of the Profit-and-Loss of the hedged contract over time, defined as the financial result of the difference between a self-financing portfolio and the derivative price, quantifying the residual risk that cannot be eliminated even under the optimal strategy. The proposed approach formalizes this by including a temporal dimension in addition to its spatial nature; we analyze the properties, its induced acceptance set, and obtain its dual representation. For numerical estimation, we solve the optimal hedging problem and introduce a version of the stochastic maximum principle applied to risk measures. We conduct a numerical analysis for the construction of the maturity risk surfaces by magnitude, moneyness, and time.

Keywords: maturity risk; optimal hedging; composed risk measures; stochastic control; market incompleteness.

*The author is grateful for the financial support of the Coordination for the Improvement of Higher Education Personnel under process number 88887.959969/2024-00.

1 Introduction

Risk measures in mathematical finance assess a position in terms of acceptability. In Artzner et al. [1999], coherent risk measures are given an axiomatic formulation and are represented through acceptance sets, so risk is read in primal terms as the minimal capital injection that renders the position acceptable. This links the concept directly to capital requirements rather than to statistical summaries. Föllmer and Schied [2002] extended the framework beyond coherence by relaxing sub-additivity and homogeneity properties, which leads to a broader class of convex risk measures while preserving the monetary interpretation.

A limitation of these risk measures discussed above is their spatial nature. When used in this acceptability framework, they assess losses at a fixed time and do not describe how risk behaves as time evolves. This issue becomes explicit in the analysis of Mahmoud [2016], who points out that the temporal dimension of risk is largely absent from standard formulations and requires separate treatment to be made explicit. In his approach, this temporal dimension is provided exogenously through transformations that map entire paths into random times.

In this context, Maturity Risk (MR) has emerged as a new object of interest. It is defined in this paper through an optimization problem under a composed risk measures, in which a convex risk measure is applied after a utility transformation. The input is the cumulative profit and loss generated by hedging a contingent claim over a remaining horizon, with evaluation being performed at the terminal date. MR is then understand as the residual exposure that remains after trade has been employed in the best admissible way, thereby quantifying what cannot be eliminated through hedging. It consequently measures the risk that remains linked to the claim after the optimal strategy is taken into account.

The evaluation criteria adhere to the framework of Righi [2024] and are constructed using risk measures that meet monotonicity, convexity, and continuity in the weak topology. Under these properties, the hedging problem is similar to a formulation in terms of variational preferences Maccheroni et al. [2006], with the equivalence given by dual representation. Time enters the structure via the stochastic evolution of the underlying process, which causes variations in profit and loss along the horizon. In contrast to strictly static formulations, this technique incorporates temporal variation through trading dynamics while the risk assess stays terminal.

Hence, as far as we are concern, MR is an open concept in quantitative risk management. Also, it can be naturally formulated in incomplete markets, where perfect replication may fail and price are intervalar Jouini and Kallal [1995], Çetin et al. [2004]. One can think of this setting as one in which the market does not deliver a single valuation rule, so ordering contracts only by price becomes ambiguous. This criterion provides a reliable alternative by ranking contingent claims by the residual exposure left after admissible trading, which is precisely the component that interval prices do not resolve.

Within this context, the main challenge is to identify the basic attributes of MR: properties, acceptance set and dual representation. Once these aspects are established, the examination shifts to how MR varies with time and moneyness, resulting in its depiction as a surface. A short note here is that the surface is not postulated exogenously, but rather induced by an optimum hedging problem that occurs as the solution to a stochastic control optimization problem. In particular, the stochastic maximum principle (SMP) is adapted to risk measures to solve. We drawn a numerical experiment for a European call contract to examine how the risk behaves. The goal of the numerical analysis is not to evaluate trading performance or actual profitability, but to better understand the structural aspects of the terminal risk provided by the continuous-time formulation. The simulations allow us to study how the magnitude and form of MR respond to changes in volatility dynamics, jump intensity, or variance persistence, as well as how the risk functional chosen affects a reference exposure.

We compare MR across different specifications of composed risk measures and stochastic dynamics. Six composed criteria are considered, obtained by combining Expected Loss (EL), Entropic Risk Measure (ENT), and Expected Shortfall (ES) with linear, exponential, and Weibull-type utilities. The comparison is conducted under three stochastic models for the underlying asset, namely Geometric Brownian Motion (GBM), Heston Stochastic Volatility and Merton Jump-Diffusion leading to eighteen models.

A first strand of literature investigates hedging and pricing under convex and coherent risk criteria, using acceptability, capital requirements, and convex duality to identify feasible methods and valuation bounds in imperfect markets. Shortfall-risk minimization and acceptability-based pricing offer operational hedging rules and interval prices after a criterion is established Follmer and Leukert [1999], Carr et al. [2001]. Risk-indifference pricing connects convex risk measurements to utility-based valuation Xu [2006], whereas coherent pricing frameworks aggregate preferences using convolution approaches Cherny and Madan [2006]. Although these systems provide consistent pricing and hedging mechanisms, their emphasis is on valuation criteria or feasibility constraints, and the analysis is largely static, focusing on terminal outcomes.

A second line relates hedging concerns to stochastic control and robustness. Convex risk evaluation is related to control problems through inf-convolution and Hamilton-Jacobi-Bellman equations Toussaint and Sircar [2011], while robust formulations penalize departures from a reference model to account for volatility uncertainty Herrmann et al. [2016]. Recent advances combine learning and data-driven approaches, combining robust control with neural SDEs to increase calibration and computing performance Gierjatowicz et al. [2020]). Carbonneau and Godin [2023]) use reinforcement learning to implement Equal Risk Pricing under non-translation-invariant criteria. In some cases, hedging is handled as a control problem with a fixed evaluation method, and the emphasis is on robustness or numerical tractability.

A third body of research addresses hedging through advanced numerical techniques, particularly deep BSDE solvers applied to quadratic or model-based criteria such as mean-variance and local risk reduction Shi et al. [2023], Gnoatto et al. [2024]. These methods considerably increase the computational scope of optimum hedging in high-dimensional environments. However, across these various threads, the exposure left over after optimal hedging is not isolated as an item of independent interest, nor is it described by acceptance sets or evaluated as a function of time and contract parameters. This gap inspires the current study, which treats residual exposure caused by hedging as a risk measure in and of itself, and investigates it systematically within a comprehensive risk framework.

From this perspective, a more direct comparison is made with the GAN-based deep hedging of Limmer and Horvath [2024]. Both frameworks optimize composed risk measures within a variational-preferences viewpoint, but they do so in different ways. He targets the numerical learning of hedging strategies at the process level via adversarial training, avoiding the explicit solution of forward–backward systems and analytical optimality constraints. Accordingly, the GAN-based method emphasizes numerical robustness and scalability for optimal hedging under composed criteria, while our focus is on the construction of the MR and on isolating the temporal persistence of residual exposure, which is not addressed in Limmer and Horvath [2024].

Building on the broader literature that treats hedging as a mechanism for reducing, but typically not eliminating, uncertainty, Rockafellar and Royset [2015] formalize residual risk as the risk that remains after optimal mitigation based on available information, providing a particularly clean convex-analytic and dual characterization. Our idea is aligned with this "risk-after-hedge" perspective, but moves the focus to a time-to-maturity dimension. A complementary strand highlights that the horizon at which risk is measured is economically nontrivial: Brigo and Nordio [2010] argue that liquidity turns the liquidation horizon into a random quantity, so tail risk should be assessed over an uncertain holding period rather than a fixed window, and Colldeforns-Papiol and Ortiz-Gracia [2018] frames this point in the context of Basel's post-crisis emphasis on tail-risk metrics and horizons that may exceed benchmark values. While we introduces time as an intrinsic dimension of the object of analysis, their work varies by utilizing VaR/ES to operationalize the horizon influence over random holding periods.

The contributions of this paper are fourfold. First, we define MR as a time-indexed risk metric that makes the temporal dimension explicit by quantifying the residual exposure from any intermediate date to maturity after optimal trading, and we characterize the associated acceptance sets. Second, we study key properties of MR within an acceptability framework and dual representation. Third, we formulate and solve the optimal hedging problem under risk-based criteria via an SMP tailored to composed risk measures, establishing well-posedness and optimality of the control. Fourth, we present a numerical study of MR as a surface over time and moneyness, comparing eighteen specifications across risk measures, utility functions, and stochastic dynamics to assess robustness and model dependence.

The paper is organized as follows. Section 2 introduces the mathematical framework, presenting the market setting and the class of composed risk measures used throughout the analysis. Section 3 defines MR, studies its main properties including the relation with acceptance sets and dual representations. Section 4 present a way to find the optimal hedging strategy through stochastic control, specifically by deriving the SMP. Section 6 describes the numerical methods and reports the results analyzing MR characteristics. All figures related to the numerical experiments are collected in the Appendix for readability.

2 Background

The market is described by a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, with $T \in \mathbb{R}_+$ and $T < \infty$, where Ω is the set of states of the world, $\mathcal{F}$ is a σ-algebra on Ω, $(\mathcal{F}_t)_{t\in[0,T]}$ is a filtration and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$. We impose the usual conditions for $(\mathcal{F}_t)$ to be right-continuous ($\mathcal{F}_t = \bigcap_{t'>t} \mathcal{F}_{t'}$) and $\mathbb{P}$-complete (i.e., $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets). We interpret $(\mathcal{F}_t)$ as the information available up to time t, increasing with t (so $\mathcal{F}_t \subseteq \mathcal{F}_{t'}$ for $t \leq t'$).

Let $S := (S_t)_{t\in[0,T]}$ denote the vector-valued price process of n underlying assets, adapted to $(\mathcal{F}_t)_{t\in[0,T]}$ and such that $S_t \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^n)$ for each t. For each fixed $\omega \in \Omega$, the map $t \mapsto S_t(\omega)$ describes the temporal price trajectory in the state ω, while for each fixed $t \in [0, T]$, the map $\omega \mapsto S_t(\omega)$ is a random variable in $(\Omega, \mathcal{F}_t, \mathbb{P})$. We work in $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ to guarantee the square-integrability of prices and to interpret differentials such as dS_t as the infinitesimal variation of S at time t, whose explicit form depends on the chosen process. Moreover, a contingent claim written on these assets is specified by its terminal payoff $H_T \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}; \mathbb{R})$, and, when a time-indexed representation is needed, by $H_T := H_0 + \int_0^T dH_t$, $H_0 > 0$ is the initial price of the claim (constant) and $(H_t)_{t\in[0,T]}$ is an adapted process whose increments dH_t depend on the dynamics adopted for S.

The hedging problem requires the definition of a control strategy $h := (h_t)_{t\in[0,T]}$, where $h_t \in \mathbb{R}^n$ denotes the vector of units held in the n risky assets at time t. We say that h is predictable (with respect to $(\mathcal{F}_t)_{t\in[0,T]}$) if, for every $t > 0$, h_t is $\mathcal{F}_{t-}$–measurable, where $\mathcal{F}_{t-} := \sigma\Big(\bigcup_{u<t} \mathcal{F}_u\Big)$; equivalently, h is measurable with respect to the predictable σ–algebra. The set of admissible strategies, denoted by $\mathcal{H}$, must satisfy predictability and L^2–integrability

conditions. Formally, admissibility is defined as

$$\mathcal{H} := \Big\{\, h = (h_t)_{t\in[0,T]} \,\Big|\, h \text{ is predictable and } \mathbb{E}\Big[\int_0^T |h_t|^2\,dt\Big] < \infty \,\Big\}. \tag{1}$$

We also equip the space of admissible strategies $\mathcal{H}$ with the norm induced by the inner product,

$$\langle h_1, h_2\rangle_{\mathcal{H}} := \mathbb{E}\left[\int_0^T h_{1,t}^{\top} h_{2,t}\,dt\right], \quad h_1, h_2 \in \mathcal{H},$$

and the induced norm $\|h\|_{\mathcal{H}} := \mathbb{E}\big[\int_0^T h_t^2\,dt\big]^{1/2}$.

The set of achievable wealth $\mathcal{V}_t$ consists of all square–integrable random variables that can be generated by a self–financing trading strategy $h \in \mathcal{H}$. Formally,

$$\mathcal{V}_t := \Big\{\, V_t \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \,\Big|\, V_t = V_0 + \int_0^t h_s \cdot dS_s,\ h \in \mathcal{H},\ V_0 \in \mathbb{R} \,\Big\}, \tag{2}$$

where V_0 denotes the initial cost.

We define the profit-and-loss (P&L) process associated with h and the claim H as the adapted process $L_t^h := (L_t^h)_{t\in[0,T]}$ given by $L_t^h := V_t^h - H_t$. Using the dynamics of the wealth and of the claim, the infinitesimal variation of the P&L can be written as $dL_t^h := h_t \cdot dS_t - dH_t$, $t \in [0,T]$, which makes explicit that the instantaneous variation depends on the triplet (S, H, h). Under our standing assumptions that S and H are square–integrable semimartingales and $h \in \mathcal{H}$ is predictable and square–integrable, all the stochastic integrals above are well defined in $L^2(\Omega, \mathcal{F}_t, \mathbb{P})2$, and L_t^h is itself a square–integrable adapted process. By construction,

$$\begin{aligned} L_T^h &= V_T^h - H_T \\ &= \Big(V_0 + \int_0^T h_t\,dS_t\Big) - \Big(H_0 + \int_0^T dH_t\Big) \\ &= L_0 + \int_0^T dL_t^h. \end{aligned}$$

where $L_0 = V_0 - H_0$.

Assumption 1 (Canonical form of the P&L)**.** For every admissible strategy $h \in \mathcal{H}$, the profit-and-loss process L_t^h admits the decomposition

$$dL_t^h = b(t, S_t, H_t, h_t)\,dt + \eta(t, S_t, H_t, h_t)^\top\,dW_t + \xi(t, S_t, H_t, h_t)\,dN_t \tag{3}$$

where W is an m-dimensional Brownian motion and N is a Poisson process. The coefficients may depend on additional latent factors, for example, ν that accounts for stochastic volatility or J for the jump component when required.

Given L_0 is constant, it does not affect the choice of an optimal strategy in the problems considered below. Therefore, we may ignore this deterministic shift and focus on the random component of the terminal P&L. Considering the previously described market, we proceed with definitions regarding the conditions under which we will operate.

Definition 1 (Market completeness)**.** *We say that the market is* complete *if every contingent claim H_T is replicable by some admissible self-financing strategy $h \in \mathcal{H}$. In our framework, this means that*

$$L_T^h = 0 \quad \mathbb{P}\text{-}a.s.$$

We also say that the market is incomplete *if this property fails, that is, if there is no $h \in \mathcal{H}$ for every contingent claim H_T.*

Since we work in incomplete markets, the terminal P&L L_T^h differs from zero. Consequently, it is relevant to analyze the structural properties of the admissible strategy set $\mathcal{H}$, as this set determines the attainable terminal wealth and the P&L generated by trading, providing the necessary conditions for optimization and stability of $h \in \mathcal{H}$ developed in the sequel.

Proposition 1. *Recall that*

$$\mathcal{H} := \left\{\, h = (h_t)_{t\in[0,T]} \,\middle|\, h \text{ is predictable and } \mathbb{E}\Big[\int_0^T |h_t|^2\,dt\Big] < \infty \,\right\}$$

endowed with the norm $\|h\|_{\mathcal{H}}$. Then $\mathcal{H}$ is convex and closed.

Proof. Convexity follows directly from the definition. Let $h_1, h_2 \in \mathcal{H}$ and $\lambda \in [0,1]$, and set $h := \lambda h_1 + (1-\lambda)h_2$. Since h_1 and h_2 are predictable, so is h. Moreover,

$$|h_t|^2 = \left|\lambda h_{t,1} + (1-\lambda)h_{t,2}\right|^2 \le 2\big(\lambda^2|h_{t,1}|^2 + (1-\lambda)^2|h_{t,2}|^2\big) \le 2\big(|h_{t,1}|^2 + |h_{t,2}|^2\big),$$

and hence

$$\mathbb{E}\Big[\int_0^T |h_t|^2\,dt\Big] \le 2\,\mathbb{E}\Big[\int_0^T |h_{t,1}|^2\,dt\Big] + 2\,\mathbb{E}\Big[\int_0^T |h_{t,2}|^2\,dt\Big] < \infty,$$

so $h \in \mathcal{H}$.

To show closedness, let $(h^n)_{n\in\mathbb{N}} \subset \mathcal{H}$ be such that $|h^n - h| \to 0$ as $n \to \infty$ for some process h. Then, by the triangle inequality,

$$|h| \le |h - h^n| + |h^n|, \quad n \in \mathbb{N}.$$

Taking the limit superior as $n \to \infty$, we obtain

$$|h| \le \limsup_{n\to\infty} |h - h^n| + \limsup_{n\to\infty} |h^n| < \infty,$$

because $|h^n - h| \to 0$ and each $h^n \in \mathcal{H}$. Thus,

$$\mathbb{E}\Big[\int_0^T |h_t|^2\,dt\Big] < \infty,$$

so h is square-integrable. Since predictability is preserved under limits in this norm (we may choose a predictable version of h), it follows that h is predictable and belongs to $\mathcal{H}$. Therefore $\mathcal{H}$ is convex and closed. □

Proposition 2. *Consider the set of achievable wealth*

$$\mathcal{V}_t := \Big\{\, V_t \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \;\Big|\; V_t = V_0 + \int_0^t h_s \cdot dS_s,\ h \in \mathcal{H},\ V_0 \in \mathbb{R} \,\Big\}.$$

Then:

(i) $\mathcal{V}_t$ is a linear subspace of $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, i.e., for any $V_{t,1}, V_{t,2} \in \mathcal{V}_t$ and $a, b \in \mathbb{R}$,

$$aV_{t,1} + bV_{t,2} \in \mathcal{V}_t.$$

(ii) $\mathcal{V}_t$ is stable under deterministic shifts: if $V_t \in \mathcal{V}_t$ and $c \in \mathbb{R}$, then $V_t + c \in \mathcal{V}_t$.

In particular, $\mathcal{V}_t$ can be viewed as the space of square–integrable payoffs that are achievable by self–financing strategies. In a complete market one has $\mathcal{V}_T = L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, whereas in an incomplete market $\mathcal{V}_T$ is a proper linear subspace of $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

Proof. (i) Let $V_{t,1}, V_{t,2} \in \mathcal{V}_t$. By definition, there exist $V_{0,1}, V_{0,2} \in \mathbb{R}$ and $h_1, h_2 \in \mathcal{H}$ such that

$$V_{t,1} = V_{0,1} + \int_0^t h_{s,1} \cdot dS_s, \qquad V_{t,2} = V_{0,2} + \int_0^t h_{s,2} \cdot dS_s.$$

For arbitrary $a, b \in \mathbb{R}$, define

$$V_0 := aV_{0,1} + bV_{0,2}, \qquad h_s := ah_{s,1} + bh_{s,2}, \quad s \in [0, t].$$

Predictability and square–integrability are preserved under linear combinations, hence $h \in \mathcal{H}$. Moreover,

$$aV_{t,1} + bV_{t,2} = aV_{0,1} + bV_{0,2} + \int_0^t \big(ah_{s,1} + bh_{s,2}\big) \cdot dS_s = V_0 + \int_0^t h_s \cdot dS_s.$$

Thus, $aV_{t,1} + bV_{t,2} \in \mathcal{V}_t$, which shows that $\mathcal{V}_t$ is a linear subspace of $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$.

