Pricing and hedging financial derivatives with reinforcement learning methods

Alexandre Carbonneau

A Thesis
in the Department
of
Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy (Mathematics) at
Concordia University
Montréal, Québec, Canada

June 2021

Concordia University
School of Graduate Studies

This is to certify that the thesis prepared

By: Alexandre Carbonneau

Entitled: Pricing and hedging financial derivatives with reinforcement learning methods

and submitted in partial fulfillment of the requirements for the degree of

Doctorate of Philosophy (Mathematics)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

______________________________ Chair
Dr. Ivan Contreras

______________________________ External Examiner
Dr. Bruno Rémillard

______________________________ External to Program
Dr. Jia Yuan Yu

______________________________ Examiner
Dr. Patrice Gaillardetz

______________________________ Examiner
Dr. Cody Hyndman

______________________________ Supervisor
Dr. Frédéric Godin

Approved by ______________________________
Dr. Cody Hyndman, Chair
Department of Mathematics and Statistics

June 29, 2021 ______________________________
Pascale Sicotte, Dean
Faculty of Arts and Science

Abstract

Title: Pricing and hedging financial derivatives with reinforcement learning methods

Alexandre Carbonneau, Ph.D.

Concordia University, 2021

This thesis studies the problem of pricing and hedging financial derivatives with reinforcement learning. Throughout all four papers, the underlying global hedging problems are solved using the deep hedging algorithm with the representation of global hedging policies as neural networks. The first paper, *"Equal Risk Pricing of Derivatives with Deep Hedging"*, shows how the deep hedging algorithm can be applied to solve the two underlying global hedging problems of the equal risk pricing framework for the valuation of European financial derivatives. The second paper, *"Deep Hedging of Long-Term Financial Derivatives"*, studies the problem of global hedging very long-term financial derivatives which are analogous, under some assumptions, to options embedded in guarantees of variable annuities. The third paper, *"Deep Equal Risk Pricing of Financial Derivatives with Multiple Hedging Instruments"*, studies derivative prices generated by the equal risk pricing framework for long-term options when shorter-term options are used as hedging instruments. The fourth paper, *"Deep equal risk pricing of financial derivatives with non-translation invariant risk measures"*, investigates the use of non-translation invariant risk measures within the equal risk pricing framework.

Keywords: Deep hedging, Equal risk pricing, Convex risk measure, Reinforcement learning, Deep learning, Neural networks, Global hedging, Risk management, Variable annuity, Lookback option, Jump risk, Volatility risk, Non-translation invariant risk measure.

Acknowledgments

First, I would like to thank Frédéric Godin, my supervisor, with whom I developed an exceptional relationship during my PhD. I thank him for the countless hours of discussions on various topics which led to exciting research work. I also thank him for his energy, his kindness and his endless source of motivation which allowed me to surpass myself during my PhD. I want to highlight the significant contribution he had in the papers of my thesis.

I would also like to thank the examining committee of my thesis, Bruno Rémillard, Jia Yuan Yu, Patrice Gaillardetz and Cody Hyndman for their valuable comments.

In addition, I would like to thank many professors I had the chance to meet during my academic adventure. First, I thank my Master's supervisor, Maciej Augustyniak, who taught me my first steps as a researcher. Furthermore, I thank Manuel Morales and Claudia Gagné for the various opportunities, both academically and professionally, they have provided me throughout the years. I would also like to thank Geneviève Gauthier for her extremely valuable comments and suggestions.

I also thank the Fonds de recherche du Québec - Nature et technologies (FRQNT), The Montreal Exchange, Autorité des Marchés Financiers, Concordia University, Fin-ML, Institut des Sciences Mathématiques as well as Frédéric Godin for their financial support throughout my thesis.

Finally, and most importantly, I am extremely thankful to my soon-to-be wife, Iunia Seni, to my family and to my friends. This thesis could not have been achieved without your love, your support and your infinite patience.

Contribution of Authors

This manuscript-based thesis is separated into four papers organized as four main chapters.

Chapter 2: Alexandre Carbonneau wrote the manuscript and conducted all numerical experiments. Frédéric Godin contributed with the draft manuscript preparation. The results were published in an article entitled "Equal Risk Pricing of Derivatives with Deep Hedging".

Chapter 3: Alexandre Carbonneau wrote the manuscript and conducted all numerical experiments. The results were published in an article entitled "Deep Hedging of Long-Term Financial Derivatives".

Chapter 4: Alexandre Carbonneau wrote the manuscript and conducted all numerical experiments. Frédéric Godin contributed with the draft manuscript preparation. The preprint version presenting the results is entitled "Deep Equal Risk Pricing of Financial Derivatives with Multiple Hedging Instruments".

Chapter 5: Alexandre Carbonneau wrote the manuscript and conducted all numerical experiments. Frédéric Godin contributed with the draft manuscript preparation. The preprint version presenting the results is entitled "Deep equal risk pricing of financial derivatives with non-translation invariant risk measures".

All authors reviewed the final manuscript and approved of the contents.

Contents

List of Tables

Introduction

The celebrated Black and Scholes (1973) and Merton (1973) model constitutes the pillar of option pricing theory. One of the most important paradigms of their work is that in the so-called Black-Scholes world, every contingent claim can be perfectly replicated by continuously rebalancing a portfolio composed of a risky asset and a risk-free asset. The Black-Scholes market is *complete*, and derivatives are redundant financial assets. Consequently, the two problems of pricing and hedging a contingent claim are both solved by finding the replicating trading strategy which completely eliminates hedging risk, and by setting the value of the contingent claim to the unique arbitrage-free price being the initial value of the replication portfolio.

Fortunately for the field of mathematical finance, in practice, financial markets are typically *incomplete*, and contingent claims cannot be perfectly replicated. Indeed, various stylized features of financial markets violate several assumptions imposed by the Black-Scholes model. For instance, the presence of salient equity risk factors such as jump and volatility risks which cannot be completely hedged away, trades occurring in discrete time as well as market impact stemming from the presence of trading costs and imperfect liquidity. Thus, in this context, perfect replication is impossible, and the problem of pricing and optimally hedging contingent claims is highly relevant.

The main theme of this thesis is the use of deep reinforcement learning methods for pricing and hedging financial derivatives in incomplete markets. More precisely, all four papers of this thesis study the problem of global hedging contingent claims from the standpoint of either pricing derivatives consistently with optimal hedges, or from the standpoint of mitigating the risk exposure of derivative positions. For the former, the pricing mechanism studied in this thesis is the *equal risk pricing* framework introduced by Guo and Zhu (2017). This derivative valuation approach is consistent with non-quadratic global hedging procedures by pricing a contingent claim as the value such that the residual global hedging risk of the long and short positions is equal under optimal hedges. Furthermore, this thesis also examines the performance of non-quadratic global hedging strategies for mitigating the risk exposure of long-term financial derivatives. Throughout this thesis, the numerical scheme considered to solve global hedging problems is from the class of *deep hedging* algorithms introduced by Buehler et al. (2019b). This approach is based on a parametric approximation of the trading policy with a neural network trained using reinforcement learning. One of the most important benefits of the deep hedging approach over typical procedures such as traditional dynamics programming is to overcome the curse of dimensionality when considering high-dimensional state and action spaces.

The thesis is separated into four papers organized as four main chapters. Chapter 2 introduces the deep reinforcement learning approach based on deep hedging to implement the equal risk pricing framework with convex risk measures under general conditions. Chapter 3 studies the application of the deep hedging algorithm in the context of global hedging very long-term financial derivatives. Chapter 4 examines the impact of including short-term options as hedging instruments for pricing longer-term financial derivatives with the equal risk pricing framework under convex risk measures. Chapter 5 investigates the use of non-translation invariant risk measures within the equal risk pricing framework for pricing and hedging financial derivatives. Chapter 6 summarizes the contributions of this thesis. The bibliography for all papers is presented at the end of the thesis.

Chapter 2

Equal Risk Pricing of Derivatives with Deep Hedging

Abstract

This article provides a universal and tractable methodology based on deep reinforcement learning to implement the equal risk pricing framework for financial derivatives pricing under very general conditions. The equal risk pricing framework entails solving for a derivative price which equates the optimally hedged residual risk exposure associated respectively with the long and short positions in the option. The solution to the hedging optimization problem considered, which is inspired from the Marzban et al. (2020) framework relying on convex risk measures, is obtained through the use of the deep hedging algorithm of Buehler et al. (2019b). Consequently, the current paper's approach allows for the pricing and the hedging of a very large number of contingent claims (e.g. vanilla options, exotic options, options with multiple underlying assets) with multiple liquid hedging instruments under a wide variety of market dynamics (e.g. regime-switching, stochastic volatility, jumps). A novel ϵ-completeness measure allowing for the quantification of the residual hedging risk associated with a derivative is also proposed. The latter measure generalizes the one presented in Bertsimas et al. (2001) based on the quadratic penalty. Monte Carlo simulations are performed under a large variety of market dynamics to demonstrate the practicability of our approach, to perform benchmarking with respect to traditional methods and to conduct sensitivity analyses. Numerical results show, among others, that equal risk prices of out-of-the-money options are significantly higher than risk-neutral prices stemming from conventional changes of measure across all dynamics considered. This finding is shown to be shared by different option categories which include vanilla and exotic options.

Keywords: Reinforcement learning, Deep learning, Option pricing, Hedging, Convex risk measures.

JEL classification: C45, G13.

2.1 Introduction

Under the complete market paradigm, for instance as in Black and Scholes (1973) and Merton (1973), all contingent claims can be perfectly replicated with some dynamic hedging strategy. In such circumstances, the unique arbitrage-free price of an option must be the initial value of the replicating portfolio. However, in reality, markets are incomplete and perfect replication is typically impossible for non-linear derivatives. Indeed, there are many sources of market incompleteness observed in practice such as discrete-time rebalancing, liquidity constraints, stochastic volatility, jumps, etc. In an incomplete market, it is often impracticable for a hedger to select a trading strategy that entirely removes risk as it would typically entail unreasonable costs. For instance, Eberlein and Jacod (1997) show that the super-replication price of a European call option under a large variety of underlying asset dynamics is the initial underlying asset price. Thus, in practice, a hedger must accept the presence of residual hedging risk that is intrinsic to the contingent claim being hedged. The determination of option prices and hedging policies therefore depend on subjective assumptions regarding risk preferences of market participants.

An incomplete market derivatives pricing approach that is extensively studied in the literature consists in the selection of a suitable equivalent martingale measure (EMM). As shown in the seminal work of Harrison and Pliska (1981), if a market is incomplete and arbitrage-free, there exists an infinite set of EMMs each of which can be used to price derivatives through a risk-neutral valuation. Some popular examples of EMMs in the literature include the Esscher transform by Gerber and Shiu (1994) and the minimal-entropy martingale measure by Frittelli (2000). Option pricing functions induced by the latter risk-neutral measures can then be used to calculate Greek letters associated with the option, which leads to the specification of hedging policies, e.g. delta-hedging. However, in that case, hedging policies are not an input of the pricing procedure, but rather a by-product. Thus, hedging policies obtained from many popular EMMs are typically not optimal, and corresponding option prices are not designed in a way that is consistent with optimal hedging strategies. Another strand of literature derives martingale measures that are designed to be consistent with optimal hedging approaches such as the minimal martingale measure by Föllmer and Schweizer (1991), the variance-optimal martingale measure by Schweizer (1995) and the Extended Girsanov Principle of Elliott and Madan (1998). However, an undesirable feature of the three previous methods is their reliance on quadratic objective functions which penalize hedging gains. The identification of a pricing procedure consistent with a non-quadratic global optimization of hedging errors, i.e. a joint optimization over hedging decisions for all time periods until the maturity of the derivative, would be desirable.

In that direction, another approach studied in the literature considers the determination of derivatives prices directly from global optimal hedging strategies without having to specify an EMM. A first example of approach among these schemes is utility indifference pricing in which a trader with a specific utility function prices a contingent claim as the value such that the utility of his portfolio remains unchanged by the inclusion of the contingent claim. For instance, Hodges and Neuberger (1989) study hedging and indifference pricing under the negative exponential utility function with transaction costs under the Black-Scholes model (BSM). Closely related is the risk indifference pricing in which a risk measure is used to characterize the risk aversion of the trader instead of a utility function. For example, Xu (2006) studies the indifference pricing and hedging in an incomplete market using convex risk measures as defined in Föllmer and Schied (2002). One notable feature of utility and risk indifference pricing is that the resulting price depends on the position (long or short) of the hedger in the contingent claim. This highlights the need to identify hedging-based pricing schemes producing a unique price that is invariant to being long or short.

Recently, Guo and Zhu (2017) introduced the concept of equal risk pricing. In their framework, the option price is set as the value such that the global risk exposure of the long and short positions is equal under optimal hedging strategies. Contrarily to utility and risk indifference pricing, equal risk pricing provides a unique transactional price. The latter paper focuses mainly on theoretical features of the equal risk pricing framework and does not provide a general approach to compute the solution of the hedging problem embedded in the methodology. Thus, equal risk prices are only provided for a limited number of specific cases. Following the work of Guo and Zhu (2017), Marzban et al. (2020) adapted the equal risk pricing framework to the case where convex risk measures are used to quantify the risk exposures of the long and short positions under optimal hedging strategies. By further imposing that the risk measures can be decomposed in a way that satisfies a Markovian property, they provide dynamic programming equations that can be used to solve the hedging problems for both European and American options.

To enhance the tractability of the equal risk approach, the current paper also considers the use of convex risk measures to quantify the global risk exposures of the long and short positions under optimal hedging strategies. Hedging under a convex risk measure has been extensively studied in the literature: Alexander et al. (2003) minimize the Conditional Value-at-Risk (CVaR, Rockafellar and Uryasev (2002)) in the context of static hedging with multiple assets, Xu (2006) studies the indifference pricing and hedging under a convex risk measure in an incomplete market and Godin (2016) develops a global hedging strategy using CVaR as the cost function in the presence of transaction costs. Recently, Buehler et al. (2019b) introduced an algorithm called deep hedging to hedge a portfolio of over-the-counter derivatives in the presence of market frictions under a convex risk measure using deep reinforcement learning (deep RL). The general framework of RL is for an agent to learn over many iterations of an environment how to select sequences of actions in order to optimize a cost function. Hedging with RL has received some attention; Kolm and Ritter (2019) demonstrate that SARSA can be used to learn the hedging strategy if the objective function is a mean-variance criteria and Halperin (2020) shows that Q-learning can be used to learn the option pricing and hedging strategy under the BSM. In the novel deep hedging algorithm of Buehler et al. (2019b), an agent is trained to learn how to optimize the hedging strategy produced by a neural network through many simulations of a synthetic market. Their deep RL approach to the hedging problem helps to counter the well-known curse of dimensionality that arises when the state space gets too large. As argued by François et al. (2014), when applying traditional dynamic programming algorithms to compute hedging strategies, the curse of dimensionality can prevent the use of a large number of features to model the different components of the financial market.

The contribution of the current study is threefold. The first contribution consists in providing a universal and tractable methodology to implement the equal risk pricing framework under very general conditions. The approach based on deep RL as in Buehler et al. (2019b) can price and optimally hedge a very large number of contingent claims (e.g. vanilla options, exotic options, options with multiple underlying assets) with multiple liquid hedging instruments under a wide variety of market dynamics (e.g. regime-switching, stochastic volatility, jumps, etc.). Results presented in this paper, which rely on Buehler et al. (2019b), demonstrate that our methodological approach to equal risk pricing can approximate arbitrarily well the true equal risk price.

The second contribution of the current study consists in performing several numerical experiments studying the behavior of equal risk prices in various contexts. Such experiments showcase the wide applicability of our proposed framework. The behavior of the equal risk pricing approach is analyzed among others through benchmarking against expected risk-neutral pricing and by conducting sensitivity analyzes determining the impact on option prices of the confidence level associated with the risk measure and of the underlying asset model choice. The conduction of such numerical experiments crucially relies on the deep RL scheme outlined in the current study. Using the latter framework allows presenting numerical examples for equal risk pricing that are more extensive, realistic and varied than in previous studies; such results would most likely have been previously inaccessible when relying on more traditional computation methods (e.g. finite difference dynamic programming). Numerical results show, among others, that equal risk prices of out-of-the-money (OTM) options are significantly higher than risk-neutral prices across all dynamics considered. This finding is shown to be shared by different option categories which include vanilla and exotic options. Thus, by using the usual risk-neutral valuation instead of the equal risk pricing framework, a risk averse participant trading OTM options might significantly underprice these contracts.

The last contribution is the introduction of an asymmetric ϵ-completeness measure based on hedging strategies embedded in the equal risk pricing approach. The purpose of the metric is to quantify the magnitude of unhedgeable risk associated with a position in a contingent claim. The ϵ-completeness measure can therefore be used to quantify the level of market incompleteness inherent to a given market model. Our contribution complements the work of Bertsimas et al. (2001); their proposed measure of market incompleteness is based on the mean-squared-error cost function, while ours has the advantage of allowing to characterize the risk aversion of the hedger with any convex risk measure. Furthermore, the current paper's proposed measure is asymmetric in the sense that the risk for the long and short positions in the derivative are quantified by two different hedging strategies, unlike in Bertsimas et al. (2001) where the single variance-optimal hedging strategy is considered.

The paper is structured as follows. Section 2.2 details and adapts the equal risk pricing framework proposed in Marzban et al. (2020) and introduces the ϵ-completeness measure. Section 2.3 describes the deep RL numerical solution to equal risk pricing. Section 2.4 presents various numerical experiments including, among others, sensitivity and benchmarking analyses. Section 2.5 concludes. All proofs are provided in Section 2.6.1.

2.2 Equal risk pricing framework

This section details the theoretical option pricing setup considered in the current study.

2.2.1 Market setup

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space where $\mathbb{P}$ is the physical measure. The financial market is in discrete time with a finite time horizon of T years and known fixed trading dates $\mathcal{T} := \{0 = t_0 < t_1 < \ldots < t_N = T\}$. Consider $D + 1$ liquid and tradable assets on the market with D risky assets and one risk-free asset. Risky assets can include for instance stocks and options. Let $\{S_n\}_{n=0}^N$ be the non-negative price process of the D risky assets where $S_n := [S_n^{(1)}, \ldots, S_n^{(D)}]$ are the prices at time $t_n \in \mathcal{T}$. Also, let $\{B_n\}_{n=0}^N$ be the price process of the risk-free asset with $B_n := \exp(rt_n)$ where $r \in \mathbb{R}$ is the annualized continuously compounded risk-free rate. For convenience, assume that all assets are not paying any cash flows during the trading dates except possibly at time T. Define the market filtration $\mathbb{F} := \{\mathcal{F}_n\}_{n=0}^N$ where $\mathcal{F}_n := \sigma(S_u | u = 0, \ldots, n), n = 0, \ldots, N$. Moreover, assume that $\mathcal{F} = \mathcal{F}_N$. Throughout this paper, suppose that a European-type contingent claim paying off $\Phi(S_N, Z_N) \geq 0$ at the maturity date T must be priced, where $\{Z_n\}_{n=0}^N$ is an $\mathbb{F}$-adapted process with Z_n being a K-dimensional random vector of relevant state variables and $\Phi : [0, \infty)^D \times \mathbb{R}^K \to \mathbb{R}$. $\{Z_n\}_{n=0}^N$ can include drivers of risky asset dynamics or information relevant to price the derivative Φ. For the rest of the paper, all assets and contingent claims prices are assumed to be well-behaved and integrable enough. Specific conditions are out-of-scope.

Our option pricing approach requires solving the two distinct problems of dynamic optimal hedging, respectively one for a long and one for a short position in the contingent claim. Let $\delta := \{\delta_n\}_{n=0}^N$ be a trading strategy used by the hedger to minimize his risk exposure to the derivative, where for $n = 1, \ldots, N$, $\delta_n := [\delta_n^{(0)}, \delta_n^{(1)}, \ldots, \delta_n^{(D)}]$ is a vector containing the number of shares held in each asset during the period $(t_{n-1}, t_n]$ in the hedging portfolio. $\delta_n^{(0)}$ and $\delta_n^{(1:D)} := [\delta_n^{(1)}, \ldots, \delta_n^{(D)}]$ are respectively the positions in the risk-free asset and in the D risky assets. Furthermore, the initial portfolio (at time 0 before the first trade) is strictly invested in the risk-free asset. For the rest of the paper, assume the absence of market impact from transactions, i.e. trading in the risky assets does not affect their prices. Here are some well-known definitions in the mathematical finance literature (see for instance Lamberton and Lapeyre (2011) for more details).

Definition 2.1. *(Discounted gain process) Let $\{G_n^\delta\}_{n=0}^N$ be the discounted gain process associated with the strategy δ where G_n^δ is the discounted gain at time t_n prior to the rebalancing.*

$G_0^\delta := 0$ *and*

$$G_n^\delta := \sum_{k=1}^{n} \delta_k^{(1:D)} \bullet (B_k^{-1} S_k - B_{k-1}^{-1} S_{k-1}), \quad n = 1, 2, \ldots, N, \tag{2.1}$$

where $\bullet$ *is the scalar product operator, i.e. for two* n*-dimensional vectors* X *and* Y, $X \bullet Y := \sum_{i=1}^{n} X_i Y_i$.

Definition 2.2. *(Self-financing) The process* δ *is said to be a self-financing trading strategy if* δ *is* $\mathbb{F}$*-predictable, i.e.* $\delta_0^{(j)} \in \mathcal{F}_0$ *and* $\delta_{n+1}^{(j)} \in \mathcal{F}_n$ *for* $j = 0, \ldots, D$ *and for* $n = 0, \ldots, N-1$, *and if*

$$\delta_{n+1}^{(1:D)} \bullet S_n + \delta_{n+1}^{(0)} B_n = \delta_n^{(1:D)} \bullet S_n + \delta_n^{(0)} B_n, \quad n = 0, 1, \ldots, N-1. \tag{2.2}$$

A self-financing strategy δ *implies the absence of cash infusions into or withdrawals from the portfolio except possibly at time* 0.

Definition 2.3. *(Hedging portfolio value) Define* $\{V_n^\delta\}_{n=0}^N$ *as the hedging portfolio value process associated with the strategy* δ, *where the time-*t_n *portfolio value is given by* $V_n^\delta := \delta_n^{(1:D)} \bullet S_n + \delta_n^{(0)} B_n$, $n = 0, \ldots, N$.

Remark 2.1. *It can be shown, see for instance Lamberton and Lapeyre (2011), that* δ *is self-financing if and only if* $V_n^\delta = B_n(V_0 + G_n^\delta)$ *for* $n = 0, 1, \ldots, N$.

Definition 2.4. *(Admissible trading strategies) Let* Π *be the convex set of admissible trading strategies which consists of all sufficiently well-behaved self-financing trading strategies.*

2.2.2 Convex risk measures

In an incomplete market, perfect replication is impossible and the hedger must accept that some risks cannot be fully hedged. As such, an optimal hedging strategy (also referred to as a global hedging strategy) is defined as one that minimizes a criterion based on the closeness between the hedging portfolio value and the payoff of the contingent claim at maturity (the difference between two such quantities is referred to as the hedging error). Many different measures of distance can be used to represent the risk aversion of the hedger. In this paper, convex risk measures as defined in Föllmer and Schied (2002) are considered. As discussed in Marzban et al. (2020) and shown in the current section, the use of a convex risk measure to characterize the risk aversion of the hedger enhances the tractability of the equal risk pricing framework.

Definition 2.5. *(Convex risk measure) Let $\mathcal{X}$ be a set of random variables representing liabilities and $X_1, X_2 \in \mathcal{X}$. As defined in Föllmer and Schied (2002), $\rho : \mathcal{X} \to \mathbb{R}$ is a convex risk measure if it satisfies the following properties:*

(i) Monotonicity: $X_1 \leq X_2 \Rightarrow \rho(X_1) \leq \rho(X_2)$. A larger liability is riskier.

(ii) Translation invariance: For $a \in \mathbb{R}$, $\rho(X + a) = \rho(X) + a$. This implies that the hedger is indifferent between an empty portfolio and a portfolio with a liability X and a cash amount of $\rho(X)$:

$$\rho(X - \rho(X)) = \rho(X) - \rho(X) = 0.$$

(iii) Convexity: For $0 \leq \lambda \leq 1$, $\rho(\lambda X_1 + (1-\lambda)X_2) \leq \lambda\rho(X_1) + (1-\lambda)\rho(X_2)$. Diversification does not increase risk.

2.2.3 Optimal hedging problem

For the rest of the paper, let ρ be the convex risk measure used to characterize the risk aversion of the hedger for both the long and short positions in the usual contingent claim. Also, assume without loss of generality (w.l.o.g.) that the position in the hedging portfolio is long for both the long and short positions in the derivative.

Definition 2.6. *(Long and short sided risk) Define $\epsilon^{(L)}(V_0)$ and $\epsilon^{(S)}(V_0)$ respectively as the measured risk exposure of a long and short position in the derivative under the optimal hedge* *if the value of the initial hedging portfolio is* V_0 :

$$\epsilon^{(L)}(V_0) := \min_{\delta \in \Pi} \rho\left(-\Phi(S_N, Z_N) - B_N(V_0 + G_N^{\delta})\right), \tag{2.3}$$

$$\epsilon^{(S)}(V_0) := \min_{\delta \in \Pi} \rho\left(\Phi(S_N, Z_N) - B_N(V_0 + G_N^{\delta})\right). \tag{2.4}$$

It is worth highlighting that the measured risk exposure of the long (resp. short) position in the derivative decreases (resp. increases) with the payoff value Φ. Furthermore, since the long (resp. short) position is assumed to buy (sell) the option at time 0, the initial portfolio value V_0 of the long (resp. short) position will be negative (resp. positive).

Remark 2.2. *An assumption implicit to Definition 2.6 is that the minimum in* (2.3) *or* (2.4) *is indeed attained by some trading strategy, i.e. that the infimum is in fact a minimum. Although the identification of sufficient conditions leading to the existence of an optimal policy is left out-of-scope, such conditions were investigated in other literature works (see Section 7.2 for additional information).*

We emphasize that the optimal risk exposures of the long and short position as defined in (2.3) and (2.4) are reached through two distinct hedging strategies. Furthermore, due to the translation invariance property of ρ, the long and short measured risk exposures have also the following representation:

$$\epsilon^{(L)}(V_0) = \epsilon^{(L)}(0) - B_N V_0, \quad \epsilon^{(S)}(V_0) = \epsilon^{(S)}(0) - B_N V_0. \tag{2.5}$$

Definition 2.7. *(Optimal hedging) Let* $\delta^{(L)}$ *and* $\delta^{(S)}$ *be respectively the optimal hedging strategies for the long and short positions in the derivative:*

$$\delta^{(L)} := \arg\min_{\delta \in \Pi} \rho\left(-\Phi(S_N, Z_N) - B_N(V_0 + G_N^{\delta})\right), \tag{2.6}$$

$$\delta^{(S)} := \arg\min_{\delta \in \Pi} \rho\left(\Phi(S_N, Z_N) - B_N(V_0 + G_N^{\delta})\right). \tag{2.7}$$

The translation invariance property of ρ also implies that the optimal hedging strategies $\delta^{(L)}$ and $\delta^{(S)}$ do not depend on the initial portfolio value as shown below for $\delta^{(L)}$:

$$\begin{aligned}\delta^{(L)} &= \underset{\delta\in\Pi}{\arg\min}\, \rho\left(-\Phi(S_N, Z_N) - B_N(V_0 + G_N^{\delta})\right)\\ &= \underset{\delta\in\Pi}{\arg\min}\, \left\{\rho\left(-\Phi(S_N, Z_N) - B_N G_N^{\delta}\right) - B_N V_0\right\}\\ &= \underset{\delta\in\Pi}{\arg\min}\, \rho\left(-\Phi(S_N, Z_N) - B_N G_N^{\delta}\right).\end{aligned}$$

Similar steps show that:

$$\delta^{(S)} = \underset{\delta\in\Pi}{\arg\min}\, \rho\left(\Phi(S_N, Z_N) - B_N G_N^{\delta}\right).$$

2.2.4 Option pricing and ϵ-completeness measure

The current section outlines the equal risk pricing criterion to determine the price of a derivative. It entails finding a price for which the risk exposure to both the long position and short position hedgers are equal. One important concept in the valuation of contingent claims is the absence of arbitrage. In this paper, the notions of super-replication and sub-replication are used to define arbitrage-free pricing.

Definition 2.8 (*Super-replication and sub-replication strategies*)**.** *A super-replication strategy for a contingent claim Φ is defined as a pair (v, δ) such that $v \in \mathbb{R}$, and δ is an admissible hedging strategy for which $V_0^{\delta} = v$ and $V_N^{\delta} = B_N(v + G_N^{\delta}) \geq \Phi(S_N, Z_N)$ $\mathbb{P}$-a.s. Super-replication is a conservative approach to hedging which can be used by a seller of Φ to remove all residual hedging risk. Let $\bar{v}$ be the greatest lower bound of the set of initial portfolio values for which a super-replication strategy exists:*

$$\bar{v} := \inf\left\{v : \exists\delta \in \Pi \text{ such that } \mathbb{P}\left[B_N(v + G_N^{\delta}) \geq \Phi(S_N, Z_N)\right] = 1\right\}. \qquad (2.8)$$

$\bar{v}$ is called the super-replication price of Φ and it represents an upper bound of the set of

arbitrage-free prices for Φ*. Similarly, a sub-replication strategy is a pair* (v, δ) *that completely removes the hedging risk exposure associated with a long position in* Φ*, i.e. for which* $B_N(v + G_N^\delta) \leq \Phi(S_N, Z_N)$ $\mathbb{P}$*-a.s. The least upper bound of the set of portfolio values such that a sub-replication strategy exists is called the sub-replication price of* Φ *and is a lower bound of the set of arbitrage-free prices for* Φ*:*

$$\underline{v} := \sup \left\{ v : \exists \delta \in \Pi \text{ such that } \mathbb{P}\left[B_N(v + G_N^\delta) \leq \Phi(S_N, Z_N) \right] = 1 \right\}. \tag{2.9}$$

The following defines the concept of arbitrage-free pricing considered in this study.

Definition 2.9 (Arbitrage-free pricing)**.** *Assume that both super-replication and sub-replication strategies for* Φ *exist. Then, the price of* Φ *is said to be arbitrage-free if it falls within the interval* $[\underline{v}, \bar{v}]$ *as defined in* (2.8) *and* (2.9).

The problem of evaluating the boundaries of the arbitrage-free interval of prices, i.e. $\underline{v}$ and $\bar{v}$, is out-of-scope of the current paper. The reader is referred to El Karoui and Quenez (1995) for a formulation of the latter problem as a stochastic control problem which can be solved for instance with dynamic programming. The price of a contingent claim under the equal risk pricing framework can now be defined.

Definition 2.10 (*Equal risk price for European-type claims*)**.** *The equal risk price* $C_0^\star$ *of the contingent claim* Φ *is defined as the initial portfolio value such that the optimally hedged measured risk exposure of both the long and short positions in the derivative are equal, i.e.* $C_0^\star := C_0$ *such that:*

$$\epsilon^{(L)}(-C_0) = \epsilon^{(S)}(C_0). \tag{2.10}$$

Remark 2.3. *Contrarily to Guo and Zhu (2017), in the current paper, the optimal hedging strategy minimizes risk under the physical measure instead of under some risk-neutral measure. Two main reasons led to this modification of the original approach found in Guo and Zhu (2017).*

First, under incomplete markets, the choice of the risk-neutral measure is arbitrary, whereas the physical measure can be more objectively determined using econometrics techniques. Having the price being determined under the physical measure removes the subjectivity associated with the choice of the martingale measure. Secondly, under a risk-neutral measure $\mathbb{Q}$, the price is already characterized by the discounted expected payoff $\mathbb{E}^{\mathbb{Q}}\left[e^{-rT}\Phi(S_N, Z_N)\right]$, which makes interpretation of the risk-neutral equal risk price questionable.

Before introducing results showing that equal risk option prices are arbitrage-free, a technical assumption on which the proofs rely is outlined.

Assumption 2.1. *As in Xu (2006) and Marzban et al. (2020), assume that the risk associated to hedging losses is bounded below across all admissible trading strategies, i.e.* $\min_{\delta \in \Pi} \rho(-B_N G_N^{\delta}) > -\infty$.

The next theorem provides a characterization of equal risk prices. It also indicates that equal risk prices of contingent claims with a finite super-replication price are arbitrage-free. The representation of $C_0^\star$ in (2.11) is analogous to results found in Marzban et al. (2020) who considers a similar setup with convex risk measures. Although the arbitrage-free result is also stated in Marzban et al. (2020), a formal proof was not given.

Theorem 2.1 (Absence of arbitrage)**.** *Assume that there exist a super-replication and sub-replication strategy for Φ with a finite super-replication price. Assume also that if $\delta, \tilde{\delta} \in \Pi$, then $-\delta, -\tilde{\delta} \in \Pi$ and $\delta + \tilde{\delta} \in \Pi$. Then, under Assumption 2.1, the equal risk price $C_0^\star$ from Definition 2.10 exists, is unique, is arbitrage-free and can be expressed as*

$$C_0^\star = \frac{\epsilon^{(S)}(0) - \epsilon^{(L)}(0)}{2B_N}. \tag{2.11}$$

Remark 2.4. *The representation of $C_0^\star$ in (2.11) is analogous to results found in Marzban et al. (2020), but their work considers two different convex risk measures to assess the global risk exposures of the long and short positions, respectively. In their Lemma 2.2, they obtain necessary conditions guaranteeing that the equal risk price is arbitrage-free, but one of such* *conditions is the existence of what they refer to as a fair price interval. However, sufficient conditions guaranteeing the existence of such an interval are not provided; therefore their Lemma 2.2 is essentially a conditional result. Using the same risk measure for both the long and short positions as in the current work allows obtaining an unconditional result as the fair price interval is guaranteed to exist in that case, see the inequality* $\zeta^{(L)} \leq \zeta^{(S)}$ *in the proof of Theorem 2.1 from the current paper. Showing this result is a novel contribution of the current paper. Thus, the choice of considering identical risk measures for both the long and short positions in the current work stems from theoretical considerations.*

We now propose measures to quantify the residual risk faced by hedgers of the contingent claim. Such measures are analogous to but more general than the one proposed in Bertsimas et al. (2001) who study the case of variance-optimal hedging.

Definition 2.11 (ϵ *market completeness measure*)**.** *Define* $\epsilon^\star$ *as the level of residual risk faced by the hedger of any of the short or long position in the contingent claim if its price is the equal risk price and optimal hedging strategies are used for both positions:*

$$\epsilon^\star := \epsilon^{(L)}(-C_0^\star) = \epsilon^{(S)}(C_0^\star). \tag{2.12}$$

$\epsilon^\star$ *and* $\epsilon^\star/C_0^\star$ *are referred to as respectively the measured residual risk exposure per derivative contract and per dollar invested.*

As shown below using (2.5) and (2.11), $\epsilon^\star$ is the average of the measured risk exposure of both long and short optimally hedged positions in Φ assuming that the initial value of the portfolio is zero:

$$
\begin{aligned}
\epsilon^\star &= \epsilon^{(L)}(-C_0^\star) \\
&= \epsilon^{(L)}(0) + B_N C_0^\star \\
&= \epsilon^{(L)}(0) + B_N \left(\frac{\epsilon^{(S)}(0) - \epsilon^{(L)}(0)}{2B_N} \right) \\
&= \frac{\epsilon^{(L)}(0) + \epsilon^{(S)}(0)}{2}.
\end{aligned} \tag{2.13}
$$

Remark 2.5. *Bertsimas et al. (2001) proposed instead the following measure of market incompleteness:*

$$\varepsilon^\star = \min_{V_0,\delta} \mathbb{E}[(\Phi(S_N, Z_N) - B_N(V_0 + G_N^\delta))^2],$$

where the expectation is taken with respect to the physical measure. Our measure $\epsilon^\star$ has the advantage of characterizing the risk aversion of the hedger with a convex risk measure, contrarily to Bertsimas et al. (2001) who are restricted to the use of a quadratic penalty. Using the latter penalty entails that hedging gains are penalized during the optimization of the hedging strategy, which is clearly undesirable. The ability to rely on convex measures in the current scheme for risk quantification allows for an asymmetric treatment of hedging gains and losses which is more consistent of actual objectives of the hedging agents.

As argued by Bertsimas et al. (2001), market incompleteness is often described in the literature as a binary concept whereas in practice, it is much more natural to consider different degrees of incompleteness implying different levels of residual hedging risk. The measure $\epsilon^\star$ allows determining where is any contingent claim situated within the spectrum of incompleteness and whether it is easily hedgeable or not. As discussed in Bertsimas et al. (2001), a single metric such as $\epsilon^\star$ might not be sufficient for a complete depiction of the level of market incompleteness associated with a contingent claim. For instance, it does not depict the entire hedging error distribution, nor does it directly indicate which scenarios are the main drivers of hedging residual risk. Nevertheless, $\epsilon^\star$ is still a good indication of the efficiency of the optimal hedging procedure for a given derivative. Moreover, sensitivity analyses over $\epsilon^\star$ with respect to various model dynamics can be done to assess the impact of the different sources of market incompleteness. Numerical experiments in Section 2.4 will attempt to provide some insight on drivers of $\epsilon^\star$.

2.3 Tractable solution to equal risk pricing

In the current section, a tractable solution is proposed to implement the equal risk pricing framework. The approach uses the recent deep hedging algorithm of Buehler et al. (2019b) to train two distinct neural networks which are used to approximate the optimal hedging strategy respectively for the long and the short position in the derivative.

2.3.1 Feedforward neural network

For convenience, a very similar notation for neural networks as the one introduced by Buehler et al. (2019b) is used (see Section 4 of their paper). The reader is referred to Goodfellow et al. (2016) for a general description of neural networks.

Definition 2.12 (*Feedforward neural network*)**.** *Let $X \in \mathbb{R}^{d_{in}\times 1}$ be a feature vector of dimensions $d_{in} \in \mathbb{N}$ and $L, d_1, \ldots, d_{L-1}, d_{out} \in \mathbb{N}$ with $L \geq 2$. Define a feedforward neural network (FFNN) as the mapping $F_\theta : \mathbb{R}^{d_{in}} \to \mathbb{R}^{d_{out}}$ with trainable parameters θ:*

$$F_\theta(X) := A_L \circ F_{L-1} \circ \ldots \circ F_1, \tag{2.14}$$
$$F_l := \sigma \circ A_l, \quad l = 1, \ldots, L-1,$$

where $\circ$ denotes the function composition operator, and for any $l = 1, \ldots, L$, the function A_l is defined through $A_l(Y) := W^{(l)}Y + b^{(l)}$ with

- $W^{(1)} \in \mathbb{R}^{d_1 \times d_{in}}$, $b^{(1)} \in \mathbb{R}^{d_1 \times 1}$ *and* $Y \in \mathbb{R}^{d_{in}\times 1}$ *if* $l = 1$,
- $W^{(l)} \in \mathbb{R}^{d_l \times d_{l-1}}$, $b^{(l)} \in \mathbb{R}^{d_l \times 1}$ *and* $Y \in \mathbb{R}^{d_{l-1}\times 1}$ *if* $l = 2, \ldots, L-1$ *and* $L \geq 3$,
- $W^{(L)} \in \mathbb{R}^{d_{out} \times d_{L-1}}$, $b^{(L)} \in \mathbb{R}^{d_{out} \times 1}$ *and* $Y \in \mathbb{R}^{d_{L-1}\times 1}$ *if* $l = L$.

The activation function $\sigma : \mathbb{R} \to \mathbb{R}$ is applied element-wise to outputs of the pre-activation functions A_l. Moreover,

$$\theta := \{W^{(1)}, \dots, W^{(L)}, b^{(1)}, \dots, b^{(L)}\} \tag{2.15}$$

is the set of trainable parameters of the FFNN.

The following definition of sets of FFNN will be used throughout the rest of the section to define, for instance, the two neural networks used for hedging the long and short position in the derivative, the tractable solution to the equal risk pricing framework and the optimization procedure of neural networks.

Definition 2.13 (*Sets of FFNN*). *Let $\mathcal{NN}^{\sigma}_{\infty,d_{in},d_{out}}$ be the set of all FFNN mapping from $\mathbb{R}^{d_{in}} \to \mathbb{R}^{d_{out}}$ as in Definition 2.12 with a fixed activation function σ and an arbitrary number of layers and neurons per layer. Since a unique activation function is considered in the numerical section, let $\mathcal{NN}_{\infty,d_{in},d_{out}} := \mathcal{NN}^{\sigma}_{\infty,d_{in},d_{out}}$. Moreover, for all $M \in \mathbb{N}$ and $R_M \in \mathbb{N}$ that depends on M, let $\Theta_{M,d_{in},d_{out}} \subseteq \mathbb{R}^{R_M}$. Define $\mathcal{NN}_{M,d_{in},d_{out}}$ as the set of neural networks F_θ as in (2.14) with $\theta \in \Theta_{M,d_{in},d_{out}}$:*

$$\mathcal{NN}_{M,d_{in},d_{out}} := \{F_\theta | \theta \in \Theta_{M,d_{in},d_{out}}\}. \tag{2.16}$$

The sequence of sets $\{\mathcal{NN}_{M,d_{in},d_{out}}\}_{M\in\mathbb{N}}$ is assumed to have the following properties:

- *For any $M \in \mathbb{N}$: $\mathcal{NN}_{M,d_{in},d_{out}} \subset \mathcal{NN}_{M+1,d_{in},d_{out}}$ where $\subset$ denotes strict inclusion,*
- *$\bigcup_{M\in\mathbb{N}} \mathcal{NN}_{M,d_{in},d_{out}} = \mathcal{NN}_{\infty,d_{in},d_{out}}$.*

This definition of sets of FFNN introduced by Buehler et al. (2019b) is very convenient as the sets $\{\mathcal{NN}_{M,d_{\text{in}},d_{\text{out}}}\}_{N\in\mathbb{N}}$ can be used to describe two cases of interest in deep learning. Here are two different possible definitions for $\mathcal{NN}_{M,d_{\text{in}},d_{\text{out}}}$.

(A) Let $\{L^{(M)}\}_{M\in\mathbb{N}}$, $\{d_1^{(M)}\}_{M\in\mathbb{N}}$, $\{d_2^{(M)}\}_{M\in\mathbb{N}}, \dots$ be non-decreasing integer sequences. Then, $\mathcal{NN}_{M,d_{\text{in}},d_{\text{out}}}$ is defined as the set of all FFNN mapping from $\mathbb{R}^{d_{\text{in}}} \to \mathbb{R}^{d_{\text{out}}}$ with a fixed structure of $L^{(M)}$ layers and of $d_1^{(M)}, \dots, d_{L^{(M)}-1}^{(M)}, d_{\text{out}}$ neurons per layer. This case is useful for the problem of fitting the trainable parameters θ with a fixed set of hyperparameters.

(B) Let $\mathcal{NN}_{M,d_{\text{in}},d_{\text{out}}}$ be the set of all FFNN mapping from $\mathbb{R}^{d_{\text{in}}} \to \mathbb{R}^{d_{\text{out}}}$ for an arbitrary number of layers and number of neurons per layer with at most M non-zero trainable parameters. This case is useful to describe the complete optimization problem of neural networks which include the selection of hyperparameters, often called hyperparameters tuning.

Unless specified otherwise, one can assume w.l.o.g. either definition for $\{\mathcal{NN}_{M,d_{\text{in}},d_{\text{out}}}\}_{M\in\mathbb{N}}$.

2.3.2 Equal risk pricing with two neural networks

To formulate how two distinct neural networks can approximate arbitrarily well the optimal hedging of the long and short position in a derivative, the following assumption is applied for the rest of the paper.