(ii) Let $V_t \in \mathcal{V}_t$ and $c \in \mathbb{R}$. Then there exist $V_0 \in \mathbb{R}$ and $h \in \mathcal{H}$ such that

$$V_t = V_0 + \int_0^t h_s \cdot dS_s.$$

Set $\widetilde{V}_0 := V_0 + c$ and keep the same strategy h. Then

$$V_t + c = \widetilde{V}_0 + \int_0^t h_s \cdot dS_s,$$

so $V_t + c \in \mathcal{V}_t$ by definition. This proves stability under deterministic shifts.

Finally, in a complete market every $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$–payoff is attainable, hence $\mathcal{V}_T = L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, while in an incomplete market $\mathcal{V}_T$ is a proper linear subspace of $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. □

Assumption 2 (Closed attainable wealth space)**.** We assume that the set of achievable wealth $\mathcal{V}_t$ is a closed linear subspace of $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$.

The results on admissible strategies $\mathcal{H}$ and achievable wealth $\mathcal{V}_t$ clarify which trading rules are allowed and which terminal payoffs can be reached by self–financing trading strategies. All optimization problems to be further formulated are written in terms of the terminal P&L, and this set is a translate of $\mathcal{V}_t$ in $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$. To ensure that optimal strategies exist and are well defined, we adopt the Assumption 2, which is consistent with our models where prices are square–integrable semimartingales and the stochastic integral satisfies an isometry property Øksendal [2003].

Remark 1. Note that our defined structure does not imply any condition, such as no-arbitrage or NFLVR, but only describes the analytical regularity of the strategy class.

Moreover, Proposition 2 ensures that the set $\mathcal{V}_t$ is a linear subspace of $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, thus for every $h \in \mathcal{H}$, the set $\mathcal{L}_t := \{L_t^h : h \in \mathcal{H}\}$ can also be viewed as a subspace of $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$. Giving $\mathcal{L}_t$ with the inner product inherited from $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$,

$$\langle V_1, V_2 \rangle_{\mathcal{V}_t} := \mathbb{E}[V_{1,t}\, V_{2,t}], \quad V_{1,t}, V_{2,t} \in \mathcal{V}_t,$$

$$\langle L_1^h, L_2^h \rangle_{\mathcal{L}_t} := \mathbb{E}[L_{1,t}^h\, L_{2,t}^h], \quad L_{1,t}^h, L_{2,t}^h \in \mathcal{L}_t,$$

and the associated norm $\|V\|_{\mathcal{V}_t} := \mathbb{E}\big[\int_0^t |V_s|^2\, ds\big]^{1/2}$, $\|L\|_{\mathcal{L}_t} := \mathbb{E}\big[\int_0^t |L_s^h|^2\, ds\big]^{1/2}$. With these choices, $\mathcal{L}_t$ is a Hilbert space.

Now, we fix $t \in [0, T]$ and consider the risk evaluation of square–integrable $\mathcal{F}_t$–measurable random variables. Accordingly, the space $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ is taken as the underlying domain, and X and Y denote generic elements of this space, which is endowed with its natural duality structure, identifying continuous linear functionals via the pairing $X \mapsto \mathbb{E}[XY]$. The weak topology $\sigma(L^2, L^2)$ is considered, and $\mathcal{Q}$ denotes the set of probability measures $\mathbb{Q}$ on $(\Omega, \mathcal{F}_t)$ that are equivalent to $\mathbb{P}$ and whose Radon–Nikodym densities $\frac{d\mathbb{Q}}{d\mathbb{P}}$ belong to it. $\mathcal{Q}$ is viewed as a convex subset of $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$.

Definition 2. *A convex risk measure is a function* $\rho : L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \to \mathbb{R}$ *that satisfies the following properties for all* $X, Y \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$*:*

i. ***Monotonicity****: If* $X \leq Y$ *almost surely, then* $\rho(X) \geq \rho(Y)$*.*

ii. ***Translation invariance****: For all* $m \in \mathbb{R}$*,* $\rho(X + m) = \rho(X) - m$*.*

iii. ***Convexity****: For all* $\lambda \in [0, 1]$*,* $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$*.*

iv. ***Lower semicontinuity in the weak topology****: If* $X_n \to X$ *in* $\sigma(L^2, L^2)$*, when* $n \to \infty$*, then* $\rho(X) \leq \liminf_{n\to\infty} \rho(X_n)$*.*

In the reflexive space $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, weak lower semicontinuity with respect to $\sigma(L^2, L^2)$ is equivalent to the usual Fatou property for convex risk measures, and this regularity is precisely what guaranties the existence of a robust dual representation. Combining Theorems 2.11 and 3.1 in Kaina and Rüschendorf [2009] yields the dual representation in the general setting of L^p for $1 < p < \infty$, and holds for $p = 2$.

$$\rho(X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \left\{ \mathbb{E}_{\mathbb{Q}}[-X] - \alpha_\rho(\mathbb{Q}) \right\}, \tag{4}$$

with penalty function

$$\alpha_\rho(\mathbb{Q}) = \sup_{X\in L^2} \left\{ \mathbb{E}_{\mathbb{Q}}[-X] - \rho(X) \right\}, \tag{5}$$

where $\mathcal{Q} := \left\{ \mathbb{Q} \sim \mathbb{P} : \frac{d\mathbb{Q}}{d\mathbb{P}} \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \right\}$.

In order to define compound risk measures, we first introduce the notion of a utility function.

Definition 3. *A function* $u : \mathbb{R} \to \mathbb{R}$ *is called a utility function if it satisfies the following properties:*

i. ***Continuity****:* u *is continuous on* $\mathbb{R}$*.*

ii. ***Monotonicity****:* u *is non-decreasing, i.e.,* $u'(x) \geq 0$ *for all* $x \in \mathbb{R}$*.*

iii. ***Concavity****:* u *is concave, i.e.,* $u''(x) \leq 0$ *for all* $x \in \mathbb{R}$*.*

Although utility functions are often normalized in the literature (for instance by imposing conditions such as $u(0) = 0$), these normalizations play no role in our analysis and do not affect the formulation of the optimization problem, so we do not impose them here.

Besides this, we consider the convex conjugate of the function $-u$ by $u^*(q) = \sup_{x\in\mathbb{R}}\{-qx - u(x)\}, \quad q \in \mathbb{R}$.

Following the theoretical framework in Righi [2024], we now define a composed risk measure obtained by combining a convex risk measure from definition 2 and a utility function from definition 3. The resulting composed risk measure is given by $\rho_u(X) := \rho(u(X))$, for all $X \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, and inherits monotonicity, convexity, and continuity in the weak topology $\sigma(L^2, L^2)$.

Moreover, the composed risk measure ρ_u admits a dual representation of the form

$$\rho_u(X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \big\{\mathbb{E}_{\mathbb{Q}}[-X] - \alpha_{\rho_u}(\mathbb{Q})\big\}, \quad X \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}), \tag{6}$$

and the associated penalty function is

$$\alpha_{\rho_u}(\mathbb{Q}) = \sup_{X\in L^2} \big\{\mathbb{E}_{\mathbb{Q}}[-X] - \rho(u(X))\big\}, \tag{7}$$

with $\mathcal{Q} := \big\{\mathbb{Q} \sim \mathbb{P} : \frac{d\mathbb{Q}}{d\mathbb{P}} \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})\big\}$. So, minimizing the composed risk ρ_u is equivalent to maximizing the associated variational-preferences functional.

Assumption 3 (Non-degeneracy)**.** There exists $\mathbb{Q} \in \mathcal{Q}$ such that $\alpha_{\rho_u}(\mathbb{Q}) < \infty$.

Assumption 3 implies that the effective domain of the penalty is non-empty. Consequently, the dual representation of ρ_u is non-degenerate and satisfies $\rho_u(X) > -\infty$ for all $X \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, ruling out the trivial case. This plays the same role as the no-irrelevance condition (e.g., $P(0) > -\infty$) used in the hedging setup of Righi [2024]. Based on this, we end the section after defining the hedging problem.

Definition 4 (Hedging problem)**.** *For a given contingent claim with terminal payoff H_T, the hedging problem consists in finding an admissible strategy $h^* \in \mathcal{H}$ such that*

$$h^* \in \operatorname{argmin}_{h\in\mathcal{H}} \rho_u(L_T^h), \tag{8}$$

where $L_T^h = V_T^h - H_T$ denotes the terminal profit-and-loss generated by the self-financing strategy h.

3 Proposed Approach

Throughout this section, we formalize the notion of MR and study the triple: properties, acceptance set and duality representation. This construction is inspired on the framework of Follmer and Schied [2016]. Since the focus is on the residual exposure carried from a given time t until maturity, rather than on instantaneous or recursive risk control, the relevant random variables, as H varies, are the residual P&L $L^h_{t,T}(H) := \int_{(t,T]} dL^h_s(H) \in \mathcal{L}t(H)$ with the convention $LT, T^h(H) = dL^h_T(H)$. For each fixed $t \in [0,T]$, the mapping $h \mapsto L^h_{t,T}(H)$ is affine.

Definition 5 (Maturity Risk)**.** *Let $\rho_u : L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \to \mathbb{R}$ be a composed risk measure. For each $t \in [0,T]$, the Maturity Risk is defined by*

$$MR(t,H) := \inf_{h \in \mathcal{H}} \rho_u(L^h_{t,T}(H)) - \inf_{h \in \mathcal{H}} \rho_u(L^h_{T,T}(H)).$$

The second term plays the role of a normalization. It isolates the purely temporal component of residual risk by removing the irreducible terminal exposure common to all strategies. As a consequence, $MR_t(H)$ measures the additional risk borne by maintaining the position from time t until maturity.

Definition 6 (Insensitivity to gains)**.** *A composed risk measure $\rho_u : L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \to \mathbb{R}$ is said to be insensitive to gains if for all $X, Y \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $Y \geq 0$ and $Y = 0$ almost surely on the scenarios that determine the risk of X, it holds that*

$$\rho_u(X + Y) = \rho_u(X).$$

Insensitivity to gains formalizes the idea that ρ_u reacts only to outcomes that are relevant for risk assessment. Positive payoffs occurring outside the risk-relevant scenarios—i.e., those that do not contribute to the evaluation of $\rho_u(X)$—have no impact on the measured risk. As a consequence, intermediate gains that do not affect the tail or adverse region identified by ρ_u cannot be exploited to artificially reduce the assessed risk.

Assumption 4 (Finite admissible risk)**.** For every $t \in [0,T]$, there exists at least one admissible strategy $\bar{h} \in \mathcal{H}$ such that $L^{\bar{h}}_{t,T}(H) \in \mathcal{L}_t$ satisfies $\rho_u(L^{\bar{h}}_{t,T}(H)) < +\infty$, and similarly $\rho_u(L^{\bar{h}}_{T,T}(H)) < +\infty$.

Theorem 1 (Properties of MR)**.** *Let $\rho_u : L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \to \mathbb{R}$ and consider $MR : L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \to \mathbb{R}$, then the following properties hold:*

i. *$MR(t, H)$ is well-defined and is finite for all $t \in [0,T]$.*

ii. *The mapping $H \to MR(t, H)$ is continuous with respect to the weak-topology $\sigma(L^2, L^2)$.*

iii. ***Normalization at maturity:*** *$MR(T, H) = 0$.*

iv. ***Monotonicity in the P&L:*** *For $L^h_{t,T}(H_1), L^h_{t,T}(H_2) \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $L^h_{t,T}(H_1) \geq L^h_{t,T}(H_2)$, for all $h \in \mathcal{H}$. If*

$$\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_2)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_1)\big) \leq \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_2)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_1)\big),$$

then

$$MR(t, H_1) \leq MR(t, H_2).$$

v. ***Monotonicity in the horizon:*** *If ρ_u satisfies insensitivity to gains, then for all $0 \leq t < t' \leq T$,*

$$MR_t(H) \geq MR_{t'}(H).$$

vi. ***Convexity:*** *Let $L^h_{t,T}(H_1), L^h_{t,T}(H_2) \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, for all $h \in \mathcal{H}$, and $\lambda \in [0,1]$. If $\inf_{h\in\mathcal{H}} \rho_u$ is affine at T, then*

$$MR(t, \lambda H_1 + (1-\lambda)H_2) \leq \lambda MR(t, H_1) + (1-\lambda) MR(t, H_2).$$

Assume in addition that $u : \mathbb{R} \to \mathbb{R}$ is affine and that $\rho : L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \to \mathbb{R}$ is a coherent risk measure. Then, ρ_u is coherent and $MR(t, H)$ satisfies:

vii. ***Cash-invariance:*** *for all $m \in \mathbb{R}$,*

$$MR(t, H + m) = MR(t, H).$$

viii. ***Sub-additivity:*** *for* $L^h_{t,T}(H_1), L^h_{t,T}(H_2) \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$,

$$MR(t, H_1 + H_2) \leq MR(t, H_1) + MR(t, H_2).$$

ix. ***Positive homogeneity:*** *for all* $\lambda \geq 0$,

$$MR(t, \lambda H) = \lambda MR(t, H).$$

Proof. i. Fix $t \in [0, T]$ and $H \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. For any $h \in \mathcal{H}$, $L^h_{t,T}(H) \in \mathcal{L}_t \subset L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, so $\rho_u(L^h_{t,T}(H))$ is well defined and, by Assumptions 3 and 4, satisfies $\inf_{h\in\mathcal{H}} \rho_u(L^h_{t,T}(H)) > -\infty$ and $\inf_{h\in\mathcal{H}} \rho_u(L^h_{t,T}(H)) < +\infty$. The same reasoning applies to $\inf_{h\in\mathcal{H}} \rho_u(L^h_{T,T}(H))$, which shows that both infima in the definition of $MR(t, H)$ are finite real numbers; hence, $MR(t, H)$ is well defined and finite.

ii. Let $H_n \rightharpoonup H$ in $\sigma(L^2, L^2)$. By Assumption 2, the infimum in the definition of $MR(t, H)$ is attained, so h^* is the optimal strategy. Then

$$MR(t, H_n) \leq \rho_u\big(L^{h^*}_{t,T}(H_n)\big) - \rho_u\big(L^{h^*}_{T,T}(H_n)\big) \longrightarrow \rho_u\big(L^{h^*}_{t,T}(H)\big) - \rho_u\big(L^{h^*}_{T,T}(H)\big) = MR(t, H),$$

since $H \mapsto L^h_{\cdot}(H)$ is affine and ρ_u is weakly continuous. Conversely, let h^*_n be optimal for H_n. By weak continuity of ρ_u,

$$MR(t, H) \leq \liminf_{n\to\infty} \Big(\rho_u\big(L^{h^*_n}_{t,T}(H)\big) - \rho_u\big(L^{h^*_n}_{T,T}(H)\big)\Big) = \liminf_{n\to\infty} MR(t, H_n).$$

Hence $MR(t, H_n) \to MR(t, H)$.

iii. From the definition of MR, we obtain

$$MR(T, H) = \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H)\big) = 0.$$

iv. From monotonicity of ρ_u, we have that

$$\begin{aligned} L^h_{t,T}(H_1) \geq L^h_{t,T}(H_2) \;&\Rightarrow\; \rho_u\big(L^h_{t,T}(H_1)\big) \leq \rho_u\big(L^h_{t,T}(H_2)\big) \\ &\Rightarrow\; \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_1)\big) \leq \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_2)\big). \end{aligned}$$

Given the theorem hypothesis of bounded variation at maturity,

$$\begin{aligned} &\inf_h \rho_u\big(L^h_{T,T}(H_2)\big) - \inf_h \rho_u\big(L^h_{T,T}(H_1)\big) \leq \inf_h \rho_u\big(L^h_{t,T}(H_2)\big) - \inf_h \rho_u\big(L^h_{t,T}(H_1)\big) \\ (\implies)\ &\inf_h \rho_u\big(L^h_{t,T}(H_1)\big) - \inf_h \rho_u\big(L^h_{T,T}(H_1)\big) \leq \inf_h \rho_u\big(L^h_{t,T}(H_2)\big) - \inf_h \rho_u\big(L^h_{T,T}(H_2)\big) \\ (\implies)\ &MR(t,H_1) \leq MR(t,H_2). \end{aligned}$$

v. Since ρ_u is convex, we can write

$$\inf_{h\in\mathcal{H}} \rho_u\big(\lambda L^h_{t,T}(H_1) + (1-\lambda)L^h_{t,T}(H_2)\big) \leq \lambda \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_1)\big) + (1-\lambda) \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_2)\big),$$

Here, the assumption that ρ_u is affine at time T ensures there is no penalization of mixture, so the cost is additive at maturity. Thus,

$$\begin{aligned} \inf_{h\in\mathcal{H}} \rho_u\big(\lambda L^h_{T,T}(H_1) + (1-\lambda)L^h_{T,T}(H_2)\big) &= \lambda \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_1)\big) + (1-\lambda) \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_2)\big), \\ (\implies) \quad MR\big(t, \lambda H_1 + (1-\lambda)H_2\big) &= \inf_{h\in\mathcal{H}} \rho_u\big(\lambda L^h_{t,T}(H_1) + (1-\lambda)L^h_{t,T}(H_2)\big) \\ &\quad - \inf_{h\in\mathcal{H}} \rho_u\big(\lambda L^h_{T,T}(H_1) + (1-\lambda)L^h_{T,T}(H_2)\big) \\ &\leq \lambda\Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_1)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_1)\big)\Big) \\ &\quad + (1-\lambda)\Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_2)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_2)\big)\Big) \\ &= \lambda MR(t,H_1) + (1-\lambda) MR(t,H_2). \end{aligned}$$

vi. Fix $0 \leq t < t' \leq T$. For any admissible strategy $h \in \mathcal{H}$, the P&L decomposition gives

$$L^h_{t,T}(H) = L^h_{t,t'}(H) + L^h_{t',T}(H), \qquad L^h_{t,t'}(H) \geq 0.$$

By the insensitivity to gains property, adding a non-negative payoff that does not affect the risk-relevant scenarios cannot reduce the risk measure. Hence,

$$\rho_u\big(L^h_{t,T}(H)\big) \geq \rho_u\big(L^h_{t',T}(H)\big), \qquad \forall\, h \in \mathcal{H}.$$

By monotonicity of ρ_u, it follows that

$$\rho_u\big(L^h_{t,T}(H)\big) \geq \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t',T}(H)\big), \qquad \forall\, h \in \mathcal{H}.$$

Taking the infimum over $h \in \mathcal{H}$ yields

$$\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H)\big) \geq \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t',T}(H)\big).$$

Subtracting the common normalization term $\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H)\big)$, we obtain

$$MR(t, H) \geq MR(t', H).$$

vii. Here, we use the cash-additivity property of ρ_u obtained by an u affine,

$$\begin{aligned} MR(t, H+m) &= \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H)+m\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H)+m\big) \\ &= \inf_{h\in\mathcal{H}} \Big(\rho_u\big(L^h_{t,T}(H)\big) - m\Big) - \inf_{h\in\mathcal{H}} \Big(\rho_u\big(L^h_{T,T}(H)\big) - m\Big) \\ &= \Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H)\big) - m\Big) - \Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H)\big) - m\Big) \\ &= \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H)\big) \\ &= MR(t, H). \end{aligned}$$

viii. Directly from sub-additivity, we get

$$\begin{aligned}
MR(t, H_1 + H_2) &= \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_1) + L^h_{t,T}(H_2)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_1) + L^h_{T,T}(H_2)\big) \\
&\le \inf_{h\in\mathcal{H}} \Big(\rho_u\big(L^h_{t,T}(H_1)\big) + \rho_u\big(L^h_{t,T}(H_2)\big)\Big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_1) + L^h_{T,T}(H_2)\big) \\
&\le \inf_{h\in\mathcal{H}} \Big(\rho_u\big(L^h_{t,T}(H_1)\big) + \rho_u\big(L^h_{t,T}(H_2)\big)\Big) - \inf_{h\in\mathcal{H}} \Big(\rho_u\big(L^h_{T,T}(H_1)\big) + \rho_u\big(L^h_{T,T}(H_2)\big)\Big) \\
&= \inf_{h\in\mathcal{H}} \Big(\rho_u\big(L^h_{t,T}(H_1)\big) + \rho_u\big(L^h_{t,T}(H_2)\big)\Big) \\
&\qquad - \Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_1)\big) + \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_2)\big)\Big) \\
&\le \Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_1)\big) + \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_2)\big)\Big) \\
&\qquad - \Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_1)\big) + \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_2)\big)\Big) \\
&= \Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_1)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_1)\big)\Big) \\
&\qquad + \Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H_2)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H_2)\big)\Big) \\
&= MR(t, H_1) + MR(t, H_2).
\end{aligned}$$

ix. Also, from definition of the property,

$$\begin{aligned}
MR(t, \lambda H) &= \inf_{h\in\mathcal{H}} \rho_u\big(\lambda L^h_{t,T}(H)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(\lambda L^h_{T,T}(H)\big) \\
&= \inf_{h\in\mathcal{H}} \big(\lambda \rho_u\big(L^h_{t,T}(H)\big)\big) - \inf_{h\in\mathcal{H}} \big(\lambda \rho_u\big(L^h_{T,T}(H)\big)\big) \\
&= \lambda \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H)\big) - \lambda \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H)\big) \\
&= \lambda\Big(\inf_{h\in\mathcal{H}} \rho_u\big(L^h_{t,T}(H)\big) - \inf_{h\in\mathcal{H}} \rho_u\big(L^h_{T,T}(H)\big)\Big) \\
&= \lambda MR(t, H).
\end{aligned}$$

□

The properties above hold for a fixed horizon t. There is no reason to expect sub-additivity or positive homogeneity to be preserved when comparing maturity risks across different horizons, $MR(t, H)$ and $MR(t', H)$ are defined on different spaces and the residual profit $L^h_{t,T}(H), L^h_{t',T}(H)$ itself depends on t and t' respectively.

Remark 2. The monotonicity in the horizon of MR is not an intrinsic property, but depends on the features of the chosen risk measure. In particular, it holds only when the evaluation criterion is insensitive gains, as is the case with tail-based risk measures, and may fail when measures react to gains outside risk-relevant scenarios. This highlights that temporal ordering is determined by the evaluation function rather than market dynamics alone. Furthermore, MR is not a dynamic risk measure in the sense of time consistency or recursivity; instead, it provides a static quantification of the irreducible exposure left after all admissible hedging strategies have been applied.

We now associate MR with a collection of acceptance sets. The goal is not to reinterpret the risk measure, but rather to determine which positions are acceptable after considering admissible hedging strategies. Acceptability is thus determined endogenously by the market and available trading opportunities.

Definition 7 (Acceptance sets). *Let $MR : L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \to \mathbb{R}$. For each level $c \in \mathbb{R}$, the induced acceptance set associated with $MR(t, H)$ is defined as*

$$\mathcal{A}(t,c)_{MR} := \Big\{ H \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}) : MR(t,H) \leq c \Big\}. \tag{9}$$

The set $\mathcal{A}(t,c) := \mathcal{A}(t,c)_{MR}$ includes all claims that can be traded from time t to maturity and whose risk, measured relative to the terminal benchmark, does not exceed the level c when hedged. In this sense, acceptability is governed by both the risk criterion and the market structure.

Theorem 2 (Properties of the acceptance sets). *Consider $MR : L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \to \mathbb{R}$ and let $\mathcal{A}(t,c)$ be the associated acceptance set.Then, the following properties are satisfied:*

i. ***Non-emptiness and closedness:*** *There exists $c_0 \in \mathbb{R}$ such that $\mathcal{A}(t, c_0) \neq \varnothing$. For any $c \in \mathbb{R}$, $\mathcal{A}(t,c)$ is weakly closed.*

ii. ***Monotonicity:*** *$MR(t, H_1) \leq MR(t, H_2)$, for all $H_1 \geq H_2$, if and only if $H_2 \in \mathcal{A}(t,c) \implies H_1 \in \mathcal{A}(t,c)$.*

iii. ***Temporal inclusion:*** *$MR(t,H) \geq MR(t', H)$, for any $0 \leq t \leq t' \leq T$, if and only if*

$\mathcal{A}(t,c)$ respects temporal inclusion, that is,

$$\mathcal{A}(t,c) \subseteq \mathcal{A}(t',c), \qquad 0 \le t < t' \le T.$$

iv. *MR can be recovered from A*

$$MR_{\mathcal{A}}(t,H) = \inf\{c \in \mathbb{R} : \; H \in \mathcal{A}(t,c)\}. \tag{10}$$

v. ***Convexity:*** *MR is convex if and only if $\mathcal{A}(t,c)$ is a convex set.*

vi. ***Positive homogeneity and sub-additivity:*** *MR is positively homogeneous if and only if $\mathcal{A}(t,0)$ is a cone. In particular, MR is sub-additive if and only if $\mathcal{A}(t,c)$ is convex cone for all $c \in \mathbb{R}$.*

vii. *If $\mathcal{A}(t,c)$ is convex, then $MR_{\mathcal{A}}$ is convex.*

viii. *If $\mathcal{A}(t,0)$ is a cone, then $MR_{\mathcal{A}}$ is positively homogeneous. In particular, if $\mathcal{A}(t,0)$ is a convex cone, then $MR_{\mathcal{A}}$ is sub-additive.*

ix. *$\mathcal{A}(t,c)_{MR_{\mathcal{A}}}$ is equal to $\overline{\mathcal{A}(t,c)}$.*

Proof. i. By finiteness of $MR(t,\cdot)$, there exists H_0 such that $MR(t,H_0) < +\infty$. Setting $c_0 := MR(t,H_0)$, $H_0 \in \mathcal{A}(t,c_0)$, thus $\mathcal{A}(t,c_0) \neq \varnothing$.

Let $H_n \in \mathcal{A}(t,c)$ with $H_n \rightharpoonup H$ in $\sigma(L^2,L^2)$. By weak continuity of $MR(t,\cdot)$, $MR(t,H) = \lim_{n\to\infty} MR(t,H_n) \le c$, thus $H \in \mathcal{A}(t,c)$. Therefore, $\mathcal{A}(t,c)$ is weakly closed.

ii. Assume $H_1 \ge H_2 \Rightarrow MR(t,H_1) \le MR(t,H_2)$. If $H_2 \in \mathcal{A}(t,c)$, $MR(t,H_1) \le MR(t,H_2) \le c \quad \Rightarrow \quad H_1 \in \mathcal{A}(t,c)$.

Conversely, suppose $H_2 \in \mathcal{A}(t,c) \Rightarrow H_1 \in \mathcal{A}(t,c)$ for all $c \in \mathbb{R}$. Taking $c := MR(t,H_2)$, $H_2 \in \mathcal{A}(t,c) \implies H_1 \in \mathcal{A}(t,c) \implies MR(t,H_1) \le MR(t,H_2)$.

iii. If $MR(t,H) \ge MR(t',H)$ for $0 \le t \le t' \le T$, then $H \in \mathcal{A}(t,c) \implies H \in \mathcal{A}(t',c)$.

Conversely, if $\mathcal{A}(t,c) \subseteq \mathcal{A}(t',c)$ for all $c \in \mathbb{R}$ when $0 \le t \le t' \le T$, then for $c := MR(t,H)$, we have $H \in \mathcal{A}(t,c) \implies MR(t',H) \le MR(t,H)$.

iv. $H \in \mathcal{A}(t,c) \iff MR(t,H) \le c$. Hence,

$$\{c \in \mathbb{R} : \; H \in \mathcal{A}(t,c)\} = \{c \in \mathbb{R} : \; MR(t,H) \le c\} = [MR(t,H), +\infty),$$

and

$$MR_{\mathcal{A}}(t,H) = \inf\{c \in \mathbb{R} : \; H \in \mathcal{A}(t,c)\} = MR(t,H).$$

v. Assume $MR(t,\cdot)$ is convex. Fix $c \in \mathbb{R}$ and let $H_1, H_2 \in \mathcal{A}(t,c)$. For $\lambda \in [0,1]$,

$$MR\big(t, \lambda H_1 + (1-\lambda)H_2\big) \le \lambda MR(t,H_1) + (1-\lambda)MR(t,H_2) \le c,$$

hence $\lambda H_1 + (1-\lambda)H_2 \in \mathcal{A}(t,c)$.

Conversely, assume $\mathcal{A}(t,c)$ is convex for all c. Fix H_1, H_2 and $\lambda \in [0,1]$. For $\varepsilon > 0$, choose

$$H_1 \in \mathcal{A}(t,c_1), \quad c_1 \le MR(t,H_1) + \varepsilon, \qquad H_2 \in \mathcal{A}(t,c_2), \quad c_2 \le MR(t,H_2) + \varepsilon.$$

Set $c := \lambda c_1 + (1-\lambda)c_2$. Then $H_1, H_2 \in \mathcal{A}(t,c)$ and, by convexity,

$$\lambda H_1 + (1-\lambda)H_2 \in \mathcal{A}(t,c) \Rightarrow MR\big(t, \lambda H_1 + (1-\lambda)H_2\big) \le c.$$

Letting $\varepsilon \downarrow 0$ yields convexity of $MR(t,\cdot)$.

vi. If $MR(t,\cdot)$ is positively homogeneous and $H \in \mathcal{A}(t,0)$,

$$MR(t,\lambda H) = \lambda MR(t,H) \le 0, \qquad \lambda \ge 0,$$

thus $\lambda H \in \mathcal{A}(t,0)$, i.e. $\mathcal{A}(t,0)$ is a cone. If MR is sub-additive, it is convex, and $\mathcal{A}(t,0)$ is a convex cone.

Conversely, if $\mathcal{A}(t,0)$ is a cone,

$$MR_{\mathcal{A}}(t,\lambda H) = \inf\{c : \; \lambda H \in \mathcal{A}(t,c)\} = \lambda MR_{\mathcal{A}}(t,H).$$

Also, sub-additivity follows from convexity.

vii. Consider $\mathcal{A}(t,c)$ is convex for all $c \in \mathbb{R}$. Fix H_1, H_2 and $\lambda \in [0,1]$. For $\varepsilon > 0$, choose

$$H_1 \in \mathcal{A}(t,c_1), \quad c_1 \le MR_{\mathcal{A}}(t,H_1) + \varepsilon, \qquad H_2 \in \mathcal{A}(t,c_2), \quad c_2 \le MR_{\mathcal{A}}(t,H_2) + \varepsilon.$$

Set $c = \lambda c_1 + (1-\lambda)c_2$. Then $H_1, H_2 \in \mathcal{A}(t,c)$ and

$$\lambda H_1 + (1-\lambda)H_2 \in \mathcal{A}(t,c) \Rightarrow MR_{\mathcal{A}}\big(t, \lambda H_1 + (1-\lambda)H_2\big) \leq c.$$

Letting $\varepsilon \downarrow 0$ yields convexity.

viii. Given $\mathcal{A}(t,0)$ is a cone. Fix H and $\lambda \geq 0$. Let $c > MR_{\mathcal{A}}(t,H)$. Then $H \in \mathcal{A}(t,c)$ and

$$MR(t,\lambda H) \leq \lambda c \Rightarrow \lambda H \in \mathcal{A}(t,\lambda c) \implies MR_{\mathcal{A}}(t,\lambda H) \leq \lambda c.$$

Taking the infimum over $c > MR_{\mathcal{A}}(t,H)$,

$$MR_{\mathcal{A}}(t,\lambda H) \leq \lambda MR_{\mathcal{A}}(t,H).$$

Conversely, let $d > MR_{\mathcal{A}}(t,\lambda H)$. Then $\lambda H \in \mathcal{A}(t,d)$ and, for $\lambda > 0$,

$$MR(t,H) \leq \frac{d}{\lambda} \implies MR_{\mathcal{A}}(t,H) \leq \frac{d}{\lambda} \implies \lambda MR_{\mathcal{A}}(t,H) \leq d.$$

Taking the infimum over d,

$$MR_{\mathcal{A}}(t,\lambda H) = \lambda MR_{\mathcal{A}}(t,H).$$

If $\mathcal{A}(t,0)$ is convex, fix H_1, H_2 and $\varepsilon > 0$. Choose

$$H_1 \in \mathcal{A}(t,c_1), \quad c_1 \leq MR_{\mathcal{A}}(t,H_1) + \varepsilon, \qquad H_2 \in \mathcal{A}(t,c_2), \quad c_2 \leq MR_{\mathcal{A}}(t,H_2) + \varepsilon.$$

Then,

$$H_1 + H_2 \in \mathcal{A}(t, c_1 + c_2) \Rightarrow MR_{\mathcal{A}}(t, H_1 + H_2) \leq c_1 + c_2.$$

Letting $\varepsilon \downarrow 0$ yields sub-additivity.

ix. Let $H \in \mathcal{A}(t,c)_{MR_{\mathcal{A}}}$. Then $MR_{\mathcal{A}}(t,H) \leq c$ and there exists $(H_n) \subset \mathcal{A}(t,c)$ such that$H_n \rightharpoonup H \quad$ in $\sigma(L^2, L^2)$. Since $\mathcal{A}(t,c)$ is weakly closed, $H \in \mathcal{A}(t,c)$.
Conversely, if $H \in \overline{\mathcal{A}(t,c)}$, there exists $(H_n) \subset \mathcal{A}(t,c)$ such that $H_n \rightharpoonup H$. By weak

continuity of $MR_{\mathcal{A}}(t,\cdot)$, $MR_{\mathcal{A}}(t,H) = \lim_{n\to\infty} MR_{\mathcal{A}}(t,H_n) \le c$, so $H \in \mathcal{A}(t,c)_{MR_{\mathcal{A}}}$. Therefore, $\mathcal{A}(t,c)_{MR_{\mathcal{A}}} = \overline{\mathcal{A}(t,c)}$.

□

Theorem 3 (Dual representation of MR)**.** *Assume $MR(t,\cdot) : L^2(\Omega,\mathcal{F}_t,\mathbb{P}) \to \mathbb{R}$. Then, when u is non-affine, non-affine utility u, $MR(t,H)$ admits the dual representation*

$$MR(t,H) = \sup_{Q\ge 0}\Big\{\mathbb{E}[HQ] - \beta(t,Q)\Big\}, \qquad \beta(t,Q) = \sup_{H\in L^2(\Omega,\mathcal{F}_t,\mathbb{P})}\Big\{\mathbb{E}[HQ] - MR(t,H)\Big\}. \quad (11)$$

In the case where u is affine, cash-invariance implies that

$$MR(t,H) = -\beta(t,0) = \inf_{H\in L^2(\Omega,\mathcal{F}_t,\mathbb{P})} MR(t,H). \quad (12)$$

Proof. Fix t. Since $MR(t,\cdot) : L^2(\Omega,\mathcal{F}_t,\mathbb{P}) \to \mathbb{R}$ is convex and weakly continuous, the Fenchel-Moreau theorem yields

$$MR(t,H) = \sup_Q\{\mathbb{E}[HQ] - \beta(t,Q)\}, \qquad \beta(t,Q) = \sup_H\{\mathbb{E}[HQ] - MR(t,H)\}.$$

Monotonicity of MR implies $\beta(t,Q) = +\infty$ whenever $\mathbb{P}(Q < 0) > 0$, hence the supremum may be restricted to $Q \ge 0$.

If u is affine, then $MR(t,H+m) = MR(t,H)$ for all $m \in \mathbb{R}$. Substituting $H+m$ in the dual form gives

$$MR(t,H) = \sup_{Q\ge 0}\{\mathbb{E}[HQ] + m\,\mathbb{E}[Q] - \beta(t,Q)\} \quad \forall m,$$

which forces $\mathbb{E}[Q] = 0$ for every effective Q. Together with $Q \ge 0$, this implies $Q = 0$ a.s., and therefore

$$MR(t,H) = -\beta(t,0) = -\sup_{H\in L^2(\Omega,\mathcal{F}_t,\mathbb{P})}\{-MR(t,H)\} = \inf_{H\in L^2(\Omega,\mathcal{F}_t,\mathbb{P})} MR(t,H).$$

□

When u is affine, cash-invariance yields $MR(t,H+m) = MR(t,H)$ for all constants m. As a result, when a deterministic shift is applied to the reward, the value remains constant. For this reason, we identify payoffs by the relation $H + m \sim H$, which means that two elements are in the same class if they differ only by a constant. Then $MR(t,\cdot)$ is well-defined on the collection of equivalence classes $[H] = \{H + m : m \in \mathbb{R}\}$, and it depends only on $[H]$, not on the specific representative H. In this interpretation, taking $\inf_{H\in L^2(\Omega,\mathcal{F}_t,\mathbb{P})} MR(t,H)$ means taking the infimum over these classes, that is, selecting the class $[H]$ that minimizes the residual risk when constants are ignored.

4 Optimization

By Definitions 4 and 5, the claim H is associated with a maturity risk $MR(t)$, defined through the minimization of the composed risk measure ρ_u evaluated on the residual profit-and-loss $L^h_{t,T}$ over admissible strategies $h \in \mathcal{H}$. Accordingly, the optimization problem is formulated on the space of attainable P&L outcomes, viewed as a subspace $\mathcal{L}_t \subset L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, and ρ_u is considered through its restriction $\rho_u : \mathcal{L}_t \to \mathbb{R}$.

From now on, we regard this minimization as a stochastic control problem and aim to reconstruct the Stochastic Maximum Principle (SMP) for risk measures. Our formulation follows Peng [1990]; for a general review of stochastic control see Yong and Zhou [1999]. We denote by X_t the state vector collecting the model factors driving the P&L, including (S_t, H_t) and, when required by the model specification, stochastic volatility ν_t or jump J. The control variable is the trading strategy $h \in \mathcal{H}$. In this formulation, the stochastic differential equation in Assumption 1 is taken as the controlled state dynamics whose evolution depends on h through the drift, diffusion and, when presented, jump coefficients.