Assumption 2.2. *For each position (long and short) in the derivative, there exists a function $f : \mathbb{R}^{\tilde{D}} \to \mathbb{R}^D$ (distinct for the long and short position) such that at each rebalancing date, the optimal hedge is of the form $\delta_{n+1}^{(1:D)} = f(S_n, V_n, \mathcal{I}_n, T - t_n)$ where $\tilde{D} := D + 2 + dim(\mathcal{I}_n)$ with $\mathcal{I}_n$ being some random vector encompassing relevant necessary information to compute the optimal hedging strategy, which depends on the market setup considered.*

Note that Assumption 2.2 typically holds for low-dimension processes $\{\mathcal{I}_n\}_{n=0}^N$ when some form of Markov dynamics common in the hedging literature is assumed. See, for example, François et al. (2014) for the case of regime-switching models.

In what follows, $\mathcal{L}$ and $\mathcal{S}$ used both as subscripts and superscripts denote respectively the long and short position hedges.

Definition 2.14 (*Hedging with two neural networks*)**.** *Let $X_n := (S_n, V_n, \mathcal{I}_n, T - t_n) \in \mathbb{R}^{\tilde{D}}$ be the feature vector for each trading time $t_n \in \{t_0, \dots, t_{N-1}\}$. For some $M_{\mathcal{L}} \in \mathbb{N}$, let*

$F_\theta^{\mathcal{L}} \in \mathcal{NN}_{M_{\mathcal{L}},\tilde{D},D}$ be a FFNN. Given X_n as an input, $F_\theta^{\mathcal{L}}$ outputs a D-dimensional vector of the number of shares of each of the D risky assets held in the hedging portfolio of the long position during the period $(t_n, t_{n+1}]$, i.e. $\delta_{n+1}^{(1:D)} = F_\theta^{\mathcal{L}}(X_n)$. Similarly, for some $M_{\mathcal{S}} \in \mathbb{N}$, $F_\theta^{\mathcal{S}} \in \mathcal{NN}_{M_{\mathcal{S}},\tilde{D},D}$ is a distinct FFNN which computes the position in the D risky assets to hedge the short position in the option at each time step. These two FFNN are referred to as the long-$\mathcal{NN}$ and short-$\mathcal{NN}$.

Remark 2.6. *In the current paper's approach, the two neural networks are trained separately to minimize different cost functions. As such, $F_\theta^{\mathcal{L}}$ and $F_\theta^{\mathcal{S}}$ will possibly have a different structure, e.g. different number of layers and number of neurons per layer, and different values of trainable parameters.*

The problem of evaluating the measured risk exposure of the long and short positions under optimal hedging can now be formulated as a classical deep learning optimization problem. Since the input and output of $F_\theta^{\mathcal{L}}$ and $F_\theta^{\mathcal{S}}$ are always respectively of dimensions $\tilde{D}$ and D, let $\Theta_{M_{\mathcal{L}}} := \Theta_{M_{\mathcal{L}},\tilde{D},D}$ and $\Theta_{M_{\mathcal{S}}} := \Theta_{M_{\mathcal{S}},\tilde{D},D}$ be the sets of trainable parameters values as in (2.16) for respectively the long-$\mathcal{NN}$ and short-$\mathcal{NN}$.

Definition 2.15 (*Long and short sided risk with two neural networks*)**.** *For $M_{\mathcal{L}}, M_{\mathcal{S}} \in \mathbb{N}$, define $\epsilon^{(M_{\mathcal{L}})}(V_0)$ and $\epsilon^{(M_{\mathcal{S}})}(V_0)$ as the measured risk exposure of the long and short position in the derivative if $F_\theta^{\mathcal{L}}$ and $F_\theta^{\mathcal{S}}$ are used to compute the hedging strategies and the initial hedging portfolio value is V_0:*

$$\epsilon^{(M_{\mathcal{L}})}(V_0) := \min_{\theta \in \Theta_{M_{\mathcal{L}}}} \rho\left(-\Phi(S_N, Z_N) - B_N(V_0 + G_N^{\delta^{(\mathcal{L},\theta)}})\right), \tag{2.17}$$

$$\epsilon^{(M_{\mathcal{S}})}(V_0) := \min_{\theta \in \Theta_{M_{\mathcal{S}}}} \rho\left(\Phi(S_N, Z_N) - B_N(V_0 + G_N^{\delta^{(\mathcal{S},\theta)}})\right), \tag{2.18}$$

where $\delta^{(\mathcal{L},\theta)}$ and $\delta^{(\mathcal{S},\theta)}$ in (2.17) and (2.18) are to be understood respectively as the trading strategies obtained through $F_\theta^{\mathcal{L}}$ and $F_\theta^{\mathcal{S}}$.

Remark 2.7. *Suppose Assumption 2.2 is satisfied. Using the universal function approximation theorem of Hornik (1991) which essentially states that a FFNN approximates multivariate* *functions arbitrarily well, Buehler et al. (2019b) show that for any well-behaved and integrable enough asset prices dynamics and contingent claims (see Proposition 4.3 of their paper*[1]*):*

$$\lim_{M_{\mathcal{S}}\to\infty} \epsilon^{(M_{\mathcal{S}})}(0) = \epsilon^{(S)}(0), \quad \lim_{M_{\mathcal{L}}\to\infty} \epsilon^{(M_{\mathcal{L}})}(0) = \epsilon^{(L)}(0). \tag{2.19}$$

Thus, this result shows that for both the long and short positions, there exists a large FFNN which can approximate arbitrarily well the optimal hedging strategy.

The equal risk pricing approach and the measure of market incompleteness described in Section 2.2 can now be restated with the use of the long-$\mathcal{NN}$ and short-$\mathcal{NN}$. Let $C_0^{(\star,\mathcal{NN})}$ and $\epsilon^{(\star,\mathcal{NN})}$ be respectively the equal risk price as in Definition 2.10 and the measure of market incompleteness as in Definition 2.11 if the risk exposure of the long and short position in Φ are measured with $\epsilon^{(M_{\mathcal{L}})}$ and $\epsilon^{(M_{\mathcal{S}})}$. Similar steps as in the proof of Theorem 2.1 and (2.13) leads to the following representation for $C_0^{(\star,\mathcal{NN})}$ and $\epsilon^{(\star,\mathcal{NN})}$:

$$C_0^{(\star,\mathcal{NN})} = \frac{\epsilon^{(M_{\mathcal{S}})}(0) - \epsilon^{(M_{\mathcal{L}})}(0)}{2B_N}, \quad \epsilon^{(\star,\mathcal{NN})} = \frac{\epsilon^{(M_{\mathcal{L}})}(0) + \epsilon^{(M_{\mathcal{S}})}(0)}{2}. \tag{2.20}$$

Moreover, a direct consequence of Remark 2.7 applied to $C_0^{(\star,\mathcal{NN})}$ and $\epsilon^{(\star,\mathcal{NN})}$ as stated in (2.20) is that the current paper's approach can approximate arbitrarily well the true equal risk price and measure of incompleteness, i.e.:

$$\lim_{M_{\mathcal{S}},M_{\mathcal{L}}\to\infty} C_0^{(\star,\mathcal{NN})} = C_0^{\star}, \quad \lim_{M_{\mathcal{S}},M_{\mathcal{L}}\to\infty} \epsilon^{(\star,\mathcal{NN})} = \epsilon^{\star}.$$

[1] Buehler et al. (2019b) consider a more general market with a filtration generated by a process $\{I_k\}$ where $I_k \in \mathbb{R}^d$ contains any new market information at time t_k. They use a distinct neural network at each trading date which can be a function of $(I_0, \ldots, I_k, \delta_k)$ at time t_k. From remarks 5 and 6 of Buehler et al. (2019b), the convergence result (2.19) holds under Assumption 2.2 by using instead a single FFNN for both the long and short position for all time steps as in Definition 2.14 of the current paper.

2.3.3 Optimization of feedforward neural networks

The training procedure of the long-$\mathcal{NN}$ and short-$\mathcal{NN}$ consists in searching for their optimal parameters $\theta^{(\mathcal{L})} \in \Theta_{M_{\mathcal{L}}}$ and $\theta^{(\mathcal{S})} \in \Theta_{M_{\mathcal{S}}}$ to minimize the measured risk exposures as in (2.17) and (2.18). The approach utilized in this paper is based on the deep hedging algorithm of Buehler et al. (2019b). The training procedure of the short-$\mathcal{NN}$ with (minibatch) stochastic gradient descent (SGD), a very popular algorithm in deep learning, is presented. It is straightforward to adapt the latter to the long-$\mathcal{NN}$ with a simple modification to the cost function (2.21) that follows. Let $J(\theta)$ be the cost function to be minimized for the short derivative position hedge[2]:

$$J(\theta) := \rho\left(\Phi(S_N, Z_N) - B_N G_N^{\delta^{(\mathcal{S},\theta)}}\right), \quad \theta \in \Theta_{M_{\mathcal{S}}}. \tag{2.21}$$

Denote $\theta_0 \in \Theta_{M_{\mathcal{S}}}$ as the initial[3] parameter values of $F_\theta^{\mathcal{S}}$. The classical SGD algorithm consists in updating iteratively the trainable parameters as follows:

$$\theta_{j+1} = \theta_j - \eta_j \nabla_\theta J(\theta_j), \tag{2.22}$$

where ∇_θ denotes the gradient operator with respect to θ and η_j is a small positive deterministic value which is typically progressively reduced through iterations, i.e. as j increases. Recall that in the current framework, a synthetic market is considered where paths of the hedging instruments can be simulated. Let $N_{\text{batch}} \in \mathbb{N}$ be the size of a simulated minibatch $\mathbb{B}_j :=$

[2] Recall from (2.18) that the relation between $J(\theta)$ and the measured risk exposure of the short position is

$$\epsilon^{(M_{\mathcal{S}})}(0) = \min_{\theta \in \Theta_{M_{\mathcal{S}}}} J(\theta).$$

[3] In this paper, the initialization of θ is always done with the Glorot uniform initialization from Glorot and Bengio (2010).

$\{\pi_{i,j}\}_{i=1}^{N_{\text{batch}}}$ with $\pi_{i,j}$ being the i^{th} hedging error if the trainable parameters are $\theta = \theta_j$:

$$\pi_{i,j} := \Phi(S_{N,i}, Z_{N,i}) - B_N G_{N,i}^{\delta^{(\mathcal{S},\theta_j)}}. \tag{2.23}$$

Moreover, let $\hat{\rho}(\mathbb{B}_j)$ be the empirical estimator of $\rho(\pi_{i,j})$. The gradient of the cost function $\nabla_\theta J(\theta_j)$ is estimated with $\nabla_\theta \hat{\rho}(\mathbb{B}_j)$ evaluated at $\theta = \theta_j$.

In the numerical section, the convex risk measure is assumed to be the Conditional Value-at-Risk (CVaR) as defined in Rockafellar and Uryasev (2002). For $\alpha \in (0, 1)$, such a risk measure can be formally defined as

$$\text{VaR}_\alpha(X) := \min_x \{x | \mathbb{P}(X \leq x) \geq \alpha\}, \quad \text{CVaR}_\alpha(X) := \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_\gamma(X) d\gamma$$

where $\text{VaR}_\alpha(X)$ is the Value-at-Risk (VaR) of confidence level α. For an absolutely continuous integrable random variable[4], the CVaR has the following representation

$$\text{CVaR}_\alpha(X) := \mathbb{E}[X | X \geq \text{VaR}_\alpha(X)], \quad \alpha \in (0, 1). \tag{2.24}$$

The CVaR has been extensively used in the risk management literature as it considers tail risk by averaging all losses larger than the VaR. For a simulated minibatch of hedging errors $\mathbb{B}_j$, let $\{\pi_{[i],j}\}_{i=1}^{N_{\text{batch}}}$ be the corresponding ordered sequence and $\tilde{N} := \lceil \alpha N_{\text{batch}} \rceil$ where $\lceil x \rceil$ is the ceiling function (i.e. the smallest integer greater or equal to x). Following the work of Hong et al. (2014) (see Section 2 of their paper), let $\widehat{\text{VaR}}_\alpha(\mathbb{B}_j)$ and $\widehat{\text{CVaR}}_\alpha(\mathbb{B}_j)$ be the estimators of the VaR and CVaR of the short hedging error at confidence level α:

$$\widehat{\text{VaR}}_\alpha(\mathbb{B}_j) := \pi_{[\tilde{N}],j},$$

[4] In Section 2.4, the only dynamics considered for the risky assets produce integrable and absolutely continuous hedging errors.

$$\widehat{\text{CVaR}}_\alpha(\mathbb{B}_j) := \widehat{\text{VaR}}_\alpha(\mathbb{B}_j) + \frac{1}{(1-\alpha)N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \max(\pi_{i,j} - \widehat{\text{VaR}}_\alpha(\mathbb{B}_j), 0).$$

Note that $\widehat{\text{CVaR}}_\alpha(\mathbb{B}_j)$ depends of every trainable parameters in θ_j as the $\pi_{i,j}$ are functions of the trading strategy produced by the output of the short-$\mathcal{NN}$. Furthermore, since the gain process and the trading strategy are linearly dependent, $\widehat{\text{CVaR}}_\alpha(\mathbb{B}_j)$ is also linearly dependent of the trading strategy. The latter implies that $\nabla_\theta \widehat{\text{CVaR}}_\alpha(\mathbb{B}_j)$ can be computed exactly as the gradient of the output of a FFNN with respect to the trainable parameters can be computed exactly (see e.g. Goodfellow et al. (2016)). Moreover, a very popular algorithm in deep learning to compute the gradient of a cost function with respect to the parameters is backpropagation (Rumelhart et al., 1986), often called backprop. Backprop leverages efficiently the structure of neural networks and the chain rule of calculus to obtain such gradient. In practice, deep learning libraries such as Tensorflow (Abadi et al., 2016) are often used to implement backprop. Moreover, sophisticated SGD algorithms such as Adam (Kingma and Ba, 2014) which dynamically adapt the η_j in (2.22) over time have been shown to improve the training of neural networks. For all of the numerical experiments in Section 2.4, Tensorflow and Adam were used.

Remark 2.8. *It can be shown that $\widehat{CVaR}_\alpha(\mathbb{B}_j)$ is biased in finite sample size, but is a consistent and asymptotically normal estimator of the CVaR (see e.g. Theorem 2 of Trindade et al. (2007)). The specific impacts of this bias on the optimization procedure presented in this paper are out-of-scope. Multiple considerations are typically used to determine the minibatch size. It is often treated as an additional hyperparameter (see e.g. Chapter 8.1.3 of Goodfellow et al. (2016) for additional details). Numerical results presented in Section 2.4 of the current paper are robust to different minibatch sizes, i.e. no significant difference was observed under different minibatch sizes.*

Remark 2.9. *As mentioned in Remark 2.4, identical risk measures for both the long and short positions are considered in the current work. However, had different risk measures been considered for the long and short positions, as in Marzban et al. (2020) for instance, the* *numerical algorithm described in the current section would have been essentially identical. Indeed, the training of the short-$\mathcal{NN}$ and long-$\mathcal{NN}$ is always done separately even with a unique convex risk measure for both positions, as the two neural networks minimize two different cost functions. Thus, one could consider two different convex risk measures in the hedging problem with no modifications to the numerical algorithm.*

2.4 Numerical results

This section illustrates the implementation of the equal risk pricing framework under different market setups. Our analysis starts off in Section 2.4.2 with a sensitivity analyses of equal risk prices and residual hedging risk in relation with the choice of convex risk measure. The assessment of the impact of different empirical properties of assets returns on the equal risk pricing framework is performed in Section 2.4.3. A comparison with benchmarks consisting in risk-neutral expected prices under commonly used EMMs is also presented. Section 2.4.4 shows that the current paper's approach is very general and is able to price exotic derivatives and assess their associated residual hedging risk. The setup for the latter numerical experiments is detailed in Section 2.4.1.

2.4.1 Numerical procedure

A single risky asset (i.e. $D = 1$) of initial price $S_0 = 100$ is considered. It can be assumed for convenience to be a non-dividend paying stock.[5] The annualized continuous risk-free rate is $r = 0.02$. Daily rebalancing with 260 business days per year is applied, i.e. $t_i - t_{i-1} = 1/260$ for $i = 1, \ldots, N$. The contingent claim to be priced is a vanilla European put option on the risky asset with maturity $T = 60/260$. Different levels of moneyness are considered: $K = 90$ for OTM, $K = 100$ for at-the-money (ATM) and $K = 110$ for in-the-money (ITM). The convex risk measure used by the hedger is assumed to be the CVaR risk measure.

[5] The case of considering a stock with dividend payments is a trivial extension. Indeed, the numerical algorithm of Section 2.3 would have been essentially identical with the incorporation of dividend payments through the modifications of the dynamics of each model for the underlying stock prices and by modifying gain processes accordingly.

2.4.1.1 Regime-switching model

For $n = 1, \dots, N$, the daily log-returns $y_n := \log(S_n/S_{n-1})$ are assumed, unless stated otherwise, to follow a Gaussian regime-switching (RS) model. RS models have the ability to reproduce broadly accepted stylized facts of asset returns such as heteroskedasticity, autocorrelation in absolute returns, leverage effect and fat tails, see, for instance Ang and Timmermann (2012). Under RS models, log-returns depend on an unobservable discrete-time process. Let $h = \{h_n\}_{n=0}^N$ be a finite state Markov chain taking values in $\{1, \dots, H\}$ for a positive integer H, where h_n is the regime or state of the market during the period $[t_n, t_{n+1})$. Let $\{\gamma_{i,j}\}_{i=1,j=1}^{H,H}$ be the homogeneous transition probabilities of the Markov chain, where for $n = 0, \dots, N-1$:

$$\mathbb{P}(h_{n+1} = j | \mathcal{F}_n, h_n, \dots, h_0) = \gamma_{h_n,j}, \quad j = 1, \dots, H, \tag{2.25}$$

with the distribution of h_0 assumed to be the stationary distribution of the Markov chain. Let $\Delta := 1/260$. The daily log-returns are assumed to have the following dynamics:

$$y_{n+1} = \mu_{h_n}\Delta + \sigma_{h_n}\sqrt{\Delta}\epsilon_{n+1}, \quad n = 0, \dots, N-1, \tag{2.26}$$

where $\{\epsilon_n\}_{n=1}^N$ are independent standard normal random variables and $\{\mu_i, \sigma_i\}_{i=1}^H$ are the yearly model parameters with $\mu_i \in \mathbb{R}$ and $\sigma_i > 0$. Following the work of Godin et al. (2019), define $\xi := \{\xi_n\}_{n=0}^N$ as the predictive probability process with $\xi_n := [\xi_{n,1}, \dots, \xi_{n,H}]$ and $\xi_{n,j}$ as the probability that the Markov chain is in the j^{th} regime during $[t_n, t_{n+1})$ conditional on the investor's filtration, i.e.:

$$\xi_{n,j} := \mathbb{P}(h_n = j | \mathcal{F}_n), \quad j = 1, \dots, H. \tag{2.27}$$

François et al. (2014) show that the optimal hedging portfolio composition at time t_n is strictly a function of $\{S_n, V_n, \xi_n\}$. Thus, in Assumption 2.2, the feature vector considered for both the long-$\mathcal{NN}$ and short-$\mathcal{NN}$ is $X_n = [S_n, V_n, \xi_n, T - t_n]$. François et al. (2014) also provide a recursion to compute the predictive probabilities processes ξ. For $k = 1, \ldots, H$, define the function ϕ_k as the Gaussian pdf with mean μ_k and standard deviation σ_k:

$$\phi_k(x) := \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left(-\frac{(x-\mu_k)^2}{2\sigma_k^2}\right).$$

Setting ξ_0 as the stationary distribution of the Markov chain, the $\xi_{n,i}$ can be recursively computed for $n = 1, \ldots, N$ as follows:

$$\xi_{n,i} = \frac{\sum_{j=1}^{H} \gamma_{j,i} \phi_j(y_n) \xi_{n-1,j}}{\sum_{j=1}^{H} \phi_j(y_n) \xi_{n-1,j}}, \quad i = 1, \ldots, H.$$

In Section 2.4.3, different dynamics for the underlying will be considered. Each model is estimated with maximum likelihood on the same time series of daily log-returns on the S&P 500 price index for the period 1986-12-31 to 2010-04-01 (5863 observations). Resulting parameters are in Section 2.6.3.

2.4.1.2 Neural network structure

The training of the long-$\mathcal{NN}$ and short-$\mathcal{NN}$ is done as described in Section 2.3.3 with 100 epochs[6], a minibatch size of 1,000 on a training set (in-sample) of 400,000 independent simulated paths and a learning rate of 0.0005 with the Adam algorithm. Numerical results presented are obtained from a test set (out-of-sample) of 100,000 independent paths. The structure of every neural network is 3 layers with 56 neurons per layer and the activation function is the rectified linear unit (ReLU) where ReLU : $\mathbb{R} \rightarrow [0, \infty)$ is defined as $\text{ReLU}(x) := \max(0, x)$.

[6] An epoch is defined as one complete iteration of SGD over the training set. For a training set of 400,000 paths and a batch size of 1,000, one epoch is equivalent to 400 iterations of SGD.

2.4.2 Sensitivity analyses

In this section, we perform sensitivity analyses of equal risk prices and residual hedging risk with respect to the confidence level of the CVaR. Three different confidence levels are considered: CVaR_α at levels $0.90, 0.95$ and 0.99. Optimizing risk exposure using a higher level α corresponds to agents with a higher risk aversion as the latter puts more relative weight on losses of larger magnitude. Thus, the choice of the confidence level is motivated by the objective of assessing the impact of the level of risk aversion of hedging agents on equal risk pricing. Table 2.1 presents the equal risk option prices and residual hedging risk exposures under the three confidence levels. Our numerical results show that under the equal risk

Table 2.1: Sensitivity analysis of equal risk prices $C_0^{(\star,\mathcal{NN})}$ and residual hedging risk $\epsilon^{(\star,\mathcal{NN})}$ for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 60/260$.

	$C_0^{(\star,\mathcal{NN})}$			$\epsilon^{(\star,\mathcal{NN})}$			$\epsilon^{(\star,\mathcal{NN})}/C_0^{(\star,\mathcal{NN})}$		
Moneyness	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
$\text{CVaR}_{0.90}$	1.40	4.19	11.14	1.36	2.61	1.68	0.97	0.62	0.15
$\text{CVaR}_{0.95}$	32%	4%	2%	35%	11%	14%	2%	7%	12%
$\text{CVaR}_{0.99}$	91%	42%	14%	99%	79%	98%	4%	26%	74%

Notes: These results are computed based on 100,000 independent paths generated from the regime-switching model under $\mathbb{P}$ (see Section 2.4.1.1 for model definition and Section 2.6.3 for model parameters). The training of neural networks is done as described in Section 2.4.1.2. Values for the $\text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ risk measures are expressed relative to $\text{CVaR}_{0.90}$ (% increase).

pricing framework, an increase in the risk aversion of hedging agents leads to increased put option prices. Indeed, under the use of the $\text{CVaR}_{0.99}$ risk measure, option prices significantly increase across all moneynesses with relative increases of respectively 91%, 42% and 14% for OTM, ATM and ITM contracts with respect to prices obtained with the $\text{CVaR}_{0.90}$. By using the $\text{CVaR}_{0.95}$ instead of $\text{CVaR}_{0.90}$, only OTM equal risk prices are significantly impacted with an increase of 32%, while for ATM and ITM, the increase seems marginal. The positive association between put option prices and the confidence level of hedgers can be explained by the fact that a put option payoff is bounded below by zero. Therefore, the hedging error of the short position has a much heavier right tail than for the long position. Thus, an increase in α often implies a larger increase for the risk exposure of the short position than for the long position. This results in a higher equal risk price to compensate the heavier increase in risk exposure of the short position.

As expected, the risk exposure per option contract ($\epsilon^{(\star,\mathcal{NN})}$) also increases with the level of risk aversion across all moneynesses. This is a direct consequence of (2.13) and the monotonicity property of CVaR_α with respect to α. Also, the risk exposure per dollar invested ($\epsilon^{(\star,\mathcal{NN})}/C_0^{(\star,\mathcal{NN})}$) for ITM and ATM contracts exhibits high sensitivity to the confidence level α, while for OTM the value of α seems much less important. This observation for OTM contracts is due to a similar relative increase in prices and residual risk exposures obtained when α is increased. From these results, we can conclude that in practice, the choice of the confidence level (or more generally of the risk measure itself) needs to be carefully analyzed as it can have a material impact on equal risk option prices.

2.4.3 Model induced incompleteness

In this section, we consider four different dynamics for the underlying: the BSM, a GARCH process, a regime-switching process and a jump-diffusion. This is motivated by the objective of assessing the impact of different empirical properties of asset returns on the equal risk pricing framework. Indeed, Monte Carlo simulations from these models enable quantifying the impact of time-varying volatility, regime risk and jump risk on equal risk prices and residual hedging risk. Moreover, risk-neutral expected prices are used as benchmarks to equal risk prices under common EMMs found in the literature. The physical dynamics of each model is described below and the associated risk-neutral dynamics are provided in Section 2.6.2.

2.4.3.1 Discrete BSM

Under the discrete Black-Scholes model, log-returns are assumed to be i.i.d. normal random variables with daily mean and variance of respectively $(\mu - \frac{\sigma^2}{2})\Delta$ and $\sigma^2\Delta$:

$$y_n = \left(\mu - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\epsilon_n, \quad n = 1, \ldots, N, \tag{2.28}$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$ are the yearly model parameters, and $\{\epsilon_n\}_{n=1}^N$ are independent standard normal random variables. The feature vector of the neural network is $X_n = [S_n, V_n, T - t_n]$.

2.4.3.2 Discrete Merton jump-diffusion (MJD) model

The jump-diffusion model of Merton (1976) generalizes the BSM by incorporating random jumps within paths. Let $\{\epsilon_n\}_{n=1}^N$ be independent standard normal random variables, $\{N_n\}_{n=0}^N$ be values of a homogeneous Poisson process of intensity λ at $t_0, \ldots, t_N$ and $\{\chi_j\}_{j=1}^\infty$ be i.i.d. normal random variables of mean $\gamma \in \mathbb{R}$ and variance ϑ^2. $\{N_n\}_{n=0}^N$, $\{\epsilon_n\}_{n=1}^N$ and $\{\chi_j\}_{j=1}^\infty$ are assumed independent. For $n = 1, \ldots, N$:

$$y_n = \left(\alpha - \lambda\left(e^{\gamma+\vartheta^2/2} - 1\right) - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\epsilon_n + \sum_{j=N_{n-1}+1}^{N_n} \chi_j, \tag{2.29}$$

where $\{\alpha, \gamma, \vartheta, \lambda, \sigma\}$ are the model parameters with $\{\alpha, \lambda, \sigma\}$ being on a yearly scale, $\alpha \in \mathbb{R}$ and $\sigma > 0$. The feature vector of the neural network is $X_n = [S_n, V_n, T - t_n]$.

2.4.3.3 GARCH model

In constrast to the BSM or MJD model, GARCH models allow for the volatility of asset returns to be time-varying. The GJR-GARCH(1,1) model of Glosten et al. (1993) assumes that the conditional variance of log-returns is stochastic and captures important features of asset returns such as the leverage effect and volatility clustering. For $n = 1, \ldots, N$:

$$\begin{aligned} y_n &= \mu + \sigma_n \epsilon_n, \\ \sigma_{n+1}^2 &= \omega + \alpha \sigma_n^2 (|\epsilon_n| - \gamma \epsilon_n)^2 + \beta \sigma_n^2, \end{aligned} \tag{2.30}$$

where the model parameters $\{\omega, \alpha, \beta\}$ are positive real values, $\mu \in \mathbb{R}, \gamma \in \mathbb{R}$ and $\{\epsilon_n\}_{n=1}^N$ is a sequence of independent standard normal random variables. Given the initial value σ_1^2, $\{\sigma_n^2\}_{n=1}^N$ is predictable with respect to $\mathbb{F}$. A common assumption which is used in this paper is to set σ_1^2 as the stationary variance: $\sigma_1^2 = \frac{\omega}{1-\alpha(1+\gamma^2)-\beta}$. The feature vector of the neural network is $X_n = [S_n, V_n, \sigma_{n+1}, T - t_n]$.

2.4.3.4 Results

Table 2.2 presents the equal risk prices and residual hedging risk exposures for the four dynamics considered based on the $\text{CVaR}_{0.95}$ risk measure. Values observed for $C_0^{(\star,\mathcal{NN})}$

Table 2.2: Equal risk prices $C_0^{(\star,\mathcal{NN})}$ and residual hedging risk $\epsilon^{(\star,\mathcal{NN})}$ for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 60/260$ under different dynamics.

	$C_0^{(\star,\mathcal{NN})}$			$\epsilon^{(\star,\mathcal{NN})}$			$\epsilon^{(\star,\mathcal{NN})}/C_0^{(\star,\mathcal{NN})}$		
Moneyness	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
BSM	0.58	3.53	10.39	0.35	0.74	0.59	0.60	0.21	0.06
MJD	5%	−2%	0%	50%	62%	41%	42%	65%	41%
GJR-GARCH	68%	−4%	−1%	165%	115%	28%	57%	124%	30%
Regime-switching	217%	23%	9%	428%	291%	223%	67%	218%	196%

Notes: These results are computed based on 100,000 independent paths generated from each of the four different models for the underlying under $\mathbb{P}$ (see Section 2.4.1.1 and Section 2.4.3 for model definitions and Section 2.6.3 for model parameters). For each model, the training of neural networks is done as described in Section 2.4.1.2. The confidence level of the CVaR risk measure is $\alpha = 0.95$. Results are expressed relative to the BSM (% increase).

indicate that the sensitivity of equal risk prices with respect to the dynamics of the underlying highly depends on the moneyness. Indeed, OTM prices are significantly impacted by the choice of dynamics; choosing the GJR-GARCH or RS models instead of the BSM leads to a price increase respectively of 68% or 217%. For ATM and ITM contracts, only the RS model seems to materially alter equal risk prices in comparison to BSM prices with respective increases of 23% and 9%. Moreover, the numerical results confirm that the increase in hedging residual risk generated by time-varying volatility, regime risk and jump risk is far from being marginal and is highly sensitive to the moneyness of the option. Values obtained for the metric $\epsilon^{(\star,\mathcal{NN})}/C_0^{(\star,\mathcal{NN})}$ show that regime risk has the most impact with an increase of the risk exposure per dollar invested of 67%, 218% and 196% respectively for the OTM, ATM and ITM contracts in comparison to the BSM. When compared to jump risk, time-varying volatility seems to have a higher impact on residual hedging risk for OTM and ATM options, while jump risk has a higher impact on ITM contracts. It is interesting to note that Augustyniak et al. (2017) evaluate the impact of the dynamics of the underlying on the risk exposure and on the price of contingent claims under a quadratic penalty. Their numerical results show that the risk exposure is highly sensitive to the dynamics, but not the price. This is in contrast to numerical results of the current study which show that under a non-quadratic penalty, prices can also vary significantly with the dynamics of the underlying.

Table 2.3: Equal risk and risk-neutral prices for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 60/260$.

	Risk-neutral prices			Equal risk prices		
Moneyness	OTM	ATM	ITM	OTM	ATM	ITM
BSM	0.53	3.51	10.36	10%	1%	0%
MJD	0.46	3.32	10.24	34%	4%	2%
GJR-GARCH	0.57	2.98	9.84	71%	14%	4%
Regime-switching	0.56	3.10	10.33	231%	40%	10%

Notes: Results for equal risk prices are computed based on 100,000 independent paths generated from each of the four different models for the underlying under $\mathbb{P}$ (see Section 2.4.1.1 and Section 2.4.3 for model definitions and Section 2.6.3 for model parameters). For each model, the training of neural networks is done as described in Section 2.4.1.2. The confidence level of the CVaR risk measure is $\alpha = 0.95$. Results for risk-neutral prices are computed under the associated risk-neutral dynamics described in Section 2.6.2. Equal risk prices are expressed relative to risk-neutral prices (% increase).

Table 2.3 compares equal risk prices to risk-neutral prices for each dynamics. These results show that except for a few cases, equal risk prices are significantly higher than risk-neutral prices across all dynamics and moneynesses. This is especially true for OTM contracts: the lowest and highest relative price increases are 10% and 231% when going from the BSM to the regime-switching model. This significant increase in option prices can be attributed to the different treatment of market scenarios by each approach. Expected risk-neutral prices consider averages of all scenarios, while equal risk prices with the CVaR risk measure coupled with a high confidence level α only consider extreme scenarios.

The latter observation has important implications for financial participants in the option market. Indeed, by using the risk-neutral valuation approach instead of the equal risk pricing framework, a risk averse participant acting as a provider of options, e.g. a market maker, might significantly underprice OTM put options. From the perspective of the equal risk pricing framework, risk-neutral prices imply more residual risk for the short position of OTM put contracts than for the long position. It is important to note that the risk-neutral dynamics considered in this paper assume that jump and regime risk are not priced in the market. Additional analyses comparing equal risk prices to risk-neutral prices under alternative EMMs embedding other forms of risk premia (see for instance Bates (1996) for jump risk premium and Godin et al. (2019) for regime risk premium) may prove worthwhile in further work.

2.4.4 Exotic contingent claims

In this section, two exotic contingent claims are considered for the equal risk pricing framework, namely an Asian average price put and lookback put with fixed strike. For $Z_n = \frac{1}{n+1}\sum_{i=0}^{n} S_i$, $n = 0, \ldots, N$, the Asian option's payoff is:

$$\Phi(S_N, Z_N) = \max(0, K - Z_N).$$

For $Z_n = \min_{i=0,\ldots,n} S_i$, $n = 0, \ldots, N$, the lookback option's payoff is:

$$\Phi(S_N, Z_N) = \max(0, K - Z_N).$$

The same assumptions as in Section 2.4.3 are imposed, and only the regime-switching model is considered. The maturity is still $T = 60/260$. The feature vector for both exotic contingent claims is $X_n = [S_n, Z_n, V_n, \xi_n, T - t_n]$. Table 2.4 presents the prices and residual hedging risk for the three contingent claims (including the vanilla put option studied in previous sections) and Table 2.5 compares the equal risk prices to risk-neutral prices. From Table 2.4, we observe that the residual hedging risk exposure ($\epsilon^{(\star,\mathcal{NN})}$) is the highest for the lookback option, followed by the vanilla put and then the Asian option. Although this was expected, our approach has the benefit of quantifying how risk varies across different option categories. Table 2.5 shows that equal risk prices under the RS model are significantly higher than risk-neutral prices across all contingent claims considered. The difference is most important for OTM contracts with respective increases of 231%, 465% and 220% for the put, Asian and lookback options. The finding that equal risk prices tend to be higher than risk-neutral prices is therefore shared by multiple option categories.

Table 2.4: Equal risk prices $C_0^{(\star,\mathcal{NN})}$ and residual hedging risk $\epsilon^{(\star,\mathcal{NN})}$ for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) vanilla put, Asian average price put and lookback put options of maturity $T = 60/260$.

	$C_0^{(\star,\mathcal{NN})}$			$\epsilon^{(\star,\mathcal{NN})}$			$\epsilon^{(\star,\mathcal{NN})}/C_0^{(\star,\mathcal{NN})}$		
Moneyness	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
Put	1.84	4.34	11.33	1.84	2.90	1.91	1.00	0.67	0.17
Asian	−66%	−36%	−7%	−65%	−32%	−55%	1%	6%	−52%
Lookback	64%	80%	64%	64%	70%	205%	0%	−5%	86%

Notes: These results are computed based on 100,000 independent paths generated from the regime-switching model under $\mathbb{P}$ (see Section 2.4.1.1 for model definition and Section 2.6.3 for model parameters). The training of neural networks is done as described in Section 2.4.1.2. The confidence level of the CVaR risk measure is $\alpha = 0.95$. Results are expressed relative to the put option (% increase).

Table 2.5: Equal risk and risk-neutral prices for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) vanilla put, Asian average price put and lookback put options of maturity $T = 60/260$.

	Risk-neutral prices			Equal risk prices		
Moneyness	OTM	ATM	ITM	OTM	ATM	ITM
Put	0.56	3.10	10.33	231%	40%	10%
Asian	0.11	1.77	9.91	465%	56%	6%
Lookback	0.94	5.61	15.57	220%	39%	19%

Notes: Results for equal risk prices are computed with 100,000 independent paths generated from the regime-switching model under $\mathbb{P}$ (see Section 2.4.1.1 for model definition and Section 2.6.3 for model parameters). The training of neural networks is done as described in Section 2.4.1.2. The confidence level of the CVaR risk measure is $\alpha = 0.95$. Results for risk-neutral prices are computed under the associated risk-neutral dynamics described in Section 2.6.2. Equal risk prices are expressed relative to risk-neutral prices (% increase).

2.5 Conclusion

This paper presents a deep reinforcement learning approach to price and hedge financial derivatives under the equal risk pricing framework. This framework introduced by Guo and Zhu (2017) sets option prices such that the optimally hedged residual risk exposure of the long and short positions in the contingent claim is equal. Adaptations to the latter scheme are used as proposed in Marzban et al. (2020) by considering convex risk measures under the physical measure to evaluate residual risk exposures. A rigorous proof that equal risk prices under these modifications are arbitrage-free in general market settings which can include an arbitrary number of hedging instruments is given in the current paper.

Moreover, a universal and tractable solution based on the deep hedging algorithm of Buehler et al. (2019b) to implement the equal risk pricing framework under very general conditions is described. Results presented in this paper, which rely on Buehler et al. (2019b), demonstrate that our methodological approach to equal risk pricing can approximate arbitrarily well the true equal risk price. This study also introduces asymmetric ϵ-completeness measures to quantify the level of unhedgeable risk associated with a position in a contingent claim. These measures complement the work of Bertsimas et al. (2001) who proposed market incompleteness measures under the quadratic penalty, while ours has the advantage of characterizing the risk aversion of the hedger with any convex risk measure. Additionally, the measures introduced in this paper are asymmetric in the sense that the risk for the long and short positions in the derivative are quantified by two different hedging strategies, unlike in Bertsimas et al. (2001) where the single variance-optimal hedging strategy is considered.

Furthermore, Monte Carlo simulations were performed to study the equal risk pricing framework under a large variety of market dynamics. The behavior of equal risk pricing is analyzed through the choice of the underlying asset model and of the confidence level associated with the risk measure, and is benchmarked against expected risk-neutral pricing. The conduction of these numerical experiments crucially relied on the deep RL algorithm presented in this study. Numerical results showed that except for a few cases, equal risk prices are significantly higher than risk-neutral prices across all dynamics and moneynesses considered. This finding is shown to be most important for OTM contracts and shared by multiple option categories. Furthermore, for a fixed model for the underlying, sensitivity analyzes show that the choice of confidence level under the CVaR risk measure has a material impact on equal risk prices. Numerical experiments also provided insight on drivers of the ϵ-completeness measures introduced in the current paper. The numerical study confirms that for vanilla put options, the increase in hedging residual risk generated by time-varying volatility, regime risk and jump risk is far from being marginal and is highly sensitive to the moneyness of the option.

Future research on equal risk pricing could prove worthwhile. First, a question which remains is whether the consistence of equal risk pricing approach with risk-neutral valuations can be made explicit. Moreover, additional analyses comparing equal risk prices to risk-neutral prices under alternative EMMs embedding other forms of risk premia may also prove worthwhile. Furthermore, a numerical study of the equal risk pricing framework under other convex measures than the CVaR could be of interest. We note that Marzban et al. (2020) provide numerical results of equal risk pricing under the worst-case risk measure in the context of robust optimization. Lastly, the financial market could be extended by including different market frictions such as transaction costs and trading constraints. The latter inclusions would require examining if equal risk prices are guaranteed to remain arbitrage-free in this context.

2.6 Appendix

2.6.1 Proof of Theorem 2.1

Using the representation (2.5) of long and short measured risk exposures:

$$\epsilon^{(L)}(-C_0^\star) = \epsilon^{(S)}(C_0^\star) \iff \epsilon^{(L)}(0) + B_N C_0^\star = \epsilon^{(S)}(0) - B_N C_0^\star \iff C_0^\star = \frac{\epsilon^{(S)}(0) - \epsilon^{(L)}(0)}{2B_N}. \tag{2.31}$$

This shows that $C_0^\star$ exists, is unique and is given by (2.11). Next, we show that the equal risk price is arbitrage-free. Some parts of the proof are inspired by the work of Xu (2006). Let $(\bar{v}, \bar{\delta})$ be a super-replication strategy of Φ, see Definition 2.8, where $\bar{v}$ is the super-replication price as in (2.8) and let $\tilde{\delta} := \underset{\delta \in \Pi}{\arg\min}\, \rho\left(-B_N G_N^{\delta}\right)$. Note[7] that for any $\breve{\delta}, \acute{\delta} \in \Pi$, $G_n^{\breve{\delta}+\acute{\delta}} = G_n^{\breve{\delta}} + G_n^{\acute{\delta}}$. Using the translation invariance and monotonicity properties of ρ:

$$\begin{aligned}
\epsilon^{(S)}(0) &= \min_{\delta \in \Pi} \rho\left(\Phi(S_N, Z_N) - B_N G_N^{\delta}\right) \\
&\leq \rho\left(\Phi(S_N, Z_N) - B_N (G_N^{\bar{\delta}+\tilde{\delta}})\right) \\
&= \rho\left(\Phi(S_N, Z_N) - B_N (G_N^{\bar{\delta}} + G_N^{\tilde{\delta}})\right) \\
&= \rho\left(\Phi(S_N, Z_N) - B_N (\bar{v} + G_N^{\bar{\delta}}) - B_N G_N^{\tilde{\delta}}\right) + B_N \bar{v} \\
&\leq \rho(-B_N G_N^{\tilde{\delta}}) + B_N \bar{v},
\end{aligned} \tag{2.32}$$

[7] $G_0^{\breve{\delta}+\acute{\delta}} = G_0^{\breve{\delta}} + G_0^{\acute{\delta}} = 0$ by definition and for $n = 1, \ldots, N$:

$$G_n^{\breve{\delta}+\acute{\delta}} = \sum_{k=1}^{n} (\breve{\delta}_k^{(1:D)} + \acute{\delta}_k^{(1:D)}) \bullet (B_k^{-1} S_k - B_{k-1}^{-1} S_{k-1}) = G_n^{\breve{\delta}} + G_n^{\acute{\delta}}.$$

where for (2.32), the monotonicity property is applied to $\Phi(S_N, Z_N) - B_N(\bar{v} + G_N^{\bar{\delta}}) \leq 0$ $\mathbb{P}$-a.s. This implies

$$\zeta^{(S)} := \frac{\epsilon^{(S)}(0) - \rho(-B_N G_N^{\tilde{\delta}})}{B_N} \leq \bar{v}. \tag{2.33}$$

Similarly, let $(\underline{v}, \underline{\delta})$ be a sub-replication strategy where $\underline{v}$ is the sub-replication price. Note[8] that for any $\delta \in \Pi$, $G_n^{\delta} = -G_n^{-\delta}$. Using the translation invariance and monotonicity properties of ρ:

$$\begin{aligned}
\epsilon^{(L)}(0) &= \min_{\delta \in \Pi} \rho\left(-\Phi(S_N, Z_N) - B_N G_N^{\delta}\right) \\
&\leq \rho\left(-\Phi(S_N, Z_N) - B_N G_N^{\tilde{\delta} - \underline{\delta}}\right) \\
&= \rho\left(B_N G_N^{\underline{\delta}} - \Phi(S_N, Z_N) - B_N G_N^{\tilde{\delta}}\right) \\
&= \rho\left(B_N(\underline{v} + G_N^{\underline{\delta}}) - \Phi(S_N, Z_N) - B_N G_N^{\tilde{\delta}}\right) - B_N \underline{v} \\
&\leq \rho(-B_N G_N^{\tilde{\delta}}) - B_N \underline{v},
\end{aligned} \tag{2.34}$$

where for (2.34), the monotonicity property is applied to $B_N(\underline{v} + G_N^{\underline{\delta}}) - \Phi(S_N, Z_N) \leq 0$ $\mathbb{P}$-a.s. This implies

$$\underline{v} \leq \zeta^{(L)} := \frac{\rho(-B_N G_N^{\tilde{\delta}}) - \epsilon^{(L)}(0)}{B_N}. \tag{2.35}$$

Using (2.31), $C_0^{\star}$ has the representation $C_0^{\star} = 0.5(\zeta^{(L)} + \zeta^{(S)})$. The last step of the proof entails showing that $\zeta^{(L)} \leq \zeta^{(S)}$, which implies that the derivative price $C_0^{\star} \in [\underline{v}, \bar{v}]$ and is

[8] For $n = 0$, the result is direct since $G_0^{\delta} = 0$ for any $\delta \in \Pi$. For $n = 1, \ldots, N$:

$$G_n^{\delta} = \sum_{k=1}^{n} \delta_k^{(1:D)} \bullet (B_k^{-1} S_k - B_{k-1}^{-1} S_{k-1}) = -\sum_{k=1}^{n} (-\delta_k^{(1:D)}) \bullet (B_k^{-1} S_k - B_{k-1}^{-1} S_{k-1}) = -G_n^{-\delta}.$$

arbitrage-free in the sense of Definition 2.9. Define the risk measure ϱ as

$$\varrho(X) := \min_{\delta\in\Pi} \rho\left(X - B_N G_N^{\delta}\right). \tag{2.36}$$

Buehler et al. (2019b) show that since ρ is a convex risk measure and Π is a convex set, ϱ is a convex risk measure (see Proposition 3.1 of their paper). Note that $\epsilon^{(S)}(0) = \varrho(\Phi(S_N, Z_N))$ and $\epsilon^{(L)}(0) = \varrho(-\Phi(S_N, Z_N))$ by definition. With the translation invariance and convexity properties of ϱ, we obtain that

$$\begin{aligned}
\min_{\delta\in\Pi} \rho\left(-B_N G_N^{\delta}\right) &= \varrho(0) = \varrho\left(\frac{1}{2}\Phi(S_N, Z_N) - \frac{1}{2}\Phi(S_N, Z_N)\right) \\
&\leq \frac{1}{2}\varrho(\Phi(S_N, Z_N)) + \frac{1}{2}\varrho(-\Phi(S_N, Z_N)) \\
&= \frac{1}{2}\epsilon^{(S)}(0) + \frac{1}{2}\epsilon^{(L)}(0) \\
\implies \frac{\min_{\delta\in\Pi} \rho\left(-B_N G_N^{\delta}\right) - \epsilon^{(L)}(0)}{B_N} &\leq \frac{\epsilon^{(S)}(0) - \min_{\delta\in\Pi} \rho\left(-B_N G_N^{\delta}\right)}{B_N} \\
\implies \zeta^{(L)} &\leq \zeta^{(S)}. \quad \square
\end{aligned}$$

2.6.2 Risk-neutral dynamics

Since the market is arbitrage-free under the models assumed for the underlying, the first fundamental theorem of asset pricing implies that their exist a probability measure $\mathbb{Q}$ such that $\{S_n e^{-rt_n}\}_{n=0}^{N}$ is an $(\mathbb{F}, \mathbb{Q})$-martingale (see, for instance, Delbaen and Schachermayer (1994)). For the rest of Section 2.6.2, let $\{\epsilon_n^{\mathbb{Q}}\}_{n=1}^{N}$ be independent standard normal random variables under $\mathbb{Q}$ and denote $P_{0,T}$ as the price at time 0 of a contingent claim of payoff $\Phi(S_N, Z_N)$ at maturity T:

$$P_{0,T} := e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_N, Z_N) \,\middle|\, \mathcal{F}_0\right]. \tag{2.37}$$

Here are the risk-neutral dynamics for each model considered.