[First-order variational equations] Consider a family of perturbations $(h^\varepsilon)_\varepsilon$ given by $h^\varepsilon = h^* + \varepsilon v$, where $h^* \in \mathcal{H}$ is an admissible control and a candidate optimal control, and $v \in \mathcal{H}$ is an arbitrary direction. Assume that the coefficients of the system are Lipschitz continuous and sufficiently smooth in the state and control variables. Then the first-order variational processes

$$Y_t := \lim_{\varepsilon\to 0} \frac{X_t^{h^\varepsilon} - X_t^{h^*}}{\varepsilon}, \qquad Z_t := \lim_{\varepsilon\to 0} \frac{L_t^{h^\varepsilon} - L_t^{h^*}}{\varepsilon}$$

exist and satisfy the linear first-order variational equations

$$
\begin{aligned}
dY_t &= \partial_x \mu(t, X_t^{h^*}, h_t^*)\, Y_t\, dt + \partial_h \mu(t, X_t^{h^*}, h_t^*)\, v_t\, dt \\
&\quad + \partial_x \sigma(t, X_t^{h^*}, h_t^*)\, Y_t\, dW_t + \partial_h \sigma(t, X_t^{h^*}, h_t^*)\, v_t\, dW_t \\
&\quad + \partial_x J(t, X_t^{h^*}, h_t^*)\, Y_t\, dN_t + \partial_h J(t, X_t^{h^*}, h_t^*)\, v_t\, dN_t, \qquad Y_0 = 0,
\end{aligned} \tag{13}
$$

and

$$
\begin{aligned}
dZ_t &= \partial_x b(t, X_t^{h^*}, h_t^*)\, Y_t\, dt + \partial_h b(t, X_t^{h^*}, h_t^*)\, v_t\, dt \\
&\quad + \partial_x \eta(t, X_t^{h^*}, h_t^*)\, Y_t\, dW_t + \partial_h \eta(t, X_t^{h^*}, h_t^*)\, v_t\, dW_t \\
&\quad + \partial_x \xi(t, X_t^{h^*}, h_t^*)\, Y_t\, dN_t + \partial_h \xi(t, X_t^{h^*}, h_t^*)\, v_t\, dN_t, \qquad Z_0 = 0.
\end{aligned} \tag{14}
$$

Proof. By the Lipschitz continuity of the coefficients in (x, h), there exists a constant $K > 0$ such that, for all t and all $(x_1, h_1), (x_2, h_2)$,

$$
\begin{aligned}
&|\mu(t, x_1, h_1) - \mu(t, x_2, h_2)| + |\sigma(t, x_1, h_1) - \sigma(t, x_2, h_2)| + |J(t, x_1, h_1) - J(t, x_2, h_2)| \\
&\qquad \le K\big(|x_1 - x_2| + |h_1 - h_2|\big), \\
&|b(t, x_1, h_1) - b(t, x_2, h_2)| + |\eta(t, x_1, h_1) - \eta(t, x_2, h_2)| + |\xi(t, x_1, h_1) - \xi(t, x_2, h_2)| \\
&\qquad \le K\big(|x_1 - x_2| + |h_1 - h_2|\big).
\end{aligned}
$$

Let $(X^{h^\varepsilon}, L^{h^\varepsilon})$ and (X^{h^*}, L^{h^*}) be the state processes associated with h^ε and h^*, respectively. Subtracting the integral forms of the corresponding SDEs and dividing by ε, we obtain for each t

$$
\begin{aligned}
\frac{X_t^{h^\varepsilon} - X_t^{h^*}}{\varepsilon} &= \int_0^t \frac{\mu(s, X_s^{h^\varepsilon}, h_s^\varepsilon) - \mu(s, X_s^{h^*}, h_s^*)}{\varepsilon}\, ds \\
&\quad + \int_0^t \frac{\sigma(s, X_s^{h^\varepsilon}, h_s^\varepsilon) - \sigma(s, X_s^{h^*}, h_s^*)}{\varepsilon}\, dW_s \\
&\quad + \int_0^t \frac{J(s, X_s^{h^\varepsilon}, h_s^\varepsilon) - J(s, X_s^{h^*}, h_s^*)}{\varepsilon}\, dN_s,
\end{aligned}
$$

and

$$\frac{L_t^{h^\varepsilon} - L_t^{h^*}}{\varepsilon} = \int_0^t \frac{b(s, X_s^{h^\varepsilon}, h_s^\varepsilon) - b(s, X_s^{h^*}, h_s^*)}{\varepsilon}\, ds + \int_0^t \frac{\eta(s, X_s^{h^\varepsilon}, h_s^\varepsilon) - \eta(s, X_s^{h^*}, h_s^*)}{\varepsilon}\, dW_s + \int_0^t \frac{\xi(s, X_s^{h^\varepsilon}, h_s^\varepsilon) - \xi(s, X_s^{h^*}, h_s^*)}{\varepsilon}\, dN_s.$$

Using $h_s^\varepsilon = h_s^* + \varepsilon v_s$ and applying the mean value theorem in (x, h), for each fixed (s, ω) we write

$$\mu(s, X_s^{h^\varepsilon}, h_s^\varepsilon) - \mu(s, X_s^{h^*}, h_s^*) = \partial_x \mu(s, \bar{X}_s^\varepsilon, \bar{h}_s^\varepsilon)\big(X_s^{h^\varepsilon} - X_s^{h^*}\big) + \partial_h \mu(s, \bar{X}_s^\varepsilon, \bar{h}_s^\varepsilon)\big(h_s^\varepsilon - h_s^*\big),$$

dividing by ε and using $h_s^\varepsilon - h_s^* = \varepsilon v_s$ yields

$$\frac{\mu(s, X_s^{h^\varepsilon}, h_s^\varepsilon) - \mu(s, X_s^{h^*}, h_s^*)}{\varepsilon} = \partial_x \mu(s, \bar{X}_s^\varepsilon, \bar{h}_s^\varepsilon) \frac{X_s^{h^\varepsilon} - X_s^{h^*}}{\varepsilon} + \partial_h \mu(s, \bar{X}_s^\varepsilon, \bar{h}_s^\varepsilon)\, v_s,$$

with analogous expressions for the remaining coefficients σ, J, b, η, ξ. By the smoothness assumptions, the partial derivatives are continuous and bounded. Therefore, for any sequence of intermediate points $\bar{X}_s^\varepsilon, \bar{h}_s^\varepsilon$ the composition of $\partial_x b(s, X_s^{h^*}, h_s^*)$ and $\partial_h b(s, X_s^{h^*}, h_s^*)$ (including the other coefficients) are uniformly bounded in ϵ and converge, as $\epsilon \to 0$, to the corresponding derivatives evaluated at $(X_s^{h^*}, h_s^*)$. Substituting these expressions into the integral equations for the differentiated quotients

$$\Delta X_t^\varepsilon := \frac{X_t^{h^\varepsilon} - X_t^{h^*}}{\varepsilon}, \qquad \Delta L_t^\varepsilon := \frac{L_t^{h^\varepsilon} - L_t^{h^*}}{\varepsilon}$$

satisfies a linear system $(\Delta X_t^{\epsilon}, \Delta L_t^{\epsilon})$. For each $t \in [0,T]$, the limits

$$Y_t := \lim_{\varepsilon \to 0} \frac{X_t^{h^\varepsilon} - X_t^{h^*}}{\varepsilon}, \qquad Z_t := \lim_{\varepsilon \to 0} \frac{L_t^{h^\varepsilon} - L_t^{h^*}}{\varepsilon}$$

exist on the appropriate space. Passing to the limit in the linearized equation system leads exactly to the stated first-order variational equations

$$\begin{aligned} dY_t &= \partial_x \mu(t, X_t^{h^*}, h_t^*)\, Y_t\, dt + \partial_h \mu(t, X_t^{h^*}, h_t^*)\, v_t\, dt \\ &\quad + \partial_x \sigma(t, X_t^{h^*}, h_t^*)\, Y_t\, dW_t + \partial_h \sigma(t, X_t^{h^*}, h_t^*)\, v_t\, dW_t \\ &\quad + \partial_x J(t, X_t^{h^*}, h_t^*)\, Y_t\, dN_t + \partial_h J(t, X_t^{h^*}, h_t^*)\, v_t\, dN_t, \qquad Y_0 = 0, \end{aligned}$$

and

$$\begin{aligned} dZ_t &= \partial_x b(t, X_t^{h^*}, h_t^*)\, Y_t\, dt + \partial_h b(t, X_t^{h^*}, h_t^*)\, v_t\, dt \\ &\quad + \partial_x \eta(t, X_t^{h^*}, h_t^*)\, Y_t\, dW_t + \partial_h \eta(t, X_t^{h^*}, h_t^*)\, v_t\, dW_t \\ &\quad + \partial_x \xi(t, X_t^{h^*}, h_t^*)\, Y_t\, dN_t + \partial_h \xi(t, X_t^{h^*}, h_t^*)\, v_t\, dN_t, \qquad Z_0 = 0. \end{aligned}$$

□

Y_t is the first–order response of the state X_t to a directional perturbation v, while Z_t is the corresponding first–order variation of the P&L. Hence Y_t depends only on v, whereas Z_t depends on both v and Y_t, reflecting the indirect influence of the control on L_t through the state.

Let $\rho_u : \mathcal{L}_t \to \mathbb{R}$ be a composed risk measure. Then, ρ_u is Gâteaux differentiable at $L_{t,T}^h$ if:

i. $u : \mathbb{R} \to \mathbb{R}$ is differentiable a.e. and $u'(L_{t,T}^h)Z, \quad \forall Z \in \mathcal{L}_t$;

ii. ρ is Gâteaux differentiable at the point $u(L_{t,T}^h)$.

Proof. Let $Z \in \mathcal{L}_t$ and consider the directional increment $\frac{\rho_u(L_{t,T}^h + \varepsilon Z) - \rho_u(L_{t,T}^h)}{\varepsilon} = \frac{\rho\big(u(L_{t,T}^h + \varepsilon Z)\big) - \rho\big(u(L_{t,T}^h)\big)}{\varepsilon}$. By condition (i), $\frac{u(L_{t,T}^h + \varepsilon Z) - u(L_{t,T}^h)}{\varepsilon} \to u'(L_{t,T}^h)Z$ in $\mathcal{L}_t$ as $\varepsilon \to 0$. Hence, by condition (ii), ρ is Gâteaux differentiable at $u(L_{t,T}^h)$, the above limit exists and is linear in Z. Thus, ρ_u is

Gâteaux differentiable at $L^h_{t,T}$. □

Assume that the mapping $h \mapsto L^h_T$ is also Gâteaux differentiable at h^* and $\rho_u : \mathcal{L}_t \to \mathbb{R}$ is Gâteaux differentiable at $L^{h^*}_T$. Then there exists a unique $\Upsilon \in \mathcal{L}_t$ such that, for every $v \in \mathcal{H}$,

$$\rho'_u(h^*)[v] = D\rho_u(L^{h^*}_T)[\Gamma v] = \langle \Upsilon, \Gamma v\rangle_{\mathcal{L}_t} = \mathbb{E}[\Upsilon\, \Gamma v]. \tag{15}$$

Here, $\Gamma v := DL_T(h^*)[v]$ denotes the directional derivative of L^h_T at h^* in the direction v.

Proof. Assume that $\rho_u : \mathcal{L}_t \to \mathbb{R}$ is Gâteaux differentiable at $L^{h^*}_T$. Then, $D\rho_u(L^{h^*}_T) : \mathcal{L}_t \to \mathbb{R}$ is a continuous linear functional. Since $\mathcal{L}_t$ is a Hilbert space with inner product $\langle Z_1, Z_2\rangle_{\mathcal{L}_t} := \mathbb{E}[Z_1 Z_2]$, the Riesz representation theorem yields the existence of a unique $\Upsilon \in \mathcal{L}_t$ such that, for all $Z_T \in \mathcal{L}_t$,

$$D\rho_u(L^{h^*}_T)[Z_T] = \langle \Upsilon, Z_T\rangle_{\mathcal{L}_t} = \mathbb{E}[\Upsilon Z_T]. \tag{16}$$

Assume next that the mapping $h \mapsto L^h_T$ is Gâteaux differentiable at h^*. Denote its derivative by $DL_T(h^*)[v] = \Gamma v$, where $\Gamma : \mathcal{H} \to \mathcal{L}_t$ is linear and continuous. For $v \in \mathcal{H}$ and $h^\varepsilon = h^* + \varepsilon v$, the terminal value admits the expansion

$$L^{h^\varepsilon}_T = L^{h^*}_T + \varepsilon\, \Gamma v + o(\varepsilon), \qquad \frac{o(\varepsilon)}{\varepsilon} \to 0 \text{ in } \mathcal{L}_t. \tag{17}$$

Since ρ_u is Gãteaux differentiable at $L^{h^*}_T$,

$$\rho_u(L^{h^\varepsilon}_T) = \rho_u(L^{h^*}_T) + D\rho_u(L^{h^*}_T)\left[L^{h^\varepsilon}_T - L^{h^*}_T\right] + o(\epsilon)\,. \tag{18}$$

Substituting the expansion of $L^{h^\varepsilon}_T - L^{h^*}_T$, and using linearity and continuity of $D\rho_u(L^{h^*}_T)$,

$$\rho_u(L^{h^\varepsilon}_T) = \rho_u(L^{h^*}_T) + \varepsilon\, D\rho_u(L^{h^*}_T)[\Gamma v] + o(\varepsilon). \tag{19}$$

Dividing by ε and letting $\varepsilon \to 0$, the directional derivative of the mapping $h \mapsto \rho_u(L^h_T)$ at h^* in the direction v exists and satisfies

$$\rho'_u(h^*)[v] := \lim_{\varepsilon\to 0} \frac{\rho_u(L^{h^\varepsilon}_T) - \rho_u(L^{h^*}_T)}{\varepsilon} = D\rho_u(L^{h^*}_T)[\Gamma v]. \tag{20}$$

Using the Riesz representation of $D\rho_u(L_T^{h^*})$,

$$\rho_u'(h^*)[v] = \langle \Upsilon, \Gamma v \rangle_{\mathcal{L}_t} = \mathbb{E}[\Upsilon\, \Gamma v]. \tag{21}$$

□

In simple terms, Υ assesses how the risk functional ρ_u reacts to tiny changes in the terminal payoff L_T, while the operator Γ shows how a perturbation of the strategy h propagates into a variation of L_T. The adjoint operator Γ^* reverses the action and returns a terminal sensitivity to the space of strategies. To achieve a local-in-time optimality requirement, the terminal sensitivity must be expressed as a time integral, which necessitates the use of adjoint processes. Since the functional $\rho_u(L_T^h)$ depends only on the terminal value L_T^h, the adjoint associated with the state variable X^h has zero terminal condition and is therefore identically zero. Hence, it is sufficient to introduce a single non-trivial adjoint associated with the equation of L^h.

Denote by $(p_t, q_t, r_t)_{t\in[0,T]}$ the triple of adapted processes solving

$$\begin{cases} dp_t = -\Big(\partial_x b(t, X_t^{h^*}, h_t^*)\, p_t + \partial_x \eta(t, X_t^{h^*}, h_t^*)\, q_t + \partial_x \xi(t, X_t^{h^*}, h_t^*)\, r_t\Big)\, dt + q_t\, dW_t + r_t\, dN_t, \\ p_T = \Xi, \end{cases} \tag{22}$$

where $\Xi \in \mathcal{L}_t$. If (Y, Z) is the solutions of the variational system of equations associated with $h^\varepsilon \in \mathcal{H}$, then, for every $v \in \mathcal{H}$,

$$\mathbb{E}[\Xi Z_T] = \mathbb{E}\int_0^T \Big(\partial_h b(t, X_t^{h^*}, h_t^*)\, p_t + \partial_h \eta(t, X_t^{h^*}, h_t^*)\, q_t + \partial_h \xi(t, X_t^{h^*}, h_t^*)\, r_t\Big)\, v_t\, dt. \tag{23}$$

In particular, if the adjoint is required to represent the derivative of ρ_u at h^*, its terminal condition can be chosen as $p_T = \Xi = \Upsilon$.

Proof. Let $v \in \mathcal{H}$ and let (Y, Z) be the first-order variations associated with the perturbation $h^\varepsilon = h^* + \varepsilon v$, as given by the variational system of equations. Fix $\Xi \in \mathcal{L}_t$ and let (p, q, r) be the adapted solution of the adjoint equation with terminal condition $p_T = \Xi$.

Consider the product process $p_t Z_t$ on the interval $[0, T]$. By Itô's formula, its differential is given by

$$d(p_t Z_t) = p_t \, dZ_t + Z_t \, dp_t + d[p, Z]_t,$$

where $d[p, Z]_t$ denotes the quadratic covariation. Substituting the dynamics of Z_t from the variational equation and the dynamics of p_t from the adjoint equation, we group the drift terms. The terms involving Y_t cancel due to the choice of the adjoint coefficients $\partial_x b$, $\partial_x \eta$ and $\partial_x \xi$. The remaining drift term is linear in the perturbation v_t and reads

$$\big(\partial_h b(t, X_t^{h^*}, h_t^*) \, p_t + \partial_h \eta(t, X_t^{h^*}, h_t^*) \, q_t + \partial_h \xi(t, X_t^{h^*}, h_t^*) \, r_t\big) \, v_t \, dt.$$

Integrating from 0 to T and taking expectations, the stochastic integral terms have zero expectation by square-integrability. Since $Z_0 = 0$, we obtain

$$\mathbb{E}[p_T Z_T] = \mathbb{E} \int_0^T \big(\partial_h b(t, X_t^{h^*}, h_t^*) \, p_t + \partial_h \eta(t, X_t^{h^*}, h_t^*) \, q_t + \partial_h \xi(t, X_t^{h^*}, h_t^*) \, r_t\big) \, v_t \, dt.$$

This establishes the duality relation with terminal condition $p_T = \Xi$.

On the other hand, the directional derivative of the risk functional at h^* satisfies $\rho_u'(h^*)[v] = \mathbb{E}[\Upsilon Z_T]$. If the adjoint is required to represent this derivative, then $\mathbb{E}[\Xi Z_T] = \mathbb{E}[\Upsilon Z_T]$ for all $Z_T = \Gamma v$. Since $\mathbb{E}\big[(\Xi - \Upsilon) \, \Gamma v\big] = 0 \implies \Gamma^* \Xi = \Gamma^* \Upsilon$. In particular, the adjoint terminal condition may be chosen as $p_T = \Xi = \Upsilon$, without loss of generality, since only $\Gamma^* \Xi$ enters the optimality condition. □

The Lemma 4 shows that the adjoint equation 22, with terminal condition chosen as $\Xi = \Upsilon$, is in perfect duality with the variational system. The product $p_t Z_t$ plays the role of a "stochastic Lagrange bracket" that turns the terminal functional $\mathbb{E}[\Upsilon Z_T]$ into a time integral whose integrand is proportional to v_t.

Theorem 4 (Stochastic Maximum Principle for ρ_u)**.** *Assume that the mapping $h \mapsto \rho_u(L_T^h)$ is convex and Gâteaux differentiable at h^*. Assume that the controlled state (X^h, L^h) is well defined and that the coefficients are Lipschitz continuous and sufficiently smooth. Define the*

Hamiltonian

$$H(t,x,h,p,q,r) := p\,b(t,x,h) + q\,\eta(t,x,h) + r\,\xi(t,x,h). \tag{24}$$

Let (p,q,r) be the adjoint process associated with (X^{h^*}, L^{h^*}) with terminal condition $p_T = \Upsilon$. If h^* is optimal, hence

$$\partial_h H(t, X_t^{h^*}, h_t^*, p_t, q_t, r_t) = 0, \quad \textit{for a.e. } t \in [0,T],\ \mathbb{P}\textit{-a.s.}. \tag{25}$$

Proof. Fix $v \in \mathcal{H}$ and consider the perturbation $h^\varepsilon = h^* + \varepsilon v$. By Propositions 4 and 4, the mapping $h \mapsto \rho_u(L_T^h)$ is Gâteaux differentiable at h^* and there exists $\Upsilon \in \mathcal{L}_t$ such that $\rho_u'(h^*)[v] = \mathbb{E}[\Upsilon\, Z_T]$, where Z_T denotes the terminal first-order variation associated with thedirection v.