2.6.2.1 Regime-switching

The change of measure considered is the so-called regime-switching mean-correcting transform, a popular choice under RS models (see, e.g. Hardy (2001) and Bollen (1998)). This change of measure preserves the model dynamics of regime-switching except for a shift to the drift in each respective regime. More precisely, during the passage from $\mathbb{P}$ to $\mathbb{Q}$, the transition probabilities of the Markov chain and the volatilities are left unchanged, but the drifts $\mu_i\Delta$ are shifted to $(r - \sigma_i^2/2)\Delta$ for regimes $i = 1, \ldots, H$. The resulting dynamics for the log-returns under $\mathbb{Q}$ is

$$y_{n+1} = \left(r - \frac{\sigma_{h_n}^2}{2}\right)\Delta + \sigma_{h_n}\sqrt{\Delta}\epsilon_{n+1}^{\mathbb{Q}}, \quad n = 0, \ldots, N-1. \tag{2.38}$$

Let $\mathbb{H} := \{\mathcal{H}_n\}_{n=0}^N$ be the filtration generated by the markov chain h:

$$\mathcal{H}_n := \sigma(h_u \mid u = 0, \ldots, n), \quad n = 0, \ldots, N. \tag{2.39}$$

Following the work of Godin et al. (2019), option prices can be developed as follow. Let $\mathbb{G} := \{\mathcal{G}_n\}_{n=0}^N$ be the filtration which contains all latent factors as well as information available to market participants, i.e. $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$. Thus, the process $\{(S_n, h_n)\}$ is Markov under $\mathbb{Q}$ with respect to $\mathbb{G}$. With the law of iterated expectations, the time-0 price of a derivative $P_{0,T}$ can be written as follows:

$$\begin{aligned} P_{0,T} &= e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_N, Z_N) \mid \mathcal{F}_0\right] \\ &= e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_N, Z_N) \mid \mathcal{G}_0\right] \mid \mathcal{F}_0\right] \\ &= e^{-rT}\sum_{j=1}^{H}\xi_{0,j}^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_N, Z_N) \mid S_0, h_0 = j\right], \end{aligned} \tag{2.40}$$

where $\xi_0^{\mathbb{Q}}$ is assumed to be the stationary distribution of the Markov chain under $\mathbb{P}$. The computation of $P_{0,T}$ can be done through Monte Carlo simulations for all contingent claims (i.e. vanilla and exotic).

2.6.2.2 Discrete BSM

By a discrete-time version of the Girsanov theorem, there exists a market price of risk process $\{\tilde{\lambda}_n\}_{n=1}^N$ such that $\epsilon_n^{\mathbb{Q}} = \epsilon_n + \tilde{\lambda}_n$, for $n = 1, \ldots, N$. Setting $\tilde{\lambda}_n := \sqrt{\Delta}\left(\frac{\mu - r}{\sigma}\right)$ and replacing $\epsilon_n = \epsilon_n^{\mathbb{Q}} - \tilde{\lambda}_n$ into (2.28), it is straightforward to obtain the $\mathbb{Q}$-dynamics of the log-returns:

$$y_n = \left(r - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\epsilon_n^{\mathbb{Q}}, \quad n = 1, \ldots, N. \tag{2.41}$$

The computation of $P_{0,T}$ can be done with the well-known closed-form solution for vanilla put options (i.e. the Black-Scholes equation) and through Monte Carlo simulations for exotic contingent claims.

2.6.2.3 Discrete MJD

The change of measure used assumes no risk premia for jumps as in Merton (1976) and simply shifts the drift in (2.29) from α to r. The $\mathbb{Q}$-dynamics is thus

$$y_n = \left(r - \lambda\left(e^{\gamma+\delta^2/2} - 1\right) - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\epsilon_n^{\mathbb{Q}} + \sum_{j=N_{n-1}+1}^{N_n} \chi_j,$$

where $\{\chi_j\}_{j=1}^{\infty}$ and $\{N_n\}_{n=0}^N$ have the same distribution than under $\mathbb{P}$. The computation of $P_{0,T}$ for vanilla put options can be quickly performed with the fast Fourier transform (see, e.g. Carr and Madan (1999)). The pricing of exotic contingent claims can be done through Monte Carlo simulations.

2.6.2.4 GARCH

The risk-neutral measure considered is often used in the GARCH option pricing literature under which the one-period ahead conditional log-return mean is shifted, but the one-period ahead conditional variance is left untouched (see e.g. Duan (1995)). For $n = 1, \ldots, N$, let $\varphi_n \in \mathcal{F}_{n-1}$ and define $\epsilon_n^{\mathbb{Q}} := \epsilon_n + \varphi_n$. Replacing $\epsilon_n = \epsilon_n^{\mathbb{Q}} - \varphi_n$ into (2.30), we obtain $y_n = \mu - \sigma_n\varphi_n + \sigma_n\epsilon_n^{\mathbb{Q}}$ for $n = 1, \ldots, N$. To ensure that $\{S_n e^{-rt_n}\}_{n=0}^N$ is an $(\mathbb{F}, \mathbb{Q})$-martingale, the one-period conditional expected return under $\mathbb{Q}$ must be equal to the daily risk-free rate, i.e.:

$$\mathbb{E}^{\mathbb{Q}}[e^{y_n} \mid \mathcal{F}_{n-1}] = e^{\mu - \sigma_n\varphi_n + \sigma_n^2/2} = e^{r\Delta} \iff \varphi_n := \frac{\mu - r\Delta + \sigma_n^2/2}{\sigma_n}, \quad n = 1, \ldots, N.$$

Thus, the $\mathbb{Q}$-dynamics of the GJR-GARCH(1,1) model is:

$$\begin{aligned} y_n &= r\Delta - \sigma_n^2/2 + \sigma_n\epsilon_n^{\mathbb{Q}}, \\ \sigma_{n+1}^2 &= \omega + \alpha\sigma_n^2(|\epsilon_n^{\mathbb{Q}} - \varphi_n| - \gamma(\epsilon_n^{\mathbb{Q}} - \varphi_n))^2 + \beta\sigma_n^2. \end{aligned}$$

The computation of $P_{0,T}$ can be done through Monte Carlo simulations for all contingent claims.

2.6.3 Maximum likelihood estimates results

This section presents estimated parameters for the various underlying asset models considered in numerical experiments from Section 2.4.

Table 2.6: Maximum likelihood parameter estimates of the Black-Scholes model.

μ	σ
0.0892	0.1952

Notes: Parameters were estimated on a time series of daily log-returns on the S&P 500 index for the period 1986-12-31 to 2010-04-01 (5863 log-returns). Both μ and σ are on an annual basis.

Table 2.7: Maximum likelihood parameter estimates of the GJR-GARCH(1,1) model.

μ	ω	α	γ	β
2.871e-04	1.795e-06	0.0540	0.6028	0.9105

Notes: Parameters were estimated on a time series of daily log-returns on the S&P 500 index for the period 1986-12-31 to 2010-04-01 (5863 log-returns).

Table 2.8: Maximum likelihood parameter estimates of the regime-switching model.

	Regime		
Parameter	1	2	3
μ	0.2040	0.0337	-0.6168
σ	0.0971	0.1865	0.5070
ν	0.4755	0.4561	0.0684
	0.9870	0.0127	0.0003
γ	0.0139	0.9807	0.0053
	0.0000	0.0380	0.9620

Notes: Parameters were estimated with the EM algorithm of Dempster et al. (1977) on a time series of daily log-returns on the S&P 500 index for the period 1986-12-31 to 2010-04-01 (5863 log-returns). ν represent probabilities associated with the stationary distribution of the Markov chain. γ is the transition matrix as in (2.25). μ and σ are on an annual basis.

Table 2.9: Maximum likelihood parameter estimates of the Merton jump-diffusion model.

α	σ	λ	γ	ϑ
0.0875	0.1036	92.3862	-0.0015	0.0160

Notes: Parameters were estimated on a time series of daily log-returns on the S&P 500 index for the period 1986-12-31 to 2010-04-01 (5863 log-returns). α, σ and λ are on an annual basis.

Chapter 3

Deep Hedging of Long-Term Financial Derivatives

Abstract

This study presents a deep reinforcement learning approach for global hedging of long-term financial derivatives. A similar setup as in Coleman et al. (2007) is considered with the risk management of lookback options embedded in guarantees of variable annuities with ratchet features. The deep hedging algorithm of Buehler et al. (2019b) is applied to optimize neural networks representing global hedging policies with both quadratic and non-quadratic penalties. To the best of the author's knowledge, this is the first paper that presents an extensive benchmarking of global policies for long-term contingent claims with the use of various hedging instruments (e.g. underlying and standard options) and with the presence of jump risk for equity. Monte Carlo experiments demonstrate the vast superiority of non-quadratic global hedging as it results simultaneously in downside risk metrics two to three times smaller than best benchmarks and in significant hedging gains. Analyses show that the neural networks are able to effectively adapt their hedging decisions to different penalties and stylized facts of risky asset dynamics only by experiencing simulations of the financial market exhibiting these features. Numerical results also indicate that non-quadratic global policies are significantly more geared towards being long equity risk which entails earning the equity risk premium.

Keywords: Reinforcement learning; Global hedging; Variable annuity; Lookback option; Jump risk.

JEL Classification: C45, C61, G32.

3.1 Introduction

Variable annuities (VAs), also known as segregated funds and equity-linked insurance, are financial products that enable investors to gain exposure to the market through cashflows that depend on equity performance. These products often include financial guarantees to protect investors against downside equity risk with benefits that can be expressed as the payoff of derivatives. For instance, a guaranteed minimum maturity benefit (GMMB) with ratchet feature is analogous to a lookback put option by providing a minimum monetary amount at the maturity of the contract equal to the maximum account value on specific dates (e.g. anniversary dates of the policy). The valuation of VAs guarantees is typically done with classical option pricing theory by computing the expected risk-neutral discounted cashflows of embedded options under an appropriate equivalent martingale measure; see, for instance, Brennan and Schwartz (1976), Boyle and Schwartz (1977), Pelsser (2003), Bauer et al. (2008) and Ng and Li (2011). A comprehensive review of pricing segregated funds guarantees literature can be found in Gan (2013).

During the subprime mortgage financial crisis, many insurers incurred large losses in segregated fund portfolios due in part to poor risk management with some insurers even stopping writing VAs guarantees in certain markets (Zhang (2010)). Two categories of risk management approaches are typically used in practice: the actuarial method and the financial engineering method (Boyle and Hardy (1997)). The former consists in providing stochastic models for the risk factors and setting a reserve held in risk-free assets to cover liabilities associated to VAs guarantees with a certain probability (e.g. the Value-at-Risk at 99%). The second approach, commonly known as *dynamic hedging*, entails solving for a self-funded sequence of positions in securities to mitigate the risk exposure of embedded options. Dynamic hedging is a popular risk management approach among insurance companies and is studied in this current paper; the reader is referred to Hardy (2003) for a detailed description of the actuarial method.

Financial markets are said to be complete if every contingent claim can be perfectly replicated with some dynamic hedging strategy. In practice, segregated funds embedded options are typically not attainable as a consequence of their many interrelated risks which are very complex to manage such as equity risk, interest rate risk, mortality risk and basis risk. For insurance companies selling VAs with guarantees, market incompleteness entails that some level of residual risk must be accepted as being intrinsic to the embedded options; the identification of optimal hedging policies in such context is thus highly relevant. Nevertheless, the attention of the actuarial literature has predominantly been on the valuation of segregated funds, not on the design of optimal hedging policies. Indeed, the hedging strategies considered are most often suboptimal and not necessarily in line with financial objectives of insurance companies. One popular hedging approach is the *greek-based policy* where assets positions depend on the sensitivities of the option value (i.e. the value of the guarantee) to different risk factors. Boyle and Hardy (1997) and Hardy (2000) delta-hedge GMMBs under market completeness for mortality risk and Augustyniak and Boudreault (2017) delta-rho hedge GMMBs and guaranteed minimum death benefits (GMDBs) in the presence of model uncertainty for both equity and interest rate. An important pitfall of greek-based policies in incomplete markets is their suboptimality by design: they are a by-product of the choice of pricing kernel (i.e. of the equivalent martingale measure) for option valuation, not of an optimization procedure over hedging decisions to minimize residual risk. Furthermore, as shown in the seminal work of Harrison and Pliska (1981), in incomplete markets, there exist an infinite set of equivalent martingale measures each of which is consistent with arbitrage-free pricing and can thus be used to compute positions in hedging instruments (i.e. the greeks).

Another strand of literature optimizes hedging policies with local and global criterions. *Local risk minimization* (Föllmer and Schweizer (1988) and Schweizer (1991)) consists in choosing assets positions to minimize the periodic risk associated with the hedging portfolio. On the other hand, *global risk minimization* procedures jointly optimize all hedging decisions with the objective of minimizing the expected value of a loss function applied to the terminal hedging error. In spite of their myopic view of the hedging problem by not necessarily minimizing the risk associated with hedging shortfalls, local risk minimization procedures are attractive for the risk mitigation of VAs guarantees as they are simple to implement and they have outperformed greek-based hedging in several studies. Coleman et al. (2006) and Coleman et al. (2007) apply local risk minimization procedures for risk mitigation of GMDBs using standard options with the foremost considering the presence of both interest rate and jump risk and the latter the presence of volatility and jump risk. Kélani and Quittard-Pinon (2017) extends the work of Coleman et al. (2007) in a general Lévy market with the inclusion of mortality risk and transaction costs, and Trottier et al. (2018b) and Trottier et al. (2018a) propose a local risk minimization scheme for guarantees in the presence of basis risk.

Within the realm of total risk minimization, *global quadratic hedging* pioneered by the seminal work of Schweizer (1995) aims at jointly optimizing all hedging decisions with a quadratic penalty for hedging shortfalls. The latter paper provides a theoretical solution to the optimal policy with a single risky asset (see Rémillard and Rubenthaler (2013) for the multidimensional asset case) and Bertsimas et al. (2001) develops a tractable solution to the optimal policy relying on stochastic dynamic programming. A major drawback of global quadratic hedging is in penalizing equally gains and losses, which is naturally not in line with the financial objectives of insurance companies. Alternatively, *non-quadratic global hedging* applies an asymmetric treatment to hedging errors by overly (and most often strictly) penalizing hedging losses. In contrast to global quadratic hedging, there is usually no closed-form solution to the optimal policy, but numerical implementations have been proposed in the literature: François et al. (2014) develops a methodology with stochastic dynamic programming algorithms for global hedging with any desired penalty function, Godin (2016) adapts the latter numerical implementation under the Conditional Value-at-Risk measure in the presence of transaction costs and Dupuis et al. (2016) studies global hedging procedures under the semi-mean-square error penalty in the context of short-term hedging for an electricity retailer. The aforementioned studies demonstrated the vast superiority of non-quadratic global hedging over popular alternative hedging schemes (e.g. greek-based policies, local risk minimization and global quadratic hedging). Yet, to the best of the author's knowledge, both quadratic and non-quadratic global hedging has seldom been applied for risk mitigation of segregated funds guarantees, or more generally, of long-term contingent claims.[1] Furthermore, numerical schemes for global hedging are computationally intensive and often rely on solving Bellman's equations, which is known to be prone to the curse of dimensionality (Powell, 2009). In the context of dynamically hedging segregated funds guarantees, the latter is a major drawback as it restrains the number of risk factors to consider for the financial market as well as prevents the use of multiple assets in the design of hedging policies. A feasible implementation of global hedging for the risk mitigation of VAs guarantees which is flexible to the choice of market features, to the hedging instruments and to the penalty for hedging errors would be desirable.

Recently, Buehler et al. (2019b) introduced a deep reinforcement learning (deep RL) algorithm called *deep hedging* to hedge a portfolio of over-the-counter derivatives in the presence of market frictions. The general framework of RL is for an agent to learn over many iterations of an environment how to select sequences of actions to optimize a cost function. RL has been applied successfully in many areas of quantitative finance such as algorithmic trading (e.g. Moody and Saffell (2001) and Deng et al. (2016)), portfolio optimization (e.g. Jiang et al. (2017) and Almahdi and Yang (2017)) and option pricing (e.g. Li et al. (2009), Becker et al. (2019) and Carbonneau and Godin (2021b)). Hedging has also received substantial attention: Halperin (2020) and Kolm and Ritter (2019) propose TD-learning approaches to the hedging problem and Cao et al. (2020) and Carbonneau and Godin (2021b) deep hedge European options under respectively the quadratic penalty and the Conditional Value-at-Risk measure. The deep hedging algorithm trains an agent to learn how to approximate optimal hedging decisions by neural networks through many simulations of a synthetic market. This approach is related to the deep learning method of Han and E (2016) by directly optimizing

[1] An exception is the work of Ankirchner et al. (2014) which considers a minimal-variance hedging strategy for VAs guarantees in continuous-time in the presence of basis risk.

policies for stochastic control problems with Monte Carlo simulations. Arguably, the most important benefit of using neural networks to approximate optimal policies is to overcome the curse of dimensionality which arises when the state and action space gets too large.

The contribution of this paper is threefold. First, this study presents a deep reinforcement learning procedure for global hedging long-term financial derivatives which are analogous under assumptions made in this study to embedded options of segregated funds. The methodological approach, which relies on the deep hedging algorithm, can be applied for the risk mitigation of any long-term European-type contingent claims (e.g. vanilla, path-dependent) with multiple hedging instruments (e.g. standard options and underlying) under any desired penalty (e.g. quadratic and non-quadratic) and in the presence of different risky assets stylized features (e.g. jump, volatility and regime risk). The second contribution consists in conducting broad numerical experiments of hedging long-term contingent claims with the optimized global policies. A similar setup as in the work of Coleman et al. (2007) is considered with the risk mitigation of ratchet GMMBs strictly for financial risks in the presence of jumps for equity. To the best of the author's knowledge, this is the first paper that presents such an extensive benchmarking of quadratic and non-quadratic global policies for long-term options with the use of various hedging instruments and by considering different risky assets dynamics. The use of neural networks to solve global hedging problems enables us to provide novel qualitative insights into long-term global hedging. Such benchmarking would have been hardly attainable when relying on more traditional optimization procedures for global hedging such as stochastic dynamic programming due to the high-dimensional continuous state and action spaces considered in this study. Numerical experiments demonstrate the vast superiority of non-quadratic global hedging as it results simultaneously in downside risk metrics two to three times smaller than best benchmarks and in significant hedging gains. Our results clearly demonstrate that non-quadratic global hedging should be prioritized over other popular dynamic hedging procedures found in the literature as it is tailor-made to match the financial objectives of the hedger by always significantly reducing the downside risk as well as earning large expected positive returns. The third contribution is in providing important insights into specific characteristics of the optimized global policies. Monte Carlo experiments indicate that on average, non-quadratic global policies are significantly more bullish than their quadratic counterpart by holding a larger average equity risk exposure which entails earning the equity risk premium. The conduction of these experiments, and thus of the finding of these novel qualitative observations into long-term global hedging policies, heavily relies on the neural-based hedging scheme considered in this paper. Key factors which contribute to this specific characteristic of non-quadratic global policies are identified. Furthermore, analyses of numerical results show that the training algorithm is able to effectively adapt hedging policies (i.e. neural networks parameters) to different stylized features of risky asset dynamics only by experiencing simulations of the financial market exhibiting these features.

The paper is structured as follows. Section 3.2 introduces the notation and the optimal hedging problem. Section 3.3 describes the numerical scheme based on deep RL to optimize global hedging policies. Section 3.4 presents benchmarking of the risk mitigation of GMMBs under various market settings. Section 3.5 concludes.

3.2 Hedging long-term contingent claims

This section details the financial market setup and the hedging problem considered in this paper.

3.2.1 Market setup

The financial market is in discrete-time with a finite time horizon of $T \in \mathbb{N}$ years and $N+1$ known observation dates $\mathcal{T} := \{t_i : t_i = i\Delta_N, i = 0, \dots, N\}$ with $\Delta_N := T/N$. The probability space $(\Omega, \mathcal{F}_T, \mathbb{P})$ with $\mathbb{P}$ as the physical measure is equipped with the filtration $\mathbb{F} := \{\mathcal{F}_{t_n}\}_{n=0}^{N}$ that defines all available information of the financial market to investors. A total of $D+2$ liquid assets are accessible to financial participants with $D+1$ risky assets and one risk-free asset. Let $\{B_{t_n}\}_{n=0}^{N}$ be the price process of the risk-free asset where $B_{t_n} := e^{rt_n}$ with $r \in \mathbb{R}$ as the annualized continuously compounded risk-free rate. The risky assets include a non-dividend paying stock and D liquid vanilla European-type options such as calls and puts on the stock which expire on observation dates in $\mathcal{T}$. In this context, the specification of two distinct price processes, one at the beginning and one at the end of each trading period, is required. Let $\{\bar{S}_{t_n}^{(b)}\}_{n=0}^{N}$ be the *beginning-of-period* risky price process, where $\bar{S}_{t_n}^{(b)} := [S_{t_n}^{(0,b)}, \ldots, S_{t_n}^{(D,b)}]$ are the prices at the beginning of $[t_n, t_{n+1})$ with $S_{t_n}^{(0,b)}$ and $S_{t_n}^{(j,b)}$ respectively as the price of the underlying and of the j^{th} option. Similarly, let $\{\bar{S}_{t_n}^{(e)}\}_{n=0}^{N-1}$ be the *end-of-period* risky price process, where $\bar{S}_{t_n}^{(e)} := [S_{t_n}^{(0,e)}, \ldots, S_{t_n}^{(D,e)}]$ are the prices at the end of $[t_n, t_{n+1})$ before the next rebalancing at t_{n+1}. For the tradable options, if the j^{th} option matures at t_{n+1}, then $S_{t_n}^{(j,e)}$ is the payoff of the derivative and $S_{t_{n+1}}^{(j,b)}$ is the price of a new contract with the same characteristics (i.e. same payoff function and time-to-maturity). For the underlying, the equality $S_{t_{n+1}}^{(0,b)} = S_{t_n}^{(0,e)}$ holds $\mathbb{P}$-a.s. for $n = 0, \ldots, N-1$.

This paper studies the problem of hedging long-term contingent claims embedded in segregated funds guarantees by means of dynamic hedging. A similar setup as in the work of Coleman et al. (2007) is considered. While the latter paper examines the presence of both jump risk and volatility risk for the equity, the current work strictly assesses the impact of jump risk on the risk management of long-term contingent claims. Note that the methodological approach presented in Section 3.3 for optimizing global policies can easily be adapted to the presence of additional risk factors for equity (e.g. volatility risk and regime risk). For the rest of the paper, assume that mortality risk can be completely diversified away and let T be the known maturity in years of the embedded guarantee to be hedged. This assumption can be motivated by the fact that in practice, insurance companies can significantly reduce the impact of mortality risk on their segregated funds portfolios by insuring additional policies. Furthermore, all VAs are assumed to be held until expiration (i.e. lapse risk is not considered) and their values are linked to a liquid index such as the S&P500, which implies no basis risk.

In this study, the option embedded in VAs is a GMMB with an annual ratchet feature that provides a payoff at time T of the maximum anniversary account value. The anniversary dates of the equity-linked insurance account are assumed to form a subset of the observation dates, i.e. $\{0, 1, \ldots, T\} \subseteq \mathcal{T}$. For $\lfloor \cdot \rfloor : \mathbb{R} \to \mathbb{R}$ denoting the floor function, i.e. $\lfloor x \rfloor$ is the largest integer smaller or equal to x, let $\{Z_{t_n}\}_{n=0}^{N}$ be the running maximum anniversary value process of the equity-linked account :

$$Z_{t_n} = \begin{cases} \max(S_0^{(0,b)}, \ldots, S_m^{(0,b)}), & \text{if } \lfloor t_n \rfloor = m \text{ and } m \in \{0, \ldots, T-1\}, \\ \max(S_0^{(0,b)}, \ldots, S_{T-1}^{(0,b)}), & \text{if } t_n = T. \end{cases}$$

The payoff of the GMMB with annual ratchet can be expressed as the account value at time T plus a lookback put option payoff

$$\begin{aligned} \max(S_0^{(0,b)}, \ldots, S_T^{(0,b)}) &= \max(\max(S_0^{(0,b)}, \ldots, S_{T-1}^{(0,b)}), S_T^{(0,b)}) \\ &= \max(Z_T - S_T^{(0,b)}, 0) + S_T^{(0,b)}. \end{aligned} \tag{3.1}$$

Thus, the assumptions of market completeness with respect to mortality risk and lapse risk considered in this paper entail that the risk exposure of the insurer selling a GMMB[2] is equivalent to holding short position in a long-term lookback option of fixed maturity T and of payoff function $\Phi : \mathbb{R} \times \mathbb{R}^T \to [0, \infty)$:

$$\Phi(S_T^{(0,b)}, Z_T) := \max(Z_T - S_T^{(0,b)}, 0). \tag{3.2}$$

The notation for trading strategies considered in this study for the mitigation of the risk exposure associated to a short position in Φ is now outlined. Let $\delta := \{\delta_{t_n}\}_{n=0}^{N}$ be a trading strategy used by the hedger to minimize his risk exposure to Φ, where for $n = 1, \ldots, N$,

[2] Coleman et al. (2007) consider the problem of hedging a ratchet GMDB with a fixed and known maturity T. The use of a fixed maturity in the latter paper is motivated by assuming market completeness under mortality risk and hedging the expected loss of the guarantee. While the current paper considers the risk mitigation of a GMMB instead of a GMDB, assumptions made in both papers (i.e. no mortality risk and lapse risk) entail that the benefits of the two guarantees are equivalent and result in the same lookback put option to hedge as in (3.2).

$\delta_{t_n} := (\delta_{t_n}^{(0)}, \ldots, \delta_{t_n}^{(D)}, \delta_{t_n}^{(B)})$ is a vector containing the number of shares held in each asset during the period $(t_{n-1}, t_n]$ with $\delta_{t_n}^{(0:D)} := (\delta_{t_n}^{(0)}, \ldots, \delta_{t_n}^{(D)})$ and $\delta_{t_n}^{(B)}$ respectively as the positions in the $D+1$ risky assets and in the risk-free asset. The initial portfolio (at time 0 before the first trade) is invested strictly in the risk-free asset (i.e. $\delta_0^{(1:D)} := [0, \ldots, 0]$). Also, for convenience, all options used as hedging instruments have one period maturity, i.e. they are traded once and held until expiration. Here is an additional assumption considered for the rest of the paper.

Assumption 3.1. *The market is liquid and trading in risky assets does not affect their prices.*

Before formally describing the global hedging optimization problem associated to a short position in Φ, some well-known concepts in the mathematical finance literature must be described. The reader is referred to Lamberton and Lapeyre (2011) for additional details. Let $\{G_{t_n}^{\delta}\}_{n=0}^{N}$ be the discounted gain process associated with the strategy δ where $G_{t_n}^{\delta}$ is the discounted gain at time t_n prior to rebalancing. $G_0^{\delta} := 0$ and

$$G_{t_n}^{\delta} := \sum_{k=1}^{n} \delta_{t_k}^{(0:D)} \bullet (B_{t_k}^{-1} \bar{S}_{t_{k-1}}^{(e)} - B_{t_{k-1}}^{-1} \bar{S}_{t_{k-1}}^{(b)}), \quad n = 1, 2, \ldots, N, \tag{3.3}$$

where $\bullet$ is the scalar product operator, i.e. for two n-dimensional vectors X and Y, $X \bullet Y := \sum_{i=1}^{n} X_i Y_i$. Also, let $\{V_{t_n}^{\delta}\}_{n=0}^{N}$ be the hedging portfolio values for a trading strategy δ where $V_{t_n}^{\delta}$ is the value prior to rebalancing at time t_n:

$$V_{t_n}^{\delta} := \delta_{t_n}^{(0:D)} \bullet \bar{S}_{t_{n-1}}^{(e)} + \delta_{t_n}^{(B)} B_{t_n}, \quad n = 1, \ldots, N, \tag{3.4}$$

and $V_0^{\delta} := \delta_0^{(B)}$ since the initial capital amount is assumed to be strictly invested in the risk-free asset. Moreover, in this paper, trading strategies considered require no cash infusion nor withdrawal except at the initialization of the contract (i.e. at time 0). Furthermore, trading strategies considered are $\mathbb{F}$-predictable, i.e. $\delta_0^{(j)} \in \mathcal{F}_0$ and $\delta_{n+1}^{(j)} \in \mathcal{F}_n$ for $j = 0, \ldots, D$. Such strategies are called *self-financing*. More precisely, the hedging strategy δ is said to be self-financing if it is $\mathbb{F}$-predictable and if

$$\delta_{t_{n+1}}^{(0:D)} \bullet \bar{S}_{t_n}^{(b)} + \delta_{t_{n+1}}^{(B)} B_{t_n} = V_{t_n}^{\delta}, \quad n = 0, 1, \ldots, N-1. \tag{3.5}$$

Lastly, let Π be the set of admissible trading strategies for the hedger which consists of all sufficiently well-behaved self-financing strategies.

Remark 3.1. *It can be shown that δ is self-financing if and only if $V_{t_n}^{\delta} = B_{t_n}(V_0^{\delta} + G_{t_n}^{\delta})$ for $n = 0, 1, \ldots, N$. See for instance Lamberton and Lapeyre (2011).*

3.2.2 Optimal hedging problem

The optimization problem of hedging the risk exposure associated to a short position in the long-term lookback option is now formally defined. For the hedger, the problem consists in the design of a trading policy which minimizes a *penalty*, also referred to as a *loss function*, of the difference between the payoff of the lookback option and the hedging portfolio value at maturity (i.e. the *hedging error* or *hedging shortfall*). Strategies embedded in such policies are called *global hedging strategies* as they are jointly optimized over all hedging decisions until the maturity of the lookback option. Let $\mathcal{L} : \mathbb{R} \to \mathbb{R}$ be a loss function for the hedging error. For the rest of the paper, assume without loss of generality that the position in the hedging portfolio is long, and that all assets and penalties are well-behaved and integrable enough. Specific conditions are beyond the scope of this study.

Definition 3.1. *(Global risk exposure) Define $\epsilon(V_0)$ as the global risk exposure of the short position in Φ under optimal hedge if the value of the initial hedging portfolio is $V_0 \in \mathbb{R}$:*

$$\epsilon(V_0) := \min_{\delta \in \Pi} \mathbb{E}\left[\mathcal{L}\left(\Phi(S_T^{(0,b)}, Z_T) - V_T^{\delta}\right)\right], \tag{3.6}$$

where the expectation is taken with respect to the physical measure.

The author wants to emphasize that the global hedging problem (3.6) is very complex to solve due to trading policies allowing for the use multiple hedging instruments, to the generality of the loss function and to the joint optimization over all trading decisions until the expiration date of the contingent claim being hedged. Numerical schemes proposed in the literature to solve this global hedging problem often rely on dynamics programming procedures (see, for instance, François et al. (2014)). However, one important contribution of this current paper is in conducting various numerical experiments studying the performance of long-term global policies and identifying important characteristics of these policies. The extensive empirical analysis of long-term global hedging carried out in Section 3.4 would be hardly reachable by traditional numerical procedures and heavily relies on the neural network hedging scheme described in the following section. One example of penalty which has been extensively studied in the global hedging literature is the *mean-square error* (MSE): $\mathcal{L}(x) = x^2$. This penalty entails that hedging gains and losses are treated equally, which as argued in Bertsimas et al. (2001), could be desirable for a financial participant who has to provide a price quote on a security prior to knowing his position (long or short). In a realistic setting, the choice of loss function should reflect the financial objectives and the risk aversion of the hedger. In the context of this paper where the position in Φ is always short, penalizing hedging gains is clearly undesirable for the hedger. The corresponding loss function to the MSE that penalizes exclusively hedging losses, not gains, is the *semi-mean-square error* (SMSE): $\mathcal{L}(x) = x^2 \mathbb{1}_{\{x>0\}}$.

The author wants to emphasize that different penalties will often result in different optimal hedging strategies. An extensive numerical study of the impact of optimizing trading policies under the MSE or SMSE loss function for the risk management of lookback options is performed in Section 3.4. Note that the methodological procedure for global hedging presented in Section 3.3 is flexible to any well-behaved penalties (see e.g. Carbonneau and Godin (2021b) for an implementation with the Conditional Value-at-Risk measure). Moreover, while this paper studies a specific example of long-term option to hedge, namely the lookback option of payoff Φ, the numerical scheme to approximate optimal hedging strategies can be applied with any European-type derivative of well-behaved payoff function, which can naturally include other VAs guarantees with payoffs analogous to financial derivatives.

Remark 3.2. *The terminology used in this paper to describe the global hedging problem and financial market setup is the one usually found in the mathematical finance literature. Note that the setup could also be formulated using terminology from the reinforcement learning literature, for instance by following the one from Sutton and Barto (2018) with the concepts of states to represent financial market observations, actions to represent the number of shares traded in hedging instruments and risk-adjusted cost functions to represent the expected hedging shortfall objective. For an example of a description of the global hedging setup with a reinforcement learning terminology, the reader is referred to the problem formulation presented in Buehler et al. (2019a).*

3.3 Methodology

This section describes the reinforcement learning procedure used to optimize global policies. The approach relies on the *deep hedging* algorithm of Buehler et al. (2019b) who showed that a *feedforward neural network* (FFNN) can be used to approximate arbitrarily well optimal hedging strategies in very general financial market conditions. At its core, a FFNN is a parameterized composite function which maps input to output vectors through the composition of a sequence of functions called *hidden layers*. Each hidden layer applies an affine and a nonlinear transformation to input vectors. For $L, d_0, \ldots, d_L, \tilde{d} \in \mathbb{N}$, a FFNN $F_\theta : \mathbb{R}^{d_0} \to \mathbb{R}^{\tilde{d}}$ with L hidden layers has the following representation:

$$F_\theta(X) := o \circ h_L \circ \ldots \circ h_1,$$

$$h_l(X) := g(W_l X + b_l), \quad l = 1, \ldots, L,$$

where $\circ$ denotes the function composition operator, $W_l \in \mathbb{R}^{d_l \times d_{l-1}}$ and $b_l \in \mathbb{R}^{d_l \times 1}$ are respectively known as the weight matrix and bias vector of the l^{th} hidden layer h_l, g is a nonlinear function applied to each scalar given as input and finally, $o : \mathbb{R}^{d_L} \to \mathbb{R}^{\tilde{d}}$ is the *output* *function* which applies an affine transformation to the output vector of the last hidden layer h_L as well as possibly a nonlinear transformation. Furthermore, the *trainable parameters* θ is the set of all weight matrices and bias vectors which are *learned* (i.e. fitted in statistical terms) by minimizing a specified cost function.

In the current study, the type of neural network considered for functions representing hedging policies is from the family of *recurrent neural networks* (RNNs, Rumelhart et al. (1986)). In contrast to FFNNs which maps input vectors to output vectors, RNNs is a class of neural networks mapping input sequences to output sequences. The architecture of RNNs is similar to FFNNs by applying successive affine and nonlinear transformations to inputs through hidden layers, but differs by having self-connections. Indeed, the RNN hidden layer is a function of both an input vector from the current time-step and an output vector from the hidden layer of the previous time-step, hence the name *recurrent.* In contrast to FFNNs, feedback loops in hidden layers entail that each output is dependent of past inputs, which makes RNNs more appropriate for time-series modeling. The type of RNN considered for dynamic hedging in this study is the *long short-term memory* (LSTM) introduced by Hochreiter and Schmidhuber (1997). This choice of neural network is motivated by recent results of Buehler et al. (2019a) who showed that LSTMs hedging policies are more effective for the risk mitigation of path-dependent contingent claims than FFNNs policies. Additional remarks are made in subsequent sections to motivate the choice of an LSTM for the specific setup considered in the current paper. For more general information about RNNs, the reader is referred to Chapter 10 of Goodfellow et al. (2016) and the many references therein.

The LSTM architecture is now formally defined. The application of LSTMs as functions representing global hedging policies is described in Section 3.3.1. In what follows, the time-steps are the same as the observation dates of the financial market.

Definition 3.2. *(LSTM) Let $F_\theta : \mathbb{R}^{N\times d_{in}} \to \mathbb{R}^{N\times d_{out}}$ be an LSTM which maps the sequence of feature vectors $\{X_{t_n}\}_{n=0}^{N-1}$ to output vectors $\{Y_{t_n}\}_{n=0}^{N-1}$, where X_{t_n} and Y_{t_n} are respectively two vectors of dimensions $d_{in}, d_{out} \in \mathbb{N}$. Let sigm$(\cdot)$ and tanh$(\cdot)$ be the sigmoid and hyperbolic* *tangent functions applied element-wise to each scalar given as input.*[3] *For* $H \in \mathbb{N}$*, the time-*t_n *computation of* F_θ *consists in* H *LSTM cells, each of which outputs a vector of* d_j *neurons denoted as* $h_{t_n}^{(j)} \in \mathbb{R}^{d_j \times 1}$ *for* $d_j \in \mathbb{N}$ *and* $j = 1, \ldots, H$*. More precisely, the computation done by the* j^{th} *LSTM cell at time* t_n *is as follows*[4]*:*

$$
\begin{aligned}
i_{t_n}^{(j)} &= sigm(W_i^{(j)}[h_{t_{n-1}}^{(j)}, h_{t_n}^{(j-1)}] + b_i^{(j)}), \\
f_{t_n}^{(j)} &= sigm(W_f^{(j)}[h_{t_{n-1}}^{(j)}, h_{t_n}^{(j-1)}] + b_f^{(j)}), \\
o_{t_n}^{(j)} &= sigm(W_o^{(j)}[h_{t_{n-1}}^{(j)}, h_{t_n}^{(j-1)}] + b_o^{(j)}), \\
c_{t_n}^{(j)} &= f_{t_n}^{(j)} \odot c_{t_{n-1}}^{(j)} + i_{t_n}^{(j)} \odot tanh(W_c^{(j)}[h_{t_{n-1}}^{(j)}, h_{t_n}^{(j-1)}] + b_c^{(j)}), \\
h_{t_n}^{(j)} &= o_{t_n}^{(j)} \odot tanh(c_{t_n}^{(j)}),
\end{aligned}
\tag{3.7}
$$

where $[\cdot\,,\cdot]$ *and* $\odot$ *denote respectively the concatenation of two vectors and the Hadamard product (i.e. the element-wise product) and*

- $W_i^{(1)}, W_f^{(1)}, W_o^{(1)}, W_c^{(1)} \in \mathbb{R}^{d_1 \times (d_1 + d_{in})}$ *and* $b_i^{(1)}, b_f^{(1)}, b_o^{(1)}, b_c^{(1)} \in \mathbb{R}^{d_1 \times 1}$.
- *If* $H \geq 2$*:* $W_i^{(j)}, W_f^{(j)}, W_o^{(j)}, W_c^{(j)} \in \mathbb{R}^{d_j \times (d_j + d_{j-1})}$ *and* $b_i^{(j)}, b_f^{(j)}, b_o^{(j)}, b_c^{(j)} \in \mathbb{R}^{d_j \times 1}$ *for* $j = 2, \ldots, H$.

At each time-step, the input of the first LSTM cell is the feature vector (i.e. $h_{t_n}^{(0)} := X_{t_n}$*) and the final output is an affine transformation of the output of the last LSTM cell:*

$$
Y_{t_n} = W_y h_{t_n}^{(H)} + b_y, \quad n = 0, \ldots, N-1,
\tag{3.8}
$$

where $W_y \in \mathbb{R}^{d_{out} \times d_H}$ *and* $b_y \in \mathbb{R}^{d_{out} \times 1}$*. Lastly, the set of trainable parameters denoted as* θ

[3] For $X := [X_1, \ldots, X_K]$, $\text{sigm}(X) := \left[\frac{1}{1+e^{-X_1}}, \ldots, \frac{1}{1+e^{-X_K}}\right]$ and $\tanh(X) := \left[\frac{e^{X_1}-e^{-X_1}}{e^{X_1}+e^{-X_1}}, \ldots, \frac{e^{X_K}-e^{-X_K}}{e^{X_K}+e^{-X_K}}\right]$.

[4] At time 0 (i.e. $n = 0$), the computation of the H LSTM cells is the same as in (3.7) with $h_{t_{-1}}^{(j)}$ and $c_{t_{-1}}^{(j)}$ as vectors of zeros of dimensions d_j for $j = 1, \ldots, H$.

consists of all weight matrices and bias vectors:

$$\theta := \left\{ \{W_i^{(j)}, W_f^{(j)}, W_o^{(j)}, W_c^{(j)}, b_i^{(j)}, b_f^{(j)}, b_o^{(j)}, b_c^{(j)}\}_{j=1}^{H}, W_y, b_y \right\}. \tag{3.9}$$

Remark 3.3. *In the deep learning literature, the* $i_{t_n}^{(j)}$*,* $f_{t_n}^{(j)}$ *and* $o_{t_n}^{(j)}$ *are known as input gates, forget gates and output gates. Their architectures have shown to help to alleviate the issue of learning long-term dependencies of time series with classical RNNs as they control the information passed through the LSTM cells. The reader is referred to Bengio et al. (1994) for more information about this latter pitfall of RNNs and to Chapter* 10.10 *of Goodfellow et al. (2016) and the many references therein for more general information about LSTMs.*

3.3.1 LSTM neural networks representing global policies

In the context of dynamic hedging, an LSTM maps a sequence of feature vectors consisting of relevant financial market observations to the sequence of positions in each asset for all time-steps. The set of trainable parameters θ is optimized to minimize the expected value of a loss function applied to the terminal hedging error obtained as a result of the trading decisions made by the LSTM. The following definition describes more formally how the LSTM computes the hedging strategy. Note that in numerical experiments presented in Section 3.4, the hedging instruments used for the risk minimization of Φ are either exclusively the underlying stock, or only standard options (i.e. vanilla calls and puts). The case of using both the stock and options is not considered because of its redundancy: option positions can always replicate investments in the underlying stock with calls and puts.

Definition 3.3. *(Hedging with an LSTM) Let* F_θ *be an LSTM as in Definition 3.2 which maps the sequence of feature vectors* $\{X_{t_n}\}_{n=0}^{N-1}$ *to the output vectors* $\{Y_{t_n}\}_{n=0}^{N-1}$*. The choice of hedging instruments (i.e. the underlying or standard options) implies differences for the* *feature vectors and output vectors*[5]*:*

1) Hedging only with the underlying: the feature vector at each time-step is[6]

$$X_{t_n} := [\log(S_{t_n}^{(0,b)}), \log(Z_{t_n}), V_{t_n}^{\delta}/V_0^{\delta}], \quad n = 0, \ldots, N-1,$$

and F_θ outputs at each rebalancing date the position in the underlying: $\delta_{t_n}^{(0)} = Y_{t_{n-1}}$.

2) Hedging only with options: the feature vector at each time-step includes option prices as well as the price of the underlying:

$$X_{t_n} := [\log(\bar{S}_{t_n}^{(b)}), \log(Z_{t_n}), V_{t_n}^{\delta}/V_0^{\delta}], \quad n = 0, \ldots, N-1,$$

and F_θ outputs at each rebalancing date the position in the D options: $[\delta_{t_n}^{(1)}, \ldots, \delta_{t_n}^{(D)}] = Y_{t_{n-1}}$.

It is important to note that the choice of dynamics for the financial market could imply that relevant necessary information to compute the time-t_n trading strategy should be added to feature vectors. For instance, Carbonneau and Godin (2021b) apply the deep hedging algorithm with GARCH models, which entails adding the volatility process to feature vectors. In the current paper, the models considered for the underlying imply that $\{S_{t_n}^{(0,b)}\}_{n=0}^{N}$ is a

[5] When δ is self-financing, the computation of $V_{t_n}^{\delta}$ for $n = 1, \ldots, N$ can be done recursively as in (3.10) below given V_0^{δ}:

$$\begin{aligned} V_{t_n}^{\delta} &= B_{t_n}(V_0^{\delta} + G_{t_n}^{\delta}) \\ &= B_{t_n}(V_0^{\delta} + G_{t_{n-1}}^{\delta} + \delta_{t_n}^{(0:D)} \bullet (B_{t_n}^{-1}\bar{S}_{t_{n-1}}^{(e)} - B_{t_{n-1}}^{-1}\bar{S}_{t_{n-1}}^{(b)})) \\ &= \frac{B_{t_n}}{B_{t_{n-1}}} V_{t_{n-1}}^{\delta} + \delta_{t_n}^{(0:D)} \bullet (\bar{S}_{t_{n-1}}^{(e)} - \frac{B_{t_n}}{B_{t_{n-1}}}\bar{S}_{t_{n-1}}^{(b)}) \\ &= e^{r\Delta_N} V_{t_{n-1}}^{\delta} + \delta_{t_n}^{(0:D)} \bullet (\bar{S}_{t_{n-1}}^{(e)} - e^{r\Delta_N}\bar{S}_{t_{n-1}}^{(b)}). \end{aligned} \tag{3.10}$$

[6] Using the transformations $\{\log(S_{t_n}^{(0,b)}), \log(Z_{t_n}), V_{t_n}^{\delta}/V_0^{\delta}\}$ instead of $\{S_{t_n}^{(0,b)}, Z_{t_n}, V_{t_n}^{\delta}\}$ in feature vectors for the numerical experiments of Section 3.4 was found to significantly improve the training of neural networks. The log transformation could not be applied for the hedging portfolio values since $V_{t_n}^{\delta}$ can theoretically take values on the real line. Note that Buehler et al. (2019b) and Buehler et al. (2019a) also represented risky asset prices in terms of their logarithms in feature vectors.

Markov process under $\mathbb{P}$ and thus that no additional variables must be added to feature vectors. The author emphasizes that the methodological approach considered for optimizing global policies with the use of LSTMs can easily be adapted to dynamics requiring the inclusion of additional state variables.

Remark 3.4. *Buehler et al. (2019a) deep hedge exotic derivatives with an LSTM with feature vectors that does not include a path-dependent state variable such as* $\{Z_{t_n}\}_{n=0}^{N-1}$*. One can observe that adding* $\{Z_{t_n}\}_{n=0}^{N-1}$ *to feature vectors as per Definition 3.3 significantly improved the performance of the optimized hedging policies when the number of trading period is large (i.e. for large values of* N*), while for less frequent trading, the gain is marginal.*

Remark 3.5. *Theoretical results from Buehler et al. (2019b) show that a FFNN could have been used to approximate arbitrarily well the optimal hedging policy in the setup considered in this study (see Proposition 4.3 of their paper). However, simulations showed that hedging with an LSTM was significantly more effective than with a FFNN for the numerical experiments conducted in Section 3.4 in terms of both computational time (i.e. faster learning with LSTMs) and hedging effectiveness, which motivated the use of LSTMs as trading policies. The justifications of the superiority of LSTMs over FFNNs in the context of this paper are out-of-scope and are left out as interesting potential future work.*

For the rest of the paper, a single set of hyperparameters for the LSTM is considered in terms of the number of LSTM cells and neurons per cell.[7] The optimization problem thus consists in searching for the optimal values of trainable parameters for this specific architecture of LSTM. The hyperparameter tuning step is not considered in this paper; the reader is referred to Carbonneau and Godin (2021b) for a complete description of the optimal hedging problem with FFNNs which includes hyperparameter tuning. The following defines formally the alternative optimization problem considered, where the optimization boils down to solving

[7] Note that as per Definition 3.3, the dimensions of the input and output of the LSTM at each time-step, i.e. d_{in} and d_{out}, are dependent of the choice of hedging instruments. Thus, while the number of neurons $d_1, \ldots, d_H$ and the number of LSTM cells H is fixed for the numerical experiments of Section 3.4, the total number of trainable parameters will vary with respect to the choice of hedging instruments.

for the trainable parameters of the LSTM trading policy.

Definition 3.4. *(Global risk exposure with an LSTM) Define $\tilde{\epsilon}(V_0)$ as the global risk exposure of the short position in Φ under optimal hedge if the hedging strategy is given by F_θ and if the value of the initial hedging portfolio is $V_0 \in \mathbb{R}$:*

$$\tilde{\epsilon}(V_0) := \min_{\theta \in \mathbb{R}^q} \mathbb{E}\left[\mathcal{L}\left(\Phi(S_T^{(0,b)}, Z_T) - V_T^{\delta^\theta}\right)\right], \tag{3.11}$$

where δ^θ is to be understood as the output vectors of F_θ and $q \in \mathbb{N}$ is the total number of trainable parameters.

3.3.2 Training of neural networks

The numerical scheme to optimize the set of trainable parameters θ is now described. For convenience, a similar notation as in the work of Carbonneau and Godin (2021b) is used. For a given loss function and an initial portfolio value, the objective is to find θ such that the risk exposure of a short position in Φ is minimized (i.e. as in (3.11)). The training procedure was originally proposed in Buehler et al. (2019b) and relies on (mini-batch) stochastic gradient descent (SGD), a very popular algorithm in the deep learning literature to train neural networks. Denote $J : \mathbb{R}^q \to \mathbb{R}$ as the cost function to minimize:

$$J(\theta) := \mathbb{E}\left[\mathcal{L}\left(\Phi(S_T^{(0,b)}, Z_T) - V_T^{\delta^\theta}\right)\right], \quad \theta \in \mathbb{R}^q.$$

Let θ_0 be the initial values for the trainable parameters.[8] The optimization procedure consists in updating iteratively the trainable parameters as follows:

$$\theta_{j+1} = \theta_j - \eta_j \nabla_\theta J(\theta_j), \tag{3.12}$$

[8] In this paper, the initial values of θ are always set as the *Glorot uniform initialization* of Glorot and Bengio (2010).

where ∇_θ is the gradient operator with respect to θ and $\{\eta_j\}_{j\geq 0}$ is a sequence of small positive real values. In the context of this paper, $\nabla_\theta J(\theta)$ is unknown analytically and is estimated with Monte Carlo sampling. Let $\mathbb{B}_j := \{\pi_{i,j}\}_{i=1}^{N_{\text{batch}}}$ be a mini-batch of simulated hedging errors of size $N_{\text{batch}} \in \mathbb{N}$ with $\pi_{i,j}$ as the i^{th} hedging error if $\theta = \theta_j$:

$$\pi_{i,j} := \Phi(S_{T,i}^{(0,b)}, Z_{T,i}) - V_{T,i}^{\delta^{\theta_j}},$$

where $S_{T,i}^{(0,b)}, Z_{T,i}$ and $V_{T,i}^{\delta^{\theta_j}}$ are to be understood as the values of the i^{th} simulated path. Moreover, denote $\hat{J} : \mathbb{R}^{N_{\text{batch}}} \to \mathbb{R}$ as the empirical estimator of $J(\theta_j)$ evaluated with the mini-batch $\mathbb{B}_j$. Mini-batch stochastic gradient descent consists in approximating $\nabla_\theta J(\theta_j)$ in the update rule (3.12) with $\nabla_\theta \hat{J}(\mathbb{B}_j)$. In Section 3.4, the MSE and SMSE penalties, defined respectively as $\mathcal{L}^{\text{MSE}}(x) := x^2$ and $\mathcal{L}^{\text{SMSE}}(x) := x^2 \mathbb{1}_{\{x>0\}}$, are considered for the numerical experiments conducted. The empirical estimator of the cost function under each penalty can be stated as follows:

$$\hat{J}^{\text{MSE}}(\mathbb{B}_j) := \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \pi_{i,j}^2,$$
$$\hat{J}^{\text{SMSE}}(\mathbb{B}_j) := \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \pi_{i,j}^2 \mathbb{1}_{\{\pi_{i,j}>0\}}.$$

One essential property of the architecture of neural networks is that the gradient of empirical cost functions (i.e. $\nabla_\theta \hat{J}(\mathbb{B}_j)$ for both penalties) can be computed exactly. Indeed, note that hedging errors are linearly dependent of the trading strategies produced as the outputs of the LSTM. Furthermore, the gradient of the outputs of an LSTM with respect to trainable parameters can be computed exactly (see e.g. Chapter 10 of Goodfellow et al. (2016)).

Remark 3.6. *The algorithm backpropagation through time (BPTT) is often used to compute exactly the gradient of a cost function with respect to the trainable parameters for recurrent type of neural networks such as an LSTM. BPTT leverages the structure of LSTMs (e.g. parameters sharing at each time-step) as well as the chain rule of calculus to obtain such* *gradients. In practice, efficient deep learning libraries such as Tensorflow (Abadi et al., 2016) are often used to implement BPTT. Moreover, algorithms such as Adam (Kingma and Ba, 2014) which dynamically adapt the terms $\{\eta_j\}_{j\geq 0}$ in (3.12) have been shown to significantly improve the training of neural networks. For the rest of the paper, Tensorflow and Adam are used to train every neural network. Pseudo-code of the algorithm for training neural networks is presented in Section 3.6.*

3.4 Numerical study

This section presents an extensive numerical study of the neural-based global hedging scheme for the mitigation of the risk exposure associated to a short position in the long-term lookback option. Section 3.4.3 examines the hedging effectiveness of both quadratic and non-quadratic global hedging strategies as well as the local risk minimization scheme of Coleman et al. (2007) with different hedging instruments and different dynamics for the financial market. The conduction of such thorough benchmarking experiments heavily relies on the methodological approach considered in this paper, namely the use of neural networks to represent long-term global trading policies with the deep hedging algorithm. The performance of these extensive numerical experiments enables us to provide novel qualitative insights into specific characteristics of the optimized long-term global hedging policies in Section 3.4.4 which have yet to be studied in the literature. The setup considered for all numerical experiments is described in Section 3.4.1 and Section 3.4.2.

3.4.1 Market setup

The market setup considered in this paper is very similar to the work of Coleman et al. (2007). The contingent claim to hedge is a lookback option of payoff Φ as in (3.2) with a time-to-maturity of 10 years (i.e. $T = 10$). The annualized continuously compounded risk-free rate is set at 3% (i.e. $r = 0.03$) and $S_0^{(0,b)} = 100$. In the design of hedging policies, the trading instruments considered are either exclusively the underlying stock, two options or six options.

All options have a time-to-maturity of 1 year, are traded once and are held until expiration. For the case of two options, the hedging instruments available at the beginning of each year consist of at-the-money (ATM) calls and puts. With six options, three calls of moneynesses $K \in \{S_{t_n}^{(0,b)}, 1.1S_{t_n}^{(0,b)}, 1.2S_{t_n}^{(0,b)}\}$ and three puts of moneynesses $K \in \{S_{t_n}^{(0,b)}, 0.9S_{t_n}^{(0,b)}, 0.8S_{t_n}^{(0,b)}\}$ are available at the beginning of each year t_n. As for the underlying, both monthly and yearly rebalancing are considered in numerical experiments. Yearly time-steps are used for all hedging instruments (i.e. $N = 10$) except when hedging is done with the underlying on a monthly basis (i.e. $N = 120$).

Remark 3.7. *The methodological approach of Section 3.3 is in no way dependent on this choice of hedging instruments.*

3.4.1.1 Global hedging penalties

The penalties studied for global hedging are the MSE and SMSE, and the respective optimization procedures are referred to as quadratic deep hedging (QDH) and semi-quadratic deep hedging (SQDH). While the MSE penalizes equally hedging gains and losses, the SMSE is more in line with the actual objectives of the hedger as it corresponds to an agent who strictly penalizes hedging losses proportionally to their squared values. It is important to note that the computational cost of the deep hedging algorithm is closed to invariant to the choice of loss function. The motivation for assessing the effectiveness of QDH is the popularity of the quadratic penalty in the global hedging literature.

3.4.1.2 LSTM training

The training procedure for neural networks is done as described in Section 3.2 with a training set of 350,000 paths with 150 epochs[9] and a mini-batch size of 1,000. More precisely, mini-batches used for the SGD procedure are sampled exclusively from this training set.

[9] One epoch is defined as a complete iteration of SGD on the training set. For a training set and mini-batch size of respectively 350,000 and 1,000, one epoch consists of a total of 350 updates of parameters as in (3.12).

Furthermore, a validation set of 75,000 additional independent paths is used to find the optimal set of trainable parameters out of the 150 epochs. Indeed, at the end of each epoch, the hedging statistic associated to the penalty being optimized (i.e. MSE for QDH and SMSE for SQDH) is evaluated on the validation set at the current values of the trainable parameters. Following the training methodology commonly considered in the literature (see, for instance, Goodfellow et al. (2016)), the optimal set of trainable parameters is approximated by the one that minimizes the empirical cost function on the validation set out of 150 epochs. The author wants to highlight that the validation set is used exclusively for selecting the trainable parameters out of the 150 epochs, not to perform the SGD procedure with the latter using only the training set.

All hedging statistics values reported in subsequent sections with QDH or SQDH procedures are from a test set (out-of-sample) of 75,000 paths. The structure of the LSTM is as in Definition 3.2 with two LSTM cells (i.e. $H = 2$) and 24 neurons per cell (i.e. $d_1 = d_2 = 24$). The Adam optimizer (Kingma and Ba (2014)) is used for all examples with a learning rate hyperparameter of 0.01 for QDH and $\frac{0.01}{6}$ for SQDH since a smaller learning rate was found to improve the training under the SMSE penalty.

Remark 3.8. *The training procedure considered in this paper could also be used for selecting hyperparameters with the validation set, for instance the number of LSTM cells or neurons per cell. The interested reader is referred to Chapter* 11.4 *of Goodfellow et al. (2016) for examples of well-known procedures for selecting hyperparameters. However, additional results not presented in this paper show that the use of an LSTM with more capacity (additional LSTM cell) or with less capacity (less neurons per cell) does not impact significantly the hedging performance of the neural networks. Consequently, qualitative conclusions presented in this section are robust to different sets of hyperparameters for the neural network, and a single architecture for LSTMs is considered throughout the rest of the paper.*

3.4.1.3 Local risk minimization

Define $\{C^\delta_{t_n}\}_{n=0}^N$ as the discounted cumulative cost process associated to a trading strategy δ:

$$C^\delta_{t_n} := B^{-1}_{t_n} V^\delta_{t_n} - G^\delta_{t_n}, \quad n = 0, \ldots, N.$$

Contrarily to global hedging procedures, local risk minimization schemes result in hedging strategies that are not necessarily self-financing. Indeed, the optimization of hedging strategies under this framework imposes the constraint that the terminal portfolio value exactly matches the payoff of the contingent claim, i.e. $V^\delta_T = \Phi(S^{(0,b)}_T, Z_T)$ $\mathbb{P}$-a.s., which can always be respected by the injection or withdrawal of capital at time T. Under this constraint, local risk minimization optimizes at each time-step starting backward from time T positions in the assets which minimize the expected squared incremental cost. More precisely, for $n = N-1, \ldots, 0$, the optimization aims at finding $(\delta^{(0:D)}_{t_{n+1}}, \delta^{(B)}_{t_{n+1}})$ that minimize $\mathbb{E}[(C^\delta_{t_{n+1}} - C^\delta_{t_n})^2 | \mathcal{F}_{t_n}]$ at time t_n with the constraint that $V^\delta_T = \Phi(S^{(0,b)}_T, Z_T)$. The optimal initial capital amount to invest, denoted as $V^\star_0$ hereafter, is also obtained as a result of this scheme. Once the trading strategy δ is optimized with the local risk minimization procedure, a self-financing strategy can be constructed by setting the initial portfolio value as $V^\delta_0 = V^\star_0$, by following the optimized trading strategy strictly for the risky assets (i.e. $\delta^{(0:D)}_{t_n}$ for $n = 1, \ldots, N$) and by adjusting positions in the risk-free asset such that the trading strategy is self-financing (i.e. respecting (3.5)). Hedging statistics reported in the numerical experiments of this section with local risk minimization are self-financing as per the latter description and are from the work of Coleman et al. (2007). For examples of numerical schemes to implement local risk procedures, the reader is referred to Coleman et al. (2006) or Augustyniak et al. (2017).

The motivation for benchmarking global hedging policies to local risk minimization is twofold. First, local risk procedures are popular for the risk mitigation of VAs guarantees in the literature due to their tractability in high-dimensional setups (e.g. Coleman et al. (2006), Coleman et al. (2007), Kélani and Quittard-Pinon (2017), Trottier et al. (2018b) and Trottier et al. (2018a)). Second, in the context of hedging European vanilla options of maturity one to three years, Augustyniak et al. (2017) showed that global quadratic hedging procedures trading exclusively the underlying stock improve upon the downside risk reduction over local risk minimization schemes. The question remains if the latter holds for longer maturities and when liquid options are used as hedging instruments.

3.4.1.4 Hedging metrics

The hedging metrics considered for the benchmarking of the different trading policies include the root-mean-square error (RMSE) and the semi-RMSE (i.e. the root of the SMSE statistic). Tail risk metrics are also studied with the Value-at-Risk (VaR) and the Conditional Value-at-Risk (CVaR, Rockafellar and Uryasev (2002)). For an absolutely continuous integrable random variable[10], the CVaR at confidence level α has the following representation:

$$\text{CVaR}_\alpha(X) := \mathbb{E}[X|X \geq \text{VaR}_\alpha(X)], \quad \alpha \in (0,1), \tag{3.13}$$

where $\text{VaR}_\alpha(X) := \min_x \{x : \mathbb{P}(X \leq x) \geq \alpha\}$ is the VaR at confidence level α. The CVaR_α represents tail risk by averaging all hedging errors larger than the α^{th} percentile of the distribution of hedging errors (i.e. the VaR_α metric). Recall that all hedging statistics presented in subsequent sections with neural networks are estimated with conventional empirical estimators on the test set. These hedging statistics are thus obtained in an out-of-sample fashion.

3.4.2 Dynamics of financial market

In this paper, the choice of risky assets dynamics is motivated by the objective of studying the optimized global policies under different stylized features of the financial market. It is important to recall that deep hedging is a model-free reinforcement learning approach: the

[10] All dynamics assumed for the underlying in Section 3.4 imply that hedging errors are absolutely continuous integrable random variables.

LSTM is never explicitly told the dynamics of the financial market during its training phase. Instead, the neural network must learn through many simulations of a market generator how to dynamically adapt its embedded policy, i.e. its trainable parameters, with the objective of minimizing the expected loss function of the resulting hedging errors. The current work studies the impact of the presence of jump risk on optimized global policies by considering the Merton jump-diffusion model (MJD, Merton (1976)) as well as the Black-Scholes model (BSM, Black and Scholes (1973)). Note that parameters values for the BSM and MJD dynamics are the same as in the ones from Coleman et al. (2007) and are presented respectively in Table 3.1 and Table 3.2. It is worth noting that while the model parameters imply somewhat similar periodic means and standard deviations for log-returns, the MJD dynamics entails large and volatile negative jumps occurring on average once over the lifetime of the lookback option to be hedged.

Moreover, both stochastic dynamics for risky assets considered in this paper imply that the market is arbitrage-free. By the first fundamental theorem of asset pricing, there exist a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that $\{e^{-rt_n} S_{t_n}^{(b,0)}\}_{n=0}^{N}$ is an $(\mathbb{F}, \mathbb{Q})$-martingale (see, for instance, Delbaen and Schachermayer (1994)). Let $y_{t_n} := \log(S_{t_n}^{(0,b)}/S_{t_{n-1}}^{(0,b)})$ be the periodic log-return of the underlying, and $\{\epsilon_{t_n}^{\mathbb{P}}\}_{n=1}^{N}$ and $\{\epsilon_{t_n}^{\mathbb{Q}}\}_{n=1}^{N}$ be sequences of independent standard normal random variables under respectively $\mathbb{P}$ and $\mathbb{Q}$. The dynamics of both models are now formally defined.

3.4.2.1 BSM under $\mathbb{P}$

The discrete BSM assumes that log-returns are i.i.d. normal random variables of periodic mean and variance of respectively $(\mu - \frac{\sigma^2}{2})\Delta_N$ and $\sigma^2 \Delta_N$:

$$y_{t_n} = \left(\mu - \frac{\sigma^2}{2}\right) \Delta_N + \sigma\sqrt{\Delta_N}\epsilon_{t_n}^{\mathbb{P}}, \quad n = 1, \ldots, N, \tag{3.14}$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$ are the yearly model parameters.

3.4.2.2 MJD under $\mathbb{P}$

The MJD model extends upon the idealized Black-Scholes dynamics with the presence of random jumps to the underlying stock price. More precisely, let $\{\zeta_k^{\mathbb{P}}\}_{k=1}^{\infty}$ be independent normal random variables of mean μ_J and variance σ_J^2, and $\{N_{t_n}^{\mathbb{P}}\}_{n=0}^{N}$ be values of a Poisson process of intensity $\lambda > 0$ where $\{\zeta_k^{\mathbb{P}}\}_{k=1}^{\infty}, \{N_{t_n}^{\mathbb{P}}\}_{n=0}^{N}$ and $\{\epsilon_{t_n}^{\mathbb{P}}\}_{n=1}^{N}$ are independent. Periodic log-returns under this model can be stated as follows[11]:

$$y_{t_n} = \left(\alpha - \lambda\left(e^{\mu_J+\sigma_J^2/2} - 1\right) - \frac{\sigma^2}{2}\right)\Delta_N + \sigma\sqrt{\Delta_N}\epsilon_{t_n}^{\mathbb{P}} + \sum_{k=N_{t_{n-1}}^{\mathbb{P}}+1}^{N_{t_n}^{\mathbb{P}}} \zeta_k^{\mathbb{P}}, \tag{3.15}$$

where $\{\alpha, \mu_J, \sigma_J, \lambda, \sigma\}$ are the model parameters with $\{\alpha, \lambda, \sigma\}$ being on a yearly scale, $\alpha \in \mathbb{R}$ and $\sigma > 0$.

3.4.2.3 BSM under $\mathbb{Q}$

By a discrete-time version of the Girsanov theorem, there exist an $\mathbb{F}$-adapted market price of risk process $\{\varphi_{t_n}\}_{n=1}^{N}$ such that

$$\epsilon_{t_n}^{\mathbb{Q}} = \epsilon_{t_n}^{\mathbb{P}} - \varphi_{t_n}, \quad n = 1, \ldots, N. \tag{3.16}$$

For $n = 1, \ldots, N$, let $\varphi_{t_n} := -\sqrt{\Delta_N}\left(\frac{\mu-r}{\sigma}\right)$. By replacing $\epsilon_{t_n}^{\mathbb{P}} = \epsilon_{t_n}^{\mathbb{Q}} + \varphi_{t_n}$ into (3.14), it is straightforward to obtain the $\mathbb{Q}$-dynamics of log-returns:

$$y_{t_n} = \left(r - \frac{\sigma^2}{2}\right)\Delta_N + \sigma\sqrt{\Delta_N}\epsilon_{t_n}^{\mathbb{Q}}, \quad n = 1, \ldots, N. \tag{3.17}$$

[11] We adopt the convention that if $N_{t_n}^{\mathbb{P}} = N_{t_{n-1}}^{\mathbb{P}}$, then:

$$\sum_{k=N_{t_{n-1}}^{\mathbb{P}}+1}^{N_{t_n}^{\mathbb{P}}} \zeta_k^{\mathbb{P}} = 0.$$

The pricing of European calls and puts used as hedging instruments under this model is done with the well-known Black-Scholes closed-form solutions.

3.4.2.4 MJD under $\mathbb{Q}$

The change of measure considered under the MJD dynamics is the same as the one from Coleman et al. (2007). Let $\{\zeta_k^{\mathbb{Q}}\}_{k=1}^{\infty}$ be independent normal random variables under $\mathbb{Q}$ of mean $\tilde{\mu}_J$ and variance $\tilde{\sigma}_J^2$, and $\{N_{t_n}^{\mathbb{Q}}\}_{n=0}^{N}$ be values of a Poisson process of intensity $\tilde{\lambda} > 0$ where $\{\zeta_k^{\mathbb{Q}}\}_{k=1}^{\infty}$, $\{N_{t_n}^{\mathbb{Q}}\}_{n=0}^{N}$ and $\{\epsilon_{t_n}^{\mathbb{Q}}\}_{n=1}^{N}$ are independent. The $\mathbb{Q}$-dynamics of log-returns can be stated as follows:

$$y_{t_n} = \left(r - \tilde{\lambda}\left(e^{\tilde{\mu}_J + \tilde{\sigma}_J^2/2} - 1\right) - \frac{\sigma^2}{2}\right)\Delta_N + \sigma\sqrt{\Delta_N}\epsilon_{t_n}^{\mathbb{Q}} + \sum_{k=N_{t_{n-1}}^{\mathbb{Q}}+1}^{N_{t_n}^{\mathbb{Q}}} \zeta_k^{\mathbb{Q}},$$

where $\tilde{\sigma}_J := \sigma_J$, $\tilde{\mu}_J := \mu_J - (1-\gamma)\sigma_J^2$, $\tilde{\lambda} := \lambda e^{-(1-\gamma)(\mu_J - \frac{1}{2}(1-\gamma)\sigma_J^2)}$ with $\gamma \leq 1$ as the risk aversion parameter which is set for all experiments as $\gamma = -1.5$. Note that the value of the risk aversion parameter implies more frequent and more negative jumps on average under $\mathbb{Q}$ than under $\mathbb{P}$ by increasing $\tilde{\lambda}$ and decreasing $\tilde{\mu}_J$. The pricing of European calls and puts used as hedging instruments under the MJD model is done with the well-known closed-form solutions.

Table 3.1: Parameters of the Black-Scholes model.

μ	σ
0.10	0.15

Notes: Both μ and σ are on an annual basis.

Table 3.2: Parameters of the Merton jump-diffusion model.

α	σ	λ	μ_J	σ_J	γ
0.10	0.15	0.10	-0.20	0.15	-1.5

Notes: α, σ and λ are on an annual basis.

3.4.3 Benchmarking of hedging policies

In this section, the hedging effectiveness of QDH, SQDH and local risk minimization is assessed under various market settings. The analysis starts off in Section 3.4.3.1 by comparing QDH and local risk minimization performance as both approaches are optimized with a quadratic criterion; the benchmarking of global hedging policies embedded in QDH and SQDH procedures is done afterwards in Section 3.4.3.2.

3.4.3.1 QDH and local risk minimization benchmark

Table 3.3 and Table 3.4 present hedging statistics of QDH and local risk minimization under respectively the BSM and MJD model.[12] For comparative purposes, the initial capital investment is set to the optimized value obtained as a result of the local risk minimization procedure of Coleman et al. (2007) for all examples. Note that this choice naturally puts QDH procedures at a disadvantage. Also, since QDH procedures optimize the MSE penalty, this global procedure is expected to outperform local risk minimization on the RMSE metric. The question remains if QDH also improves upon the downside risk captured by the $\text{VaR}_{0.95}$ and $\text{CVaR}_{0.95}$ statistics.

Numerical results under both dynamics demonstrate that QDH outperforms local risk minimization across all downside risk metrics and all hedging instruments. Indeed, the risk reduction obtained with QDH over local risk minimization is most impressive with six options: the percentage decrease for respectively the RMSE, $\text{VaR}_{0.95}$ and $\text{CVaR}_{0.95}$ statistics are of $33\%, 52\%$ and 36% under the BSM and of $27\%, 38\%$ and 30% under the MJD model. As for hedging exclusively with the underlying stock on a monthly and yearly basis or when trading only two options, the improvement of QDH over local risk minimization for the three hedging statistics ranges between 5% to 13% under the BSM and 8% to 20% under the MJD model, except for the $\text{VaR}_{0.95}$ metric with the stock on a monthly basis under the MJD dynamics

[12] The choice of hedging statistics presented in Table 3.3 and Table 3.4 are the ones considered in Coleman et al. (2007). Additional hedging statistics for QDH are presented in Section 3.4.3.2.

Table 3.3: Benchmarking of quadratic deep hedging (QDH) and local risk minimization to hedge the lookback option of $T = 10$ years under the BSM.

		Local risk minimization				QDH			
Statistics	V_0^δ	Mean	RMSE	$\text{VaR}_{0.95}$	$\text{CVaR}_{0.95}$	Mean	RMSE	$\text{VaR}_{0.95}$	$\text{CVaR}_{0.95}$
Stock (year)	13.9	0.3	15.9	28.5	43.2	−0.2	14.8	25.5	41.2
Stock (month)	17.3	0.1	5.5	8.9	13.0	0.0	4.9	7.7	12.2
Two options	17.4	−0.1	4.6	7.0	11.9	−0.1	4.2	6.1	11.2
Six options	17.7	0.0	1.6	2.4	3.8	−0.2	1.1	1.2	2.4

Notes: Hedging statistics under the BSM with $\mu = 0.1, \sigma = 0.15, r = 0.03$ and $S_0^{(0,b)} = 100$ (see Section 3.4.2.1 for model description under $\mathbb{P}$ and Section 3.4.2.3 for the risk-neutral dynamics used for option pricing). *Hedging instruments*: monthly and yearly underlying, yearly ATM call and put options (*two options*) or three yearly calls and puts of strikes $K = \{S_{t_n}^{(0,b)}, 1.1S_{t_n}^{(0,b)}, 1.2S_{t_n}^{(0,b)}\}$ and $K = \{S_{t_n}^{(0,b)}, 0.9S_{t_n}^{(0,b)}, 0.8S_{t_n}^{(0,b)}\}$ (*six options*). Results for local risk minimization and initial portfolio values V_0^δ are from Table 3 of Coleman et al. (2007). Results for QDH are computed based on 75,000 independent paths generated from the BSM under $\mathbb{P}$. Training of the neural networks is done as described in Section 3.4.1.2.

which achieves 30% reduction. These results clearly demonstrate that in the risk management of the lookback option, the use of global procedures rather than local procedures leads to significantly better trading strategies. The conclusion that global procedures outperform local procedures for long-term contingent claims when trading exclusively the underlying or multiple shorter-term options for the setting considered in this study is a novel qualitative contribution of this paper and motivates the use of neural networks to represent global hedging policies for the risk mitigation of long-term options.

3.4.3.2 QDH and SQDH benchmark

Having shown the outperformance of QDH over local risk minimization in the previous section, the benchmarking of QDH and SQDH policies is now examined. The exact same setup as in the previous section is considered, except for the initial capital investment of trading strategies which is set as the risk-neutral price of the lookback option under both

Table 3.4: Benchmarking of quadratic deep hedging (QDH) and local risk minimization to hedge the lookback option of $T = 10$ years under the MJD model.

		Local risk minimization				QDH			
Statistics	V_0^δ	Mean	RMSE	$\text{VaR}_{0.95}$	$\text{CVaR}_{0.95}$	Mean	RMSE	$\text{VaR}_{0.95}$	$\text{CVaR}_{0.95}$
Stock (year)	19.5	0.4	21.4	38.4	60.5	0.1	19.5	33.1	55.8
Stock (month)	22.8	0.1	13.0	23.5	38.4	0.2	11.0	16.3	33.5
Two options	24.6	−0.1	6.0	8.4	15.2	0.1	5.2	6.7	12.9
Six options	25.2	0.0	1.9	2.8	4.6	0.2	1.3	1.7	3.2

Notes: Hedging statistics under the MJD model with $\alpha = 0.1, \sigma = 0.15, \lambda = 0.1, \mu_J = -0.2, \sigma_J = 0.15, \gamma = -1.5, r = 0.03$ and $S_0^{(0,b)} = 100$ (see Section 3.4.2.2 for model description under $\mathbb{P}$ and Section 3.4.2.4 for the risk-neutral dynamics used for option pricing). *Hedging instruments*: monthly and yearly underlying, yearly ATM call and put options (*two options*) or three yearly calls and puts of strikes $K = \{S_{t_n}^{(0,b)}, 1.1S_{t_n}^{(0,b)}, 1.2S_{t_n}^{(0,b)}\}$ and $K = \{S_{t_n}^{(0,b)}, 0.9S_{t_n}^{(0,b)}, 0.8S_{t_n}^{(0,b)}\}$ (*six options*). Results for local risk minimization and initial portfolio values V_0^δ are from Table 4 of Coleman et al. (2007). Results for QDH are computed based on 75,000 independent paths generated from the MJD model under $\mathbb{P}$. Training of the neural networks is done as described in Section 3.4.1.2.

dynamics for all hedging instruments: 17.7\$ for BSM and 25.3\$ for MJD.[13] This choice is motivated by the objective of comparing on common grounds the results obtained across the different hedging instruments for both global hedging approaches. Table 3.5 and Table 3.6 present descriptive statistics of the hedging shortfall obtained with QDH and SQDH under respectively the BSM and MJD model.

Numerical results indicate that as compared to QDH, SQDH policies result in downside risk metrics two to three times smaller for almost all examples and earn significant gains (i.e. negative mean hedging errors) across all hedging instruments. While QDH minimizes the RMSE statistic, the downside risk captured by the semi-RMSE, VaR_α and CVaR_α statistics for α equal to 0.95 and 0.99 are always significantly reduced with SQDH policies as compared to QDH policies. Indeed, the downside risk reduction with SQDH over QDH in the latter hedging statistics ranges between 51% to 85% under the BSM and 45% to 76% under the MJD model. These impressive gains in risk reduction can be attributed to the fact that

[13] Risk-neutral prices of the lookback option were estimated with simulations for both dynamics.

Table 3.5: Benchmarking of quadratic deep hedging (QDH) and semi-quadratic deep hedging (SQDH) to hedge the lookback option of $T = 10$ years under the BSM.

Statistics	Mean	RMSE	semi-RMSE	$\text{VaR}_{0.95}$	$\text{VaR}_{0.99}$	$\text{CVaR}_{0.95}$	$\text{CVaR}_{0.99}$	Skew
QDH								
Stock (year)	−0.5	14.8	12.0	25.8	49.8	41.4	69.9	1.9
Stock (month)	0.2	4.9	3.6	7.9	14.5	12.1	19.4	0.4
Two options	0.0	4.2	3.2	6.5	14.0	11.5	20.9	1.9
Six options	0.0	1.1	0.8	1.2	2.8	2.4	5.1	9.9
SQDH								
Stock (year)	−32.1	43.8	4.4	6.4	21.1	16.0	32.7	−1.0
Stock (month)	−10.1	15.0	1.5	2.5	6.1	4.9	9.4	−1.5
Two options	−5.4	10.0	1.6	1.3	5.5	4.1	9.8	−2.3
Six options	−0.9	2.2	0.4	0.2	1.0	0.8	2.0	−5.0

Notes: Hedging statistics under the BSM with $\mu = 0.1, \sigma = 0.15, r = 0.03, S_0^{(0,b)} = 100$ and $V_0^\delta = 17.7$ for all examples (see Section 3.4.2.1 for model description under $\mathbb{P}$ and Section 3.4.2.3 for the risk-neutral dynamics used for option pricing). *Hedging instruments*: monthly and yearly underlying, yearly ATM call and put options (*two options*) or three yearly calls and puts of strikes $K = \{S_{t_n}^{(0,b)}, 1.1S_{t_n}^{(0,b)}, 1.2S_{t_n}^{(0,b)}\}$ and $K = \{S_{t_n}^{(0,b)}, 0.9S_{t_n}^{(0,b)}, 0.8S_{t_n}^{(0,b)}\}$ (*six options*). Results for each penalty are computed based on 75,000 independent paths generated from the BSM under $\mathbb{P}$. Training of the neural networks is done as described in Section 3.4.1.2.

QDH schemes penalize equally upside and downside risk, whereas SQDH procedures strictly penalize hedging losses proportionally to their squared values. Furthermore, hedging statistics also indicate that SQDH policies achieve significant gains under both models and across all hedging instruments with a lesser extend for six options. Hedging with the underlying stock on a yearly basis results in the most expected gains, followed by the stock on a monthly basis, two options and six options. All of these results clearly demonstrate that SQDH policies should be prioritized over QDH policies as they are tailor-made to match the financial objectives of the hedger by always significantly reducing the downside risk as well as earning positive returns on average. Section 3.4.4 that follows will shed some light on specific characteristics of the SQDH policies which result in these large average hedging

Table 3.6: Benchmarking of quadratic deep hedging (QDH) and semi-quadratic deep hedging (SQDH) to hedge the lookback option of $T = 10$ years under the MJD model.

Statistics	Mean	RMSE	semi-RMSE	$\text{VaR}_{0.95}$	$\text{VaR}_{0.99}$	$\text{CVaR}_{0.95}$	$\text{CVaR}_{0.99}$	Skew
QDH								
Stock (year)	−1.6	19.8	15.6	32.3	66.4	54.5	95.4	2.1
Stock (month)	0.2	11.2	9.4	15.7	42.8	32.6	64.6	3.2
Two options	0.0	5.2	3.8	6.7	15.4	12.7	25.1	1.6
Six options	−0.1	1.3	0.9	1.4	3.6	2.9	6.2	2.3
SQDH								
Stock (year)	−35.2	49.7	6.7	11.4	31.7	24.6	47.7	−0.8
Stock (month)	−22.8	33.8	4.2	6.5	18.3	14.3	29.6	−1.1
Two options	−5.9	11.2	1.7	2.2	7.1	5.5	12.2	−2.5
Six options	−1.3	3.1	0.5	0.3	1.4	1.1	2.9	−4.8

Notes: Hedging statistics under the MJD model with $\alpha = 0.1, \sigma = 0.15, \lambda = 0.1, \mu_J = -0.2, \sigma_J = 0.15, \gamma = -1.5, r = 0.03, S_0^{(0,b)} = 100$ and $V_0^{\delta} = 25.3$ for all examples (see Section 3.4.2.2 for model description under $\mathbb{P}$ and Section 3.4.2.4 for the risk-neutral dynamics used for option pricing). *Hedging instruments*: monthly and yearly underlying, yearly ATM call and put options (*two options*) or three yearly calls and puts of strikes $K = \{S_{t_n}^{(0,b)}, 1.1S_{t_n}^{(0,b)}, 1.2S_{t_n}^{(0,b)}\}$ and $K = \{S_{t_n}^{(0,b)}, 0.9S_{t_n}^{(0,b)}, 0.8S_{t_n}^{(0,b)}\}$ (*six options*). Results for each penalty are computed based on 75,000 independent paths generated from the MJD model. Training of the neural networks is done as described in Section 3.4.1.2.

gains and downside risk reduction. Moreover, it is also interesting to note that the distinct treatment of hedging shortfalls by each penalty has a direct implication on the skewness statistic. Indeed, by strictly optimizing squared hedging losses, SQDH effectively minimize the right tail of hedging errors, which entails negative skewness. As for QDH, the positive skewness for all examples can be explained by the fact that the payoff of the lookback option is highly positively asymmetric, since it is bounded below at zero and has no upper bound.

Lastly, Coleman et al. (2007) observed with local risk minimization that while hedging with six options always results in more effective trading strategies, the relative performance of using yearly ATM call and put options (i.e. two options) or the underlying on a monthly basis depends on the dynamics of the risky asset. The same conclusions can be made from the benchmarking results obtained with both global hedging procedures presented in this study. Indeed, hedging statistics of both QDH and SQDH policies under the Black-Scholes dynamics in Table 3.5 show that downside risk metrics are for the most part only slightly better when trading two options as compared to hedging with the underlying stock on a monthly basis. On the other hand, values from Table 3.6 indicate that hedging with two options under the MJD model result in downside risk metrics at least two times smaller than with the stock on a monthly basis for both QDH and SQDH. This observation stems from the fact that hedging jump risk with options is significantly more effective than with the underlying. Thus, hedging statistics reported in this paper show that the observation made by Coleman et al. (2007) with respect to the significant improvement in hedging effectiveness of local risk minimization with options in the presence of jump risk also holds for both QDH and SQDH policies.