Let $(p_t, q_t, r_t)_{t\in[0,T]}$ be the adjoint process defined in Proposition 4 with terminal condition $p_T = \Upsilon$. So, by the duality relation of Proposition 4, for every $v \in \mathcal{H}$,

$$\mathbb{E}[\Upsilon\, Z_T] = \mathbb{E}\int_0^T \Big(\partial_h b(t, X_t^{h^*}, h_t^*)\, p_t + \partial_h \eta(t, X_t^{h^*}, h_t^*)\, q_t + \partial_h \xi(t, X_t^{h^*}, h_t^*)\, r_t\Big)\, v_t\, dt.$$

Combining the two identities gives us

$$\rho_u'(h^*)[v] = \mathbb{E}\left[\int_0^T \Big(\partial_h b(t, X_t^{h^*}, h_t^*)\, p_t + \partial_h \eta(t, X_t^{h^*}, h_t^*)\, q_t + \partial_h \xi(t, X_t^{h^*}, h_t^*)\, r_t\Big)\, v_t\, dt\right].$$

From the definition of the Hamiltonian, the integrand coincides with $\partial_h H(t, X_t^{h^*}, h_t^*, p_t, q_t, r_t)\, v_t$. Hence,

$$\rho_u'(h^*)[v] = \mathbb{E}\left[\int_0^T \partial_h H\big(t, X_t^{h^*}, h_t^*, p_t, q_t, r_t\big)\, v_t\, dt\right]$$

Assume now that h^* minimizes the convex functional $h \mapsto \rho_u(L_T^h)$ over $\mathcal{H}$. Hence, $\rho_u'(h^*)[v] \geq 0$ and since the Gâteaux derivative is linear in the direction, $\rho_u'(h^*)[-v] = -\rho_u'(h^*)[v] \implies \rho_u'(h^*)$. We substitute this into the previous representation to obtain

$$\mathbb{E}\left[\int_0^T \partial_h H\big(t, X_t^{h^*}, h_t^*, p_t, q_t, r_t\big)\, v_t\, dt\right] = 0$$

$$(\implies) \partial_h H\big(t, X_t^{h^*}, h_t^*, p_t, q_t, r_t\big) = 0, \quad \text{for a.e. } t \in [0,T],\ \mathbb{P}\text{-a.s.}.$$

□

5 Examples

Next, we showcase three models for the underlying asset price: the Merton Jump-Diffusion model, the Heston Stochastic Volatility model, and Geometric Brownian Motion (GBM). In these examples, we take into consideration a smooth payoff function $H(t, S_t)$ (e.g., $H(T, S_T) = \max(S_T - K, 0), \quad K > 0$, for a European call), which guarantees the application of Itô's lemma and permits a canonical representation of the $P\&L$ dynamics in accordance with Assumption 1. The notation and general structure remain identical across the three examples, with differences arising only from the characteristics of the underlying stochastic process. For each model, the corresponding first–order optimality conditions are derived directly within the example.

The composed risk measures are then created and used for all stochastic specifications. The linked terminal condition is common to all three processes and only needs to be specified once because the evaluation criterion is applied to the terminal P&L and is independent of the asset dynamics. Three risk measures — Expected Loss (EL), Entropic Risk Measure (ERM), and Expected Shortfall (ES) — as well as linear, exponential, and Weibull-type utility functions are taken into account in the study. Because they don't meet the necessary requirements, compositions containing ES with exponential or Weibull-type utilities are discarded. EL with linear utility is a special case among the remaining specifications, resulting in a degenerate optimal control. In total, the combination of three stochastic models with six composed risk measures gives rise to eighteen distinct problems.

Assumption 5 (Regularity of coefficients)**.** The drift and diffusion coefficients are assumed to be locally Lipschitz and to satisfy linear growth conditions. These assumptions ensure the well-posedness of the forward Stochastic Differential Equations (SDEs).

Example 1 (GBM)**.** Consider the one-dimensional case in which the underlying asset follows a GBM,

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t, \tag{26}$$

with $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$ constants, and $(W_t)_{t\in[0,T]}$ a standard Brownian motion. Applying

Itô's lemma to $H(t, S_t)$ yields

$$dH_t = \left(\frac{\partial H}{\partial t} + \mu S_t \frac{\partial H}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 H}{\partial S^2}\right) dt + \sigma S_t \frac{\partial H}{\partial S}\, dW_t. \tag{27}$$

Using $dL_t^h = h_t dS_t - dH_t$, the P&L dynamics can be written directly in canonical form as

$$dL_t^{h,\mathrm{GBM}} = \underbrace{\left[(h_t - \tfrac{\partial H}{\partial S})\, \mu S_t - \left(\tfrac{\partial H}{\partial t} + \tfrac{1}{2}\sigma^2 S_t^2 \tfrac{\partial^2 H}{\partial S^2}\right)\right]}_{b^{\mathrm{GBM}}(t,S_t,H_t,h_t)} dt + \underbrace{(h_t - \tfrac{\partial H}{\partial S})\sigma S_t}_{\eta^{\mathrm{GBM}}(t,S_t,H_t,h_t)}\, dW_t. \tag{28}$$

and the derivatives of the coefficients with respect to the control are therefore given by

$$\begin{cases} \partial_h b^{\mathrm{GBM}}(t, S_t, H_t, h_t) = \mu S_t, \\ \partial_h \eta^{\mathrm{GBM}}(t, S_t, H_t, h_t) = \sigma S_t. \end{cases}$$

The associated Hamiltonian reads

$$H(t, S_t, H_t, h_t, p_t, q_t) = b^{\mathrm{GBM}}(t, S_t, H_t, h_t)\, p_t + \eta^{\mathrm{GBM}}(t, S_t, H_t, h_t)\, q_t,$$

so that

$$\partial_h H(t, S_t, H_t, h_t, p_t, q_t) = \mu S_t p_t + \sigma S_t q_t.$$

The first-order condition for an interior minimizer h^* takes the form

$$0 = \mu S_t p_t + \sigma S_t q_t, \tag{29}$$

where (p_t, q_t) solves the adjoint Backward SDE (BSDE) associated with the chosen composed risk measure. Together with $(S_t, H_t, L_t^{h^*})$, this condition yields a coupled Forward-Backward Stochastic Differential Equation (FBSDE) system. In particular, the adjoint process satisfies

$$\begin{cases} dp_t = \left[\dfrac{\partial^2 H}{\partial t \partial S} + \dfrac{1}{2}\sigma^2 \left(2S_t \dfrac{\partial^2 H}{\partial S^2} + S_t^2 \dfrac{\partial^3 H}{\partial S^3}\right)\right] p_t\, dt - \dfrac{\mu}{\sigma} p_t\, dW_t, \qquad q_t = -\dfrac{\mu}{\sigma} p_t, \\ p_T = \Upsilon\big(L_{t,T}^{h^*}\big). \end{cases} \tag{30}$$

Example 2 (Heston Stochastic Volatility)**.** Take the two-dimensional process (S_t, ν_t), introduced in Heston1993, defined by

$$\begin{cases} dS_t = \mu S_t \, dt + \sqrt{\nu_t} \, S_t \, dW_t^S \\ d\nu_t = \kappa(\theta - \nu_t) \, dt + \sigma_\nu \sqrt{\nu_t} \, dW_t^\nu, \end{cases} \tag{31}$$

where $\mu \in \mathbb{R}$, $\kappa, \theta, \sigma_\nu \in \mathbb{R}_+$, and (W_t^S, W_t^ν) is a two-dimensional Brownian motion with correlation $dW_t^S dW_t^\nu = \rho \, dt$, $\rho \in [-1, 1]$. The variance process ν_t follows a Cox-Ingersoll-Ross (CIR) diffusion and remains strictly positive with probability one whenever the Feller condition

$$2\kappa\theta \geq \sigma_\nu^2$$

is satisfied.

Applying Itô's lemma to $H(t, S_t)$ yields

$$\begin{aligned} dH_t = & \left(\frac{\partial H}{\partial t} + \mu S_t \frac{\partial H}{\partial S} + \kappa(\theta - \nu_t) \frac{\partial H}{\partial \nu} + \frac{1}{2} \nu_t S_t^2 \frac{\partial^2 H}{\partial S^2} + \frac{1}{2} \sigma_\nu^2 \nu_t \frac{\partial^2 H}{\partial \nu^2} \right) dt \\ & + \rho \sigma_\nu \nu_t S_t \frac{\partial^2 H}{\partial S \partial \nu} \, dt + \sqrt{\nu_t} S_t \frac{\partial H}{\partial S} \, dW_t^S + \sigma_\nu \sqrt{\nu_t} \frac{\partial H}{\partial \nu} \, dW_t^\nu. \end{aligned} \tag{32}$$

The P&L dynamics thus can be written directly in canonical form as

$$\begin{aligned} dL_t^{h,\text{Heston}} = & \underbrace{\left(h_t \mu S_t + h_t \kappa(\theta - \nu_t) - \frac{\partial H}{\partial t} - \mu S_t \frac{\partial H}{\partial S} - \kappa(\theta - \nu_t) \frac{\partial H}{\partial \nu} - \tfrac{1}{2} \nu_t S_t^2 \frac{\partial^2 H}{\partial S^2} - \tfrac{1}{2} \sigma_\nu^2 \nu_t \frac{\partial^2 H}{\partial \nu^2} - \rho \sigma_\nu \nu_t S_t \frac{\partial^2 H}{\partial S \partial \nu} \right)}_{b^{Heston}(t, S_t, \nu_t, H_t, h_t)} dt \\ & + \underbrace{\left(h_t^S - \tfrac{\partial H}{\partial S} \right) \sqrt{\nu_t} \, S_t}_{\eta_1^{Heston}(t, S_t, \nu_t, H_t, h_t)} \, dW_t^S + \underbrace{\left(h_t \sigma_\nu - \sigma_\nu \tfrac{\partial H}{\partial \nu} \right) \sqrt{\nu_t}}_{\eta_2^{Heston}(t, \nu_t, H_t, h_t)} \, dW_t^\nu. \end{aligned} \tag{33}$$

with derivatives of the coefficients with respect to the control given by

$$\begin{cases} \partial_h b^{\text{Heston}}(t, S_t, \nu_t, H_t, h_t) = \mu S_t + \kappa(\theta - \nu_t), \\ \partial_h \eta_1^{\text{Heston}}(t, S_t, \nu_t, H_t, h_t) = \sqrt{\nu_t} \, S_t, \\ \partial_h \eta_2^{\text{Heston}}(t, \nu_t, H_t, h_t) = \sigma_\nu \sqrt{\nu_t}. \end{cases}$$

The Hamiltonian is

$$H(t,S_t,\nu_t,H_t,h_t,p_t,q_t,r_t) = b^{\text{Heston}}(t,S_t,\nu_t,H_t,h_t)\,p_t + \eta_1^{\text{Heston}}(t,S_t,\nu_t,H_t,h_t)\,q_t + \eta_2^{\text{Heston}}(t,\nu_t,H_t,h_t)\,r_t,$$

so that

$$\partial_h H(t,S_t,\nu_t,H_t,h_t,p_t,q_t,r_t) = \big(\mu S_t + \kappa(\theta-\nu_t)\big)p_t + \sqrt{\nu_t}\,S_t\,q_t + \sigma_\nu\sqrt{\nu_t}\,r_t.$$

Hence, the first-order condition for h^* is obtained by setting this derivative equal to zero:

$$0 = \big(\mu S_t + \kappa(\theta-\nu_t)\big)p_t + \sqrt{\nu_t}\,S_t\,q_t + \sigma_\nu\sqrt{\nu_t}\,r_t, \tag{34}$$

where (p_t,q_t,r_t) solves the adjoint BSDE, coupled with forward dynamics of $(S_t,\nu_t,H_t,L_t^{h^*})$, this yields a FBSDE system. The adjoint process satisfies

$$\begin{cases} dp_t = \left[\dfrac{\partial^2 H}{\partial t\partial S} + \mu S_t\dfrac{\partial^2 H}{\partial S^2} + \kappa(\theta-\nu_t)\dfrac{\partial^2 H}{\partial S\partial\nu} + \dfrac{1}{2}\nu_t S_t^2\dfrac{\partial^3 H}{\partial S^3}\right. \\ \qquad\left. + \dfrac{1}{2}\sigma_\nu^2\nu_t\dfrac{\partial^3 H}{\partial S\partial\nu^2} + \rho\sigma_\nu\nu_t S_t\dfrac{\partial^3 H}{\partial S^2\partial\nu}\right]p_t\,dt \\ \qquad - \sqrt{\nu_t}\left(\dfrac{\partial H}{\partial S} + S_t\dfrac{\partial^2 H}{\partial S^2}\right)q_t\,dt + \sigma_\nu\sqrt{\nu_t}\dfrac{\partial^2 H}{\partial S\partial\nu}r_t\,dt + q_t\,dW_t^S + r_t\,dW_t^\nu, \\ p_T = \Upsilon\big(L_{t,T}^{h^*}\big). \end{cases} \tag{35}$$

Example 3 (Merton Jump-Diffusion)**.** Consider the Jump-Diffusion process, originally proposed by Merton1976, whose dynamics are described by

$$dS_t = \mu S_t\,dt + \sigma S_t\,dW_t + S_{t-}(J-1)\,dN_t, \tag{36}$$

where $\mu\in\mathbb{R}$, $\sigma>0$, $(W_t)_{t\in[0,T]}$ is a standard Brownian motion, and $(N_t)_{t\in[0,T]}$ is a Poisson process with intensity $\lambda>0$. The jump size J follows a lognormal distribution, with $\log J\sim\mathcal{N}(\mu_J,\sigma_J^2)$, and the Brownian and jump components are independent. $H(t,S_t)$ can be characterized by

$$dH_t = \left(\frac{\partial H}{\partial t} + \mu S_t \frac{\partial H}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 H}{\partial S^2} \right) dt + \sigma S_t \frac{\partial H}{\partial S}\, dW_t \\ + \Big[H\big(t, JS_{t-}\big) - H\big(t, S_{t-}\big) \Big] dN_t. \tag{37}$$

Substituting the above expression into the definition of the P&L, one obtains the canonical decomposition

$$\begin{aligned} dL_t^{h,\text{Jump}} &= \underbrace{\Big[\big(h_t - \tfrac{\partial H}{\partial S}\big)\, \mu S_t - \big(\tfrac{\partial H}{\partial t} + \tfrac{1}{2}\sigma^2 S_t^2 \tfrac{\partial^2 H}{\partial S^2} \big) \Big]}_{b^{\text{Jump}}(t,S_t,H_t,h_t)}\, dt \\ &\quad + \underbrace{\big(h_t - \tfrac{\partial H}{\partial S}\big)\sigma S_t}_{\eta^{\text{Jump}}(t,S_t,H_t,h_t)}\, dW_t \\ &\quad + \underbrace{\Big(h_t S_{t-}(J-1) - \big[H(t, JS_{t-}) - H(t, S_{t-}) \big] \Big)}_{\xi^{\text{Jump}}(t,S_t,H_t,h_t)}\, dN_t. \end{aligned} \tag{38}$$

From this representation, the sensitivity of the coefficients with respect to the control variable is

$$\begin{cases} \partial_h b^{\text{Jump}}(t, S_t, H_t, h_t) = \mu S_t, \\ \partial_h \eta^{\text{Jump}}(t, S_t, H_t, h_t) = \sigma S_t, \\ \partial_h \xi^{\text{Jump}}(t, S_t, H_t, h_t) = S_{t-}(J-1). \end{cases}$$

The corresponding Hamiltonian is constructed as

$$H(t, S_t, H_t, h_t, p_t, q_t, r_t) = b^{\text{Jump}}(t, S_t, H_t, h_t)\, p_t + \eta^{\text{Jump}}(t, S_t, H_t, h_t)\, q_t + \xi^{\text{Jump}}(t, S_t, H_t, h_t)\, r_t,$$

which implies

$$\partial_h H(t, S_t, H_t, h_t, p_t, q_t, r_t) = \mu S_t\, p_t + \sigma S_t\, q_t + S_{t-}(J-1)\, r_t.$$

Accordingly, the first-order optimality condition associated with h^* is given by

$$0 = \mu S_t\, p_t + \sigma S_t\, q_t + S_{t-}(J-1)\, r_t, \tag{39}$$

when combined with the forward evolution $(S_t, H_t, L_t^{h^*})$. This condition gives rise to the coupled FBSDE, where the (p_t, q_t, r_t) satisfies the BSDE below:

$$\begin{cases} dp_t = \left[\frac{\partial^2 H}{\partial t \partial S} + \mu S_t \frac{\partial^2 H}{\partial S^2} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^3 H}{\partial S^3} + \sigma^2 S_t \frac{\partial^2 H}{\partial S^2} \right] p_t \, dt \\ \qquad - \sigma \left(\frac{\partial H}{\partial S} + S_t \frac{\partial^2 H}{\partial S^2} \right) q_t \, dt - \left[(J-1)\, r_t - \Big(J \frac{\partial H}{\partial S}(t, JS_t) - \frac{\partial H}{\partial S}(t, S_t) \Big) r_t \right] dt \\ \qquad + q_t \, dW_t + r_t \, dN_t, \\ p_T = \Upsilon\big(L_{t,T}^{h^*}\big). \end{cases} \tag{40}$$

Example 4 (EL with linear utility)**.** Consider EL as the risk measure, defined by $\rho(X) = -\mathbb{E}[X]$, and a linear utility function $u(x) = x$. The resulting composed risk measure is

$$\rho_u(X) = -\mathbb{E}[X]. \tag{41}$$

When applied to the $L_{t,T}^h$, this yields $\rho_u(L_{t,T}^h) = -\mathbb{E}[L_{t,T}^h]$. The corresponding Gâteaux derivative satisfies

$$D\rho_u(X)[Z] = -\mathbb{E}[Z] = \mathbb{E}[(-1)Z], \tag{42}$$

which implies that the adjoint processes are given by $p_t = -1$ and $q_t = 0$.

It is easy to see that the objective function is affine and therefore does not admit an interior minimizer. Whenever the drift of $L_{T,t}^h$ is non-zero, the value of $-\mathbb{E}[L_{t,T}^h]$ can be made arbitrarily small by sending $|h_t|$ to infinity in the appropriate direction, leading to explosive solutions. Due to this reason, EL with linear utility does not fit coherently into the general convex-analytic framework and should be regarded as a degenerate limiting case.