3.4.4 Qualitative characteristics of global policies

While the previous section examined the hedging performance of QDH and SQDH with various hedging instruments and different market scenarios, the current section provides novel insights into specific characteristics of the optimized global policies. The analysis starts off by comparing the *average equity risk exposure* of QDH and SQDH policies, also called *average exposure* for convenience, with the same dynamics for the underlying as in previous sections (i.e. BSM and MJD model). The motivation of the latter is to assess if either the MSE or SMSE penalty result in hedging policies more geared towards being long equity risk and are thus earning the equity risk premium. In this paper, the equity risk exposure is measured as the average portfolio delta over one complete path of the financial market. More formally, for $(\delta_{t_{n+1}}^{(0:D)}, \delta_{t_{n+1}}^{(B)})$ given and fixed, the portfolio delta at the beginning of year t_n denoted as $\tilde{\Delta}_{t_n}^{(pf)}$

is defined as

$$\begin{aligned}
\tilde{\Delta}_{t_n}^{(pf)} &:= \frac{\partial V_{t_n}^{\delta}}{\partial S_{t_n}^{(0,b)}} \\
&= \frac{\partial}{\partial S_{t_n}^{(0,b)}} \left(\delta_{t_{n+1}}^{(0:D)} \bullet \bar{S}_{t_n}^{(b)} + \delta_{t_{n+1}}^{(B)} B_{t_n} \right) \\
&= \delta_{t_{n+1}}^{(0)} + \sum_{j=1}^{D} \delta_{t_{n+1}}^{(j)} \tilde{\Delta}^{(j)},
\end{aligned}$$

where $\tilde{\Delta}^{(j)}$ is the j^{th} option delta (i.e. $\tilde{\Delta}^{(j)} = \frac{\partial S_{t_n}^{(j,b)}}{\partial S_{t_n}^{(0,b)}}$ for $j = 1, \ldots, D$). Note that $\tilde{\Delta}^{(j)}$ is time-independent since the calls and puts used for hedging are always of the same characteristics at each trading date (i.e. same moneyness and maturity) and both risky asset models are homoskedastic, which entails that the underlying returns have the same conditional distribution for all time-steps. The $\tilde{\Delta}^{(j)}$ can be computed with the well-known closed form solutions under both models. Average exposure values reported hereafter are computed as the average portfolio delta over $\tilde{N}$ simulated paths:

$$\bar{\Delta}^{(pf)} := \frac{1}{\tilde{N} N} \sum_{k=1}^{\tilde{N}} \sum_{n=0}^{N-1} \tilde{\Delta}_{t_n,k}^{(pf)},$$

where $\tilde{\Delta}_{t_n,k}^{(pf)}$ is the time-t_n portfolio delta of the k^{th} simulated path. Note that all average exposure results presented below are from a test set of $\tilde{N} = 75{,}000$ paths.

3.4.4.1 Average exposure results

Table 3.7 presents average exposures of QDH and SQDH policies with the same market setup as in previous sections with respect to hedging instruments, model parameters and lookback option to hedge. The initial capital investments are again set as the risk-neutral price of the lookback option under each dynamics (i.e. 17.7$ and 25.3$ respectively for BSM and MJD). Numerical results indicate that on average, SQDH policies are significantly more bullish than QDH policies under both dynamics and for all hedging instruments with a lesser extend

Table 3.7: Average equity exposures with quadratic deep hedging (QDH) and semi-quadratic deep hedging (SQDH) for the lookback option of $T = 10$ years under the BSM and MJD model.

	BSM		MJD	
	QDH	SQDH	QDH	SQDH
Stock (year)	-0.10	0.18	-0.14	0.17
Stock (month)	-0.10	-0.01	-0.15	0.07
Two options	-0.12	-0.06	-0.10	-0.04
Six options	-0.12	-0.11	-0.10	-0.08

Notes: Average equity exposures under the BSM and MJD model with $S_0^{(0,b)} = 100$ and $r = 0.03$. Both models dynamics under $\mathbb{P}$ and $\mathbb{Q}$ are described in Section 3.4.2 (see Table 3.1 and Table 3.2 for parameters values). Initial capital investments are respectively of 17.7\$ and 25.3\$ under BSM and MJD. *Hedging instruments*: monthly and yearly underlying, yearly ATM call and put options (*two options*) or three yearly calls and puts of strikes $K = \{S_{t_n}^{(0,b)}, 1.1S_{t_n}^{(0,b)}, 1.2S_{t_n}^{(0,b)}\}$ and $K = \{S_{t_n}^{(0,b)}, 0.9S_{t_n}^{(0,b)}, 0.8S_{t_n}^{(0,b)}\}$ (*six options*). Results for QDH and SQDH are computed based on 75,000 independent paths generated from the BSM and MJD model under $\mathbb{P}$. Training of the neural networks is done as described in Section 3.4.1.2.

when trading six options. This characteristic of SQDH policies to be more geared towards being long equity risk through a larger average exposure is most important when trading the underlying stock on a yearly basis, followed by hedges with the stock on a monthly basis, with two options and with six options. The observation that the average exposure of SQDH policies is only slightly larger than the average exposure of QDH policies when hedging with six options is consistent with benchmarking results presented in previous sections. Indeed, values from Table 3.5 and Table 3.6 show that the absolute difference between the hedging statistics obtained with QDH and SQDH policies is by far the smallest with six options. The latter naturally implies that the hedging positions of both global hedging procedures are on average more similar with six options than with the other hedging instruments, which thus results in relatively closer average equity exposure.

One direct implication of the larger average exposure of SQDH policies as compared to QDH policies is that in the risk management of the lookback option, SQDH should result in positive expected gains. This was in fact observed in the benchmarking of global policies presented in Table 3.5 and Table 3.6, where SQDH resulted in negative mean hedging error statistics (i.e. mean hedging gains) under both risky assets dynamics. It is worth noting that the work of Trottier et al. (2018a) proposed local risk minimization strategies for long-term options which also exhibited positive average returns and important downside risk reduction as compared to delta-hedging strategies through larger equity risk exposures.

3.4.4.2 Analysis of SQDH bullishness

The distinctive feature of SQDH policies to hold a larger average equity exposure than with QDH can firstly be explained by the impact of hedging gains and losses on the optimized policies as measured by each penalty. On the one hand, by minimizing the MSE statistic in a market with positive expected log-returns for the underlying stock as implied by both models parameters values, QDH policies have to be less bullish whenever the hedging portfolio value at maturity is expected to be larger than the lookback option payoff. On the other hand, SQDH policies are strictly penalized for hedging losses proportionally to their squared values, not for hedging gains. The latter entails that SQDH policies are not constrained to reduce their equity risk exposure when the hedging portfolio value is expected to be larger than the lookback option payoff. The second important factor which contributes to SQDH bullishness specifically when hedging is done with the stock is the capacity of deep agents to learn to benefit from *time diversification of risk*. In the context of this study, time diversification of risk refers to the fact that investing in stocks over a long-term horizon reduces the risk of observing large losses as compared to short-term investments. Average exposure values in Table 3.7 indicate that deep agents hedging with the underlying and penalized with the SMSE have learned to hold a larger equity risk exposure than under the MSE penalty to benefit simultaneously from the positive expected returns of the underlying and from the downside risk reduction with time diversification of risk. This observation is most important when trading the underlying stock with yearly rebalancing, where SQDH policies obtained average exposures of 0.18 and 0.17 under respectively the Black-Scholes and the MJD dynamics as compared to -0.10 and -0.14 with QDH.

Moreover, it is very interesting to note that the deep agents rely more on time diversification of risk in the presence of jump risk, i.e. with the MJD dynamics. Indeed, the average exposure difference between SQDH and QDH policies when trading the stock is significantly larger under the MJD dynamics with a difference of 0.31 and 0.22 for yearly and monthly trading as compared to 0.28 and 0.09 under the BSM.[14] The latter observations can be explained by the fact that as shown in Section 3.4.3.2, hedging exclusively with the underlying in the presence of jump risk is inefficient as compared to trading options. Thus, in the presence of jump risk, SQDH agents learn to rely more on time diversification of risk by having on average larger positions in the underlying as compared to SQDH agents trained on a Black-Scholes dynamics. These findings thus provide additional evidence that the deep hedging algorithm is in fact model-free in the sense that the neural networks are able to effectively adapt their trading policies to different stylized facts of risky asset dynamics only by experiencing simulations of the financial market exhibiting these features.

3.5 Conclusion

This paper studies global hedging strategies of long-term financial derivatives with a reinforcement learning approach. A similar financial market setup to the work of Coleman et al. (2007) is considered by studying the impact of equity jump risk on the hedging effectiveness of global procedures for segregated funds GMMBs. In the context of this paper, the latter guarantee is equivalent to holding a short position in a long-term lookback option of fixed maturity. The deep hedging algorithm of Buehler et al. (2019b) is applied to optimize long short-term memory networks representing global hedging policies with the mean-square error (MSE) and semi-mean-square error (SMSE) penalties and with various hedging instruments

[14] For instance, the average exposure difference between SQDH and QDH with the underlying on a yearly basis under the MJD model is $0.17 - (-0.14) = 0.31$.

(e.g. standard options and the underlying). Monte Carlo simulations are performed under the Black-Scholes model (BSM) and the Merton jump-diffusion (MJD) model to compare the hedging effectiveness of quadratic deep hedging (QDH) and semi-quadratic deep hedging (SQDH).

Numerical results demonstrate that under both dynamics and across all trading instruments, SQDH results in hedging policies which simultaneously reduce downside risk and increase expected returns as compared to QDH procedures. The downside risk reduction achieved with SQDH over QDH ranges between 51% to 85% under the BSM and 45% to 76% under the MJD model. Numerical experiments also indicate that QDH outperforms the local risk minimization scheme of Coleman et al. (2007) across all downside risk metrics and all hedging instruments. These results clearly show that SQDH policies should be prioritized over other dynamic hedging schemes (e.g. QDH, local risk minimization and greek-based hedging) as they are tailor-made to match the financial objectives of the hedger by significantly reducing downside risk as well as resulting in large expected positive returns.

Monte Carlo experiments are also done to provide novel qualitative insights into specific characteristics of the optimized long-term global policies. Numerical results show that on average, SQDH policies are significantly more bullish than QDH policies for every example considered. Analysis presented in this paper indicate that the bullishness of SQDH policies stems from the impact of hedging gains and losses on the optimized policies as measured by each penalty. Furthermore, an additional factor which contributes to the larger average equity exposure of SQDH policies when hedging exclusively with the underlying stock is the capacity of deep agents to learn to benefit from time diversification of risk. The latter is shown to be most important in the presence of equity jump risk, where deep agents penalized with the SMSE learned by experiencing many simulations of the financial market to rely more on time diversification risk through larger positions in the underlying as compared to training on the Black-Scholes dynamics due to the lesser efficiency of hedging with the stock in the presence of jumps.

Further research in the area of global hedging for long-term contingent claims with the deep hedging algorithm would prove worthwhile. The analysis of the impact of additional equity risk factors (e.g. volatility risk and regime risk) on the optimized policies would be of interest. The same methodological approach presented in this paper could be applied with the addition of the latter equity risk factors with closed to no modification to the algorithm. Moreover, robustness analysis of the optimized policies when dynamics experienced slightly differ from the ones used to train the neural networks would prove worthwhile. Also, the impact of the inclusion of realistic transaction costs for trading hedging instruments could be examined by following the methodology of the original work of Buehler et al. (2019b).

3.6 Pseudo-code deep hedging

Algorithm 3.1 presents pseudo-code of the training procedure for neural networks formally introduced in Section 3.3.2. More precisely, this pseudo-code presents the procedure for a one-step update of the set of trainable parameters, i.e. from θ_j to θ_{j+1}. For convenience, the approach presented is for the case of trading exclusively the underlying, but it is trivial to generalize for the case of trading shorter-term options. Note that the pseudo-code is applicable under both the BSM and MJD dynamics (see step 6 below). Also, subscript i represents the i^{th} simulated path among the mini-batch of size N_{batch}. Lastly, recall that a GitHub repository with samples of codes in Python for the training procedure of neural networks is available online: github.com/alexandrecarbonneau. The implementation replicates Table 3.5 and Table 3.7 under the BSM, and can easily be adapted to reproduce all results presented in Section 3.4.

Algorithm 3.1 Pseudo-code deep hedging when trading underlying stock
Input: θ_j
Output: θ_{j+1}

for $i = 1, \ldots, N_{\text{batch}}$ **do** ▷ Loop over each path of mini-batch
  $X_{t_0,i} = [\log(S^{(0,b)}_{0,i}), \log(Z_{0,i}), V^{\delta}_{0,i}/V^{\delta}_{0,i}]$ ▷ Time-0 feature vector of F_θ
  for $n = 0, \ldots, N-1$ **do**
    $Y_{t_n,i} \leftarrow$ time-t_n output of LSTM F_θ with $\theta = \theta_j$
    $\delta^{(0)}_{t_{n+1},i} = Y_{t_n,i}$
    $y_{t_{n+1},i} \sim$ (3.14) or (3.15) ▷ Sample next log-return
    $S^{(0,b)}_{t_{n+1},i} = S^{(0,b)}_{t_n,i} e^{y_{t_{n+1},i}}$
    if $\lfloor t_{n+1} \rfloor = m$ and $m \in \{0, \ldots, T-1\}$ **then**
      $Z_{t_{n+1},i} = \max(S^{(0,b)}_{0,i}, \ldots, S^{(0,b)}_{m,i})$
    else if $t_{n+1} = T$ **then**
      $Z_{t_{n+1},i} = \max(S^{(0,b)}_{0,i}, \ldots, S^{(0,b)}_{T-1,i})$
    end if
    $V^{\delta}_{t_{n+1},i} = e^{r\Delta_N} V^{\delta}_{t_n,i} + \delta^{(0)}_{t_{n+1},i}(S^{(0,b)}_{t_{n+1},i} - e^{r\Delta_N} S^{(0,b)}_{t_n,i})$ ▷ See (3.10) for details
    $X_{t_{n+1},i} = [\log(S^{(0,b)}_{t_{n+1},i}), \log(Z_{t_{n+1},i}), V^{\delta}_{t_{n+1},i}/V^{\delta}_0]$ ▷ Time-t_{n+1} feature vector of F_θ
  end for
  $\Phi(S^{(0,b)}_{T,i}, Z_{T,i}) = \max(Z_{T,i} - S^{(0,b)}_{T,i}, 0)$
  $\pi_{i,j} = \Phi(S^{(0,b)}_{T,i}, Z_{T,i}) - V^{\delta}_{T,i}$
end for
if $\mathcal{L}$ is MSE **then**
  $\hat{J} = \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \pi^2_{i,j}$
else if $\mathcal{L}$ is SMSE **then**
  $\hat{J} = \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \pi^2_{i,j} \mathbb{1}_{\{\pi_{i,j}>0\}}$
end if
$\eta_j \leftarrow$ Adam algorithm
$\theta_{j+1} = \theta_j - \eta_j \nabla_\theta \hat{J}$ ▷ $\nabla_\theta \hat{J}$ computed with Tensorflow

Notes: Subscript i represents the i^{th} simulated path among the mini-batch of size N_{batch}. Also, the time-0 feature vector is fixed for all paths, i.e. $S^{(0,b)}_{0,i} = S^{(0,b)}_0, Z^{(0,b)}_{0,i} = Z^{(0,b)}_0$ and $V^{\delta}_{0,i} = V^{\delta}_0$.

Chapter 4

Deep Equal Risk Pricing of Financial Derivatives with Multiple Hedging Instruments

Abstract

This paper studies the equal risk pricing (ERP) framework for the valuation of European financial derivatives. This option pricing approach is consistent with global trading strategies by setting the premium as the value such that the residual hedging risk of the long and short positions in the option are equal under optimal hedging. The ERP setup of Marzban et al. (2020) is considered where residual hedging risk is quantified with convex risk measures. The main objective of this paper is to assess through extensive numerical experiments the impact of including options as hedging instruments within the ERP framework. The reinforcement learning procedure developed in Carbonneau and Godin (2021b), which relies on the deep hedging algorithm of Buehler et al. (2019b), is applied to numerically solve the global hedging problems by representing trading policies with neural networks. Among other findings, numerical results indicate that in the presence of jump risk, hedging long-term puts with shorter-term options entails a significant decrease of both equal risk prices and market incompleteness as compared to trading only the stock. Monte Carlo experiments demonstrate the potential of ERP as a fair valuation approach providing prices consistent with observable market prices. Analyses exhibit the ability of ERP to span a large interval of prices through the choice of convex risk measures which is close to encompass the variance-optimal premium.

Keywords: Equal risk pricing, Deep hedging, Convex risk measure, Reinforcement learning.

4.1 Introduction

In the famous setup of Black and Scholes (1973) and Merton (1973), every contingent claim can be perfectly replicated through continuous trading in the underlying stock and a risk-free asset. These markets are said to be *complete*, and derivatives are redundant securities with a unique arbitrage-free price equal to the initial value of the replicating portfolio. However, the gigantic size of the derivatives market demonstrates unequivalently that options are non-redundant and provide additional value above the exclusive trading of the underlying asset from the standpoint of speculation, risk management and arbitraging (Hull, 2003). Such value-added of derivatives in the real world stems from market incompleteness which arises from several stylized features of market dynamics such as discrete-time trading, equity risk (e.g. jump and volatility risks) and market impact (e.g. trading costs and imperfect liquidity). In contrast to the complete market paradigm, in incomplete markets, the price of a derivative cannot be uniquely specified by a no-arbitrage argument since perfect replication is not always possible.

The problem of determining the value of a derivative is intrinsically intertwined with its corresponding hedging strategy. On the spectrum of derivative valuation procedures in incomplete markets, one extreme possibility is the so-called *super-hedging strategy*, where the derivative premium is set as the value such that the residual hedging risk of the seller is nullified. However, the super-hedging premium is in general very large and is thus most often deemed impractical (Gushchin and Mordecki, 2002). On the other hand, a more reasonable and practical derivative premium entails that some level of risk cannot be hedged away and is thus intrinsic to the contingent claim. An additional layer of complexity to the hedging problem in incomplete markets is in selecting not only the sequence of investments in trading instruments, but also the *category* of hedging instruments in the design of optimal hedges. Indeed, some categories of instruments are more effective to mitigate certain risk factors than others. For instance, it is well-known that in the presence of random jumps, option hedges are much more effective than trading exclusively the underlying stock due to the convex property of derivatives prices (see, for instance, Coleman et al. (2007) and Carbonneau (2021)). More generally, the use of option hedges dampens tail risk stemming from different risk factors (e.g. jump and volatility risks). The focus of this paper lies precisely on studying a derivative valuation approach called *equal risk pricing* (ERP) for pricing European derivatives consistently with optimal hedging strategies trading in various categories of hedging instruments (e.g. vanilla calls and puts as well as the underlying stock).

The ERP framework introduced by Guo and Zhu (2017) determines the *equal risk price* (i.e. the premium) of a financial derivative as the value such that the long and short positions in the contingent claim have the same residual hedging risk under optimal trading strategies. An important application of ERP in the latter paper is for pricing derivatives in the presence of short-selling restrictions for the underlying stock. Various studies have since extended this approach: Ma et al. (2019) provide Hamilton-Jacobi-Bellman equations for the optimization problem and establish additional analytical pricing formulas for equal risk prices, He and Zhu (2020) generalize the problem of pricing derivatives with short-selling restriction for the underlying by allowing for short trades in a correlated asset and Alfeus et al. (2019) perform an empirical study of equal risk prices when short selling is banned. One crucial pitfall of the Guo and Zhu (2017) framework considered in all of the aforementioned papers is that the optimization problem required to be solved for the computation of equal risk prices is very complex. Consequently, closed-form solutions are restricted to very specific setups (e.g. Black-Scholes market) and no numerical scheme has been proposed to account for more realistic market assumptions.

Marzban et al. (2020) recently extended the ERP framework by considering the use of convex risk measures under the physical measure to quantify residual hedging risk. A major benefit of the ERP setup of the latter paper is that it does not require the specification of an equivalent martingale measure (EMM), which is arbitrary in incomplete markets since there is an infinite set of EMMs (Harrison and Pliska, 1981). Also, using convex measures to quantify residual risk is shown in Marzban et al. (2020) to significantly reduce the complexity of computing equal risk prices; the optimization problem essentially boils down to solving two distinct non-quadratic global hedging problems, one for the long and one for the short position in the option. Dynamic programming equations are provided in Marzban et al. (2020) for the aforementioned global hedging problems. However, it is well-known that traditional dynamics programming procedures are prone to the curse of dimensionality when the state and action spaces gets too large (Powell, 2009). The main objective of this current paper consists in studying the impact of trading different and possibly multiple hedging instruments on the ERP framework, which thus necessitates large action spaces. Furthermore, a specific focus of this study is on assessing the interplay between different equity risk factors (e.g. jump and volatility risks) and the use of options as hedging instruments. Consequently, large state spaces are also required to model the dynamics of the underlying stock and to characterize the physical measure dynamics of the implied volatility of options used as hedging instruments. A feasible numerical procedure in high-dimensional state and action spaces is therefore essential to this paper.

Carbonneau and Godin (2021b) expanded upon the work of Marzban et al. (2020) by developing a tractable solution with reinforcement learning to compute equal risk prices in high-dimensional state and action spaces. The approach of the foremost study relies on the deep hedging algorithm of Buehler et al. (2019b) to represent the long and short optimal trading policy with two distinct neural networks. One of the most important benefits of parameterizing trading policies as neural networks is that the computational complexity increases marginally with the dimension of the state and action spaces. Carbonneau and Godin (2021b) also introduce novel ϵ-completeness metrics to quantify the level of market incompleteness which will be used throughout this current study. Several papers have studied different aspects of the class of deep hedging algorithms: Buehler et al. (2019a) extend upon the work of Buehler et al. (2019b) by hedging path-dependent contingent claims with neural networks, Carbonneau (2021) presents an extensive benchmarking of global policies parameterized with neural networks to mitigate the risk exposure of very long-term contingent claims, Cao et al. (2020) show that the deep hedging algorithm provides good approximations of optimal initial capital investments for variance-optimal hedging problems and Horvath et al. (2021) deep hedge in a non-Markovian framework with rough volatility models for risky assets.

The main objective of this paper consist in the assessment of the impact of using multiple hedging instruments on the ERP framework through exhaustive numerical experiments. To the best of the authors' knowledge, this is the first study within the ERP literature that considers trades involving options in the design of optimal hedges. The performance of these numerical experiments heavily relies on the use of reinforcement learning procedures to train neural networks representing trading policies and would be hardly reachable with other numerical methods. The first key contribution of this paper consists in providing a broad analysis of the impact of jump and volatility risks on equal risk prices and on our ϵ-completeness metrics. These assessments expand upon the work of Carbonneau and Godin (2021b) in two ways. First, the latter paper conducted sensitivity analyses of the ERP framework under different risky assets dynamics by trading exclusively with the underlying stock, not with options. However, the use of options as hedging instruments in the presence of such risk factors allows for the mitigation of some portion of unattainable residual risk when trading exclusively with the stock. Second, this current paper examines the sensitivity of equal risk prices and residual hedging risk to different levels of jump and volatility risks through a range of empirically plausible model parameters for asset prices dynamics (e.g. frequent small jumps and rare extreme jumps). The motivation is to provide new qualitative insights into the interrelation of different stylized features of jump and volatility risks on the ERP framework that are more extensive than in previous work studies. The main conclusions of these experiments of pricing 1-year European puts are summarized below.

1) In the presence of downward jump risk, numerical values indicate that hedging with options entails significant reduction of both equal risk prices and on the level of

market incompleteness as compared to hedging solely with the underlying stock. The latter stems from the fact that while the residual hedging risk of both the long and short positions in 1-year puts decreases when short-term option trades are used for mitigating the presence of jump risk, a larger decrease is observed for the short position due to jump risk dynamics entailing predominantly negative jumps. These results further demonstrate that options are non-redundant securities as it is the case in the Black-Scholes world.

2) In the presence of volatility risk, numerical experiments demonstrate that while the use of options as hedging instruments can entail smaller derivative premiums, the impact can also be marginal and is highly sensitive to the moneyness level of the put option being priced as well as to the maturity of the traded options. This observation stems from the fact that contrarily to jump risk, volatility risk impacts both upside and downside risk. Thus, the use of option hedges does not necessarily benefit more the short position with a larger decrease of residual hedging risk as observed in the presence of jump risk.

3) The average price level of short-term options (i.e. average implied volatility level) used as hedging instruments is effectively reflected into the equal risk price of longer-term options. This demonstrates the potential of the ERP framework as a fair valuation approach consistent with observable market prices, which could be used, for instance, to price over-the-counter or long-term less liquid derivatives with short-term highly liquid options.

The last contribution of this paper is in benchmarking equal risk prices to derivative premiums obtained with *variance-optimal hedging* (Schweizer, 1995). Variance-optimal hedging procedures solve jointly for the initial capital investment and a self-financing strategy minimizing the expected squared hedging error. The optimized initial capital investment can be viewed as the production cost of the derivative, since the resulting dynamic trading strategy replicates the derivative's payoff as closely as possible in a quadratic sense.[1] The main motivation for these experiments is the popularity of variance-optimal hedging procedures in the literature for pricing derivatives. Furthermore, while these two derivative valuation procedures are both consistent with optimal trading criteria, the underlying global hedging problem of each approach treats hedging shortfall through a radically different scope. Indeed, equal risk prices obtained under the Conditional Value-at-Risk measure with large confidence level values, as considered in this paper, are the result of joint optimizations over hedging decisions to minimize tail risk of hedging shortfalls which penalize mainly (and most often exclusively) hedging losses, not gains. Conversely, variance-optimal procedures penalize equally hedging gains and losses, not solely losses. This benchmarking of equal risk prices to variance-optimal premiums highlights the flexibility of ERP procedures for derivatives valuation through the choice of convex risk measure. Indeed, numerical values show that the range of equal risk prices obtained with several convex measures can be very large and is close to encompass the variance-optimal premium.

The rest of the paper is as follows. Section 4.2 details the equal risk pricing framework considered in this study. Section 4.3 presents the numerical scheme to solve the optimization problem with the use of neural networks. Section 4.4 performs extensive numerical experiments studying the equal risk pricing framework. Section 4.5 concludes.

4.2 Equal risk pricing framework

This section details the equal risk pricing (ERP) framework considered in this paper, which is an extension of the derivative valuation scheme introduced in Marzban et al. (2020) with the addition of multiple hedging instruments.

[1] Note that derivatives premiums prescribed by variance-optimal procedures coincide with the risk-neutral price obtained under the so-called *variance-optimal martingale measure* (Schweizer, 1996).

4.2.1 Specification of the financial market

The financial market is in discrete-time with a finite time horizon of T years and $N+1$ observation dates characterized by the set $\mathcal{T} := \{t_n : t_n = n\Delta_N, n = 0, \dots, N\}$ where $\Delta_N := T/N$. The probability space $(\Omega, \mathbb{P}, \mathcal{F})$ is equipped with the filtration $\mathbb{F} := \{\mathcal{F}_n\}_{n=0}^N$ satisfying the usual conditions, where $\mathcal{F}_n$ contains all information available to market participants at time t_n. Assume $\mathcal{F} = \mathcal{F}_N$. $\mathbb{P}$ is referred to as the physical probability measure. On each observation date, a total of $D+2$ financial securities can be traded on the market, which includes a risk-free asset, a non-dividend paying stock and D standard European calls and puts on the latter stock whose maturity dates fall within $\mathcal{T}$. Let $\{B_n\}_{n=0}^N$ be the price process of the risk-free asset, where $B_n := e^{rt_n}$ for $n = 0, \dots, N$ with $r \in \mathbb{R}$ being the annualized continuously compounded risk-free rate. The definition of the price process for the risky securities is now outlined. Since some of the tradable options can mature before the final time horizon T, the set of options that can be traded at the beginning of two different observation periods could differ. To reflect this modeling feature and properly represent gains of trading strategies, two different stochastic processes are defined, namely the price of tradable assets at the beginning and at the end of each period. First, let $\{\bar{S}_n^{(b)}\}_{n=0}^N$ be the *beginning-of-period risky price process* whose element $\bar{S}_n^{(b)}$ contains the time-t_n price of all risky assets traded at time t_n. More precisely, $\bar{S}_n^{(b)} := [S_n^{(0,b)}, \dots, S_n^{(D,b)}]$ with $S_n^{(0,b)}$ and $S_n^{(j,b)}$ respectively being the time-t_n price of the underlying stock and of the j^{th} option that can be traded at time t_n for $j = 1, \dots, D$. Similarly, let $\{\bar{S}_n^{(e)}\}_{n=0}^{N-1}$ be the *end-of-period risky price process* where $\bar{S}_n^{(e)} := [S_n^{(0,e)}, \dots, S_n^{(D,e)}]$ with $S_n^{(0,e)}$ and $S_n^{(j,e)}$ respectively being the time t_{n+1} price of the underlying stock and j^{th} option that can be traded at time t_n. Since the underlying asset is denoted as the risky asset with index 0, $S_n^{(0,e)} = S_{n+1}^{(0,b)}$ for $n = 0, \dots, N-1$. Also, if the j^{th} option that can be traded at t_n matures at time t_{n+1}, then $S_n^{(j,e)}$ is the payoff of that option. In that case, $S_{n+1}^{(j,b)}$ is the price of a new contract with the same characteristics in terms of payoff function, moneyness level and time-to-maturity. Otherwise, $S_{n+1}^{(j,b)} = S_n^{(j,e)}$ holds for all time steps and all risky assets (i.e. for $j = 0, \dots, D$ and $n = 0, \dots, N-1$). An implicit assumption stemming from the latter equality is that trading in risky assets does not impact their prices. Moreover, for convenience, it is assumed throughout the current work that only options with a single-period time-to-maturity are traded, i.e. options are traded once and held until expiry.[2]

This paper studies the problem of pricing a simple European-type derivative providing a time-T payoff denoted by $\Phi(S_N^{(0,b)}) \geq 0$.[3] For such purposes, the equal risk pricing scheme (Guo and Zhu, 2017) is considered, which entails optimizing two distinct self-financing dynamic trading strategies separately for both the long and short positions on the derivative, and then determining the premium which equates the residual hedging risk of the two hedged positions. The mathematical formalism used for trading strategies in the current study is now outlined. A *trading strategy* $\{\delta_n\}_{n=0}^N$ is an $\mathbb{F}$-predictable process[4] where $\delta_n := [\delta_n^{(0)}, \ldots, \delta_n^{(D)}, \delta_n^{(B)}]$ with $\delta_n^{(B)}$ and $\delta_n^{(j)}$, $j = 0, \ldots, D$, respectively denoting the number of shares of the risk-free asset and the j^{th} risky asset traded at time t_{n-1} held in the hedging portfolio throughout the period $(t_{n-1}, t_n]$, except for the case $n = 0$ which represents the hedging portfolio composition exactly at time t_0. The notation $\delta_n^{(0:D)} := [\delta_n^{(0)}, \ldots, \delta_n^{(D)}]$ is used to define the vector containing exclusively positions in the risky assets. Furthermore, the initial capital investment of the trading strategy is always assumed to be completely invested in the risk-free asset, i.e. $\delta_0^{(B)}$ is the initial investment amount and $\delta_0^{(0:D)} := [0, \ldots, 0]$.

In this work, the trading strategies considered to hedge Φ are obtained through a joint optimization over all trading decisions to minimize global risk exposure. Before formally describing the optimization problem, some well-known prerequisites from the mathematical finance literature are now provided; the reader is referred to Lamberton and Lapeyre (2011)

[2] Note that the optimization procedure for global policies described in Section 4.3 can naturally be generalized for the case of rebalancing multiple times option contracts prior to their expiry.

[3] The derivative valuation approach presented in this paper can easily be adapted for European options whose payoff is of the form $\Phi(S_N^{(0,b)}, Z_N) \geq 0$ with $\{Z_n\}_{n=0}^N$ as some $\mathbb{F}$-adapted potentially multidimensional random process encompassing the path-dependence property of the payoff function. For examples of such exotic derivatives, the reader is referred to Carbonneau and Godin (2021b).

[4] A process $X = \{X_n\}_{n=0}^N$ is said to be $\mathbb{F}$-predictable if X_0 is $\mathcal{F}_0$-measurable and X_n is $\mathcal{F}_{n-1}$-measurable for $n = 1, \ldots, N$.

for additional details. Let $\{V_n^\delta\}_{n=0}^N$ be the hedging portfolio value process associated with the trading strategy δ, where V_n^δ is the time-t_n portfolio value prior to rebalancing with $V_0^\delta := \delta_0^{(B)}$ and

$$V_n^\delta := \delta_n^{(0:D)} \bullet \bar{S}_{n-1}^{(e)} + \delta_n^{(B)} B_n, \quad n = 1, \ldots, N, \tag{4.1}$$

where $\bullet$ is the scalar product operator, i.e. for two n-dimensional vectors X and Y, $X \bullet Y := \sum_{i=1}^n X_i Y_i$. Furthermore, denote as $\{G_n^\delta\}_{n=0}^N$ the discounted gain process associated with δ where G_n^δ is the time-t_n discounted gain prior to rebalancing with $G_0^\delta := 0$ and

$$G_n^\delta := \sum_{k=1}^n \delta_k^{(0:D)} \bullet (B_k^{-1} \bar{S}_{k-1}^{(e)} - B_{k-1}^{-1} \bar{S}_{k-1}^{(b)}), \quad n = 1, \ldots, N. \tag{4.2}$$

The trading strategies considered in this paper are always *self-financing*: they require no cash infusion nor withdrawal at intermediate times except possibly at the initialization of the strategy. More formally, a trading strategy is said to be self-financing if it is predictable and if the following equality holds $\mathbb{P}$-a.s. for $n = 0, \ldots, N-1$:

$$\delta_{n+1}^{(0:D)} \bullet \bar{S}_n^{(b)} + \delta_{n+1}^{(B)} B_n = V_n^\delta. \tag{4.3}$$

Lastly, denote Π as the set of accessible trading strategies, which includes all trading strategies that are self-financing and sufficiently well-behaved.

Remark 4.1. *It can be shown that $\delta \in \Pi$ is self-financing if and only if $V_n^\delta = B_n(V_0^\delta + G_n^\delta)$ holds $\mathbb{P}$-a.s. for $n = 0, \ldots, N$; see for instance Lamberton and Lapeyre (2011). The latter representation of portfolio values implies the following useful recursive equation (4.4) to* *compute* V_n^δ *for* $n = 1, \ldots, N$ *given* V_0^δ*:*

$$\begin{aligned}
V_n^\delta &= B_n(V_0^\delta + G_n^\delta) \\
&= B_n(V_0^\delta + G_{n-1}^\delta + \delta_n^{(0:D)} \bullet (B_n^{-1}\bar{S}_{n-1}^{(e)} - B_{n-1}^{-1}\bar{S}_{n-1}^{(b)})) \\
&= \frac{B_n}{B_{n-1}} V_{n-1}^\delta + \delta_n^{(0:D)} \bullet (\bar{S}_{n-1}^{(e)} - \frac{B_n}{B_{n-1}}\bar{S}_{n-1}^{(b)}) \\
&= e^{r\Delta_N} V_{n-1}^\delta + \delta_n^{(0:D)} \bullet (\bar{S}_{n-1}^{(e)} - e^{r\Delta_N}\bar{S}_{n-1}^{(b)}).
\end{aligned} \tag{4.4}$$

4.2.2 Equal risk pricing framework

The financial market setting considered in this paper implies incompleteness stemming from discrete-time trading and equity risk factors (e.g. jump risk and volatility risk). For the hedger, these many sources of incompleteness entail that most contingent claims are not attainable through dynamic hedging. Following the work of Marzban et al. (2020) and Carbonneau and Godin (2021b), this study quantifies the level of residual hedging risk with convex risk measures as defined in Föllmer and Schied (2002).

Definition 4.1. *(Convex risk measure) For a set of random variables* $\mathcal{X}$ *representing liabilities and* $X_1, X_2 \in \mathcal{X}$*,* $\rho : \mathcal{X} \to \mathbb{R}$ *is a convex risk measure if it satisfies the following properties:*

1) *Monotonicity:* $X_1 \leq X_2 \implies \rho(X_1) \leq \rho(X_2)$ *(larger liability is riskier).*

2) *Translation invariance: for* $c \in \mathbb{R}$ *and* $X \in \mathcal{X}$*,* $\rho(X + c) = \rho(X) + c$ *(borrowing amount* c *increases the risk by that amount).*

3) *Convexity: for* $c \in [0, 1]$*,* $\rho(cX_1 + (1 - c)X_2) \leq c\rho(X_1) + (1 - c)\rho(X_2)$ *(diversification does not increase risk).*

The hedging problem underlying the ERP framework is now formally defined.

Definition 4.2. *(Long- and short-sided risk) For a given convex risk measure* ρ*, define* $\epsilon^{(\mathcal{L})}(V_0)$ *and* $\epsilon^{(\mathcal{S})}(V_0)$ *respectively as the measured risk exposure of a long and short position* *in* Φ *under the optimal hedge if the value of the initial hedging portfolio is* $V_0 \in \mathbb{R}$*:*

$$\epsilon^{(\mathcal{L})}(V_0) := \min_{\delta \in \Pi} \rho \left(-\Phi(S_N^{(0,b)}) - B_N(V_0 + G_N^{\delta}) \right), \tag{4.5}$$

$$\epsilon^{(\mathcal{S})}(V_0) := \min_{\delta \in \Pi} \rho \left(\Phi(S_N^{(0,b)}) - B_N(V_0 + G_N^{\delta}) \right). \tag{4.6}$$

Note that the same risk measure ρ is used for both the long and short positions global hedging problems. The rationale for this choice is threefold. First, considering the same convex risk measure for long and short positions is in line with the trading activities of some market participants that both buy and sell options with no directional view of the market. One example of such participant is a market maker of derivatives which typically expects to make a profit on bid-ask spreads, not by speculating (Basak and Chabakauri, 2012). Another motivation for using the same convex measure is for cases where a price quote must be given prior to knowing if the derivative is being purchased or sold. For instance, a client asks his broker to provide a quote for a derivative without revealing his intention of buying or selling the option. A similar argument is made in Bertsimas et al. (2001) to motivate the use of a quadratic loss function for hedging shortfalls, which entails the same derivative price for the long and short position. Lastly, as shown in Carbonneau and Godin (2021b), using the same risk measure for both positions guarantees, under some specific conditions, that the ERP derivative premium is arbitrage-free.[5]

It is interesting to note that the translation invariance property of ρ entails that the optimal strategies solving (4.5)-(4.6), denoted respectively by $\delta^{(\mathcal{L})}$ and $\delta^{(\mathcal{S})}$, are invariant to the initial capital investment amount V_0. The latter significantly enhances the tractability of the solution:

[5] Nevertheless, the authors want to emphasize that the numerical scheme developed in Section 4.3 for the global hedging problems (4.5) and (4.6) could easily be extended to include two distinct convex measures respectively for the long and short position hedges (see Remark 3.4 of Carbonneau and Godin (2021b) for additional details).

$$\delta^{(\mathcal{L})} := \underset{\delta\in\Pi}{\arg\min}\, \rho\left(-\Phi(S_N^{(0,b)}) - B_N(V_0 + G_N^{\delta})\right) = \underset{\delta\in\Pi}{\arg\min}\, \rho\left(-\Phi(S_N^{(0,b)}) - B_N G_N^{\delta}\right), \quad (4.7)$$

$$\delta^{(\mathcal{S})} := \underset{\delta\in\Pi}{\arg\min}\, \rho\left(\Phi(S_N^{(0,b)}) - B_N(V_0 + G_N^{\delta})\right) = \underset{\delta\in\Pi}{\arg\min}\, \rho\left(\Phi(S_N^{(0,b)}) - B_N G_N^{\delta}\right). \quad (4.8)$$

Based on the aforementioned global hedging problems, the *equal risk price* of a derivative is defined as the initial hedging portfolio value equating the measured risk exposures for both the long and short positions.

Definition 4.3. *(Equal risk price) The equal risk price $C_0^{\star}$ of Φ is defined as the real number C_0 such that*

$$\epsilon^{(\mathcal{L})}(-C_0) = \epsilon^{(\mathcal{S})}(C_0). \quad (4.9)$$

As shown for instance in Marzban et al. (2020), equal risk prices have the following representation which is used throughout the rest of the paper:

$$C_0^{\star} = \frac{\epsilon^{(\mathcal{S})}(0) - \epsilon^{(\mathcal{L})}(0)}{2B_N}. \quad (4.10)$$

Carbonneau and Godin (2021b) introduced the market incompleteness metric $\epsilon^{\star}$ defined as the level of residual risk faced by the hedgers of Φ if the hedged derivative price is set to $C_0^{\star}$ and optimal trading strategies are used by both the long and short position hedgers:[6]

$$\epsilon^{\star} := \epsilon^{(\mathcal{L})}(-C_0^{\star}) = \epsilon^{(\mathcal{S})}(C_0^{\star}) = \frac{\epsilon^{(\mathcal{L})}(0) + \epsilon^{(\mathcal{S})}(0)}{2}. \quad (4.11)$$

Consistently with the terminology of Carbonneau and Godin (2021b), $\epsilon^{\star}$ and $\epsilon^{\star}/C_0^{\star}$ are referred respectively as the measured residual risk exposure per derivative contract and

[6] The last equality of (4.11) can easily be obtained with the translation invariance property of ρ, see equation (8) of Carbonneau and Godin (2021b) for the details.

per dollar invested. These $\epsilon^{\star}$-metrics will be extensively studied in numerical experiments conducted in Section 4.4 to assess, for instance, the impact of the use of options as hedging instruments on the level of market incompleteness.

4.3 Deep equal risk pricing

The problem of solving the ERP framework, that is evaluating equal risk prices and ϵ-completeness measures, boils down to the computation of the measured risk exposures $\epsilon^{(\mathcal{S})}(0)$ and $\epsilon^{(\mathcal{L})}(0)$. This section presents a reinforcement learning method to compute such quantities. The approach was first proposed in Carbonneau and Godin (2021b) and relies on approximating optimal trading strategies with the *deep hedging* algorithm of Buehler et al. (2019b) through the representation of the long and short global trading policy with two distinct neural networks. In its essence, neural networks are a class of composite functions mapping *feature vectors* (i.e. input vectors) to *output vectors* through multiple *hidden layers*, with the latter being functions applying successive affine and nonlinear transformations to input vectors. In this paper, the type of neural network considered to represent global hedging policies is the *long short-term memory* (LSTM, Hochreiter and Schmidhuber (1997)). LSTMs belong to the class of *recurrent neural networks* (RNNs, Rumelhart et al. (1986)), which have self-connections in hidden layers: the output of the time-t_n hidden layer is a function of both the time-t_n feature vector as well as the output of the time-t_{n-1} hidden layer. The periodic computation of long short-term memory neural networks is done with so-called *LSTM cells*, which are similar to but more complex than the typical hidden layer of RNNs. LSTMs have recently been applied with success to approximate global hedging policies in several studies: Buehler et al. (2019a), Cao et al. (2020) and Carbonneau (2021). Additional remarks are made in subsequent sections to motivate this choice of neural networks for the specific setup of this paper.

4.3.1 Neural networks representing trading policies

The following formally defines the architecture of long-short term memory neural networks. For convenience, a very similar notation for neural networks as the one of Carbonneau (2021) is used. Note that the time steps of the feature and output vectors coincide with the set of financial market trading dates $\mathcal{T}$. For additional general information about LSTMs, the reader is referred to Chapter 10.10 of Goodfellow et al. (2016) and the many references therein.

Definition 4.4. *(LSTM) For $H, d_0, \ldots, d_{H+1} \in \mathbb{N}$, let $F_\theta : \mathbb{R}^{N \times d_0} \to \mathbb{R}^{N \times d_{H+1}}$ be an LSTM which maps the sequence of feature vectors $\{X_n\}_{n=0}^{N-1}$ to output vectors $\{Y_n\}_{n=0}^{N-1}$ where $X_n \in \mathbb{R}^{d_0}$ and $Y_n \in \mathbb{R}^{d_{H+1}}$ for $n = 0, \ldots, N-1$. The computation of Y_n, the subset of outputs of F_θ associated with time t_n, is achieved through H LSTM cells, each of which outputs a vector of d_j neurons denoted as $h_n^{(j)} \in \mathbb{R}^{d_j \times 1}$ for $j = 1, \ldots, H$. More precisely, the computation applied by the j^{th} LSTM cell for the time-t_n output is as follows:*[7]

$$\begin{aligned}
i_n^{(j)} &= sigm(U_i^{(j)} h_n^{(j-1)} + W_i^{(j)} h_{n-1}^{(j)} + b_i^{(j)}), \\
f_n^{(j)} &= sigm(U_f^{(j)} h_n^{(j-1)} + W_f^{(j)} h_{n-1}^{(j)} + b_f^{(j)}), \\
o_n^{(j)} &= sigm(U_o^{(j)} h_n^{(j-1)} + W_o^{(j)} h_{n-1}^{(j)} + b_o^{(j)}), \\
c_n^{(j)} &= f_n^{(j)} \odot c_{n-1}^{(j)} + i_n^{(j)} \odot tanh(U_c^{(j)} h_n^{(j-1)} + W_c^{(j)} h_{n-1}^{(j)} + b_c^{(j)}), \\
h_n^{(j)} &= o_n^{(j)} \odot tanh(c_n^{(j)}),
\end{aligned} \tag{4.12}$$

where $\odot$ denotes the Hadamard product (the element-wise product), $sigm(\cdot)$ and $tanh(\cdot)$ are the sigmoid and hyperbolic tangent functions applied element-wise to each scalar given as input[8] *and*

- $U_i^{(j)}, U_f^{(j)}, U_o^{(j)}, U_c^{(j)} \in \mathbb{R}^{d_j \times d_{j-1}}$, $W_i^{(j)}, W_f^{(j)}, W_o^{(j)}, W_c^{(j)} \in \mathbb{R}^{d_j \times d_j}$ *and* $b_i^{(j)}, b_f^{(j)}, b_o^{(j)}, b_c^{(j)} \in$

[7] At time 0 (i.e. $n = 0$), the computation of the LSTM cells is the same as in (4.12) with $h_{-1}^{(j)}$ and $c_{-1}^{(j)}$ defined as vectors of zeros of dimension d_j for $j = 1, \ldots, H$.

[8] For $X := [X_1, \ldots, X_K]$, $\text{sigm}(X) := \left[\frac{1}{1+e^{-X_1}}, \ldots, \frac{1}{1+e^{-X_K}}\right]$ and $\tanh(X) := \left[\frac{e^{X_1}-e^{-X_1}}{e^{X_1}+e^{-X_1}}, \ldots, \frac{e^{X_K}-e^{-X_K}}{e^{X_K}+e^{-X_K}}\right]$.

$\mathbb{R}^{d_j \times 1}$ *for* $j = 1, \ldots, H$.

At each time-step, the input of the first LSTM cell is the feature vector (i.e. $h_n^{(0)} := X_n$*) and the final output is an affine transformation of the output of the last LSTM cell:*

$$Y_n = W_y h_n^{(H)} + b_y, \quad n = 0, \ldots, N-1, \tag{4.13}$$

where $W_y \in \mathbb{R}^{d_{H+1} \times d_H}$ *and* $b_y \in \mathbb{R}^{d_{H+1} \times 1}$*. Lastly, the set of trainable parameters denoted as* θ *consists of all weight matrices and bias vectors:*

$$\theta := \left\{ \{U_i^{(j)}, U_f^{(j)}, U_o^{(j)}, U_c^{(j)}, W_i^{(j)}, W_f^{(j)}, W_o^{(j)}, W_c^{(j)}, b_i^{(j)}, b_f^{(j)}, b_o^{(j)}, b_c^{(j)}\}_{j=1}^{H}, W_y, b_y \right\}. \tag{4.14}$$

In this study, the computation of hedging positions is done through the mapping of a sequence of relevant financial market observations into the periodic number of shares held in each hedging instrument with an LSTM. One of the main objectives of this paper is to analyze the impact of including vanilla options as hedging instruments on the ERP framework. For the numerical experiments conducted in the subsequent Section 4.4, the hedging instruments consist of either only the underlying asset (without options) or exclusively options (without the underlying asset). The case of using both the stock and options is not considered since the options can always replicate positions in the underlying asset with calls and puts by relying on the put-call parity. In what follows, let $\{X_n\}_{n=0}^{N-1}$ and $\{Y_n\}_{n=0}^{N-1}$ be respectively the sequence of feature vectors and output vectors of an LSTM as in Definition 4.4. When hedging only with the underlying, the time-t_n feature vector considered is[9]

$$X_n = [\log(S_n^{(0,b)}/K), V_n^{\delta}, \varphi_n], \quad n = 0, \ldots, N-1, \tag{4.15}$$

[9] The use of $\log(S_n^{(0,b)}/K)$ instead of $S_n^{(0,b)}$ in feature vectors was found to improve the learning speed of the neural networks (i.e. time taken to find a good set of trainable parameters). Note that log transformation for risky asset prices was also considered in Carbonneau (2021), Buehler et al. (2019b) and Buehler et al. (2019a).

where K is the strike price of Φ and $\{\varphi_n\}_{n=0}^{N-1}$ is a sequence of additional relevant state variables associated with the dynamics of asset prices. For instance, if the underlying log-returns are modeled with a GARCH process, it is well-known that the bivariate process of the underlying price and the GARCH volatility has the Markov property under $\mathbb{P}$ with respect to the market filtration $\mathbb{F}$. The time-t_n volatility of the GARCH process is thus added to the feature vectors through φ_n. Furthermore, in that same case where the underlying stock is considered as the only hedging instrument, the output vectors of the LSTM consist of the number of underlying asset shares to be held in the portfolio for all time steps, i.e. $Y_n = \delta_{n+1}^{(0)}$ for $n = 0, \ldots, N-1$.

Conversely, when hedging is performed with options as hedging instruments, the *implied volatilities* (IVs) of such options denoted as $\{IV_n\}_{n=0}^{N-1}$ are added to feature vectors with IV_n encompassing every implied volatilities needed to price the D options used for hedging:[10]

$$X_n = [\log(S_n^{(0,b)}/K), V_n^{\delta}, \varphi_n, IV_n], \quad n = 0, \ldots, N-1. \tag{4.16}$$

In that case, the output vectors are the number of option contracts held in the portfolio for the various time steps: $Y_n = [\delta_{n+1}^{(1)}, \ldots, \delta_{n+1}^{(D)}]$ for $n = 0, \ldots, N-1$. Recall that when options are used as hedging instruments, $\delta_{n+1}^{(0)} = 0$ for $n = 0, \ldots, N-1$.

Remark 4.2. *Although the portfolio value V_n^{δ} is in theory a redundant feature in the context of LSTMs since it can be retrieved as a function of previous times inputs and outputs of the neural network (see (4.4)), incorporating it to feature vectors was found to significantly improve upon the hedging effectiveness of the LSTMs in the numerical experiments conducted in Section 4.4.*

[10] Note that the bijection relation between implied volatilities and option prices entails that either values could theoretically be used in feature vectors as one is simply a nonlinear transformation of the other.

4.3.2 Equal risk pricing with neural networks

To numerically solve the underlying global hedging problems of the ERP framework, Carbonneau and Godin (2021b) propose to use two distinct neural networks denoted as $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$ to approximate the global trading policies of respectively the long and short positions in Φ. This is the approach considered in the current paper. As illustrated below, the procedure consists in solving the alternative problems of optimizing the neural networks trainable parameters so as to minimize the corresponding hedging shortfall:

$$\epsilon^{(\mathcal{L})}(V_0) \approx \min_{\theta \in \mathbb{R}^q} \rho \left(-\Phi(S_N^{(0,b)}) - B_N(V_0 + G_N^{\delta^{(\mathcal{L},\theta)}}) \right), \tag{4.17}$$

$$\epsilon^{(\mathcal{S})}(V_0) \approx \min_{\theta \in \mathbb{R}^q} \rho \left(\Phi(S_N^{(0,b)}) - B_N(V_0 + G_N^{\delta^{(\mathcal{S},\theta)}}) \right), \tag{4.18}$$

where $\delta^{(\mathcal{L},\theta)}$ and $\delta^{(\mathcal{S},\theta)}$ are to be understood respectively as the output sequences of $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$, and $q \in \mathbb{N}$ is the total number of trainable parameters of $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$. The approximated measured risk exposures obtained through (4.17) and (4.18) are subsequently used to compute equal risk prices and ϵ-completeness measures with (4.10) and (4.11). One implicit assumption associated with (4.17) and (4.18) is that the architecture of all neural networks in terms of the number of LSTM cells and neurons per cell is always fixed; the hyperparameter tuning step of the optimization problem is not considered in this paper. Section 4.3.3 that follows presents the procedure considered in this study to optimize the trainable parameters of the LSTMs.

Remark 4.3. *Carbonneau and Godin (2021b) show that when relying on feedforward neural networks (FFNNs[11]) instead of LSTMs, the alternative problems* (4.17)-(4.18) *allow for arbitrarily precise approximations of the measured risk exposures* (4.5)-(4.6) *due to results from Buehler et al. (2019b). Despite this theoretical ability of FFNNs to approximate arbitrarily well global hedging policies in such context, the authors of the current paper found*

[11] FFNNs are another class of neural networks which map input vectors into output vectors, in contrast to LSTMs which map input vector sequences to output vector sequences.

that LSTMs are able to learn significantly better trading policies than FFNNs in the numerical experiments carried out in Section 4.4, which motivates their use over FFNNs. The theoretical justifications for the outperformance of LSTMs over FFNNs in the financial market settings of this paper are out-of-scope and are left-out as interesting potential future research.

4.3.3 Training neural networks

The numerical scheme to optimize the trainable parameters of neural networks as entailed by the global hedging optimization problems (4.17)-(4.18) is now described. The procedure first proposed in Buehler et al. (2019b) uses minibatch stochastic gradient descent (SGD) to approximate the gradient of the cost function with Monte Carlo sampling. For convenience, the notation used for the optimization procedure is similar to the one from Carbonneau and Godin (2021b). Without loss of generality, the numerical procedure is only presented for the short measured risk exposure; the corresponding procedure for the long position is simply obtained through modifying the objective function (4.19) that follows. Let $J : \mathbb{R}^q \rightarrow \mathbb{R}$ be the cost function to be minimized for the short position in Φ, where θ is the set of trainable parameters of $F_\theta^{(\mathcal{S})}$:[12]

$$J(\theta) := \rho\left(\Phi(S_N^{(0,b)}) - B_N G_N^{\delta^{(\mathcal{S},\theta)}}\right), \quad \theta \in \mathbb{R}^q. \tag{4.19}$$

A typical stochastic gradient descent procedure entails adapting the trainable parameters iteratively and incrementally in the opposite direction of the cost function gradient with respect to θ:

$$\theta_{j+1} = \theta_j - \eta_j \nabla_\theta J(\theta_j), \tag{4.20}$$

[12] Minimizing J with respect to θ corresponds to the alternative problem (4.18) with zero initial capital. Recall that $\epsilon^{(\mathcal{L})}(0)$ and $\epsilon^{(\mathcal{S})}(0)$ are required for the computation of $C_0^\star$ and $\epsilon^\star$. Consequently, hedging portfolio values used in LSTM feature vectors are equal to hedging gains, i.e. $V_n^\delta = B_n G_n^\delta$.

where θ_0 is the initial values for the trainable parameters, η_j is a small deterministic positive real value commonly called the *learning rate* and ∇_θ denotes the gradient operator. In the current study, the *Glorot uniform initialization* of Glorot and Bengio (2010) is always used to select initial parameters in θ_0. Since closed-form solutions for the gradient of the cost function with respect to the trainable parameters are unavailable in the general market setting considered in this work, the approach relies instead on Monte Carlo sampling to provide an estimate. Thus, let $\mathbb{B}_j := \{\pi_{i,j}\}_{i=1}^{N_{\text{batch}}}$ be a minibatch of simulated hedging errors of size $N_{\text{batch}} \in \mathbb{N}$ where $\pi_{i,j}$ is the i^{th} simulated hedging error when $\theta = \theta_j$:

$$\pi_{i,j} := \Phi(S_{N,i}^{(0,b)}) - B_N G_{N,i}^{\delta^{(\mathcal{S},\theta_j)}}, \tag{4.21}$$

where $S_{N,i}^{(0,b)}$ and $G_{N,i}^{\delta^{(\mathcal{S},\theta_j)}}$ are the i^{th} random realization among the minibatch of the terminal underlying asset price and discounted hedging portfolio gains, respectively. Furthermore, denote $\hat{\rho} : \mathbb{R}^{N_{\text{batch}}} \rightarrow \mathbb{R}$ as the empirical estimator of $\rho(\Phi(S_N^{(0,b)}) - B_N G_N^{\delta^{(\mathcal{S},\theta)}})$ evaluated with minibatches of hedging errors. Minibatch SGD consists in approximating the gradient of the cost function $\nabla_\theta J(\theta_j)$ with $\nabla_\theta \hat{\rho}(\mathbb{B}_j)$ in the update rule for trainable parameters:

$$\theta_{j+1} = \theta_j - \eta_j \nabla_\theta \hat{\rho}(\mathbb{B}_j). \tag{4.22}$$

For the numerical experiments conducted in Section 4.4, the convex risk measure considered is the Conditional Value-at-Risk (CVaR, Rockafellar and Uryasev (2002)). For $\alpha \in (0,1)$, this risk measure can be formally defined as

$$\text{VaR}_\alpha(X) := \min_x \{x | \mathbb{P}(X \leq x) \geq \alpha\}, \quad \text{CVaR}_\alpha(X) := \frac{1}{1-\alpha}\int_\alpha^1 \text{VaR}_\gamma(X) d\gamma$$

where $\text{VaR}_\alpha(X)$ is the Value-at-Risk (VaR) of confidence level α. Furthermore, for an absolutely continuous integrable random variable[13], the CVaR has the representation

$$\mathrm{CVaR}_\alpha(X) := \mathbb{E}[X|X \geq \mathrm{VaR}_\alpha(X)], \quad \alpha \in (0,1). \tag{4.23}$$

Let $\{\pi_{[i],j}\}_{i=1}^{N_{\text{batch}}}$ be the order statistics (i.e. values sorted by increasing order) of $\mathbb{B}_j$. For $\tilde{N} := \lceil \alpha N_{\text{batch}} \rceil$ where $\lceil x \rceil$ is the ceiling function (i.e. the smallest integer greater or equal to x), the empirical estimator of the CVaR used in this study is from the work of Hong et al. (2014) and has the representation

$$\widehat{\mathrm{VaR}}_\alpha(\mathbb{B}_j) := \pi_{[\tilde{N}],j},$$

$$\widehat{\mathrm{CVaR}}_\alpha(\mathbb{B}_j) := \widehat{\mathrm{VaR}}_\alpha(\mathbb{B}_j) + \frac{1}{(1-\alpha)N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \max(\pi_{i,j} - \widehat{\mathrm{VaR}}_\alpha(\mathbb{B}_j), 0).$$

The gradient of the empirical estimator of the Conditional Value-at-Risk with respect to the trainable parameters (i.e. $\nabla_\theta \widehat{\mathrm{CVaR}}_\alpha(\mathbb{B}_j)$) required for the update rule (4.22) can be computed exactly without discretization or other numerical approximations. Such computations can be implemented with modern deep learning libraries such as Tensorflow (Abadi et al., 2016). Furthermore, algorithms which dynamically adapt the learning rate η_j in (4.22) such as *Adam* (Kingma and Ba, 2014) have been shown to improve upon the effectiveness of SGD procedures for neural networks. For all numerical experiments conducted in Section 4.4, an implementation of Tensorflow with the Adam algorithm is used to optimize neural networks; the reader is referred to the online Github repository for samples of codes in Python.[14] Also, Section 4.6.2 presents a pseudo-code of the training procedure for $F_\theta^{(\mathcal{S})}$.

[13] In Section 4.4, the only dynamics considered for the risky assets produce integrable and absolutely continuous hedging errors.

[14] github.com/alexandrecarbonneau.

4.4 Numerical experiments

This section performs various numerical experimentations of the ERP approach for derivatives valuation. The main goal is to study the impact of including options as hedging instruments on equal risk prices and on the level of market incompleteness. A special case assessed throughout this section is the trading of short-term vanilla options for the pricing and hedging of longer-term derivatives. The conduction of these experiments heavily relies on the neural network scheme described in Section 4.3 to solve the underlying global hedging problems of the ERP framework. Such exhaustive numerical study would have been hardly accessible with traditional methods (e.g. conventional dynamic programming algorithms) due to the high-dimensional continuous state and action spaces of the hedging problem stemming from the use of multiple short-term options as hedging instruments and from the asset price dynamics considered. As a result, the use of neural networks enables us to provide novel qualitative insights into the ERP framework.

The analysis begins in Section 4.4.2 and Section 4.4.3 with the assessment of the sensitivity of equal risk prices and residual hedging risk to the presence of two salient equity stylized features: jump and volatility risks. The impact of the choice of convex risk measure on the ERP framework when trading exclusively options is examined in Section 4.4.4. Lastly, Section 4.4.5 presents the benchmarking of equal risk prices to derivative premiums obtained with variance-optimal hedging. The specific financial market setup and asset dynamics models considered for all numerical experiments are described in Section 4.4.1 that follows.

4.4.1 Market setup and asset dynamics models

For the rest of the paper, the derivative to price is a European vanilla put option of payoff function $\Phi(S_N^{(0,b)}) = \max(K - S_N^{(0,b)}, 0)$ with $K = 90, 100$ and 110 corresponding respectively to an out-of-the-money (OTM), an at-the-money (ATM) and an in-the-money (ITM) option. The maturity of the derivative is set to 1 year (i.e. $T = 1$) with 252 days. The annualized continuously compounded risk-free rate is $r = 0.03$. In addition to the risk-free asset, the hedging instruments consist of either only the underlying stock, or exclusively shorter-term ATM European calls and puts. When hedging is performed with the underlying stock, daily and monthly rebalancing are considered, corresponding to respectively $N = 252$ and $N = 12$ trading periods per year. When hedging with options, all options are assumed to have a single-period time-to-maturity, i.e. they are traded once and held until expiration. We consider either 1-month or 3-months maturities ATM calls and puts as hedging instruments, which respectively entails $N = 12$ or $N = 4$. Less frequent rebalancing when hedging with options rather than only with the underlying stock is consistent with market practices; such hedging instruments are commonly embedded in semi-static type of trading strategies, see for instance Carr and Wu (2014). Lastly, note that daily variations for the underlying log-returns and implied volatilities are always considered throughout the rest of the paper, even with non-daily rebalancing periods (i.e. when hedging with the underlying stock on a monthly basis or with 1-month and 3-months maturities options) by aggregating daily variations over the rebalancing period.

4.4.1.1 Asset price dynamics

The asset price dynamics models considered in stochastic simulations are now introduced. To characterize jump risk, the Merton jump-diffusion model (MJD, Merton (1976)) is considered. Furthermore, the impact of volatility risk is assessed with the GJR-GARCH model of Glosten et al. (1993). Several sets of parameters are tested for each model to conduct a sensitivity analysis and highlight the impact of various model features on both equal risk prices and residual hedging risk.

Denote $y_n := \log(S_n^{(0,b)}/S_{n-1}^{(0,b)})$ as the periodic underlying stock log-return between the trading periods t_{n-1} and t_n. Since our modeling framework assumes daily variations for asset prices and possibly non-daily rebalancing, let $\{\tilde{y}_{j,n}\}_{j=1}^{M}$ be the M daily stock log-returns in the time interval $[t_{n-1}, t_n]$ such that[15]

$$y_n = \sum_{j=1}^{M} \tilde{y}_{j,n}, \quad n = 1, \ldots, N, \quad N \times M = 252, \tag{4.24}$$

where N corresponds to the number of trading dates to hedge the 1 year maturity derivative Φ and M to the number of days between two trading dates. Thus, daily stock hedges corresponds to the case of $N = 252$ and $M = 1$, monthly stock and 1-month option hedges to $N = 12$ and $M = 21$, and 3-months option hedges to $N = 4$ and $M = 63$.

The asset price dynamics are now formally defined for the daily log-returns. For the rest of the section, let $\{\epsilon_{j,n}\}_{j=1,n=1}^{M,N}$ be a sequence of independent standardized Gaussian random variables where the subsequence $\{\epsilon_{j,n}\}_{j=1}^{M}$ will be used to model the M daily innovations of log-returns in the time interval $[t_{n-1}, t_n]$.

4.4.1.2 Discrete-time Merton-Jump diffusion model (Merton, 1976)

The Merton-jump diffusion dynamics expands upon the ideal market conditions of the Black-Scholes model by incorporating random Gaussian jumps along stock paths. Let $\{N_{j,n}\}_{j=0,n=1}^{M,N}$ be a discrete-time sampling from a Poisson process of intensity parameter $\lambda > 0$ where the subsequence $\{N_{j,n}\}_{j=0}^{M}$ corresponds to the $M+1$ daily values of the Poisson process occurring during the time interval $[t_{n-1}, t_n]$. $N_{0,1} := 0$ is the initial value of the process and $N_{0,n+1} := N_{M,n}$ for $n = 1, \ldots, N-1$. Furthermore, denote $\{\xi_k\}_{k=1}^{\infty}$ as a sequence of random Gaussian variables corresponding to the jumps of mean μ_J and variance σ_J^2. $\{N_{j,n}\}_{j=0,n=1}^{M,N}$, $\{\xi_k\}_{k=1}^{\infty}$ and $\{\epsilon_{j,n}\}_{n=1,j=1}^{N,M}$ are independent. For $n = 1, \ldots, N$ and $j = 1, \ldots, M$, the daily

[15] For completeness, let $\{\tilde{S}_{j,n}^{(0,b)}\}_{j=0,n=1}^{M,N}$ be the daily underlying stock prices where $\{\tilde{S}_{j,n}^{(0,b)}\}_{j=0}^{M}$ corresponds to the $M+1$ daily prices during the period $[t_{n-1}, t_n]$. Also, let $\mathbb{G} := \{\mathcal{G}_{j,n}\}_{j=0,n=1}^{M,N}$ be a filtration satisfying the usual conditions with $\mathcal{G}_{j,n}$ containing all information available to market participants at the j^{th} day of the time period $[t_{n-1}, t_n]$. The filtration used to optimize trading strategies $\mathbb{F}$ with time steps $t_0, t_1, \ldots, t_N$ is a subset of $\mathbb{G}$ by construction. However, since the risky asset dynamics considered in this paper have the Markov property, optimizing trading strategies with the filtration $\mathbb{F}$ or $\mathbb{G}$ results in the same trading policy.

log-return dynamics can be specified as[16]

$$\tilde{y}_{j,n} = \frac{1}{252}\left(\nu - \lambda\left(e^{\mu_J+\sigma_J^2/2} - 1\right) - \frac{\sigma^2}{2}\right) + \sigma\sqrt{\frac{1}{252}}\epsilon_{j,n} + \sum_{k=N_{j-1,n}+1}^{N_{j,n}} \xi_k, \tag{4.25}$$

where $\{\nu, \mu_J, \sigma_J, \lambda, \sigma\}$ are the model parameters with $\{\nu, \lambda, \sigma\}$ being on a yearly scale, $\nu \in \mathbb{R}$ and $\sigma > 0$. Furthermore, since $\{S_n^{(0,b)}\}_{n=0}^N$ has the Markov property with respect to the filtration $\mathbb{F}$ generated by the trading dates observations, no additional state associated to the risky asset dynamics is required to be added to the feature vectors of neural networks (i.e. $\varphi_n = 0$ for all time steps n in (4.15) and (4.16)).

4.4.1.3 GJR-GARCH(1,1) model (Glosten et al., 1993)

GARCH processes also expand upon the Black-Scholes ideal framework by exhibiting well-known empirical features of risky assets such as time-varying volatility, volatility clustering and the leverage effect (i.e. negative correlation between underlying returns and its volatility). Daily log-returns modeled with a GJR-GARCH(1,1) dynamics have the representation

$$\begin{aligned} \tilde{y}_{j,n} &= \mu + \tilde{\sigma}_{j,n}\epsilon_{j,n}, \\ \tilde{\sigma}_{j+1,n}^2 &= \omega + \upsilon\tilde{\sigma}_{j,n}^2(|\epsilon_{j,n}| - \gamma\epsilon_{j,n})^2 + \beta\tilde{\sigma}_{j,n}^2, \end{aligned} \tag{4.26}$$

where $\{\tilde{\sigma}_{j,n}^2\}_{j=1,n=1}^{M+1,N}$ are the daily conditional variances of log-returns. More precisely, $\{\tilde{\sigma}_{j,n}^2\}_{j=1}^{M+1}$ are the $M+1$ daily conditional variances in the time interval $[t_{n-1}, t_n]$. Also, $\tilde{\sigma}_{1,n+1}^2 := \tilde{\sigma}_{M+1,n}^2$ for $n = 1, \ldots, N-1$. Model parameters consist of $\{\mu, \omega, \upsilon, \gamma, \beta\}$ with $\{\omega, \upsilon, \beta\}$ being positive real values and $\gamma, \mu \in \mathbb{R}$. Note that if the starting value of the GARCH process $\tilde{\sigma}_{1,1}^2$ is deterministic, then $\{\tilde{\sigma}_{j,n}^2\}_{j=1,n=1}^{M+1,N}$ can be computed recursively with the observed

[16] This paper adopts the convention that if $N_{j,n} = N_{j-1,n}$, i.e. that no jumps occurred on that day, then:

$$\sum_{k=N_{j-1,n}+1}^{N_{j,n}} \xi_k = 0.$$

daily log-returns. In this paper, $\tilde{\sigma}^2_{1,1}$ is set as the stationary variance: $\tilde{\sigma}^2_{1,1} := \frac{\omega}{1-\upsilon(1+\gamma^2)-\beta}$. Also, contrarily to the MJD model, the GJR-GARCH(1,1) requires adding at each trading time t_n the current stochastic volatility value to the feature vectors of the neural networks, i.e. $\varphi_n = \tilde{\sigma}_{1,n+1}$ for $n = 0, \dots, N-1$.

4.4.1.4 Implied volatility dynamics

This work proposes to model the daily variations of the logarithm of ATM implied volatilities with 1-month and 3-months maturities as a discrete-time version of the Ornstein-Uhlenbeck (OU) process.[17] The choice of an OU type of dynamics for IVs is motivated by the work of Cont and Da Fonseca (2002) which shows that for S&P 500 index options, the first principal component of the daily variations of the logarithm of the IV surface accounts for the majority of its variance and can be interpreted as a level effect. Also, this first principal component can be well represented by a low-order autoregressive (AR) model. The OU dynamics considered in this study therefore has the representation of an AR model of order 1.

The dynamics for the daily evolution of IVs is now formally defined. For convenience, this paper assumes that 1-month and 3-months IVs are the same.[18] Using a similar notation as for daily log-returns, let $\{\widetilde{IV}_{j,n}\}_{j=0,n=1}^{M,N}$ be the daily ATM IV process for both 1-month and 3-months maturities where $\{\widetilde{IV}_{j,n}\}_{j=0}^{M}$ are the $M+1$ daily observations during the time interval $[t_{n-1}, t_n]$ with $\widetilde{IV}_{0,n+1} := \widetilde{IV}_{M,n}$ for $n = 1, \dots, N-1$. Furthermore, let $\{Z_{j,n}\}_{j=1,n=1}^{M,N}$ be an additional sequence of independent standardized Gaussian random variables characterizing shocks in the IV dynamics. In order to incorporate the stylized feature of strong negative correlation between implied volatilities and asset returns (Cont and Da Fonseca (2002)), the modeling framework assumes that the daily innovations of log-returns and IVs are correlated with parameter $\varrho := corr(\epsilon_{j,n}, Z_{j,n})$ set at -0.6 for all time steps. The dynamics for the

[17] It is important to note that implied volatilities are used strictly for pricing options used as hedging instruments. They are not used to price the derivative Φ.

[18] It is worth highlighting that since trading strategies allow for the use of either 1-month or 3-months maturities ATM calls and puts, but not both maturities within the same strategy, 1-month and 3-months IVs are never used at the same time.

evolution of the logarithm of IVs, which is referred from now on as the log-AR(1) model, has the following representation for $n = 1, \dots, N$ and $j = 0, \dots, M-1$:

$$\log \widetilde{IV}_{j+1,n} = \log \widetilde{IV}_{j,n} + \kappa(\vartheta - \log \widetilde{IV}_{j,n}) + \sigma_{IV} Z_{j+1,n}, \tag{4.27}$$

where $\{\kappa, \vartheta, \sigma_{IV}\}$ are the model parameters with $\kappa, \vartheta \in \mathbb{R}$ and $\sigma_{IV} > 0$. The initial value of the process is set as $\log \widetilde{IV}_{0,1} = \vartheta$. Also, recall that when trading options, their corresponding implied volatilities at each trading date are added to the feature vectors of neural networks, i.e. $IV_{n-1} = \widetilde{IV}_{0,n}$ in (4.16) for $n = 1, \dots, N$.

The pricing of calls and puts used as hedging instruments is done with the well-known Black-Scholes formula hereby stated with the annual volatility term set at the implied volatility value. For the underlying price S, implied volatility IV, strike price K and time-to-maturity ΔT, the Black-Scholes pricing formulas for calls and puts are respectively

$$C(S, IV, \Delta T, K) := S\mathcal{N}(d_1) - e^{-r\Delta T} K \mathcal{N}(d_2), \tag{4.28}$$

$$P(S, IV, \Delta T, K) := e^{-r\Delta T} K \mathcal{N}(-d_2) - S\mathcal{N}(-d_1), \tag{4.29}$$

where $\mathcal{N}(\cdot)$ denotes the standard normal cumulative distribution function and

$$d_1 := \frac{\log(\frac{S}{K}) + (r + \frac{IV^2}{2})\Delta T}{IV\sqrt{\Delta T}}, \quad d_2 := d_1 - IV\sqrt{\Delta T}.$$

4.4.1.5 Hyperparameters

The set of hyperparameters for the LSTMs are two LSTM cells (i.e. $H = 2$) and 24 neurons per cell (i.e. $d_j = 24$ for $j = 1, 2$). A training set of 400,000 paths is used to optimize the trainable parameters with a total of 50 epochs and a minibatch size of 1,000 sampled exclusively from the training set.[19] The deep learning library Tensorflow (Abadi et al., 2016) is used to implement the stochastic gradient descent procedure with the Adam optimizer of Kingma and Ba (2014) with a learning rate hyperparameter value of 0.01/6. All numerical results presented throughout this section are computed based on a test set (i.e. out-of-sample dataset) of 100,000 paths. Lastly, unless specified otherwise, the convex risk measure chosen for all experiments is the CVaR with confidence level $\alpha = 0.95$. Sensitivity analyses of equal risk prices and residual hedging risk with respect to the confidence level parameter of the CVaR measure are performed in Section 4.4.4 and Section 4.4.5.

4.4.2 Sensitivity of equal risk pricing to jump risk

This section examines the sensitivity of the ERP solution to equity jump risk. The analysis is carried out by considering three different sets of parameters for the MJD dynamics which induce different levels of jump frequency and severity. While maintaining empirical plausibility, this is done by modifying the intensity parameter λ controlling the expected frequency of jumps as well as parameters μ_J and σ_J controlling the severity component of jumps. In order to better isolate the impact of different stylized features of jump risk on the ERP framework, the diffusion parameter[20] is fixed for all three sets of parameters. Also, the parameters $\{\lambda, \mu_J, \sigma_J, \nu\}$ are chosen such that the yearly expected value and standard deviation of log-returns are respectively 10% and 15% for all three cases. To facilitate the analysis, the three sets of parameters are referred to as scenario 1, scenario 2 and scenario 3 for jump risk. Model parameter values for the three scenarios are presented in Table 4.1. Scenario 1 represents relatively smaller but more frequent jumps with on average one jump per year of mean -5% and standard deviation 5%. Scenario 2 entails more severe, but less frequent jumps with on average one jump every four years of mean -10% and standard deviation 10%. Lastly, scenario 3 depicts the most extreme case with rare but very severe jumps with on

[19] One epoch consists of a complete iteration of SGD on the training set. For a training set of 400,000 paths and a minibatch of size 1,000, a total of 400 updates of the trainable parameters as in (4.22) is performed within an epoch.

[20] The parameter σ in (4.25) corresponds to the diffusion parameter of the MJD dynamics.

Table 4.1: Parameters of the Merton jump-diffusion model for the three scenarios.

	ν	σ	λ	μ_J	σ_J
Scenario 1	0.1112	0.1323	1	-0.05	0.05
Scenario 2	0.1111	0.1323	0.25	-0.10	0.10
Scenario 3	0.1110	0.1323	0.08	-0.20	0.15

Notes: ν, σ and λ are on an annual basis.

average one jump every twelve and a half years of mean -20% and standard deviation 15%. Moreover, parameter values for the log-AR(1) implied volatility model are kept fixed for all three scenarios and are presented in Table 4.2. Note that the long-run parameter ϑ is set at the logarithm of the yearly standard deviation of log-returns with $\vartheta = \log 0.15$, and other parameters are chosen in an ad hoc fashion so as to produce reasonable values for implied volatilities.

Table 4.2: Parameters of the log-AR(1) model for the evolution of implied volatilities.

κ	ϑ	σ_{IV}	ϱ
0.15	$\log(0.15)$	0.06	-0.6

4.4.2.1 Benchmarking results in the presence of jump risk

Table 4.3 presents equal risk prices $C_0^\star$ and residual hedging risk $\epsilon^\star$ across the three scenarios of jump parameters and different trading instruments. Numerical values indicate that in the presence of jump risk, hedging with options entails significant reduction of both equal risk prices and market incompleteness as compared to hedging solely with the underlying stock across all moneyness levels and jump risk scenarios. The reduction in hedging residual risk by trading options is obtained despite less frequent rebalancing than when only the stock is used. These results add additional evidence that options are indeed non-redundant as prescribed by the Black-Scholes world: the equal risk pricing framework dictates that hedging with options in the presence of jump risk can significantly impact both derivative premiums and hedging risk as quantified by our incompleteness metrics.

Table 4.3: Sensitivity analysis of equal risk prices $C_0^\star$ and residual hedging risk $\epsilon^\star$ to jump risk for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	OTM			ATM			ITM		
Jump Scenario	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
$C_0^\star$									
Daily stock	1.89	2.58	3.36	5.21	6.01	6.81	10.81	11.68	12.13
Monthly stock	1.97	2.60	3.31	5.04	5.77	6.38	10.73	11.44	11.86
1-month options	1.82	2.24	2.55	4.99	5.36	5.60	10.48	10.86	10.83
3-months options	1.74	2.08	2.39	4.87	5.12	5.28	10.43	10.51	10.57
$\epsilon^\star$									
Daily stock	1.09	1.98	2.67	1.76	2.74	3.54	1.82	2.78	3.27
Monthly stock	1.82	2.52	3.26	3.00	3.88	4.57	3.07	3.91	4.37
1-month options	0.76	1.18	1.52	1.14	1.53	1.78	1.17	1.56	1.54
3-months options	1.03	1.37	1.68	1.59	1.82	2.02	1.70	1.79	1.88
$\epsilon^\star/C_0^\star$									
Daily stock	0.58	0.77	0.79	0.34	0.46	0.52	0.17	0.24	0.27
Monthly stock	0.92	0.97	0.99	0.60	0.67	0.72	0.29	0.34	0.37
1-month options	0.42	0.53	0.60	0.23	0.28	0.32	0.11	0.14	0.14
3-months options	0.59	0.66	0.70	0.33	0.36	0.38	0.16	0.17	0.18

Notes: Results are computed based on 100,000 independent paths generated from the Merton Jump-Diffusion model for the underlying (see Section 4.4.1.2 for model description). Three different sets of parameters values are considered with $\lambda = \{1, 0.25, 0.08\}$, $\mu_J = \{-0.05, -0.10, -0.20\}$ and $\sigma_J = \{0.05, 0.10, 0.15\}$ respectively for jump scenario 1, 2, and 3 (see Table 4.1 for all parameters values). *Hedging instruments*: daily or monthly rebalancing with the underlying stock and 1-month or 3-months options with ATM calls and puts. Options used as hedging instruments are priced with implied volatility modeled with a log-AR(1) dynamics (see Section 4.4.1.4 for model description and Table 4.2 for parameters values). The training of neural networks is done as described in Section 4.4.1.5. The confidence level of the CVaR measure is $\alpha = 0.95$.

The relative reduction achieved in $C_0^\star$ with 1-month and 3-months options as compared to hedging with the stock is most important for OTM puts, followed by ATM and ITM contracts. For instance, the relative reduction obtained with 3-months options hedging over daily stock hedging ranges across the three jump risk scenarios between 8% to 29% for OTM, 6% to 22% for ATM and 4% to 13% for ITM puts.[21] This reduction in $C_0^\star$ when using options as hedging instruments can be explained by the following observations. As pointed out in Carbonneau and Godin (2021b), the fact that a put option payoff is bounded below at zero entails that the short position hedging error has a thicker right tail than the long position hedging error. Also, it is widely documented in the literature that hedging jump risk with options significantly dampens tail risk as compared to using only the underlying stock (see for instance Coleman et al. (2007) and Carbonneau (2021)).[22] Consequently, the choice of trading options to mitigate jump risk reduces the measured risk exposure of both the long and short positions, but the thicker right tail for the short position hedging error entails a larger decrease for the latter than for the long position. In such situations, the ERP framework dictates that the long position should be compensated with a lower derivative premium $C_0^\star$ to equalize residual hedging risk of both positions.

Moreover, values for both $\epsilon^\star$-metrics indicate that in the presence of jump risk, the use of options contributes significantly to the reduction of market incompleteness as both the long and short position hedges achieve risk reduction when compared to trading only with the stock. The latter conclusion is in itself not novel, and is widely documented in the literature (see, for instance, Cont and Tankov (2003) and the many references therein). Indeed, this is a consequence of the well-known convex property of put option prices, which implies that

[21] If $C_0^\star$(daily stock) and $C_0^\star$(3-months options) are equal risk prices obtained respectively by hedging with the stock on a daily basis and with 3-months options, the relative reduction is computed as $1 - \frac{C_0^\star(\text{3-months options})}{C_0^\star(\text{daily stock})}$ for all examples.

[22] Horvath et al. (2021) deep hedge derivatives under a rough Bergomi volatility model by trading the underlying stock and a variance swap. The latter paper shows that this dynamics exhibits jump-like behaviour when discretized. As results presented in this current paper highlights the fact that global hedging jump risk with option hedges is very effective, deep hedging with options could also potentially be effective under such rough volatility models.

hedging random jumps solely with the underlying stock is ineffective. Our $\epsilon^\star$-metrics have the advantage of allowing for a precise quantification of such reduction in residual hedging risk achieved through the use of options as hedging instruments.

The sensitivity of equal risk prices and residual hedging risk across the three jump risk scenarios for each set of hedging instruments is now examined. Numerical results presented in Table 4.3 indicate that for a fixed set of hedging instruments, both the equal risk price and the level of incompleteness increases with the severity of jumps across all moneyness levels. Indeed, the relative increase of equal risk prices observed under scenario 3 as compared to scenario 1 respectively for OTM, ATM and ITM puts is $78\%, 31\%$ and 12% with the daily stock, 68%, 27% and 11% with the monthly stock, 40%, 12% and 3% with 1-month options and 38%, 8% and 1% with 3-months options.[23] Similar observations can be made for both incompleteness metrics: increases in jump severity leads to larger $\epsilon^\star$ and $\epsilon^\star/C_0^\star$. This positive association between both equal risk prices and the level of market incompleteness to jump severity can be explained by the following observations. For a fixed hedging instrument and moneyness level, the long measured risk exposure is closed to invariant to jump severity (i.e. similar values across the three jump risk scenarios). The latter stems from the fact that jump dynamics considered in this paper predominantly entail negative jumps, which result in a thicker left tail for the long position hedging error (i.e. hedging gains) as jump severity increases, but in close to no impact on the right tail of the long position hedging error. In contrast, since the right tail weight of the short position hedging error increases with the expected (negative) magnitude and volatility of jumps, the short measured risk exposure always increases going from scenario 1 to scenario 3, which consequently increases both the equal risk price and the level of market incompleteness.

[23] For a fixed hedging instrument and moneyness level, if $C_0^\star$(scenario 1) and $C_0^\star$(scenario 3) are respectively the equal risk price obtained under jump risk scenario 1 and 3, the relative increase is computed as $\frac{C_0^\star(\text{scenario 3})}{C_0^\star(\text{scenario 1})} - 1$.

4.4.3 Sensitivity of equal risk pricing to volatility risk

Having examined the impact of jump risk on the ERP framework, the impact of volatility risk is now studied. In the same spirit as analyses done for jump risk, three different sets of parameters are considered for the GARCH dynamics which imply annualized stationary (expected) volatilities of $10\%, 15\%$ and 20%.[24] The three sets of parameters are presented in Table 4.4. Note that every parameter is fixed for all three sets, except for the level parameter ω, which is adjusted to attain the wanted stationary volatility. The value of the drift parameter μ is set such that the yearly expected value of log-returns is 10%. Also, values for $\{\upsilon, \gamma, \beta\}$ are inspired from parameters estimated with maximum likelihood on a time series of daily log-returns on the S&P 500 index for the period 1986-12-31 to 2010-04-01 used in Carbonneau and Godin (2021b). The same setup is considered as in Section 4.4.2 in terms of the derivative to be priced (1-year maturity European puts) and for the choice of hedging instruments (underlying stock traded on a daily or monthly basis and 1-month or 3-months maturities ATM calls and puts). The same parameters as in the study of jump risk conducted in Section 4.4.2 are used for $\{\kappa, \sigma_{\text{IV}}, \varrho\}$ of the log-AR(1) dynamics for the evolution of 1-month and 3-months ATM IVs (i.e. $\kappa = 0.15, \sigma_{\text{IV}} = 0.06$ and $\varrho = -0.6$), except for the long-run parameter ϑ, which is set to be in line with the underlying GARCH process as $\log(0.10), \log(0.15)$ and $\log(0.20)$ when the stationary volatility is 10%, 15% and 20%, respectively. It is worth highlighting that the choice of modeling implied volatilities for short-term options with higher and smaller average levels enables us to assess the impact of larger and smaller average costs for trading options on the equal risk price and residual hedging risk of longer-term options.

[24] The annualized stationary volatility with 252 days per year is computed as

$$\sqrt{\frac{252\omega}{1 - \upsilon(1 + \gamma^2) - \beta}}.$$

Table 4.4: Parameters of the GJR-GARCH model for $10\%, 15\%$ and 20% stationary yearly volatilities.

Stationary volatility	μ	ω	υ	γ	β
10%	3.968e-04	8.730e-07	0.05	0.6	0.91
15%	3.968e-04	1.964e-06	0.05	0.6	0.91
20%	3.968e-04	3.492e-06	0.05	0.6	0.91

4.4.3.1 Benchmarking results with volatility risk

Table 4.5 presents equal risk prices $C_0^\star$ and $\epsilon^\star$-metrics for put options of 1 year maturity across the three sets of volatility risk parameters and hedging instruments. Numerical results indicate that in the presence of volatility risk, the use of options as hedging instruments can reduce $C_0^\star$ as compared to daily stock hedging. However, this impact on $C_0^\star$ when trading options can be marginal and is highly sensitive to the moneyness level of the put option being priced as well as to the maturity of the traded options. Furthermore, the impact on $C_0^\star$ of the use of options within hedges tends to diminish when traded options are more costly (i.e. as the average level of implied and GARCH volatility increases). Indeed, the relative reduction in equal risk prices achieved with 1-month options hedging as compared to daily stock hedging with $10\%, 15\%$ and 20% stationary volatility is respectively $44\%, 26\%$ and 15% for OTM puts, $12\%, 9\%$ and 5% for ATM and $1\%, 2\%$ and 1% for ITM options.[25] However, the relative reduction in $C_0^\star$ with 3-months option hedges as compared to using the stock on a daily basis is overall much more marginal, with the notable exceptions of OTM and ATM puts with 10% stationary volatility which achieve respectively 25% and 7% reduction as well as for the OTM moneyness under 15% stationary volatility with a 12% reduction. Also, as expected, values presented in Table 4.5 confirm that the level of market incompleteness as measured by the $\epsilon^\star$ metric has a positive relationship with the average level of stationary volatility for all hedging instruments.