Example 5 (EL with exponential utility)**.** Consider the EL $\rho(X) = -\mathbb{E}[X]$ and the exponential utility function $u(x) = 1 - e^{-ax}$, $a > 0$. The composed risk measure is

$$\rho_u(X) = \mathbb{E}\big[e^{-aX}\big] - 1. \tag{43}$$

For $L^h_{t,T}$, we have

$$
\begin{aligned}
\rho_u(L^h_{t,T}) &= \mathbb{E}\big[e^{-aL^h_{t,T}}\big] \\
D\rho_u(X)[Z] &= \mathbb{E}\big[-ae^{-aX}Z\big], \\
\Upsilon(X) &= -ae^{-aX}, \\
p_T = \Upsilon\big(L^{h^*}_{t,T}\big) &= -a\,e^{-aL^{h^*}_{t,T}}.
\end{aligned}
\tag{44}
$$

Example 6 (EL with Weibull utility)**.** Let $\rho(X) = -\mathbb{E}[X]$ and the Weibull-type utility function $u(x) = 1 - \exp\big((-\min(x,0))^k\big)$, $k > 1$. The composed risk measure is

$$\rho_u(X) = \mathbb{E}\big[e^{(-\min(X,0))^k}\big] - 1. \tag{45}$$

Substituting $L^h_{t,T}$, we obtain

$$
\begin{aligned}
\rho_u(L^h_{t,T}) &= \mathbb{E}\big[e^{(-\min(L^h_{t,T},0))^k}\big], \qquad k > 1, \\
D\rho_u(X)[Z] &= \mathbb{E}\big[-k(-\min(X,0))^{k-1}e^{(-\min(X,0))^k}Z\big], \\
\Upsilon(X) &= -k(-\min(X,0))^{k-1}e^{(-\min(X,0))^k}, \\
p_T = \Upsilon\big(L^{h^*}_{t,T}\big) &= -k(-\min(L^{h^*}_{t,T},0))^{k-1}e^{(-\min(L^{h^*}_{t,T},0))^k}.
\end{aligned}
\tag{46}
$$

Remark 3 (Weibull-type)*.* For Weibull-type utilities, Gâteaux differentiability at $L^h_{t,T}$ requires $\mathbb{P}(L^h_{t,T} = 0) = 0$, since the utility is non-differentiable at zero. If this condition fails, differentiability is lost and the composed criterion admits only sub-differentiability.

Example 7 (ERM with linear utility)**.** Take the *ERM* with risk aversion parameter $\gamma > 0$, $\rho(X) = \frac{1}{\gamma}\log\mathbb{E}[e^{-\gamma X}]$, together with the linear utility function $u(x) = x$. The composed risk measure is

$$\rho_u(X) = \frac{1}{\gamma}\log\mathbb{E}[e^{-\gamma X}]. \tag{47}$$

Thus, the terminal condition of the adjoint can be written as

$$\begin{aligned}
\rho_u(L^h_{t,T}) &= \frac{1}{\gamma}\log\mathbb{E}\big[e^{-\gamma L^h_{t,T}}\big], \qquad \gamma>0,\\
D\rho_u(X)[Z] &= -\frac{\mathbb{E}\big[Ze^{-\gamma X}\big]}{\mathbb{E}\big[e^{-\gamma X}\big]},\\
\Upsilon(X) &= -\frac{e^{-\gamma X}}{\mathbb{E}\big[e^{-\gamma X}\big]},\\
p_T = \Upsilon\big(L^{h^*}_{t,T}\big) &= -\frac{e^{-\gamma L^{h^*}_{t,T}}}{\mathbb{E}\big[e^{-\gamma L^{h^*}_{t,T}}\big]}.
\end{aligned} \tag{48}$$

Example 8 (ERM with exponential utility)**.** Start from the *ERM* with parameter $\gamma>0$, $\rho(X)=\frac{1}{\gamma}\log\mathbb{E}[e^{-\gamma X}]$, and choose the exponential utility function $u(x)=1-e^{-ax}$, $a>0$. The corresponding composed risk measure is

$$\rho_u(X)=\frac{1}{\gamma}\log\mathbb{E}\Big[e^{\gamma e^{-aX}}\Big]-1. \tag{49}$$

For $L^h_{t,T}$, we have

$$\begin{aligned}
\rho_u(L^h_{t,T}) &= \frac{1}{\gamma}\log\mathbb{E}\big[e^{\gamma e^{-aL^h_{t,T}}}\big], \qquad a>0,\ \gamma>0,\\
D\rho_u(X)[Z] &= -a\,\frac{\mathbb{E}\big[e^{-aX}e^{\gamma e^{-aX}}Z\big]}{\mathbb{E}\big[e^{\gamma e^{-aX}}\big]},\\
\Upsilon(X) &= -a\,\frac{e^{-aX}e^{\gamma e^{-aX}}}{\mathbb{E}\big[e^{\gamma e^{-aX}}\big]},\\
p_T = \Upsilon\big(L^{h^*}_{t,T}\big) &= -a\,\frac{e^{-aL^{h^*}_{t,T}}e^{\gamma e^{-aL^{h^*}_{t,T}}}}{\mathbb{E}\big[e^{\gamma e^{-aL^{h^*}_{t,T}}}\big]}
\end{aligned} \tag{50}$$

Example 9 (ERM with Weibull utility)**.** Let $\rho(X)=\frac{1}{\gamma}\log\mathbb{E}[e^{-\gamma X}]$, $\gamma>0$, be the *ERM*, and let $u(x)=1-e^{(-\min(x,0))^k}$, $k>1$, be the Weibull-type utility function. The induced composed risk measure is

$$\rho_u(X)=\frac{1}{\gamma}\log\mathbb{E}\Big[e^{\gamma e^{(-\min(X,0))^k}}\Big]-1. \tag{51}$$

When applied to the $L^h_{t,T}$, we get

$$\begin{aligned}
\rho_u(L^h_{t,T}) &= \frac{1}{\gamma}\log \mathbb{E}\Big[e^{\gamma e^{(-\min(L^h_{t,T},0))^k}}\Big], \qquad \gamma>0,\ k>1,\\
D\rho_u(X)[Z] &= \mathbb{E}\left[-\frac{k(-\min(X,0))^{k-1}e^{(-\min(X,0))^k}e^{\gamma e^{(-\min(X,0))^k}}}{\mathbb{E}\Big[e^{\gamma e^{(-\min(X,0))^k}}\Big]}Z\right],\\
\Upsilon(X) &= -\frac{k(-\min(X,0))^{k-1}e^{(-\min(X,0))^k}e^{\gamma e^{(-\min(X,0))^k}}}{\mathbb{E}\Big[e^{\gamma e^{(-\min(X,0))^k}}\Big]},\\
p_T = \Upsilon\big(L^{h^*}_{t,T}\big) &= -\frac{k(-\min(L^{h^*}_{t,T},0))^{k-1}e^{(-\min(L^{h^*}_{t,T},0))^k}e^{\gamma e^{(-\min(L^{h^*}_{t,T},0))^k}}}{\mathbb{E}\Big[e^{\gamma e^{(-\min(L^{h^*}_{t,T},0))^k}}\Big]}.
\end{aligned} \tag{52}$$

Example 10 (ES with linear utility)**.** We define ES $\rho(X)\mathrm{ES}_\beta(X) := \inf_{\alpha\in\mathbb{R}}\Big\{\alpha + \frac{1}{1-\beta}\mathbb{E}\big[(X-\alpha)^+\big]\Big\}$, with $\beta\in(0,1)$, following Rockafellar2000. Using the linear utility function $u(x)=x$, we have

$$\rho_u(X) = \inf_{\alpha\in\mathbb{R}}\left\{\alpha + \frac{1}{1-\beta}\mathbb{E}\big[(X-\alpha)^+\big]\right\}. \tag{53}$$

For the terminal profit and loss $L^h_{t,T}$, the terminal condition of the adjoint is

$$\begin{aligned}
\rho_u(L^h_{t,T}) &= \inf_{\alpha\in\mathbb{R}}\left\{\alpha + \frac{1}{1-\beta}\mathbb{E}\big[(L^h_{t,T}-\alpha)^+\big]\right\},\\
D\rho_u(X)[Z] &= \frac{1}{1-\beta}\,\mathbb{E}\big[\mathbf{1}_{\{X\geq\alpha^*\}}Z\big],\\
\Upsilon(X) &= \frac{1}{1-\beta}\,\mathbf{1}_{\{X\geq\alpha^*\}},\\
p_T = \Upsilon\big(L^{h^*}_{t,T}\big) &= \frac{1}{1-\beta}\,\mathbf{1}_{\{L^{h^*}_{t,T}\geq\alpha^*\}}.
\end{aligned} \tag{54}$$

Remark 4 (ES)*.* For ES, Gâteaux differentiability at $L^h_{t,T}$ holds provided that a minimizer $\alpha^*\in\arg\min_{\alpha\in\mathbb{R}}\Big\{\alpha+\frac{1}{1-\beta}\mathbb{E}\big[(L^h_{t,T}-\alpha)^+\big]\Big\}$ exists and satisfies $\mathbb{P}(L^{h^*}_{t,T}=\alpha^*)=0$. If the minimizer does not exist or if the quantile is not unique, ES is in general only sub-differentiable.

6 Numerical Analysis

In order to estimate $MR(t)$ for each example in Section 5, we must compute an approximately optimal hedging trajectory, i.e., a control h^* minimizing the risk of the residual profit. This requires solving the coupled FBSDE system derived in Examples 1-3 with terminal conditions $p_T = \Upsilon(L^h_{t,T})$ as specified in Examples 5-10.

The quantity encoding the risk criterion appears only at the terminal time, which renders the coupling implicit and distribution-dependent. Furthermore, this structure is singular in the sense that distinct controls may generate terminal distributions that lead to identical terminal adjoint values, thereby causing the associated optimality system to admit multiple solutions at the discrete level. It should be noted that numerical schemes relying on strong regularity or monotone contraction properties may fail to converge or may converge to spurious fixed points, when Υ is steep, non-smooth, or highly sensitive to tail events.

The algorithm adopted below is based on successive approximations tailored to the maximum principle structure of the problem; Øksendal [2003], Kloeden and Platen [1992] serve as the basis for its implementation. Starting from an initial control, one iterates the cycle of *forward simulation* → *backward adjoint resolution* → *control update* using the discrete first-order condition. This choice is motivated by the fact that it separates the forward propagation of the state and P&L from the backward propagation of sensitivities, while keeping the coupling confined to the terminal map Υ.

The method is subject to standard limitations of iterative schemes in stochastic control, including sensitivity to step sizes, Monte Carlo noise in conditional expectations, and possible convergence to non-optimal stationary points. Nevertheless, in the present setting it provides a direct procedure to generate, for each example, a consistent family of near-optimal hedging trajectories required to evaluate the MR surface.

The numerical experiments follow a unified computational framework common to all market models considered. While each model is characterized by its own forward dynamics for the state variables, the numerical objective is identical across specifications. Given a fixed time grid $(t_n)_{n=0}^N$ with $\Delta t = T/N$ and a Monte Carlo (MC) sample of size M, we require a consistent procedure to (i) simulate S_t (ii) obtain H_t, meaning the derivatives price, and its partial derivatives. This modular design isolates model-specific components from those shared across all settings, enabling the computation of the MR on a common discretization and under comparable numerical conditions.

The first pre-processing step consists in simulate S_t forward. For each model, the state is sampled on the grid (t_n) and the resulting paths are kept fixed throughout the optimization iterations. In the GBM and Jump-Diffusion models, the asset price is propagated via a multiplicative Euler-Maruyama scheme, expressed directly in terms of the current state and the sampled increments. In the Heston model, the state includes both the asset price and the stochastic variance, and the Euler discretization is modified to ensure numerical stability of the variance component. Specifically, a truncation rule is applied to preserve non-negativity at the discrete level, while the correlated Brownian increments driving the asset and variance dynamics are generated jointly. These modifications are introduced solely for numerical stability.

The second pre-processing step is the approximation of the contingent claim price H_t along the simulated state trajectories, together with the partial derivatives required by the discrete adjoint equation. As closed-form expressions are not available uniformly across models, H_t is estimated by a least-squares Monte Carlo (LSMC) regression. The regression targets the terminal payoff, with regressors defined as functions of time and the state variables. For numerical stability, the regression is carried out on a compact domain obtained by scaling each regressor to $[-1, 1]$, and a Chebyshev polynomial basis is employed. This choice is convenient, because the resulting approximation admits analytic differentiation with respect to each regressor, allowing all required derivatives to be computed consistently from the same fitted coefficients.

In multi-factor settings, the approximation space is constructed as a tensor product of univariate Chebyshev bases in the relevant state coordinates. When the state is (t, x, v) with $x = \log S$, the polynomial degrees are selected independently along each dimension, allowing for controlled approximation complexity and balanced scaling. The regression coefficients are estimated using ridge regularization to stabilize the fit in the presence of collinearity and to improve conditioning. This stabilization is crucial, since the adjoint-based iteration depends not only on the fitted values of H, but also on its derivatives, which are more sensitive to coefficient variability. The output of this stage is a fixed set of simulated paths together with fitted representations of H and its derivatives along the same grid, which serve as inputs to the subsequent FBSDE optimization.

For each ρ_u, the optimization problem is formulated via the necessary conditions of the SMP. Given $h \in \mathcal{H}$, the forward equation governs $L^h_{t,T}$, while the BSDE part is defined accordingly. The first-order condition yields a pathwise residual $G(h)$, computed from the adjoint variables (p, q, r), whose vanishing characterizes stationarity. The numerical objective is to construct h such that $|G(h)|$ is small under the simulation measure.

The resulting system is coupled in a terminal-to-initial manner. The control h determines $L^h_{t,T}$ and the terminal adjoint condition, while the adjoint variables (p, q, r) define the residual $G(h)$ that drives the update of h. As a consequence, the coupled FBSDE system cannot be solved directly as a boundary value problem. Instead, this justify our adoption of an iterative scheme based on successive approximations. Starting from an initial guess $h^{(0)}$, each iterate $h^{(k)}$ induces a full forward–backward pass and yields a residual $G^{(k)} = G(h^{(k)})$, from which the next control is obtained via a gradient-type correction.

In practice, the residual $G(h)$ is evaluated from MC samples and from regression-based conditional expectations used in the backward propagation, which introduces variability and approximation error in the map $h \mapsto G(h)$. This effect is exacerbated when the LSMC becomes unstable or when the control-to-terminal-state map is stiff. To mitigate these effects, the implementation incorporates several stabilization mechanisms: (i) a normalized update direction to control scaling across time and state variables, (ii) an acceptance rule based on the reduction of the scalar residual $g = |G|_{\mathrm{RMS}}$; and (iii) a rollback/backtracking strategy. If a trial update fails to decrease g, the procedure reverts to the last accepted control and progressively shrinks the step size until a decrease is achieved or a minimal step size is reached. The iteration is terminated when $g < \mathrm{tol}$ or when no improving step can be found under backtracking.

In view of these considerations, the algorithm[1] presented below is best interpreted as a numerical fixed-point solver with controlled descent on the maximum principle residual,

[1] For reproducibility, the code used to generate all numeric results is available at `https://github.com/JoaoJungblut/Maturity-Risk-for-Variational-Heding`.

rather than as a globally convergent optimization method.

Algorithm 1 Gradient-based SMP iteration

Input: Simulated trajectories of S_t, and fitted approximations of H_t and its partial derivatives; initial control $h^{(0)}$; parameters max_iter, tol, $\alpha > 0$, $\eta \in (0,1)$, bt_max, $\alpha_{\min}$.

Output: Approximate optimal control $\hat{h}^*$ and iteration history.

Subroutine $\textsc{Evaluate}(h, t_{\text{start}})$**.**

(i) Simulate $L^h_{t,T}$ on $[t_{\text{start}}, T]$.

(ii) Set $p_T \leftarrow \Upsilon(L^h_T)$ (pathwise).

(iii) Solve the backward adjoint and obtain (p, q, r).

(iv) Compute the optimality residual $G(h)$ and

$g \leftarrow \|G(h)\|_{\text{RMS}}$.

return $(G(h), g)$.

Main procedure.

$h_{\text{curr}} \leftarrow h^{(0)}$

$h_{\text{prev}} \leftarrow \varnothing$

history $\leftarrow \emptyset$

for $k = 0, 1, \ldots, \text{max\_iter} - 1$ **do**

(a) Evaluate current control.

$(G_{\text{curr}}, g_{\text{curr}}) \leftarrow \textsc{Evaluate}(h_{\text{curr}}, t_{\text{start}})$

Append $(k, \alpha, g_{\text{curr}})$ to history

(b) Stopping criterion.

if $g_{\text{curr}} < \text{tol}$ **then**

$\hat{h}^* \leftarrow h_{\text{curr}}$

return $(\hat{h}^*, \text{history})$

end if

(c) First trial.

$h_{\text{try}} \leftarrow \texttt{Update}(h_{\text{curr}}, G_{\text{curr}}, \alpha)$

$(\cdot, g_{\text{try}}) \leftarrow \textsc{Evaluate}(h_{\text{try}}, t_{\text{start}})$

(d) Accept/reject.

if $g_{\text{try}} < g_{\text{curr}}$ **then**
  $h_{\text{prev}} \leftarrow h_{\text{curr}}$
  $h_{\text{curr}} \leftarrow h_{\text{try}}$
  continue
end if
(e) Early safeguard.
if $k < 2$ **then**
  continue
end if
(f) Rollback anchor.
if $h_{\text{prev}} \neq \varnothing$ **then**
  $h_{\text{anchor}} \leftarrow h_{\text{prev}}$
else
  $h_{\text{anchor}} \leftarrow h_{\text{curr}}$
end if
$(G_{\text{a}}, g_{\text{a}}) \leftarrow \textsc{Evaluate}(h_{\text{anchor}}, t_{\text{start}})$
$\alpha_{\text{try}} \leftarrow \eta\, \alpha$
(g) Backtracking.
for $j = 0, 1, \ldots, \text{bt\_max} - 1$ **do**
  if $\alpha_{\text{try}} < \alpha_{\text{min}}$ **then**
    break
  end if
  $h_{\text{new}} \leftarrow \texttt{Update}(h_{\text{anchor}}, G_{\text{a}}, \alpha_{\text{try}})$
  $(\cdot, g_{\text{new}}) \leftarrow \textsc{Evaluate}(h_{\text{new}}, t_{\text{start}})$
  if $g_{\text{new}} < g_{\text{a}}$ **then**
    $h_{\text{prev}} \leftarrow h_{\text{anchor}}$
    $h_{\text{curr}} \leftarrow h_{\text{new}}$
    $\alpha \leftarrow \alpha_{\text{try}}$
    break
  else

$\alpha_{\text{try}} \leftarrow \eta\, \alpha_{\text{try}}$

end if

end for

end for

$\hat{h}^* \leftarrow h_{\text{curr}}$

return $(\hat{h}^*, \text{history})$

In the sequel, three-dimensional MR surfaces are computed under GBM, Heston and Jump–Diffusion specifications. In all cases, asset paths are initialized at $S_0 = 10$ and simulated via MC with $M = 1000$ trajectories and $N = 252$ time steps over the unit horizon $T = 1$. The drift and diffusion parameters are fixed at $\mu = 0.05$ and $\sigma = 0.20$. In the Heston model, the variance dynamics are specified by $\kappa = 1.0$, $\theta = 0.1$, $\sigma_\nu = 0.1$, and correlation $\rho = 0.7$. For the Jump–Diffusion setting, jumps arrive according to a Poisson process with intensity $\lambda = 0.5$, and jump sizes are normally distributed with mean $\mu_J = -0.03$ and standard deviation $\sigma_J = 0.12$.