[25] If $C_0^\star$(daily stock) and $C_0^\star$(1-month options) are respectively the equal risk price obtained by hedging with the stock on a daily basis and with 1-month options, the relative reduction is computed as $1 - \frac{C_0^\star(\text{1-month options})}{C_0^\star(\text{daily stock})}$ for all examples.

Table 4.5: Sensitivity analysis of equal risk prices $C_0^\star$ and residual hedging risk $\epsilon^\star$ to volatility risk for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	OTM			ATM			ITM		
Stationary volatility	10%	15%	20%	10%	15%	20%	10%	15%	20%
$C_0^\star$									
Daily stock	1.01	2.35	3.85	3.23	5.36	7.24	8.56	10.55	12.51
Monthly stock	1.17	2.65	4.23	3.37	5.44	7.58	8.85	10.82	12.88
1-month options	0.56	1.74	3.27	2.86	4.87	6.89	8.46	10.32	12.35
3-months options	0.76	2.07	3.65	3.01	5.08	7.16	8.51	10.38	12.44
$\epsilon^\star$									
Daily stock	0.77	1.51	2.21	1.28	2.12	2.67	1.12	1.98	2.65
Monthly stock	1.15	2.44	3.67	2.15	3.41	4.62	1.92	3.29	4.56
1-month options	0.26	0.65	1.06	0.59	1.00	1.36	0.62	1.03	1.39
3-months options	0.59	1.32	2.02	1.10	1.77	2.41	1.04	1.68	2.31
$\epsilon^\star / C_0^\star$									
Daily stock	0.77	0.64	0.57	0.40	0.40	0.37	0.13	0.19	0.21
Monthly stock	0.99	0.92	0.87	0.64	0.63	0.61	0.22	0.30	0.35
1-month options	0.45	0.37	0.32	0.21	0.20	0.20	0.07	0.10	0.11
3-months options	0.77	0.64	0.55	0.36	0.35	0.34	0.12	0.16	0.19

Notes: Results are computed based on 100,000 independent paths generated from the GJR-GARCH(1,1) model for the underlying with three sets of parameters implying stationary yearly volatilities of $10\%, 15\%$ and 20% (see Section 4.4.1.3 for model description and Table 4.4 for parameters values). *Hedging instruments*: daily or monthly rebalancing with the underlying stock and 1-month or 3-months options with ATM calls and puts. Options used as hedging instruments are priced with implied volatility modeled as a log-AR(1) dynamics with $\kappa = 0.15, \sigma_{\text{IV}} = 0.06$ and $\varrho = -0.6$ for all cases, and ϑ set to $\log(0.10), \log(0.15)$ and $\log(0.20)$ when the GARCH stationary volatility is $10\%, 15\%$ and 20%, respectively (see Section 4.4.1.4 for the log-AR(1) model description). The training of neural networks is done as described in Section 4.4.1.5. The confidence level of the CVaR measure is $\alpha = 0.95$.

The previously described observations about the impact of option hedges on both equal risk prices and residual hedging risk all stem from the realized reduction in measured risk exposure by the long and short positions. However, contrarily to results obtained with jump risk, the reduction in measured risk exposure when hedging volatility risk with options can be very similar for both the long and short positions, whereas with jump risk, the reduction is asymmetric by always favoring the short position with a larger reduction. The latter can be explained by the fact that volatility risk impacts both upside and downside risk, while the impact of jump risk dynamics considered in this paper is very asymmetric by entailing significantly more weight on the right (resp. left) tail of the short (resp. long) hedging error with predominantly negative jumps. Values presented in Table 4.5 confirm this analysis of the interrelation between volatility risk and the choice of hedging instruments. For instance, for ITM puts, the measured risk exposure of the long and short positions decreases by a similar amount when trading 1-month or 3-months options as compared to daily stock hedges, which explains the significant decrease in $\epsilon^{\star}$, but also the insensitivity of $C_0^{\star}$ to the choice of hedging instruments and rebalancing frequency. On the other hand, for OTM puts, 1-month and 3-months option hedges results in larger decreases of measured risk exposure for the short position than for the long position, which explains the reduction in $C_0^{\star}$ and $\epsilon^{\star}$ as compared to daily stock hedges.

Lastly, it is very interesting to observe that the average price level of short-term options used as hedging instruments is effectively reflected into the equal risk price of longer-term options. Indeed, numerical results for $C_0^{\star}$ presented in Table 4.5 highlight the fact that higher hedging options implied volatilities for 1-month and 3-months ATM calls and puts leads to higher equal risk prices for 1-year maturity puts. Furthermore, to isolate the idiosyncratic contribution of the variations of option prices used as hedging instruments on the equal risk price from the impact of the stationarity volatility of the GARCH process, the authors also tested fixing the stationarity volatility of the GARCH process to 15% and setting the long-run parameter of the IV process to 14% and 16%. These results presented in Section 4.6.3, Table 4.9, confirm that higher implied volatilities for options used as hedging instruments leads to higher equal risk prices. All of these benchmarking results demonstrate the potential of the ERP framework as a fair valuation approach consistent with observable market prices. For instance, the ERP framework could be used to price and optimally hedge over-the-counter derivatives with vanilla options. An additional potential application is the marking-to-market of less liquid long-term derivatives (e.g. Long-Term Equity AnticiPation Securities (LEAPS)) consistently with highly liquid shorter-term option hedges. The ERP framework could also be used for the fair valuation of segregated funds guarantees, which are equivalent to very long-term (up to 40 years) derivatives sold by insurers.[26] Indeed, International Financial Reporting Standards 17 (IFRS 17, IASB (2017)) mandates a market consistent valuation of options embedded in segregated funds guarantees with readily available observable market prices at the measurement date. The ERP framework could potentially be applied to price such very long-term options consistently with shorter-term implied volatility surface dynamics, with the latter being much less challenging to calibrate due to the higher liquidity of short-term options.[27]

4.4.4 Sensitivity analyses to the confidence level of $CVaR_\alpha$

This section conducts sensitivity analyses with respect to the choice of convex risk measure on the ERP framework when trading exclusively options. Similarly to the work of Carbonneau and Godin (2021b), values for equal risk prices and $\epsilon^\star$-metrics are examined across the confidence levels $0.90, 0.95$ and 0.99 for the CVaR_α measure. As argued in the latter paper, higher confidence levels corresponds to more risk averse agents by concentrating more relative

[26] Note that Carbonneau (2021) demonstrates the potential of the deep hedging algorithm for global hedging long-term lookback options embedded in segregated funds guarantees with multiple hedging instruments. It is also worth highlighting that Barigou et al. (2020) developed a pricing scheme consistent with local non-quadratic hedging procedures for insurance liabilities which relies on neural networks.

[27] In the context of segregated funds, the short position of the embedded option is assumed to be held by an insurance company who has to provide a quote and mitigate its risk exposure. The long position is held by an unsophisticated investor who will not be hedging his risk exposure. Nevertheless, as IFRS 17 mandates the use of a fair valuation approach for embedded options consistent with observable market prices, the ERP framework could potentially be used in this context.

weight on losses of larger magnitude. The main finding of the sensitivity analysis conducted in Carbonneau and Godin (2021b) is that when trading exclusively the underlying stock, higher confidence levels leads to larger values for $C_0^\star$ and $\epsilon^\star$ metrics. The objective of this section is to assess if this finding is robust to the use of short-term option hedges instead of the underlying stock. For each confidence level, the authors of the current paper computed both equal risk prices and residual hedging risk obtained by trading 3-months ATM calls and puts with the same setup as in Section 4.4.2 and Section 4.4.3, i.e. for all three jump and volatility scenarios of parameters.[28] Overall, the main conclusions are found to be qualitatively similar for all of the different setups. Thus, to save space, values for equal risk prices and residual hedging risk are only reported under the MJD dynamics with jump risk scenario 2 by trading 3-months options; these results are presented in Table 4.6.

Numerical values reported in Table 4.6 indicate that with option hedges, an increase in the confidence level parameter of the CVaR_α measure leads to larger equal risk prices $C_0^\star$ and residual hedging risk $\epsilon^\star$ across all examples. These results confirm that the finding of Carbonneau and Godin (2021b) with respect to the sensitivity of $C_0^\star$ and $\epsilon^\star$ to the risk aversion of the hedger is robust to using exclusively options as hedging instruments. Furthermore, values for equal risk prices $C_0^\star$ show a largest increase when using $\text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ as compared to $\text{CVaR}_{0.90}$ for OTM puts, followed by ATM and ITM moneyness levels; the same conclusion was observed in Carbonneau and Godin (2021b) when trading the underlying stock. The increase in $C_0^\star$ with the risk aversion level of the hedger stems from the thicker right tail of the short position hedging error than for the long position hedging error. The latter observation is consistent with previous analyses: while option hedges are more effective

[28] Unreported tests performed by the authors show that values lower than 0.90 for the confidence level of CVaR_α with 1-month and 3-months option hedges lead to trading policies with significantly larger tail risk in a way which would deem such policies as inadmissible by hedgers. Using the $\text{CVaR}_{0.90}$ measure with 1-month options also resulted in trading policies with significantly larger tail risk. However, this large increase in tail risk was not observed with the $\text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ measures when trading 1-month options, nor with $\text{CVaR}_{0.90}$, $\text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ when trading 3-months options. These observations motivated the choice of performing sensitivity analysis for CVaR_α with $\alpha = 0.90, 0.95$ and 0.99 exclusively when trading 3-months options.

Table 4.6: Sensitivity analysis of equal risk prices $C_0^\star$ and residual hedging risk $\epsilon^\star$ for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$ under the MJD dynamics with jump risk scenario 2.

	$C_0^\star$			$\epsilon^\star$			$\epsilon^\star/C_0^\star$		
Moneyness	OTM	ATM	ITM	OTM	ATM	ITM	OTM	ATM	ITM
$\text{CVaR}_{0.90}$	1.86	4.93	10.40	0.99	1.43	1.50	0.53	0.29	0.14
$\text{CVaR}_{0.95}$	12%	4%	1%	39%	28%	20%	24%	23%	18%
$\text{CVaR}_{0.99}$	40%	10%	4%	116%	76%	65%	54%	60%	59%

Notes: Results are computed based on 100,000 independent paths generated from the Merton Jump-Diffusion model for the underlying (see Section 4.4.1.2 for model description) with parameters $\nu = 0.1111, \sigma = 0.1323, \lambda = 0.25, \mu_J = -0.10$ and $\sigma_J = 0.10$ corresponding to jump risk scenario 2 of Table 4.1. Hedging instruments consist of 3-months ATM calls and puts priced with implied volatility modeled with a log-AR(1) dynamics (see Section 4.4.1.4 for model description and Table 4.2 for parameters values). The training of neural networks is done as described in Section 4.4.1.5. Values for the $\text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ measures are expressed relative to $\text{CVaR}_{0.90}$ (% increase).

than stock hedges in the presence of equity jump risk as demonstrated in Section 4.4.2, their inclusion within hedging portfolios does not fully mitigate the asymmetry in tail risk of the residual hedging error.

4.4.5 Benchmarking of equal risk prices to variance-optimal premiums

This section presents the benchmarking of equal risk prices to derivative premiums obtained with *variance-optimal hedging* procedures (VO, Schweizer (1995)), also commonly called *global quadratic hedging*. Variance-optimal hedging solves jointly for the initial capital investment and a self-financing strategy minimizing the expected value of the squared hedging error:

$$\min_{\delta \in \Pi, V_0 \in \mathbb{R}} \mathbb{E}\left[\left(\Phi(S_N^{(0,b)}) - B_N(V_0 + G_N^\delta)\right)^2\right]. \tag{4.30}$$

The optimized initial capital investment denoted hereafter as $C_0^{(VO)}$ can be viewed as the production cost of Φ, since the resulting dynamic trading strategy replicates the derivative's payoff as closely as possible in a quadratic sense. The optimization problem (4.30) can also be solved in a similar fashion as the non-quadratic global hedging problems embedded in the ERP framework, but with two distinctions: the initial capital investment is treated as an additional trainable parameter and a single neural network is considered since the optimal trading strategy is the same for the long and short position due to the quadratic penalty.[29] The reader is referred to Section 4.6.1 for a complete description of the numerical scheme for variance-optimal hedging implemented in this study.

The setup considered for the examination of this benchmarking is the same as in Section 4.4.2 with the MJD dynamics under the three jump risk scenarios, with the exception of the confidence level of the CVaR_α measure, which is studied at first with $\alpha = 0.95$ fixed as in Section 4.4.2 and Section 4.4.3, and subsequently across $\alpha = 0.90, 0.95$ and 0.99 as in Section 4.4.4. Note that the authors also conducted the same experiments under the setup of Section 4.4.3 with volatility risk, and found that the main qualitative conclusions are very similar. The reader is referred to Table 4.12 and Table 4.13 of Section 4.6.3 for the benchmarking of ERP to VO procedures in the presence of volatility risk.

4.4.5.1 Benchmarking results

Table 4.7 presents benchmarking results of equal risk prices $C_0^\star$ to variance-optimal prices $C_0^{(VO)}$ under the MJD dynamics with the $\text{CVaR}_{0.95}$ measure. Numerical experiments show that $C_0^\star$ is at least larger than $C_0^{(VO)}$ for all examples, but the relative increase is always smaller and less sensitive to jump severity when trading options. Furthermore, the relative increase in derivative premiums observed with the ERP framework over VO hedging is the largest for OTM puts, followed by ATM and ITM options across all jump risk scenarios and hedging instruments. For instance, the relative increase in $C_0^\star$ as compared to $C_0^{(VO)}$ when trading the daily stock ranges from jump scenario 1 to scenario 3 between 17% to 75% for

[29] Cao et al. (2020) showed that the deep hedging algorithm for variance-optimal hedging problems provides good approximations of optimal initial capital investments by comparing the optimized values to known formulas.

Table 4.7: Equal risk prices $C_0^\star$ and variance-optimal (VO) prices $C_0^{(VO)}$ with jump risk for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	OTM			ATM			ITM		
Jump Scenario	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
$C_0^{(VO)}$									
Daily stock	1.62	1.77	1.92	4.71	4.79	4.80	10.20	10.18	10.12
Monthly stock	1.55	1.73	1.86	4.62	4.72	4.72	10.14	10.11	10.05
1-month options	1.71	2.04	2.38	4.82	5.11	5.31	10.27	10.45	10.52
3-months options	1.58	1.83	2.08	4.64	4.82	4.97	10.11	10.15	10.21
$C_0^\star$									
Daily stock	17%	45%	75%	11%	25%	42%	6%	15%	20%
Monthly stock	27%	51%	78%	9%	22%	35%	6%	13%	18%
1-month options	6%	10%	7%	3%	5%	6%	2%	4%	3%
3-months options	10%	14%	15%	5%	6%	6%	3%	4%	4%

Notes: Results are computed based on 100,000 independent paths generated from the Merton Jump-Diffusion model for the underlying (see Section 4.4.1.2 for model description). Three different sets of parameters values are considered with $\lambda = \{1, 0.25, 0.08\}$, $\mu_J = \{-0.05, -0.10, -0.20\}$ and $\sigma_J = \{0.05, 0.10, 0.15\}$ respectively for the jump scenario 1, 2, and 3 (see Table 4.1 for all parameters values). *Hedging instruments*: daily or monthly rebalancing with the underlying stock and 1-month or 3-months options with ATM calls and puts. Options used as hedging instruments are priced with implied volatility modeled with a log-AR(1) dynamics (see Section 4.4.1.4 for model description and Table 4.2 for parameters values). The training of neural networks for ERP and VO hedging is done as described in Section 4.4.1.5 and Section 4.6.1, respectively. The confidence level of the CVaR measure is $\alpha = 0.95$. $C_0^\star$ are expressed relative to $C_0^{(VO)}$ (% increase).

OTM puts, 11% to 42% for ATM and 6% to 20% for ITM options.[30] On the other hand, the relative increase in $C_0^\star$ as compared to $C_0^{(VO)}$ is much less sensitive to jump severity when trading 1-month options by ranging from scenario 1 to scenario 3 between 6% to 10% for OTM puts, 3% to 6% for ATM and 2% to 4% for ITM. Based on these results, we can assert that although both derivative valuation schemes are consistent with optimal trading criteria, the choice of hedging instrument and pricing procedure (hence implicitly of the treatment of hedging gains and losses) has a material impact on resulting derivative premiums and must

[30] The relative increase is computed as $\frac{C_0^\star}{C_0^{(VO)}} - 1$ for all examples.

thus be carefully chosen.

This smaller disparity between equal risk and variance-optimal prices with option hedges is in line with previous analyses: in the presence of jump or volatility risk, hedging with options entails significant reduction of the level market incompleteness as compared to trading solely the underlying stock. In such cases, premiums obtained with both derivative valuation approaches should be closer with the limiting case of being the same in a complete market.[31] These observations expand upon the work of Carbonneau and Godin (2021b), which shows that equal risk prices of puts obtained by hedging solely with the underlying stock are always larger than risk-neutral prices computed under convential change of measures. Indeed, benchmarking results presented in this current paper provide important novel insights into this price inflation phenomenon observed with the ERP framework: the disparity between equal risk and variance-optimal prices is always significantly smaller and less sensitive to stylized features of risky assets (e.g. jump or volatility risk) when option hedges are considered instead of trading exclusively the underlying stock.

Moreover, Table 4.8 presents benchmarking results of $C_0^{\star}$ to $C_0^{(VO)}$ with $\text{CVaR}_{0.90}, \text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ measures with 3-months option hedges. Values presented in this benchmarking demonstrate the ability of ERP, through the choice of convex risk measures, to span a large interval of prices which is close to encompass the variance-optimal premium. Indeed, under the $\text{CVaR}_{0.90}$ measure, we observe that $C_0^{\star}$ values are very close to $C_0^{(VO)}$ where the relative difference ranges between 0% and 3% across all moneynesses and jump risk scenarios. On the other hand, optimizing trading policies with more risk averse agents, i.e. with $\text{CVaR}_{0.95}$ or $\text{CVaR}_{0.99}$, provides a very wide range of derivative premiums with the ERP framework, especially for the OTM moneyness level. It is very interesting to note that this added flexibility

[31] To further illustrate this phenomenon, the authors also performed the same benchmarking with the Black-Scholes dynamics under which market incompleteness solely stems from discrete-time trading. The latter results are presented in the Section 4.6.3. Numerical values show that under the Black-Scholes dynamics, trading the underlying stock on a daily basis leads for most combinations of moneyness level and yearly volatility to the closest derivative premiums between ERP and VO procedures as compared to the other hedging instruments (see Table 4.11). Also, as expected under the Black-Scholes dynamics, daily stock hedging entails the smallest level of residual hedging risk across the different hedging instruments (see Table 4.10).

Table 4.8: Sensitivity analysis of equal risk prices $C_0^\star$ with $\text{CVaR}_{0.90}, \text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ measures to variance-optimal (VO) prices $C_0^{(VO)}$ under jump risk for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	OTM			ATM			ITM		
Jump Scenario	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
$C_0^{(VO)}$	1.58	1.83	2.08	4.64	4.82	4.97	10.11	10.15	10.21
$C_0^\star(\text{CVaR}_{0.90})$	3%	2%	0%	2%	2%	2%	2%	2%	2%
$C_0^\star(\text{CVaR}_{0.95})$	10%	14%	15%	5%	6%	6%	3%	4%	4%
$C_0^\star(\text{CVaR}_{0.99})$	32%	43%	55%	10%	12%	16%	6%	6%	8%

Notes: Results are computed based on 100,000 independent paths generated from the Merton Jump-Diffusion model for the underlying (see Section 4.4.1.2 for model description). Three different sets of parameters values are considered with $\lambda = \{1, 0.25, 0.08\}$, $\mu_J = \{-0.05, -0.10, -0.20\}$ and $\sigma_J = \{0.05, 0.10, 0.15\}$ respectively for jump scenario 1, 2, and 3 (see Table 4.1 for all parameters values). Hedging instruments consist of 3-months ATM calls and puts priced with implied volatility modeled as a log-AR(1) dynamics (see Section 4.4.1.4 for model description and Table 4.2 for parameters values). The training of neural networks for ERP and VO hedging is done as described in Section 4.4.1.5 and Section 4.6.1, respectively. $C_0^\star$ with $\text{CVaR}_{0.90}, \text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ are expressed relative to $C_0^{(VO)}$ (% increase).

of ERP procedures for pricing derivatives does not come at the expense of less effective hedging policies. Indeed, a major drawback of variance-optimal hedging lies in penalizing equally gains and losses through a quadratic penalty for hedging shortfalls. Conversely, the long and short trading policies solving the non-quadratic global hedging problems of the ERP framework are optimized to minimize a loss function which is possibly more in line with the financial objectives of the hedger by mainly (and most often exclusively) penalizing hedging losses, not gains.

4.5 Conclusion

This paper studies the equal risk pricing (ERP) framework for pricing and hedging European derivatives in discrete-time with multiple hedging instruments. The ERP approach sets derivative prices as the value such that the optimally hedged residual risk of the long and short positions in the contingent claim are equal. The ERP setup of Marzban et al. (2020) is considered where residual hedging risk is quantified through convex measures. The main objective of this current paper is in assessing the impact of including options within hedges on the equal risk price $C_0^\star$ and on the level of market incompleteness quantified by our $\epsilon^\star$-metrics. A specific focus is on the examination of the interplay between different stylized features of equity jump and volatility risks and the use of options as hedging instruments within the ERP framework. The numerical scheme of Carbonneau and Godin (2021b), which relies on the deep hedging algorithm of Buehler et al. (2019b), is used to solve the embedded global hedging problems of the ERP framework through the representation of the long and short trading policies with two distinct long-short term memory (LSTM) neural networks.

Sensitivity analyses with Monte Carlo simulations are performed under several empirically plausible sets of parameters for the jump and volatility risk models in order to highlight the impact of different stylized features of the models on $C_0^\star$ and $\epsilon^\star$. Numerical values indicate that in the presence of jump risk, hedging with options entails a significant reduction of both equal risk prices and market incompleteness as compared to hedging solely with the underlying stock. The latter stems from the fact that using options as hedging instruments rather than only the underlying stock shrinks the asymmetry of tail risk, which tends to both shrink option prices and reduce market incompleteness. On the other hand, in the presence of volatility risk, while option hedges can reduce equal risk prices as compared to stock hedges, the impact can be marginal and is highly sensitive to the moneyness level of the put option being priced as well as to the maturity of traded options. This can be explained by the fact that while the impact of jump risk dynamics considered in this paper is asymmetric by entailing significantly more weight on the right (resp. left) tail of the short (resp. long) hedging error through predominantly negative jumps, volatility risk impacts both upside and downside risk. Furthermore, additional experiments conducted show that the average price level of short-term options used as hedging instruments is effectively reflected into the equal risk price of longer-term options. The latter highlights the potential of the ERP framework as a fair valuation approach providing prices consistent with observable market prices. Thus, ERP could be applied for instance in the context of pricing over-the-counter derivatives with vanilla calls and puts hedges or pricing less liquid long-term derivatives (e.g. LEAPS contracts) with shorter-term liquid options.

Moreover, the benchmarking of equal risk prices to variance-optimal derivative premiums $C_0^{(VO)}$ is performed. The deep hedging algorithm is also used as the numerical scheme to solve the variance-optimal hedging problems. Numerical results show that while $C_0^{\star}$ tends to be larger than $C_0^{(VO)}$, trading options entails much smaller disparity between equal risk and variance-optimal prices as compared to trading only the underlying stock in the presence of jump or volatility risk. The latter is due to the market incompleteness being significantly smaller when option hedges are used to mitigate jump and volatility risks. Furthermore, additional experiments conducted demonstrate the ability of ERP to span a large interval of prices through the choice of convex risk measures, which is close to encompass the variance-optimal premium.

4.6 Appendix

4.6.1 Variance-optimal hedging

Denote $J^{(VO)} : \mathbb{R}^q \times \mathbb{R} \rightarrow \mathbb{R}$ as the cost function to be minimized for variance-optimal procedures:

$$J^{(VO)}(\theta, V_0) := \mathbb{E}\left[\left(\Phi(S_N^{(0,b)}) - B_N(V_0 + G_N^{\delta^\theta})\right)^2\right], \quad (\theta, V_0) \in \mathbb{R}^q \times \mathbb{R}, \tag{4.31}$$

where θ is the set of trainable parameters of the LSTM F_θ, V_0 is the initial capital investment and δ^θ is to be understood as the output sequence of F_θ. Let $\tilde{\theta} := \{\theta, V_0\}$ be the augmented set of trainable parameters which includes the initial portfolio value. Minibatch SGD with Monte Carlo sampling can naturally also be used to minimize (4.31) jointly for the trainable parameters and the initial capital investment by updating iteratively the augmented set $\tilde{\theta}$:

$$\tilde{\theta}_{j+1} = \tilde{\theta}_j - \eta_j \nabla_{\tilde{\theta}} \hat{J}^{(VO)}(\mathbb{B}_j, V_{0,j}), \tag{4.32}$$

where $\tilde{\theta}_0 := \{\theta_0, V_{0,0}\}$ is the initial set[32] and $\hat{J}^{(VO)}(\mathbb{B}_j, V_{0,j})$ is the empirical estimator of $J^{(VO)}(\theta, V_0)$ evaluated with the minibatch of hedging errors $\mathbb{B}_j = \{\Phi(S_{N,i}^{(0,b)}) - B_N(V_{0,j} + G_{N,i}^{\delta^{\theta_j}})\}_{i=1}^{N_{\text{batch}}}$ when $\tilde{\theta} = \tilde{\theta}_j$ (i.e. $\theta = \theta_j$ and $V_0 = V_{0,j}$):

$$\hat{J}^{(VO)}(\mathbb{B}_j, V_{0,j}) := \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \left(\Phi(S_{N,i}^{(0,b)}) - B_N(V_{0,j} + G_{N,i}^{\delta^{\theta_j}}) \right)^2. \tag{4.33}$$

4.6.2 Pseudo-code deep hedging

Algorithm 4.1 presents the pseudo-code to perform a one-step update of the trainable parameters as in (4.22) for the global hedging problems of the ERP framework, i.e. updating θ_j to θ_{j+1}. For convenience, the pseudo-code is presented for the case of trading exclusively the underlying stock and for the short position trading policy, but it is trivial to generalize to the case of trading other hedging instruments (e.g. short-term options) and for the long position trading policy. Note that the pseudo-code is described for the MJD dynamics, but it can be generalized to the GARCH dynamics by sampling log-returns from (4.26) in line (6), and adding the stochastic volatilities to feature vectors as described in Section 4.4.1.3. Furthermore, the pseudo-code can also easily be extended to variance-optimal hedging by updating the augmented set $\tilde{\theta}_j$ to $\tilde{\theta}_{j+1}$ with (4.32) instead of θ_j to θ_{j+1} in line (17) and by adapting the empirical cost function in line (15) to (4.33). Lastly, recall that a GitHub repository with samples of codes in Python for the training procedure of neural networks is available online: github.com/alexandrecarbonneau. The implementation replicates results of Table 4.3 with jump risk scenario 2, and can easily be adapted to reproduce all results presented in Section 4.4.

[32] As described in Section 4.4.1.4, an implied volatility dynamics is considered to price options used as hedging instruments. In numerical experiments of Section 4.4, $V_{0,0}$ is set at the price obtained with the time-0 implied volatility. The authors also tested the naive initialization scheme $V_{0,0} = 0$ as a robustness test, and found that the resulting variance-optimal premiums were marginally affected by this choice. Also, the Glorot uniform initialization of Glorot and Bengio (2010) is used to select θ_0.

Algorithm 4.1 Pseudo-code short trading policy with stock hedges under the MJD model
Input: θ_j
Output: θ_{j+1}

for $i = 1, \ldots, N_{\text{batch}}$ **do** $\triangleright$ Loop over each path of minibatch
  $X_{0,i} = [\log(S_{0,i}^{(0,b)}/K), V_{0,i}^{\delta}]$ $\triangleright$ Time-0 feature vector of $F_\theta^{(\mathcal{S})}$ with $V_{0,i}^{\delta} = 0$
  for $n = 0, \ldots, N-1$ **do**
    $Y_{n,i} \leftarrow$ time-t_n output of LSTM $F_\theta^{(\mathcal{S})}$ with $\theta = \theta_j$
    $\delta_{n+1,i}^{(0)} = Y_{n,i}$
    $y_{n+1,i} \sim$ (4.25) $\triangleright$ Sample next log-return
    $S_{n+1,i}^{(0,b)} = S_{n,i}^{(0,b)} e^{y_{n+1,i}}$
    $V_{n+1,i}^{\delta} = e^{r\Delta_N} V_{n,i}^{\delta} + \delta_{n+1,i}^{(0)} (S_{n+1,i}^{(0,b)} - e^{r\Delta_N} S_{n,i}^{(0,b)})$ $\triangleright$ See (4.4) for details
    $X_{n+1,i} = [\log(S_{n+1,i}^{(0,b)}/K), V_{n+1,i}^{\delta}]$ $\triangleright$ Time-t_{n+1} feature vector for $F_\theta^{(\mathcal{S})}$
  end for
  $\Phi(S_{N,i}^{(0,b)}) = \max(K - S_{N,i}^{(0,b)}, 0)$
  $\pi_{i,j} = \Phi(S_{N,i}^{(0,b)}) - V_{N,i}^{\delta}$
end for
$\widehat{\text{VaR}}_\alpha = \pi_{[\tilde{N}],j}$ $\triangleright$ $\tilde{N}^{\text{th}}$ ordered hedging error with $\tilde{N} := \lceil \alpha N_{\text{batch}} \rceil$
$\widehat{\text{CVaR}}_\alpha = \widehat{\text{VaR}}_\alpha + \frac{1}{(1-\alpha)N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \max(\pi_{i,j} - \widehat{\text{VaR}}_\alpha, 0)$
$\eta_j \leftarrow$ Adam algorithm
$\theta_{j+1} = \theta_j - \eta_j \nabla_\theta \widehat{\text{CVaR}}_\alpha$ $\triangleright$ $\nabla_\theta \widehat{\text{CVaR}}_\alpha$ computed with Tensorflow

Notes: Subscript i represents the i^{th} simulated path among the minibatch of size N_{batch}. Also, the time-0 feature vector is fixed for all paths, i.e. $S_{0,i}^{(0,b)} = S_0^{(0,b)}$ and $V_{0,i}^{\delta} = V_0^{\delta} = 0$.

4.6.3 Additional numerical experiments

This section contains supplementary material of the paper with additional numerical experiments of the ERP framework presented in Table 4.9 to Table 4.13.

Table 4.9: Sensitivity analysis of equal risk prices $C_0^\star$ and residual hedging risk $\epsilon^\star$ to implied volatility (IV) risk for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	OTM			ATM			ITM		
Long-run IV	14%	15%	16%	14%	15%	16%	14%	15%	16%
$\underline{C_0^\star}$									
1-month options	1.52	1.74	1.98	4.52	4.87	5.20	10.01	10.32	10.66
3-months options	1.86	2.07	2.28	4.75	5.08	5.39	10.08	10.38	10.66
$\underline{\epsilon^\star}$									
1-month options	0.61	0.65	0.66	0.96	1.00	0.96	0.99	1.03	1.01
3-months options	1.28	1.32	1.36	1.72	1.77	1.78	1.63	1.68	1.72
$\underline{\epsilon^\star / C_0^\star}$									
1-month options	0.40	0.37	0.34	0.21	0.20	0.18	0.10	0.10	0.09
3-months options	0.69	0.64	0.60	0.36	0.35	0.33	0.16	0.16	0.16

Notes: Results are computed based on 100,000 independent paths generated from the GJR-GARCH(1,1) model for the underlying with $\mu = 3.968\text{e-}04$, $\omega = 1.964\text{e-}06$, $\upsilon = 0.05$, $\gamma = 0.6$ and $\beta = 0.91$ which entails stationary yearly volatility of 15% (see Section 4.4.1.3 for model description). Hedging instruments consist of 1-month or 3-months options with ATM calls and puts. The latter options are priced with implied volatility modeled as a log-AR(1) dynamics with $\kappa = 0.15, \sigma_{\text{IV}} = 0.06$ and $\varrho = -0.6$ for all cases, and the long-run parameter ϑ set to $\log(0.14), \log(0.15)$ or $\log(0.16)$ (see Section 4.4.1.4 for the log-AR(1) model description). The training of neural networks is done as described in Section 4.4.1.5. The confidence level of the CVaR measure is $\alpha = 0.95$.

Table 4.10: Equal risk prices $C_0^\star$ and residual hedging risk $\epsilon^\star$ under the Black-Scholes model for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	OTM			ATM			ITM		
Annual volatility (σ)	10%	15%	20%	10%	15%	20%	10%	15%	20%
$C_0^\star$									
Daily stock	0.43	1.48	2.83	2.65	4.60	6.68	8.37	10.19	12.08
Monthly stock	0.49	1.67	3.15	2.71	4.70	6.72	8.56	10.42	12.39
1-month options	0.44	1.56	3.00	2.74	4.70	6.71	8.48	10.30	12.27
3-months options	0.44	1.59	3.08	2.81	4.80	6.82	8.58	10.47	12.51
$\epsilon^\star$									
Daily stock	0.19	0.41	0.65	0.38	0.64	1.00	0.39	0.67	0.88
Monthly stock	0.50	1.52	2.59	1.54	2.53	3.48	1.51	2.59	3.60
1-month options	0.19	0.57	0.94	0.57	0.91	1.30	0.64	1.01	1.37
3-months options	0.29	0.93	1.55	0.96	1.55	2.11	1.09	1.76	2.35
$\epsilon^\star/C_0^\star$									
Daily stock	0.44	0.28	0.23	0.14	0.14	0.15	0.05	0.07	0.07
Monthly stock	1.01	0.91	0.82	0.57	0.54	0.52	0.18	0.25	0.29
1-month options	0.44	0.37	0.31	0.21	0.19	0.19	0.07	0.10	0.11
3-months options	0.67	0.59	0.50	0.34	0.32	0.31	0.13	0.17	0.19

Notes: Results are computed based on 100,000 independent paths generated from the Black-Scholes model for the underlying with yearly parameters $\mu = 0.1$ and $\sigma = 0.1, 0.15$ and 0.20. *Hedging instruments*: daily or monthly rebalancing with the underlying stock and 1-month or 3-months options with ATM calls and puts. Options used as hedging instruments are priced with implied volatility modeled as a log-AR(1) dynamics with $\kappa = 0.15, \sigma_{\text{IV}} = 0.06$ and $\varrho = -0.6$ for all cases, and ϑ set to $\log(0.10), \log(0.15)$ and $\log(0.20)$ when $\sigma = 0.10, 0.15$ and 0.20, respectively (see Section 4.4.1.4 for log-AR(1) model description). The training of neural networks is done as described in Section 4.4.1.5. The confidence level of the CVaR measure is $\alpha = 0.95$.

Table 4.11: Equal risk prices $C_0^\star$ and variance-optimal (VO) prices $C_0^{(VO)}$ under the Black-Scholes model for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	OTM			ATM			ITM		
Annual volatility (σ)	10%	15%	20%	10%	15%	20%	10%	15%	20%
$C_0^{(VO)}$									
Daily stock	0.38	1.39	2.77	2.62	4.53	6.46	8.34	10.12	12.04
Monthly stock	0.35	1.36	2.73	2.55	4.47	6.39	8.27	10.07	11.98
1-month options	0.42	1.47	2.84	2.67	4.58	6.52	8.35	10.13	12.06
3-months options	0.39	1.44	2.84	2.65	4.56	6.50	8.32	10.12	12.04
$C_0^\star$									
Daily stock	12%	6%	2%	1%	2%	3%	0%	1%	0%
Monthly stock	40%	23%	15%	6%	5%	5%	3%	4%	3%
1-month options	5%	6%	6%	3%	3%	3%	2%	2%	2%
3-months options	11%	10%	9%	6%	5%	5%	3%	3%	4%

Notes: Results are computed based on 100,000 independent paths generated from the Black-Scholes model for the underlying with yearly parameters $\mu = 0.1$ and $\sigma = 0.1, 0.15$ and 0.20. *Hedging instruments*: daily or monthly rebalancing with the underlying stock and 1-month or 3-months options with ATM calls and puts. Options used as hedging instruments are priced with implied volatility modeled as a log-AR(1) dynamics with $\kappa = 0.15, \sigma_{\text{IV}} = 0.06$ and $\varrho = -0.6$ for all cases, and ϑ set to $\log(0.10), \log(0.15)$ and $\log(0.20)$ when $\sigma = 0.10, 0.15$ and 0.20, respectively (see Section 4.4.1.4 for log-AR(1) model description). The training of neural networks for ERP and VO hedging is done as described in Section 4.4.1.5 and Section 4.6.1, respectively. The confidence level of the CVaR measure is $\alpha = 0.95$. $C_0^\star$ are expressed relative to $C_0^{(VO)}$ (% increase).

Table 4.12: Equal risk prices $C_0^{\star}$ and variance-optimal (VO) prices $C_0^{(VO)}$ with volatility risk for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	OTM			ATM			ITM		
Stationary volatility	10%	15%	20%	10%	15%	20%	10%	15%	20%
$C_0^{(VO)}$									
Daily stock	0.78	1.88	3.24	2.94	4.75	6.60	8.29	9.98	11.81
Monthly stock	0.81	1.91	3.27	2.97	4.80	6.62	8.30	10.01	11.83
1-month options	0.50	1.60	3.01	2.75	4.66	6.60	8.32	10.10	12.04
3-months options	0.61	1.74	3.19	2.79	4.72	6.68	8.23	10.05	11.99
$C_0^{\star}$									
Daily stock	29%	25%	19%	10%	13%	10%	3%	6%	6%
Monthly stock	45%	39%	30%	13%	13%	14%	7%	8%	9%
1-month options	12%	9%	9%	4%	4%	4%	2%	2%	3%
3-months options	24%	19%	14%	8%	8%	7%	3%	3%	4%

Notes: Results are computed based on 100,000 independent paths generated from the GJR-GARCH(1,1) model for the underlying with three sets of parameters implying stationary yearly volatilities of $10\%, 15\%$ and 20% (see Section 4.1.3 for model description and Table 4 for parameters values). *Hedging instruments*: daily or monthly rebalancing with the underlying stock and 1-month or 3-months options with ATM calls and puts. Options used as hedging instruments are priced with implied volatility modeled as a log-AR(1) dynamics with $\kappa = 0.15, \sigma_{\text{IV}} = 0.06$ and $\varrho = -0.6$ for all cases, and ϑ set to $\log(0.10), \log(0.15)$ and $\log(0.20)$ when the GARCH stationary volatility is $10\%, 15\%$ and 20%, respectively (see Section 4.1.4 for log-AR(1) model description). The training of neural networks for ERP and VO hedging is done as described in Section 4.1.5 and Appendix A, respectively. The confidence level of the CVaR measure is $\alpha = 0.95$. $C_0^{\star}$ are expressed relative to $C_0^{(VO)}$ (% increase).

Table 4.13: Sensitivity analysis of equal risk prices $C_0^\star$ with $\text{CVaR}_{0.90}$, $\text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ measures to variance-optimal (VO) prices $C_0^{(VO)}$ under volatility risk for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	OTM			ATM			ITM		
Stationary volatility	10%	15%	20%	10%	15%	20%	10%	15%	20%
$C_0^{(VO)}$	0.61	1.74	3.19	2.79	4.72	6.68	8.23	10.05	11.99
$C_0^\star(\text{CVaR}_{0.90})$	6%	8%	6%	4%	4%	4%	2%	2%	2%
$C_0^\star(\text{CVaR}_{0.95})$	24%	19%	14%	8%	8%	7%	3%	3%	4%
$C_0^\star(\text{CVaR}_{0.99})$	81%	49%	37%	17%	16%	15%	6%	6%	7%

Notes: Results are computed based on 100,000 independent paths generated from the GJR-GARCH(1,1) model for the underlying with three sets of parameters implying stationary yearly volatilities of 10%, 15% and 20% (see Section 4.1.3 for model description and Table 4 for parameters values). Hedging instruments consist of 3-months ATM calls and puts priced with implied volatility modeled as a log-AR(1) dynamics with $\kappa = 0.15, \sigma_{\text{IV}} = 0.06$ and $\varrho = -0.6$ for all cases, and ϑ set to $\log(0.10), \log(0.15)$ and $\log(0.20)$ when the GARCH stationary volatility is 10%, 15% and 20%, respectively (see Section 4.1.4 for log-AR(1) model description). The training of neural networks for ERP and VO hedging is done as described in Section 4.1.5 and Appendix A, respectively. $C_0^\star$ with $\text{CVaR}_{0.90}$, $\text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$ are expressed relative to $C_0^{(VO)}$ (% increase).

Chapter 5

Deep equal risk pricing of financial derivatives with non-translation invariant risk measures

Abstract

The use of non-translation invariant risk measures within the equal risk pricing (ERP) methodology for the valuation of financial derivatives is investigated. The ability to move beyond the class of convex risk measures considered in several prior studies provides more flexibility within the pricing scheme. In particular, suitable choices for the risk measure embedded in the ERP framework such as the semi-mean-square-error (SMSE) are shown herein to alleviate the price inflation phenomenon under Tail Value-at-Risk based ERP as documented for instance in Carbonneau and Godin (2021b). The numerical implementation of non-translation invariant ERP is performed through deep reinforcement learning, where a slight modification is applied to the conventional deep hedging training algorithm (see Buehler et al., 2019b) so as to enable obtaining a price through a single training run for the two neural networks associated with the long and short hedging strategies. The accuracy of the neural network training procedure is shown in simulation experiments not to be materially impacted by such modification of the training algorithm.

Keywords: Option pricing, Optimal hedging, Reinforcement learning, Deep learning.

5.1 Introduction

The equal risk pricing (ERP) methodology for derivatives valuation, which was initially proposed by Guo and Zhu (2017), entails setting the price of a contingent claim as the initial hedging portfolio value which leads to equal residual hedging risk for both the long and short positions under optimal hedges. This pricing procedure is associated with numerous advantageous properties, such as the production of prices that are arbitrage-free under some technical conditions (see Guo and Zhu, 2017; Marzban et al., 2020; Carbonneau and Godin, 2021b), consistency with non-myopic global dynamic optimal hedging strategies, invariance of the price with respect to the position considered (i.e. long versus short), and the ability to consider general risk measures[1] for the objective function of the hedging optimization problem.

To further improve the ERP framework, several subsequent studies proposed some modifications to the original scheme. For instance, Marzban et al. (2020) and Carbonneau and Godin (2021b) use the physical probability measure rather than the risk-neutral one to perform hedging optimization; this has the advantage of improved interpretability of resulting prices on top of removing the subjectivity associated with the choice of the risk-neutral measure in an incomplete market setting. Furthermore, to enhance the computational tractability of the ERP approach, these two studies also considered the set of convex risk measures to represent the risk exposure of hedged transaction for both long and short parties.[2] Indeed, when convex measures are used, the translation invariance property leads to a useful characterization of equal risk prices which removes the need to perform a joint optimization over all possible values of the initial hedging portfolio.

The most natural convex risk measure to consider within the ERP approach is arguably

[1]For instance, the ability to depart from the quadratic penalty considered in the celebrated variance-optimal approach of Schweizer (1995) enables avoiding adverse behavior associated with the penalization of hedging gains.

[2]The original work from Guo and Zhu (2017) considers expected penalties as risk measures, which do not possess all properties of convex risk measures (e.g. most lack the translation invariance property). For instance, the Tail-Value-at-Risk (TVaR) is not a particular case of an expected penalty.

the Conditional Value-at-Risk (CVaR), which is equivalent to the Expected Shortfall (ES) or Tail-Value-at-Risk under the assumption that underlying loss variables are absolutely continuous. See Rockafellar and Uryasev (2002) for a formal definition of the CVaR and a description of its properties. The CVaR_α can be interpreted as the operator computing a probability weighted average of worst-case risks occurring within an event of probability below or exactly $1 - \alpha$, which is very intuitive. Moreover, it is a coherent risk measure in the sense of Artzner et al. (1999), which implies favorable properties from a risk measurement standpoint.[3] Furthermore, the CVaR measure is used extensively in practice by the financial sector to quantify capital requirements, see for instance BCBS (2016).