MR is evaluated on discrete grids of times to maturity $(t_1, \ldots, t_{21}) \subset [0, T]$ and strike levels $(K_1, \ldots, K_{21}) \subset [8, 16]$, yielding $21 \times 21 = 441$ values per surface. For each pair (t_i, K_j), MR is obtained by solving its defining optimization problem under fixed preference parameters; hence, differences across panels reflect only the choice of risk functional and the underlying price dynamics. In truncated variants, the parameter U caps the exponential growth to prevent numerical overflow. Across all surfaces, parameters are held fixed at $\gamma = 1$, $a = 1$, $k = 2$, $\beta = 0.95$, and $U = 10$, and are not varied within a given surface.

Figures 1 reports the MR surfaces. The image is organized into six rows, corresponding to ρ_u presented from Examples 5-10, and into three columns indexed by the model specifications of Examples 1-3. The horizontal axes show time t and the moneyness level S_0/K, while the vertical axis accounts for MR, the color bar gives its scale.

The resulting present a stable dependence on moneyness across all models, with MR highest close to the at the money region and lower when the strike is below the spot price, which corresponds to moneyness greater than one. The overall magnitude increases when moving from GBM to stochastic volatility or jump dynamics, which indicates higher marginal risk once additional uncertainty sources are introduced.

The temporal profile depends of ρ_u and it changes in a systematic way across specifications. ES exhibits temporal monotonicity in all models, with a decreasing in risk as the time approaches maturity. This is aligned with insensitivity to gains property 6. EL with exponential utility and the ERM with linear and exponential utility display non-monotonic patterns in horizon, including regions where MR increases with time. In such cases, the functional is driven by the central mass of the exponential transformation rather than by the losses. Weibull type specifications behave regularly under GBM, where MR tends to decrease as maturity approaches, but under Heston and Jump-Diffusion they show localized peaks at intermediate times and specific moneyness levels. These peaks are not consistent with the smooth structure of the models and they point to numerical limitations, since truncation can push the optimization into a saturated region where gradients become weak and the algorithm returns artificially high MR values.

In general, the qualitative structure reflects two components that must be separated when interpreting the plots. One component comes from the analytical form of ρ_u and its sensitivity to gains. The other component is numerical and comes from non-linear transformations combined with truncation mechanisms required for stability. On the other hand, The dependence on moneyness is robust across models and specifications, while the temporal behavior and the presence of extreme values require joint reading with the properties of the functional.

The sensitivity analysis is reported in Figures 4 to 7 through heatmaps and parameter response curves for ΔMR. Heatmaps describe how MR varies over the process parameters space, and they are evaluated at two representative times, $t = 0.2$ and $t = 0.8$. The line plots describe how MR changes when a single preference coefficient is modified, while the price dynamics are kept fixed. All figures follow the same baseline configuration used in the surface analysis, so the reported differences are driven by controlled changes.

Each heatmap is built on a two-dimensional uniform grid with nine points per axis, which gives eighty-one points per model. Under GBM, the axes correspond to $\mu \in [0.025, 0.075]$ and $\sigma \in [0.10, 0.40]$. For Heston model, the grid spans with $\kappa \in [1.0, 4.0]$ and $\theta \in [0.1, 0.3]$. Jump-Diffusion spans a grid for $\mu_J \in [-0.3, 0.3]$ and $\lambda \in [0, 1]$. ΔMR measures how MR changes at a fixed (t, K), computed when a given model parameter is perturbed around the baseline used in the surfaces.

Looking at the GBM, heatmaps are smooth across the (μ, σ) plane. The magnitude of ΔMR is larger at shorter horizons and it decreases as maturity approaches, which indicates that temporal sensitivity is gradually absorbed by the accumulation of diffusion. The main driver is the volatility coefficient σ, because increasing σ increases the magnitude of ΔMR across most of the grid. Changes in μ have a weaker effect and they rarely change the sign of ΔMR. ES differs from the other measures by keeping non-negligible temporal variation even at longer horizons.

Under Heston and jump dynamics, the heatmaps display localized regions of high magnitude and abrupt sign changes. These effects are more visible for exponential and Weibull-type utilities and they persist across the two evaluation times. Since the grids are uniform, the irregular patterns are not attributed to the discretization itself. They are linked to the interaction between non-linear transformations of the $L^h_{t,T}$ distribution and the optimization used to compute MR.

A more detail inspection shows ΔMR varies non-uniformly over the (κ, θ) plane. Variations along the mean-reversion speed κ generate larger changes than variations along the long-run variance level θ. For intermediate values of θ, increasing κ produces a marked increase in the magnitude of ΔMR, while for small κ temporal variation remains limited. Furthermore, MR is more sensitive to the μ_J than to λ. Shifts along the μ_J axis change both the magnitude and the sign of ΔMR in a clear way. Changes in λ mainly shift levels without changing the dominant pattern.

Across all models, ΔMR is typically larger at earlier evaluation times and it attenuates as maturity increases. For the remaining specifications, the comparison between $t = 0.2$ and $t = 0.8$ shows a flattening of temporal variation at longer horizons. Lastly, ES appears to have the best results by preserving sign and scale across parameters and horizons, which indicates stable temporal behavior relative to the other measures.

The aversion curves are constructed by varying one preference coefficient at a time on a one-dimensional uniform grid, and by reporting ΔMR at $t = 0.2$ and $t = 0.8$. For ERM, exponential and Weibull-type utilities specifications, respectively, we set $\gamma \in [0.2, 2.0]$, $a \in [0.2, 2.0]$ and $k \in [1.2, 3.0]$. These three cases use fifteen grid points. For ES, the confidence level ranges over $\beta \in \{0.90, 0.95, 0.975, 0.99\}$. These plots separate preference sensitivity from market dynamic's changes.

The aversion curves complement the heatmaps by isolating preference effects. For ERM and EL risk measures, changes in aversion produce non-linear adjustments in ΔMR, and near flat regions appear in some ranges. This is consistent with the possibility of negative MR when the distribution is centered on gains, so higher aversion does not necessarily amplify temporal variation. The effect is mild under GBM and it becomes more visible under Heston and Jump-Diffusion models. For Weibull-type utility, the response becomes unstable, with rapid growth in MR, which matches the extreme regions seen in the heatmaps and points to numerical limitations induced by truncation (U) in the optimization.

Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999. doi: https://doi.org/10.1111/1467-9965.00068. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9965.00068.

Damiano Brigo and Claudio Nordio. Liquidity-adjusted market risk measures with stochastic holding period, 2010. URL https://arxiv.org/abs/1009.3760.

Alexandre Carbonneau and Frédéric Godin. Deep equal risk pricing of financial derivatives with non-translation invariant risk measures. *Risks*, 11(8), 2023. ISSN 2227-9091. URL https://www.mdpi.com/2227-9091/11/8/140.

Peter Carr, Helyette Geman, and Dilip B. Madan. Pricing and hedging in incomplete markets. *Journal of Financial Economics*, 62(1):131–167, 2001. ISSN 0304-405X. doi: https://doi.org/10.1016/S0304-405X(01)00075-7. URL https://www.sciencedirect.com/science/article/pii/S0304405X01000757.

Alexander S. Cherny and Dilip B. Madan. Pricing and hedging in incomplete markets with coherent risk, 2006. URL https://arxiv.org/abs/math/0605064.

Gemma Colldeforns-Papiol and Luis Ortiz-Gracia. Computation of market risk measures with stochastic liquidity horizon. *Journal of Computational and Applied Mathematics*, 342:431–450, 2018. ISSN 0377-0427. doi: https://doi.org/10.1016/j.cam.2018.03.038. URL https://www.sciencedirect.com/science/article/pii/S0377042718302188.

Hans Follmer and Peter Leukert. Efficient hedging: Cost versus shortfall risk. (1999,18), 1999. URL https://hdl.handle.net/10419/61739. urn:nbn:de:kobv:11-10056108.

Hans Follmer and Alexander Schied. *Stochastic finance: an introduction in discrete time.* Walter de Gruyter GmbH, 2016.

Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002. URL https://EconPapers.repec.org/RePEc:spr:finsto:v:6:y:2002:i:4:p:429-447.

Patryk Gierjatowicz, Marc Sabate-Vidales, David Šiška, Lukasz Szpruch, and Žan Žurič. Robust pricing and hedging via neural sdes. 2020. URL https://arxiv.org/abs/2007.04154.

Alessandro Gnoatto, Silvia Lavagnini, and Athena Picarelli. Deep quadratic hedging. 2024. URL https://arxiv.org/abs/2212.12725.

Sebastian Herrmann, Johannes Muhle-Karbe, and Frank Thomas Seifried. Hedging with small uncertainty aversion. 2016. URL https://arxiv.org/abs/1605.06429.

Elyégs Jouini and Hédi Kallal. Arbitrage in securities markets with short-sales constraints. *Mathematical Finance*, 5(3):197–232, 1995. doi: https://doi.org/10.1111/j.1467-9965.1995.tb00065.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9965.1995.tb00065.x.

Marcel Kaina and Ludger Rüschendorf. On convex risk measures on L p -spaces. *Mathematical Methods of Operations Research*, 69(3):475–495, July 2009.

doi: 10.1007/s00186-008-0248-3. URL https://ideas.repec.org/a/spr/mathme/v69y2009i3p475-495.html.

Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23 of *Stochastic Modelling and Applied Probability*. Springer, Berlin, Heidelberg, 1992. ISBN 978-3-540-54062-5. doi: 10.1007/978-3-662-12616-5.

Yannick Limmer and Blanka Horvath. Robust hedging gans: Towards automated robustification of hedging strategies. *Applied Mathematical Finance*, 31(3):164–201, May 2024. ISSN 1466-4313. doi: 10.1080/1350486x.2024.2440661. URL http://dx.doi.org/10.1080/1350486X.2024.2440661.

Fabio Maccheroni, Massimo Marinacci, and Aldo Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498, 2006. doi: https://doi.org/10.1111/j.1468-0262.2006.00716.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0262.2006.00716.x.

Ola Mahmoud. The temporal dimension of risk, 2016. URL https://arxiv.org/abs/1501.01573.

Shige Peng. A general stochastic maximum principle for optimal control problems. *SIAM Journal on Control and Optimization*, 28(4):966–979, 1990. doi: 10.1137/0328054. URL https://doi.org/10.1137/0328054.

Marcelo B. Righi. Optimal hedging with variational preferences under convex risk measures. *Quantitative Finance*, 24(11):1703–1709, November 2024. doi: 10.1080/14697688.2024.2416981.

R. Tyrrell Rockafellar and Johannes O. Royset. Measures of residual risk with connections to regression, risk tracking, surrogate models, and ambiguity. *SIAM Journal on Optimization*, 25(2):1179–1208, 2015. doi: 10.1137/151003271.

Xiaofei Shi, Daran Xu, and Zhanhao Zhang. Deep learning algorithms for hedging with frictions. *Digital Finance*, 5(1):113–147, March 2023. doi: 10.1007/

s42521-023-00075-z. URL https://ideas.repec.org/a/spr/digfin/v5y2023i1d10.1007_s42521-023-00075-z.html.

Antoine Toussaint and Ronnie Sircar. A framework for dynamic hedging under convex risk measures. In Robert Dalang, Marco Dozzi, and Francesco Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 429–451, Basel, 2011. Springer Basel. ISBN 978-3-0348-0021-1.

Mingxin Xu. Risk measure pricing and hedging in incomplete markets. *Annals of Finance*, 2(1):51–71, January 2006. doi: 10.1007/s10436-005-0023-x.

Jiongmin Yong and Xun Yu Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations*, volume 43 of *Applications of Mathematics*. Springer, 1999. ISBN 978-0387985536.

Umut Çetin, Robert Jarrow, and Philip Protter. Liquidity risk and arbitrage pricing theory. *Finance and Stochastics*, 8(3):311–341, August 2004. doi: 10.1007/s00780-004-0123-x.

Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer, 6 edition, 2003. ISBN 978-3540047584.

Appendix

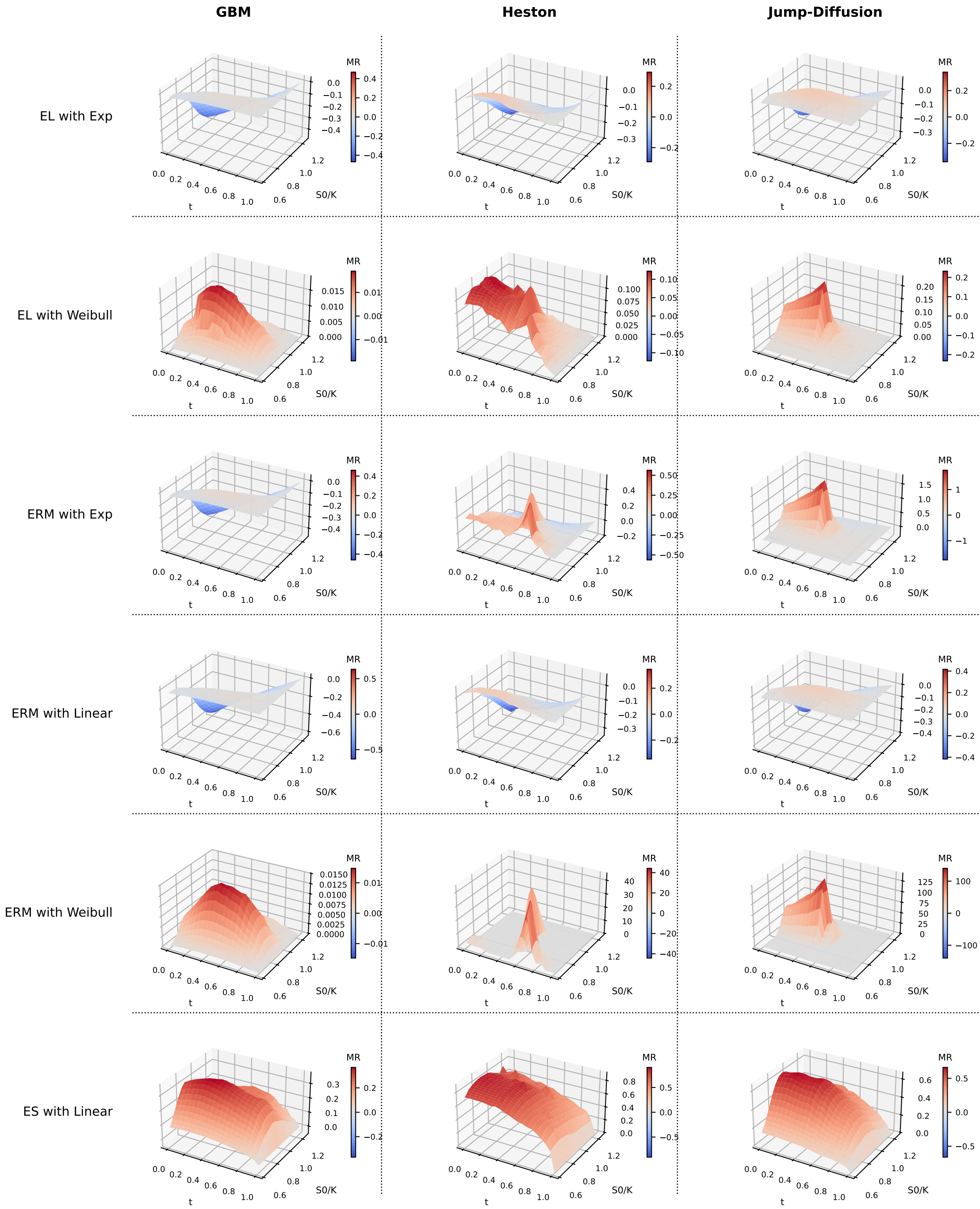

Figure 1: MR surfaces across all models.

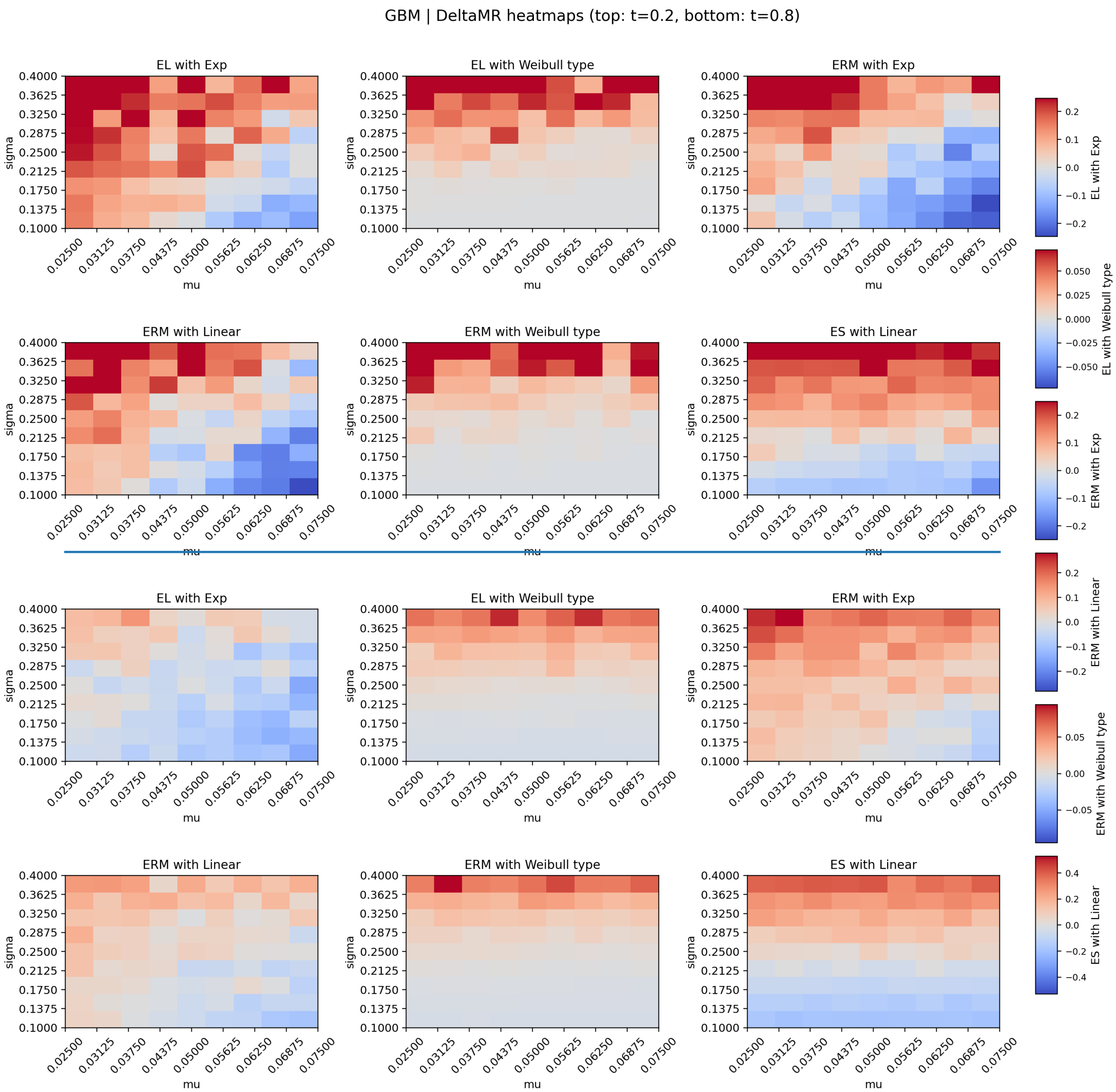

Figure 2: MR sensitivity to μ and σ parameters.

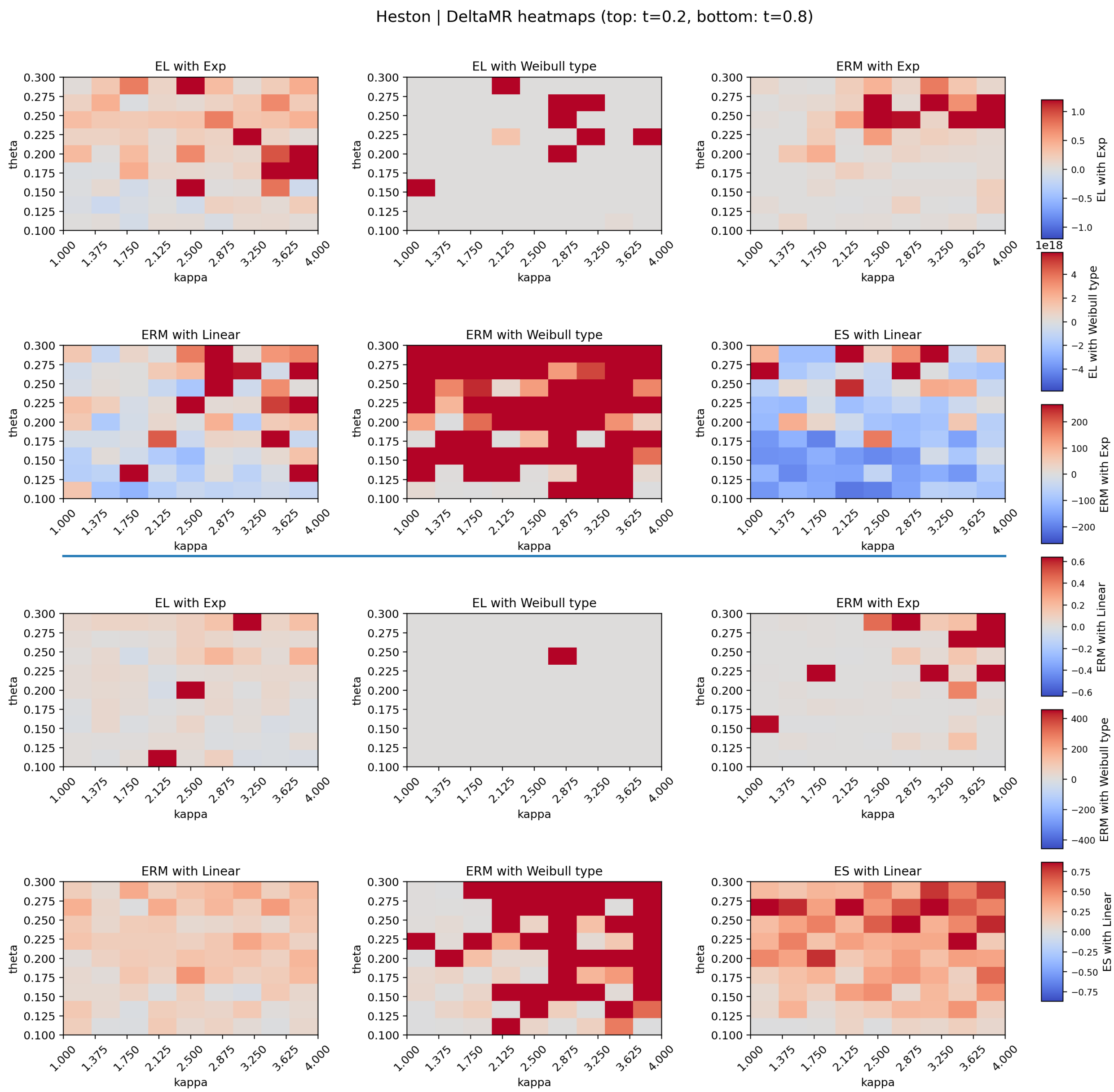

Figure 3: MR sensitivity to κ and θ parameters.

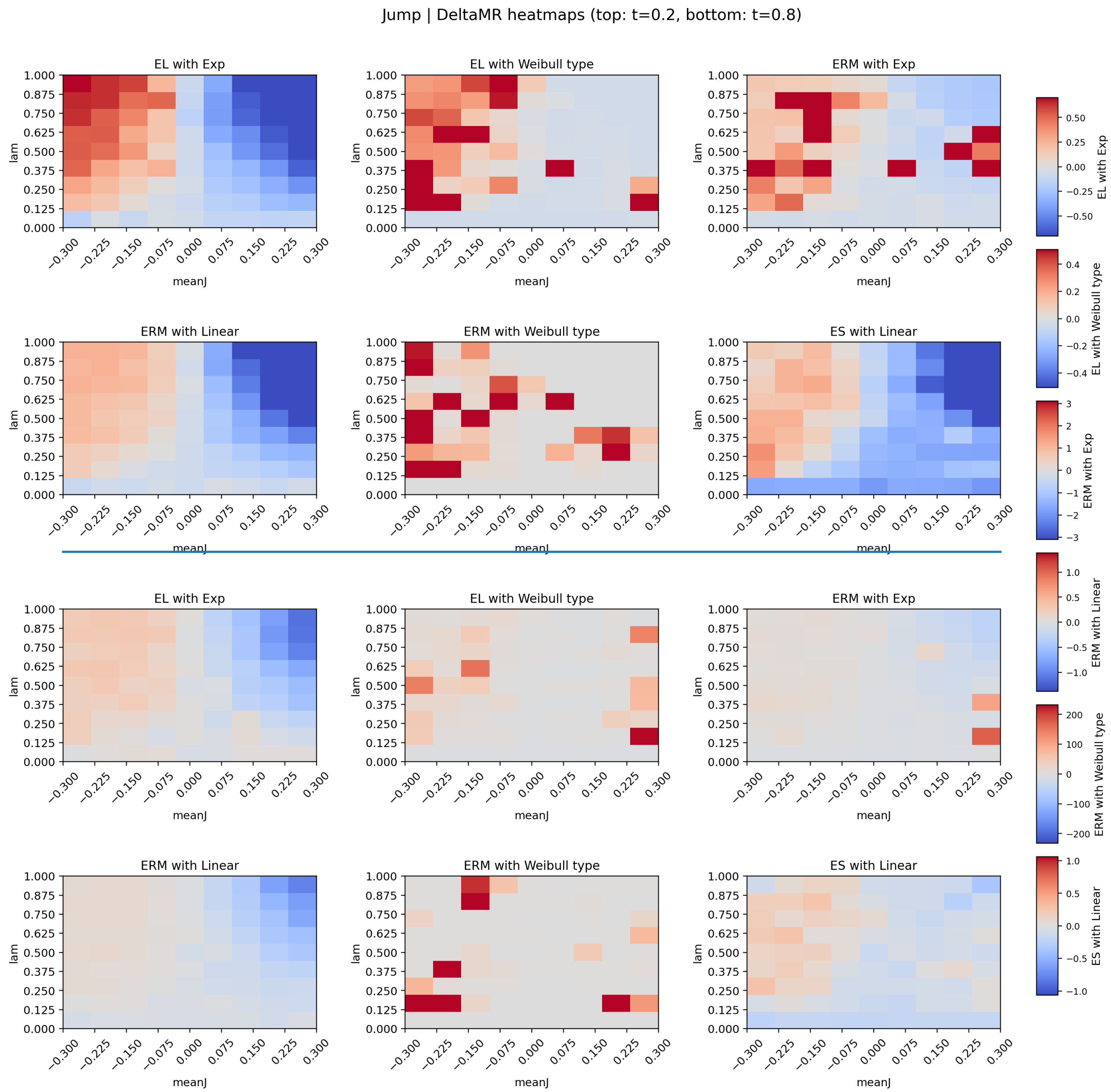

Figure 4: MR sensitivity to μ_J and λ parameters.

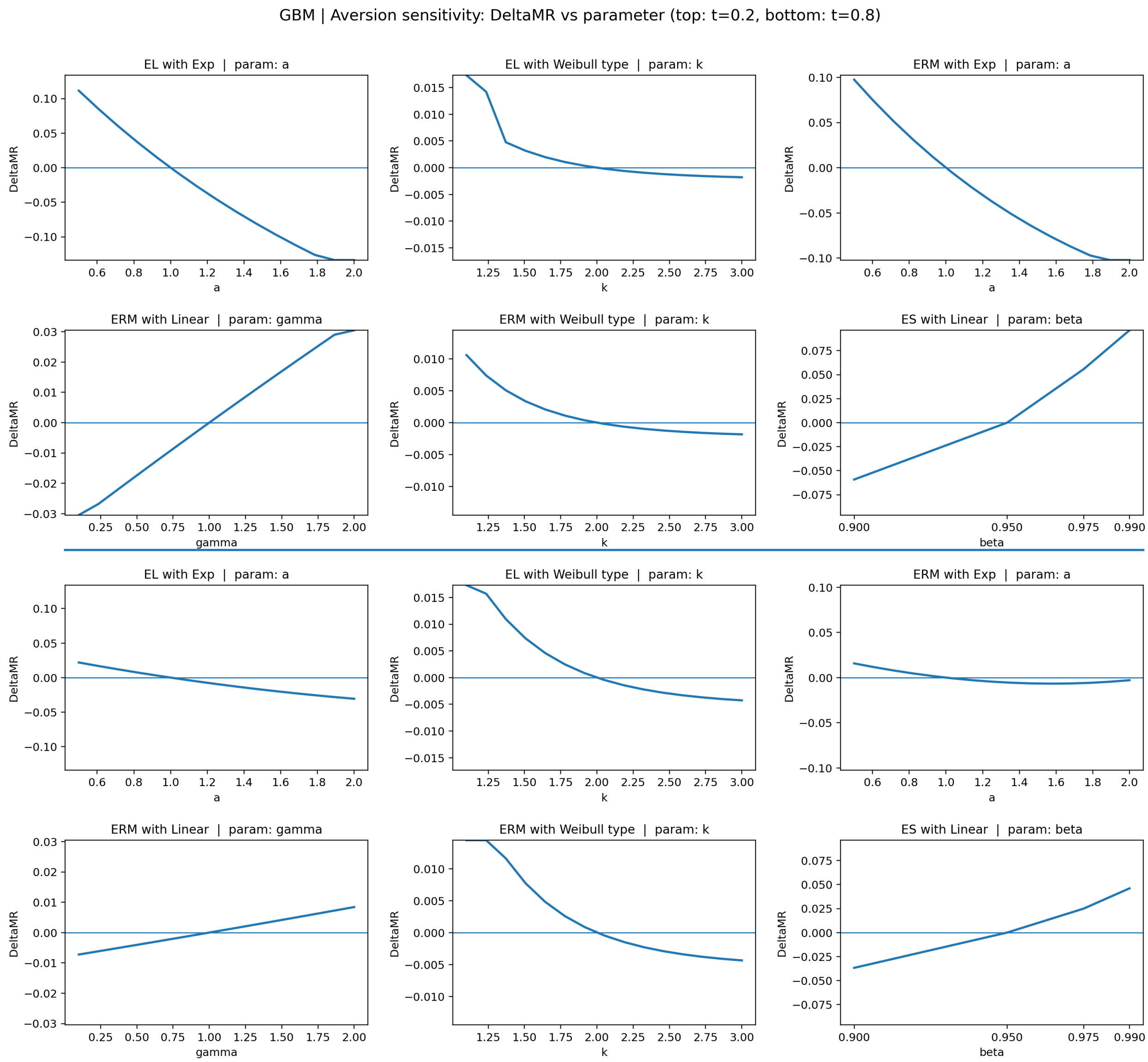

Figure 5: MR sensitivity to risk-aversion coeficients in GBM model.

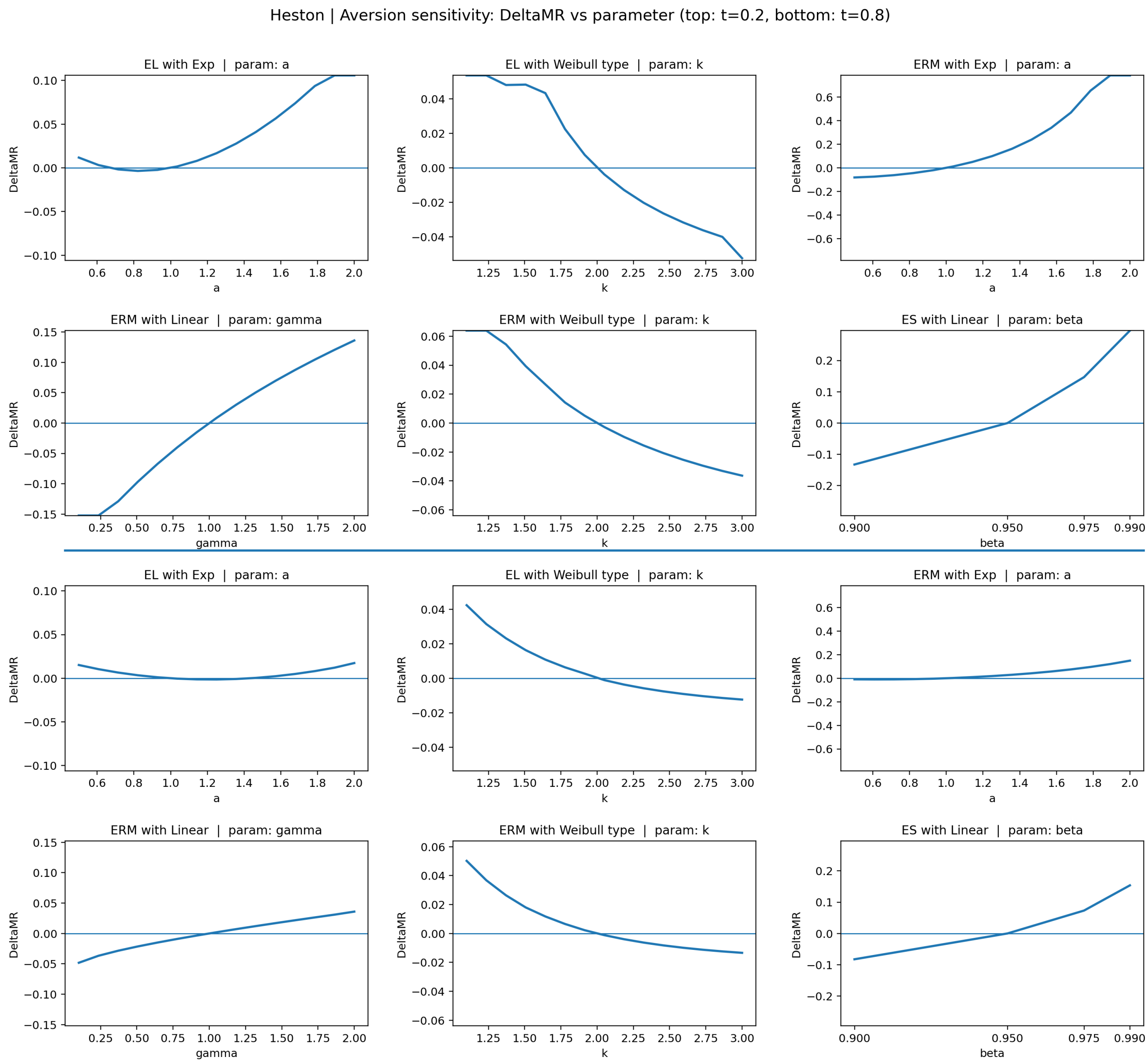

Figure 6: MR sensitivity to risk-aversion coeficients in Heston model.

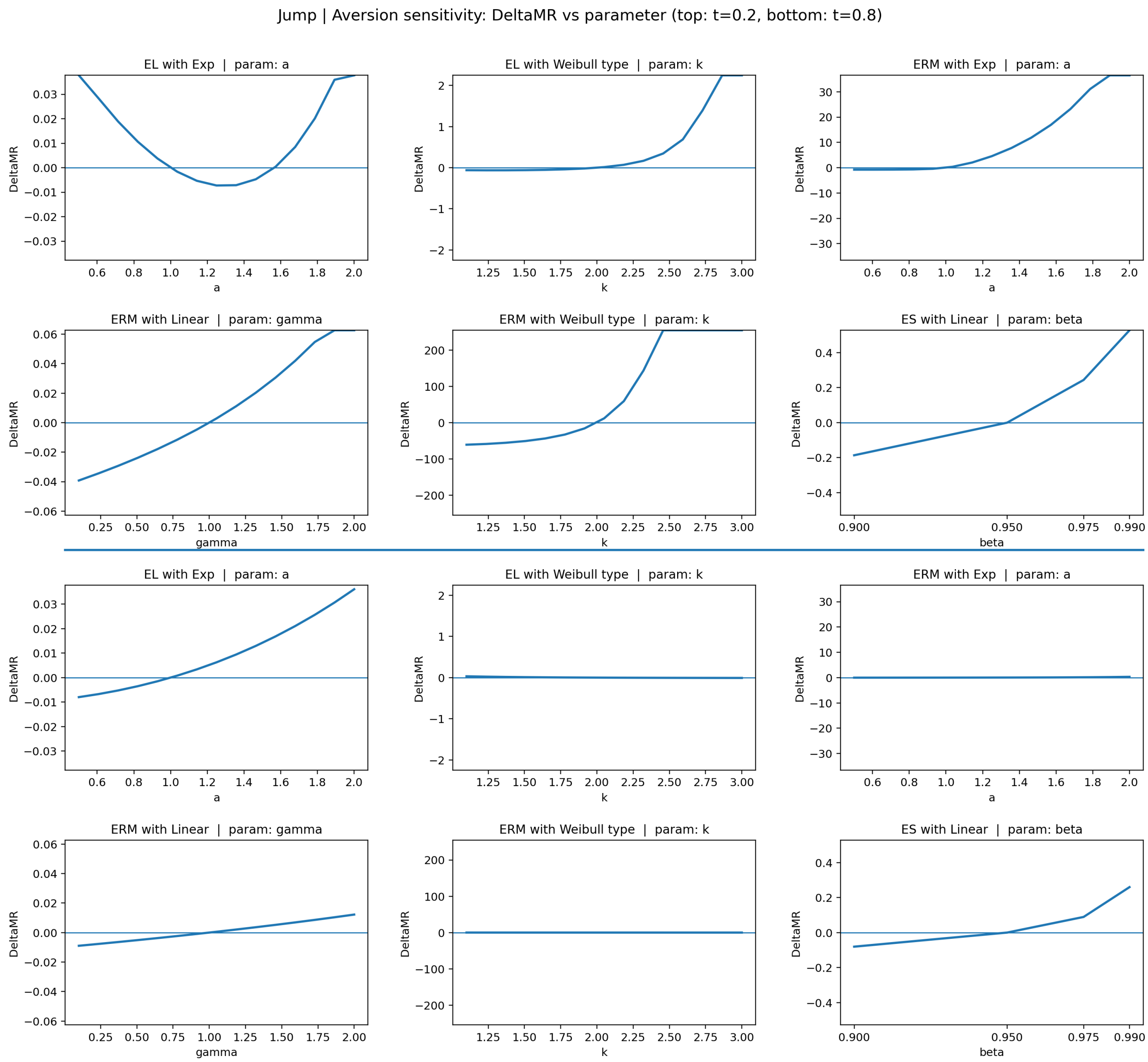

Figure 7: MR sensitivity to risk-aversion coeficients in Jump-Diffusion model.