Due to its favorable properties, several studies used the CVaR within the ERP framework: see Carbonneau and Godin (2021b) and Carbonneau and Godin (2021a). It was observed in the foremost that when only the underlying asset is used to hedge put options and conventional risk-neutral measures are used to determine the initial capital for hedging, the tail risk is much more pronounced for the short position than for the long one, especially for out-of-the-money puts. This leads to equal risk prices that are substantially higher than their risk-neutral counterparts when the confidence level α of the CVaR_α is high, to an extent that sheds doubts on the applicability of the method in practice. An avenue that was explored to remedy this drawback is to reduce the confidence level as prices were shown numerically to be positively related to the latter. Unfortunately, as shown in this current paper, reducing the confidence level to obtain smaller option prices becomes quickly impractical since the resulting hedging strategies exhibit poor risk mitigation performance with speculative behavior magnifying tail losses for very high quantiles above the CVaR confidence level. This approach should therefore not be pursued in practice. A second possible solution to the inflated ERP prices issue which is explored in Carbonneau and Godin (2021a) consists in incorporating other hedging instruments (e.g. short-term options) within dynamic hedging schemes. That approach is

[3]The class of coherent risk measures is a subset of the class of convex risk measure which assumes for instance the subadditivity and positive homogeneity properties; the latter are more stringent than the convexity property satisfied by all convex risk measures.

shown therein to produce prices that are often still higher than the traditional risk-neutral ones, but much closer to them. This avenue was thus deemed successful when applicable. However, it requires a more sophisticated model to represent the price dynamics of hedging instruments, which complicates its implementation in practice. Furthermore, hedges relying on option trades might not be feasible or desirable under some circumstances (e.g. lack of liquidity).

The aforementioned simulation-based results on ERP prices highlight the need to identify an ERP approach which can strictly rely on the underlying asset for hedging transactions and, at the same time, which can alleviate the price inflation in comparison to prices obtained from conventional pricing approaches. A straightforward route to explore so as to attempt obtaining a satisfactory ERP method respecting the above constraints is to modify the risk measure acting as the objective function in the optimal hedging problems underlying the ERP framework. For instance, risk measures putting less relative weight on tail risk and more on more moderate risk scenarios should produce lesser option prices. However, such risk measures (e.g. the semi-variance, semi-root-mean-square-error (SRMSE), etc.) do not necessarily satisfy properties of convex risk measures, in particular the translation invariance property. Equal risk prices stemming from such risk measure choices therefore do not have the convenient characterization associated with convex risk measures, which highlights the need of tailor-made numerical procedures handling this additional complexity.

The main contribution of this manuscript is twofold. The first is to propose a modification of the deep reinforcement learning approach illustrated in Carbonneau and Godin (2021b) and Carbonneau and Godin (2021a) to handle non-translation invariant risk measures within ERP naturally and without excessive additional computational burden. This modification essentially consists in feeding varying initial hedging portfolio values to simulated underlying asset paths to the deep hedging algorithm from Buehler et al. (2019b), and then coupling the trained neural network output with a bisection search to seek the initial hedging portfolio value equating risks for both the long and short positions. The latter bisection method search has previously been suggested in a similar context for instance in Marzban et al. (2020). The training algorithm modification is shown in the current work not to lead to a material deterioration in the hedging performance of the neural network underlying the numerical approach. The second contribution consists in exploring equal risk prices of options generated when using typical non-translation invariant risk measures. It is seen that the use of the class of semi-$\mathbb{L}^p$ risk measures of the form $L(x) = x^p \mathbb{1}_{\{x>0\}}$ for $p > 0$ is able to reduce ERP prices to more natural levels better in line with these of existing methodologies while simultaneously resulting in effective trading policies. Indeed, numerical results indicate that equal risk prices generated by the class of semi-$\mathbb{L}^p$ risk measures can span much more than the interval of prices obtained under the CVaR_α risk measures with conventional confidence α level values. The latter observation is shown to hold across all moneyness levels for puts, and is robust to all risky asset dynamics considered. Furthermore, the benchmarking of neural networks trading policies hedging performance demonstrates that optimized policies under the semi-$\mathbb{L}^p$ objective functions are effective for mitigating hedging risk across all values of p considered, where p is shown to control the relative weight associated to extreme hedging losses. This is in contrast with the CVaR_α objective function where hedging policies optimized with relatively small confidence level α exhibit poor risk mitigation for quantile losses larger than α. Lastly, our results show that the use of the semi-$\mathbb{L}^2$ objective function to price long-term European puts with trades involving exclusively the underlying stock is almost as successful to reduce equal risk price values as compared to values obtained by trading shorter-term options with the CVaR_α risk measure. All of these results clearly demonstrate the benefit of using the class of semi-$\mathbb{L}^p$ risk measures within the ERP framework by simultaneously alleviating the price inflation phenomenon observed under the class of CVaR measures as well as resulting in effective trading policies for risk management.

This paper is divided as follows. Section 5.2 provides a literature review about incomplete market derivatives pricing, hedging methods and reinforcement learning in finance. The theoretical setting used for the ERP approach in the current work is presented in Section 5.3.

Section 5.4 explains the reinforcement learning methodology for neural networks embedded in the ERP approach with the modified training algorithm proposed in this paper. Section 5.5 displays results of numerical experiments associated with ERP based semi-$\mathbb{L}^p$ risk measures. Section 5.6 concludes.

5.2 Literature review

Financial derivatives pricing in incomplete markets has received an extensive amount of attention in the literature. Numerous papers approach this problem through the selection of a suitable risk-neutral measure based on various considerations such as shifting of the drift to achieve risk-neutrality and model invariance, see Hardy (2001) and Christoffersen et al. (2010), consistency with equilibrium models, see Gerber and Shiu (1994) and Duan (1995), or minimum entropy distance between the physical and risk-neutral measures, see Frittelli (2000). Another strand of literature considers pricing methods consistent with optimal hedging strategies. At first, quadratic hedging methods were considered in Föllmer and Schweizer (1988), Schweizer (1995), Elliott and Madan (1998) and Bertsimas et al. (2001) due to their tractability. However, as a consequence of the limitations associated to the quadratic penalty (e.g. penalizing equally gains and losses), other objective functions were considered in alternative dynamic hedging schemes such as quantile hedging (Föllmer and Leukert, 1999), expected penalty minimization (Föllmer and Leukert, 2000) or VaR and CVaR optimization as in Melnikov and Smirnov (2012) and Godin (2016). Some pricing schemes were also developed to enable consistency with non-quadratic hedging methods, for instance utility indifference (Hodges and Neuberger, 1989) or risk indifference (Xu, 2006). An issue with the latter approaches is that different prices are obtained depending on if a long or short position is considered in the derivative. The ERP approach developed by Guo and Zhu (2017) identifying the derivative price equating hedged risk exposure of both long and short positions remedies this drawback by providing a unique price invariant to the direction (i.e. long versus short) of the position. Several additional papers have used or expanded on the initial ERP methodology. One problem often considered by that methodology is the tackling of market incompleteness arising from short-selling bans on the underlying asset: Alfeus et al. (2019), Ma et al. (2019) and He and Zhu (2020). Marzban et al. (2020) propose to substitute the risk-neutral measure for the physical measure during the determination of the equal risk price and to replace expected loss functions by convex risk measures within the objective function. Carbonneau and Godin (2021b) provide a tractable methodology based on deep reinforcement learning to implement the ERP framework with convex risk measures under very general conditions. Carbonneau and Godin (2021a) examine the impact of introducing options as hedging instruments within the ERP framework under convex risk measures.

The computation of equal risk prices for derivatives is a highly non-trivial endeavor requiring advanced numerical schemes in most cases. Marzban et al. (2020) propose to use dynamic programming which they apply on a robust optimization setting. Conversely, Carbonneau and Godin (2021b) and Carbonneau and Godin (2021a) use the deep reinforcement learning approach of Buehler et al. (2019b) coined as *deep hedging*. Other papers have relied on the deep hedging methodology for the hedging of financial derivatives: Cao et al. (2020), Carbonneau (2021) and Horvath et al. (2021). Deep reinforcement learning is a very favorable technique for multistage optimization and decision-making in financial contexts: it allows tackling high-dimensional settings with multiple state variables, underlying asset dynamics and trading instruments. For this reason, it was used in multiple other works on derivatives pricing and hedging. Various techniques were considered such as Q-learning in Halperin (2020) and Cao et al. (2021), least squares policy iteration and fitted Q-iteration for American option pricing in Li et al. (2009), or batch policy gradient in Buehler et al. (2019b). Moreover, various other financial problems were tackled through reinforcement learning procedures in the literature, for instance portfolio management as in Moody and Wu (1997), Jiang et al. (2017), Pendharkar and Cusatis (2018), García-Galicia et al. (2019), Wang and Zhou (2020), Ye et al. (2020) and Betancourt and Chen (2021), optimal liquidation, see Bao and Liu (2019), or trading optimization as in Hendricks and Wilcox (2014), Lu (2017) and Ning et al. (2018).

5.3 Financial market setup

This section details the mathematical framework for the financial market considered along with the theoretical setup for the ERP approach for derivatives valuation.

A discrete set of equally spaced time points spanning a horizon of T years $\mathcal{T} \equiv \{0 = t_0 < t_1 < \ldots < t_N = T\}$ with $t_n \equiv n\Delta$, $n = 0, \ldots, N$ is considered. Δ corresponds to the length of a time period in years. Unless specified otherwise, the current study uses either $\Delta = 1/260$ or $\Delta = 1/12$ corresponding to daily or monthly periods. Moreover, consider the probability space $(\Omega, \mathcal{F}_N, \mathbb{P})$ endowed with a filtration $\mathbb{F} \equiv \{\mathcal{F}_n\}_{n=0}^N$ satisfying the usual conditions, with $\mathcal{F}_n$ being the sigma-algebra characterizing the information available to the investor at time t_n. Multiple traded assets are introduced in the financial market. First, a risk-free asset grows at a constant periodic risk-free rate $r \in \mathbb{R}$: its time t_n price is given by $B_n \equiv e^{rt_n}$. The $D+1$ other non-dividend paying risky asset prices are characterized by the vectorial stochastic processes $\{S_n^{(b)}\}_{n=0}^N$ and $\{S_n^{(e)}\}_{n=0}^{N-1}$ where $S_n^{(b)} \equiv \left[S_n^{(0,b)}, \ldots, S_n^{(D,b)}\right]$ and $S_n^{(e)} \equiv \left[S_n^{(0,e)}, \ldots, S_n^{(D,e)}\right]$ respectively represent the *beginning-of-period* and *end-of-period* prices of risky assets $0, \ldots, D$ available for trading at time t_n. This implies $S_n^{(b)}$ is $\mathcal{F}_n$-measurable (i.e. observable at time t_n) whereas $S_n^{(e)}$ is $\mathcal{F}_{n+1}$-measurable. Due to traded instruments changing on every time period (for example, some traded options mature contracts need to be rolled-over), it is possible to have $S_n^{(j,e)} \neq S_{n+1}^{(j,b)}$, $j = 1, \ldots, D$. However, the risky asset $j = 0$ is assumed to be an underlying asset with no maturity such as a stock, thus available for trading on all periods. Hence, $S_n^{(0,e)} = S_{n+1}^{(0,b)}$. For simplicity, an absence of market frictions is assumed throughout the paper. Correspondingly, it is assumed all positions in a given portfolio are liquidated at the end of any period, and are repurchased at the beginning of the next if needed.

A European-type derivative of time t_N payoff $\Phi\left(S_N^{(0,b)}\right)$ is considered. A suitable price for that contract and corresponding hedging strategies must be determined. We define a trading strategy $\delta \equiv \{\delta_n\}_{n=0}^N$ as an $\mathbb{F}$-predictable process, i.e. δ_0 is $\mathcal{F}_0$-measurable and δ_n is $\mathcal{F}_{n-1}$-

measurable for $n = 1, \ldots, N$, where $\delta_n \equiv \left[\delta_n^{(0)}, \ldots, \delta_n^{(D)}, \delta_n^{(B)}\right]$. The latter comprises $\delta_n^{(0:D)} \equiv \left[\delta_n^{(0)}, \ldots, \delta_n^{(D)}\right]$ which contains the positions in all respective risky assets $0, \ldots, D$ within the portfolio between time t_{n-1} and time t_n, and $\delta_n^{(B)}$ which contains the portfolio investment in the risk-free asset for the same period. For a trading strategy δ, the corresponding time t_n portfolio value is defined as

$$V_n^\delta \equiv \begin{cases} \delta_0^{(0:D)} \bullet S_0^{(b)} + \delta_0^{(B)} B_0, & n = 0, \\ \delta_n^{(0:D)} \bullet S_{n-1}^{(e)} + \delta_n^{(B)} B_n, & n = 1, \ldots, N, \end{cases}$$

where $\bullet$ is the conventional scalar product, i.e. for two n-dimensional vectors X and Y, $X \bullet Y := \sum_{i=1}^n X_i Y_i$. Also, a trading strategy δ is said to be *self-financing* if

$$\delta_{n+1}^{(0:D)} \bullet S_n^{(b)} + \delta_{n+1}^{(B)} B_n = V_n^\delta, \quad n = 0, \ldots, N-1.$$

Denote by Π the set of all self-financing trading strategies that are sufficiently well-behaved mathematically.[4] It turns out that the portfolio value process of self-financing trading strategies can be expressed conveniently in terms of so-called *discounted gains*. For a trading strategy $\delta \in \Pi$, the latter is defined as

$$G_0^\delta \equiv 0, \quad G_n^\delta \equiv \sum_{j=1}^{n} \delta_j^{(0:D)} \bullet \left(B_j^{-1} S_{j-1}^{(e)} - B_{j-1}^{-1} S_{j-1}^{(b)}\right), \quad n = 1, \ldots, N.$$

Using standard arguments outlined for instance in Lamberton and Lapeyre (2011), for any self-financing trading strategy $\delta \in \Pi$,

$$V_n^\delta = B_n \left(V_0^\delta + G_n^\delta\right).$$

Such representation is convenient as it allows avoiding calculating $\delta_n^{(B)}$ for $n = 0, \ldots, N$

[4]Details characterizing well-behavedness in the context of the current study are omitted to avoid lengthy discussions straying us away from the main research objectives of the current work.

explicitly when calculating the portfolio value.

Aforementioned definitions allow posing the main optimization problems underlying the ERP methodology, which consist in finding the best self-financing trading strategies leading to optimal hedges in terms of penalized hedging errors at the maturity of the derivative. Such problems are referred to as *global hedging procedures* due to their measurement of hedging efficiency in terms of risk at maturity rather than on a period by period basis. Consider a given risk measure ρ characterizing the risk aversion of the hedger. A risk measure is a mapping taking a random variable representing a random loss as input, and return a real number representing its perceived risk as an output. Specific examples of risk measures considered in this study are formally defined subsequently. For a given value of $V_0 \in \mathbb{R}$, define mappings $\epsilon^{(\mathcal{L})} : \mathbb{R} \to \mathbb{R}$ and $\epsilon^{(\mathcal{S})} : \mathbb{R} \to \mathbb{R}$ representing optimal residual hedging risk respectively for a long or short position in the derivative when the initial portfolio value is $V_0^\delta = V_0$ as

$$\epsilon^{(\mathcal{L})}(V_0) \equiv \min_{\delta \in \Pi} \rho\left(-\Phi(S_N^{(0,b)}) - V_N^\delta\right), \quad \epsilon^{(\mathcal{S})}(V_0) \equiv \min_{\delta \in \Pi} \rho\left(\Phi(S_N^{(0,b)}) - V_N^\delta\right). \tag{5.1}$$

Optimal hedging strategies are the minimizing arguments of such optimization problems:

$$\delta^{(\mathcal{L})}(V_0) \equiv \underset{\delta \in \Pi}{\arg\min}\, \rho\left(-\Phi(S_N^{(0,b)}) - V_N^\delta\right), \quad \delta^{(\mathcal{S})}(V_0) \equiv \underset{\delta \in \Pi}{\arg\min}\, \rho\left(\Phi(S_N^{(0,b)}) - V_N^\delta\right).$$

This leads to the definition of the *equal risk price* C_0^* of the derivative Φ as the initial portfolio value V_0 such that the optimal residual hedging risk is equal for both the long and short positions, i.e.

$$\epsilon^{(\mathcal{L})}(-C_0^*) = \epsilon^{(\mathcal{S})}(C_0^*). \tag{5.2}$$

Such optimal residual risk exposure when the initial portfolio value is the equal risk price, i.e. $V_0 = C_0^*$, is referred to as the *measured residual risk exposure* and denoted as $\epsilon^* \equiv \epsilon^{(\mathcal{L})}(-C_0^*) = \epsilon^{(\mathcal{S})}(C_0^*)$. Conditions on ρ have to be imposed to guarantee the existence and uniqueness of the equal risk price (e.g. monotonicity of ρ). Under the assumption that ρ is a convex risk measure, Carbonneau and Godin (2021b) provide sufficient conditions to obtain existence and uniqueness of the solution to (5.2), see Theorem 2.1 of the latter paper.

Remark 5.1. *Under a convex measure ρ, Marzban et al. (2020) and Carbonneau and Godin (2021b) also obtain the following characterization of the equal risk price*

$$C_0^* = 0.5 B_N \left(\epsilon^{(\mathcal{S})}(0) - \epsilon^{(\mathcal{L})}(0)\right). \tag{5.3}$$

Representation (5.3) is very convenient as it requires to only obtain the optimal residual risk exposure when the initial portfolio is null instead of having to iteratively try multiple initial portfolio values. However, when ρ is not translation invariant, such representation does not hold anymore, and a tailor-made numerical scheme must thus be developed to solve for the root-finding problem (5.2).

The current work aims among others at examining a class of non-translation invariant risk measures. The main class of risk measures under study in the current paper will be referred to as the *semi-$\mathbb{L}^p$ risk measures*, which are defined as

$$\rho(X) \equiv \mathbb{E}\left[X^p \mathbb{1}_{\{X>0\}}\right], \quad p > 0. \tag{5.4}$$

The latter risk measure is clearly monotonous (i.e. $X \geq Y$ almost surely implies $\rho(X) \geq \rho(Y)$), but lacks the translation invariance property. Furthermore, the parameter p acts as a risk aversion barometer as higher values of p put more relative weight on higher losses.

The CVaR measure is also considered in some experiments of the current paper for benchmarking purposes as it is used in Carbonneau and Godin (2021b) and Carbonneau and Godin (2021a). Such a risk measure can be formally defined as

$$\text{VaR}_\alpha(X) \equiv \inf\{x : \mathbb{P}[X \leq x] \geq \alpha\}, \quad \text{CVaR}_\alpha(X) \equiv \frac{1}{1-\alpha}\int_\alpha^1 \text{VaR}_\gamma(X) d\gamma$$

for a confidence level α in $(0, 1)$. Whenever X is an absolutely continuous random variable, the CVaR admits the intuitive representation $\text{CVaR}_\alpha(X) = \mathbb{E}[X|X \geq \text{VaR}_\alpha(X)]$.

5.4 Methodology

The current section details the reinforcement learning approach followed to solve the optimization problems underlying the ERP methodology. The approach consists in applying the deep hedging algorithm of Buehler et al. (2019b) by representing hedging policies with neural networks. A slight modification to the latter paper's training methodology is required to solve the ERP global hedging problems when the risk measure is not translation invariant. An accuracy assessment is performed for the modified training algorithm.

5.4.1 Neural network approximation of the optimal solution

The approach followed to obtain a numerical solution to the optimization problems (5.1) is based on a parametric approximation of the trading policy with a neural network trained using reinforcement learning. The general idea is as follows. In multiple setups, especially those involving Markovian dynamics, the optimal trading strategies $\delta^{(\mathcal{S})}(V_0)$ and $\delta^{(\mathcal{L})}(V_0)$ often admit the following functional representation for some functions $\tilde{\delta}^{(\mathcal{L})}$ and $\tilde{\delta}^{(\mathcal{S})}$

$$\delta^{(\mathcal{L})}_{n+1}(V_0) = \tilde{\delta}^{(\mathcal{L})}\left(T - t_n, S^{(b)}_n, V_n, \mathcal{I}_n\right), \quad \delta^{(\mathcal{S})}_{n+1}(V_0) = \tilde{\delta}^{(\mathcal{S})}\left(T - t_n, S^{(b)}_n, V_n, \mathcal{I}_n\right), \quad n = 0, \ldots, N-1, \tag{5.5}$$

where $\delta^{(\mathcal{L})}_{n+1}(V_0)$ and $\delta^{(\mathcal{S})}_{n+1}(V_0)$ are to be understood as the optimal time t_n hedge for the long and short position with time 0 capital investment V_0 and $\mathcal{I}_n$ is a $\mathcal{F}_n$-measurable random vector containing a set of additional state variables summarizing all necessary information to make the optimal portfolio rebalancing decision. For instance, $\mathcal{I}_n$ can contain underlying asset volatilities if the latter has a GARCH dynamics (see Augustyniak et al., 2017), current probabilities of being in the various respective regimes when in a regime-switching setup (see François et al., 2014), implied volatilities when options are used as hedging instruments (see Carbonneau and Godin, 2021a), current assets positions when in the presence of transaction costs (see Breton and Godin, 2017), and so on.

The functional representation (5.5) enables the approximation of the optimal policies as parameterized functions. The class of functions considered in this paper is the classical *feedforward neural network* (FFNN) class, which is formally defined subsequently. Indeed, two distinct FFNNs are used to approximate the optimal trading policy of the long and short parties by mapping inputs $\{T - t_n, S_n^{(b)}, V_n, \mathcal{I}_n\}$ into the respective (long or short) portfolio positions of risky assets $\delta_{n+1}^{(0:D)}$ for any $n = 1, \ldots, N-1$.[5] More precisely, denote by $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$ the neural network mappings for respectively the long and short trading positions where $\theta \in \mathbb{R}^q$ is the q-dimensional set of parameters of the FFNNs.[6] For a given parameter set θ, the associated trading strategies are given by

$$\delta_{n+1}^{(\mathcal{L},\theta)}(V_0) \equiv F_\theta^{(\mathcal{L})}\left(T - t_n, S_n^{(b)}, V_n, \mathcal{I}_n\right), \quad \delta_{n+1}^{(\mathcal{S},\theta)}(V_0) \equiv F_\theta^{(\mathcal{S})}\left(T - t_n, S_n^{(b)}, V_n, \mathcal{I}_n\right), \quad n = 0, \ldots, N-1.$$

The optimization of trading strategy in problem (5.1) is thus replaced by the optimization of neural network parameters θ according to

$$\tilde{\epsilon}^{(\mathcal{L})}(V_0) \equiv \min_{\theta \in \mathbb{R}^q} \rho\left(-\Phi(S_N^{(0,b)}) - V_N^{\delta^{(\mathcal{L},\theta)}}\right), \quad \tilde{\epsilon}^{(\mathcal{S})}(V_0) \equiv \min_{\theta \in \mathbb{R}^q} \rho\left(\Phi(S_N^{(0,b)}) - V_N^{\delta^{(\mathcal{S},\theta)}}\right). \qquad (5.6)$$

Note that the set of optimal parameters θ will be different for the long and the short trading strategies. Furthermore, problems (5.6) only lead to an approximate solution to the initial problems (5.1) since the FFNNs are approximations of the true functional representation $\tilde{\delta}^{(\mathcal{L})}$ and $\tilde{\delta}^{(\mathcal{S})}$. Nevertheless, by relying on the universal approximation property of FFNNs (see for instance Hornik, 1991), Buehler et al. (2019b) show that there exist neural networks such that the solution $\tilde{\epsilon}^{(\mathcal{L})}, \tilde{\epsilon}^{(\mathcal{S})}$ from (5.6) can be made arbitrarily close to the solution $\epsilon^{(\mathcal{L})}, \epsilon^{(\mathcal{S})}$

[5]Recall that since the trading strategy is self-financing, $\delta_{n+1}^{(B)}$ is characterized by $\delta_{n+1}^{(0:D)}$ and V_n.

[6] It is worth highlighting that while the neural network architecture of $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$ considered in this paper is the same for both neural networks in terms of the number of hidden layers and neurons per hidden layer, and thus the total number of parameters to fit q is the same for both neural networks, one could also consider two different architectures for $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$ with no additional complexity.

from (5.1).

The mathematical definition of FFNNs architecture is now provided. For $L, d_0, \ldots, d_{L+1} \in \mathbb{N}$, let $F_\theta : \mathbb{R}^{d_0} \to \mathbb{R}^{d_{L+1}}$ be a FFNN:

$$\begin{aligned} F_\theta(X) &\equiv o \circ h_L \circ \ldots \circ h_1, \\ h_l(X) &\equiv g(W_l X + b_l), \quad l = 1, \ldots, L, \\ o(X) &\equiv W_{L+1} X + b_{L+1}, \end{aligned}$$

where $\circ$ denotes the function composition operator. Thus, F_θ is a composite function of $h_1, \ldots, h_L$ commonly known as *hidden layers* which each apply successively an affine and a nonlinear transformation to input vectors, and also of the *output function* o applying an affine transformation to the last hidden layer. The set of parameters θ to be optimized consists of all weight matrices $W_l \in \mathbb{R}^{d_l \times d_{l-1}}$ and bias vectors $b_l \in \mathbb{R}^{d_l}$ for $l = 1, \ldots, L + 1$.

5.4.2 Calibration of neural networks through reinforcement learning

As in Buehler et al. (2019b), the training of neural networks in this paper relies on a stochastic policy gradient algorithm, also known as actor-based reinforcement learning. This class of procedures optimizes directly the policy (i.e. the actor) parameterized as a neural network with minibatch stochastic gradient descent (SGD) so as to minimize a cost function as in (5.6). Without loss of generality, the training algorithm is only provided for the neural network $F_\theta^{(\mathcal{S})}$ associated with the short position, as steps for the long position are entirely analogous.

5.4.2.1 Fixed and given V_0 case

The training procedure to calibrate θ is first described for a fixed and given initial capital investment V_0 as originally considered in Buehler et al. (2019b). A slight modification to the algorithm will subsequently be presented in Section 5.4.2.2 to tackle the non-translation invariant risk measure case studied in this paper. Let $J : \mathbb{R}^q \times \mathbb{R} \to \mathbb{R}$ be the cost function for the short position hedge:

$$J(\theta, V_0) \equiv \rho\left(\Phi(S_N^{(0,b)}) - V_N^{\delta^{(\mathcal{S},\theta)}}\right), \quad \theta \in \mathbb{R}^q, V_0 \in \mathbb{R}. \tag{5.7}$$

The parameters set θ is sequentially refined to produce a sequence of estimates $\{\theta_j\}_{j\geq 1}$ minimizing the cost function J over time. This iterative procedure is as follows. First, parameters of the neural network are initialized with the Glorot uniform initialization of Glorot and Bengio (2010), which gives the initial value of the sequence θ_0. Then, to start refining the parameters, a set of $M = 400{,}000$ paths containing traded asset values and other exogenous variables associated with the assets dynamics is generated with a Monte-Carlo simulation. The set of such paths is referred to as a *training set*. On each iteration of SGD, i.e. on each update of θ_j to θ_{j+1}, a minibatch consisting in a subset of size $N_{\text{batch}} = 1{,}000$ of paths from the training set is used to estimate the cost function in (5.7). More precisely, for $\theta = \theta_j$, $F_\theta^{(\mathcal{S})}$ is used to compute the assets positions at each rebalancing date and for each path within the minibatch. Let $\mathbb{B}_j \equiv \{\pi_{i,j}\}_{i=1}^{N_{\text{batch}}}$ be the resulting set of hedging errors from this minibatch, where $\pi_{i,j}$ is the ith hedging error when $\theta = \theta_j$. Then, for $\hat{\rho} : \mathbb{R}^{N_{\text{batch}}} \to \mathbb{R}$ the empirical estimator of $\rho(\pi)$ evaluated with $\mathbb{B}_j$, the update rule for θ_j to θ_{j+1} is

$$\theta_{j+1} = \theta_j - \eta_j \nabla_\theta \widehat{\rho}(\mathbb{B}_j),$$

where $\{\eta_j\}_{j\geq 1}$ are small positive real values and ∇_θ denotes the gradient operator with respect to θ. For instance, under the semi-$\mathbb{L}^p$ class of risk measures which is extensively studied in the numerical section, the empirical estimator has the representation

$$\widehat{\rho}\,(\mathbb{B}_j) \equiv \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \pi_{i,j}^p \mathbb{1}_{\{\pi_{i,j}>0\}}.$$

Lastly, the computation of the gradient of the empirical cost function with respect to θ can be done explicitly with modern deep learning libraries such as Tensorflow (Abadi et al., 2016) and with the Adam optimizer (Kingma and Ba, 2014) which dynamically update the η_j values. The following section presents the modification to the training algorithm proposed in this paper to compute equal risk prices under non-translation invariant risk measures.

5.4.2.2 Non-translation invariant risk measures case

The main objective of this paper is to study the valuation of financial derivatives with the ERP framework under non-translation invariant risk measures. This requires solving the root-finding problem of the initial portfolio value V_0 that equates $\tilde{\epsilon}^{(\mathcal{L})}(-V_0)$ and $\tilde{\epsilon}^{(\mathcal{S})}(V_0)$; this study considers a bisection scheme for such a purpose. However, one important drawback of the bisection algorithm in the context of this paper is the requirement to obtain multiple evaluations of $\tilde{\epsilon}^{(\mathcal{L})}(-V_0)$ and $\tilde{\epsilon}^{(\mathcal{S})}(V_0)$ for different values of V_0, which can be very costly from a computational standpoint. One naive approach to implement the bisection algorithm is to proceed as follows:

1) For a given value of V_0, train the long and short neural networks $F_\theta^{(\mathcal{S})}$ and $F_\theta^{(\mathcal{L})}$ on the training set.
2) Evaluate the optimal residual hedging risk $\tilde{\epsilon}^{(\mathcal{S})}(V_0)$ and $\tilde{\epsilon}^{(\mathcal{L})}(-V_0)$ with $F_\theta^{(\mathcal{S})}$ and $F_\theta^{(\mathcal{L})}$ on a *test set* of 100,000 additional independent simulated paths.
3) If $\Delta(V_0) \equiv \tilde{\epsilon}^{(\mathcal{S})}(V_0) - \tilde{\epsilon}^{(\mathcal{L})}(-V_0) \approx 0$ according to some closeness criterion, then $C_0^\star = V_0$ is the equal risk price. Otherwise, update V_0 with the bisection algorithm and go back to step 1).

The important drawback of this naive approach lies in the necessity to retrain $F_\theta^{(\mathcal{S})}$ and $F_\theta^{(\mathcal{L})}$ for each iteration of the bisection algorithm in step 1. To circumvent the latter pitfall, this study proposes to slightly modify the training algorithm such that the neural networks learn the optimal mappings not only for *a unique fixed* initial capital investment, but rather for an *interval* of values for V_0. This provides the important benefit of only having to train once $F_\theta^{(\mathcal{S})}$ and $F_\theta^{(\mathcal{L})}$, which thus circumvent the previously described computational burden.

The slight modification made to the training algorithm described in Section 5.4.2.1 is now described. At the beginning of each SGD step, on top of sampling a minibatch of paths of risky assets, the value of V_0 is also sampled within the initial interval of values used for the bisection algorithm. For instance, in numerical experiments conducted in Section 5.5, the initial interval considered for the bisection algorithm is $[0.75C_0^{\mathbb{Q}}, 1.50C_0^{\mathbb{Q}}]$ where $C_0^{\mathbb{Q}}$ is the risk-neutral price of Φ under a chosen conventional equivalent martingale measure $\mathbb{Q}$.[7] This modification is simple to implement as it naturally leverages the fact that portfolio values are already used within input vectors of the neural networks. However, it should be noted that learning the optimal hedge for various initial capital investments is more complex, and thus a more challenging task for neural networks as compared to learning the optimal trading policy for a fixed V_0. Nevertheless, Monte Carlo experiments provided in Section 5.8 show that incorporating this slight modification to the training algorithm does not materially impact the optimized neural networks performance.

Note that pseudo-codes of the training and bisection procedures are presented respectively in Algorithm 5.1 and Algorithm 5.2 of Section 5.7. An implementation in Python and Tensorflow to replicate numerical experiments presented in Section 5.5 can also be found online at github.com/alexandrecarbonneau.

Remark 5.2. *In numerical experiments of Section 5.5, the benchmarking of equal risk prices generated under the class of semi-*$\mathbb{L}^p$ *risk measures to the ones obtained with a class of convex risk measures, namely the CVaR, is performed. The numerical scheme used to obtain equal risk prices under the* $CVaR_\alpha$ *risk measure follows the methodology of Carbonneau and Godin (2021b) by evaluating* $C_0^\star$ *with* (5.3) *where* $\tilde{\epsilon}^{(\mathcal{L})}(0)$ *and* $\tilde{\epsilon}^{(\mathcal{S})}(0)$ *are computed with the steps of*

[7] If the equal risk price is outside the initial search interval $[0.75C_0^{\mathbb{Q}}, 1.50C_0^{\mathbb{Q}}]$, the bisection algorithm must be applied once again with a new initial search interval, and the neural networks $F_\theta^{(\mathcal{S})}$ and $F_\theta^{(\mathcal{L})}$ must be trained once again on this new interval.

Section 5.4.2.1 with $V_0 = 0$ and with the empirical estimator of $\rho(\pi)$ as

$$\widehat{\rho}(\mathbb{B}_j) = \widehat{VaR}_\alpha(\mathbb{B}_j) + \frac{1}{(1-\alpha)N_{batch}} \sum_{i=1}^{N_{batch}} \max(\pi_{i,j} - \widehat{VaR}_\alpha(\mathbb{B}_j), 0),$$

where $\widehat{VaR}_\alpha(\mathbb{B}_j)$ is the usual empirical estimator of the Value-at-Risk statistic with the sample $\mathbb{B}_j$ at level α.

Lastly, it is worth highlighting an additional advantage from a computational standpoint of the class of semi-$\mathbb{L}^p$ objective functions described in this paper over the CVaR_α measures as considered for instance in Carbonneau and Godin (2021b) and Carbonneau and Godin (2021a) when relying on the neural network-based hedging scheme. Indeed, under the CVaR_α objective function, the use of minibatch stochastic gradient descent procedures to train neural networks restrain the use of extremely large quantiles for the CVaR_α (for instance, larger values than 0.99). The latter stems from the following observations. From a statistical standpoint, the estimation variance of CVaR_α increases with α. Furthermore, the empirical estimator of CVaR_α is biased in finite sample size, whereas the empirical estimator of the semi-$\mathbb{L}^p$ risk measure is unbiased for any sample size. However, while larger minibatches would provide a more accurate estimate of the gradient, i.e. reduce the variance and the bias of the CVaR estimator, this is not necessarily a favorable avenue for training neural networks. Indeed, as noted in Goodfellow et al. (2016), the amount of memory required by hardware setups can be a limiting factor to increasing minibatches size. Furthermore, most SGD algorithms converge faster in terms of total computation when allowed to approximate gradients faster (i.e. with smaller samples and more SGD steps). The interested reader is referred to Chapter 8.1.3 of Goodfellow et al. (2016) for additional information about the implications of the minibatch sizes on SGD procedures. This computational pitfall of pairing stochastic gradient descent with extreme values of α under the CVaR_α measure is not present under the semi-$\mathbb{L}^p$, which further motivates its use in the context of equal risk pricing as well as in the context of hedging.

5.5 Numerical experiments

This section presents several numerical experiments conducted to investigate prices produced by the ERP methodology under different setups. The common theme of all experiments is to examine option prices generated by the ERP framework under the class of semi-$\mathbb{L}^p$ risk measures. The analysis starts in Section 5.5.2 with a sensitivity analysis of equal risk prices with respect to the choice of objective function. This is carried out by comparing $C_0^\star$ generated with the CVaR_α and semi-$\mathbb{L}^p$ across different values of α and p controlling the risk aversion of the hedger. The hedging performance of embedded neural networks hedging policies obtained under these objective functions is also assessed. Moreover, a sensitivity analysis with respect to the choice of underlying asset price dynamics is carried out in Section 5.5.3 so as to test the impact of the inclusion of jump or volatility risk. Lastly, Section 5.5.4 presents the benchmarking of equal risk prices for long maturity options obtained with the semi-$\mathbb{L}^p$ risk measures with trades involving exclusively the underlying stock against these generated with option hedges under the CVaR_α objective function.

5.5.1 Experiments setup

Unless specified otherwise, the option to price and hedge is a European put with payoff $\Phi(S_N^{(0,b)}) \equiv \max(K - S_N^{(0,b)}, 0)$ of maturity of $T = 60/260$ and strike price K. Daily hedges with the underlying stock are used (i.e. $N = 60$). The use of option hedges and different maturities for Φ is considered exclusively is Section 5.5.4. Furthermore, the stock has an initial price of $S_0^{(0,b)} = 100$ and the annualized continuous risk-free rate is set at $r = 0.03$. Different moneyness levels are considered with $K = 90, 100$ and 110 for respectively out-of-the-money (OTM), at-the-money (ATM), and in-the-money (ITM) puts.

Moreover, as described in Section 5.4, two distinct feedforward neural networks are considered as the functional representation of the long and short hedging policies. The architecture of every neural networks is a FFNN of two hidden layers ($L = 2$) with 56 neurons per layer ($d_1 = d_2 = 56$). The activation function considered is the well-known rectified linear activation function (ReLU) with $g(x) \equiv \max(x, 0)$. For the training procedure, a training set of 400,000 paths is simulated with the $\mathbb{P}$-dynamics of the underlying stock. A total of 200 epochs[8] is used with a minibatch size of 1,000 sampled exclusively from the training set. The Adam optimizer with a learning rate hyperparameter of 0.0005 is used with Tensorflow for the implementation of the stochastic gradient descent procedure. Also, all numerical results presented in subsequent sections are obtained in an out-of-sample fashion by using exclusively a test set of 100,000 additional simulated paths.

5.5.2 Sensitivity analysis to risk measures

This section studies equal risk price values obtained under the semi-$\mathbb{L}^p$ and CVaR_α risk measures across different levels of risk aversion, i.e. different values for p and α. The main motivation is the following. Carbonneau and Godin (2021b) observed that when hedging exclusively with the underlying stock, ERP under the CVaR_α measure produces option prices which are systematically inflated in comparison to those obtained under conventional risk-neutral measures, especially for OTM puts. This inflation phenomenon is significantly magnified with fat tails dynamics such as with a regime-switching (RS) model to an extend that can cast doubt on the applicability of ERP in practice. Furthermore, while the latter paper observed a positive relation between the risk aversion level α and equal risk prices $C_0^\star$, as shown in subsequent sections of this current paper, using smaller values for α leads to trading policies exhibiting poor risk mitigation performance with speculative behavior magnifying tail risk. Consequently, the main motivation of this current section is to assess if the use of the semi-$\mathbb{L}^p$ class of risk measures helps to alleviate this price inflation phenomenon through its choice of risk measure while simultaneously resulting in optimized trading policies providing effective risk mitigation. Thus, a critical aspect of the sensitivity analysis performed in this section is the benchmarking of not only equal risk prices generated under different

[8] An *epoch* is defined as a complete iteration of the training set with stochastic gradient descent. For example, for a training set of 400,000 paths and a minibatch size of 1,000, one epoch consists of 400 updates of the set of trainable parameters θ.

objective functions, but also the assessment of the effectiveness of the resulting global trading policies.

5.5.2.1 Regime-switching model

The conduction of a sensitivity analysis with respect to the objective function within the ERP framework necessitates the selection of a suitable dynamics for the underlying stock. Indeed, the model should incorporate salient stylized features of financial markets with a specific importance on exhibiting fat tails due to the assessment of the impact of objective functions within the ERP framework allowing more or less weights on extreme scenarios through their respective risk aversion parameter (i.e α and p respectively for the CVaR_α and semi-$\mathbb{L}^p$ measures). Unless specified otherwise, this study considers a RS model for the risky asset dynamics. This class of model introduced in finance by Hamilton (1989) exhibits, among others, fat tails, the leverage effect (i.e. negative correlation between assets returns and volatility) and heteroscedasticity. The examination of the impact of the presence of jump and volatility risk on $C_0^\star$ values generated with the semi-$\mathbb{L}^p$ objective functions is done in subsequent sections. Furthermore, unless specified otherwise, model parameters for the RS model (as well as for other dynamics considered subsequently) are estimated with maximum likelihood procedures on the same time series of daily log-returns on the S&P 500 price index for the period 1986-12-31 to 2010-04-01 (5863 observations).

The description of the regime-switching model for the underlying stock is now formally defined. For $n = 1, \ldots, N$, let $y_n \equiv \log(S_n^{(0,b)}/S_{n-1}^{(0,b)})$ be the time t_n log-return and $\{\epsilon_n\}_{n=1}^N$ be a sequence of independent and identically distributed (iid) standardized Gaussian random variables. The RS model assumes that the dynamics of the underlying stock changes between different regimes representing different economical states of the financial market. These regime changes are abrupt and they drastically impact the behavior of the dynamics of financial markets for a significant time period, i.e. these regimes are persistent (Ang and Timmermann, 2012). For instance, a two regime RS model as considered in this study usually has a more bullish regime with positive expected returns and relatively small volatility, and a more bearish regime with negative expected returns and relatively large volatility. Prevalent examples of such regime changes are financial crises and important economical reforms.

From a mathematical standpoint, the class of RS models characterizes regimes by an unobservable discrete-time Markov chain with a finite number of states, and models the conditional distribution of log-returns given the current regime as a Gaussian distribution with known parameters. More formally, denote the regimes as $\{h_n\}_{n=0}^N$ where $h_n \in \{1, \dots, H\}$ is the regime in-force during the time interval $[t_n, t_{n+1})$. The model specification for the transition probabilities of the Markov Chain can be stated as

$$\mathbb{P}(h_{n+1} = j | \mathcal{F}_n, h_n, \dots, h_0) = \gamma_{h_n,j}, \quad j = 1, \dots, H, \tag{5.8}$$

where $\Gamma \equiv \{\gamma_{i,j}\}_{i=1,j=1}^{H,H}$ is the transition matrix with $\gamma_{i,j}$ being the time-independent probability of moving from regime i to regime j. Furthermore, the dynamics of log-returns have the representation

$$y_{n+1} = \mu_{h_n}\Delta + \sigma_{h_n}\sqrt{\Delta}\epsilon_{n+1}, \quad n = 0, \dots, N-1,$$

where $\{\mu_i, \sigma_i\}_{i=1}^H$ are model parameters representing the means and volatilities on a yearly basis of each regime. The use of a RS model entails that additional state variables related to the regimes must be added to feature vectors of neural networks through the vectors $\mathcal{I}_n$. Indeed, while regimes are unobservable, useful information can be filtered from the observed stock path prices. Let $\{\xi_n\}_{n=0}^N$ be the *predictive probability process* where $\xi_n \equiv [\xi_{n,1}, \dots, \xi_{n,H}]$ and $\xi_{n,j} \equiv \mathbb{P}(h_n = j | \mathcal{F}_n)$. Under the RS model, $\mathcal{I}_n = \xi_n$ for $n = 0, \dots, N-1$. Following the work of François et al. (2014), the predictive probabilities can be computed recursively for

$n = 0, \dots, N-1$ as

$$\xi_{n+1,j} = \frac{\sum_{i=1}^{H} \gamma_{i,j} \phi_i(y_{n+1}) \xi_{n,i}}{\sum_{i=1}^{H} \phi_i(y_{n+1}) \xi_{n,i}}, \quad j = 1, \dots, H,$$

where ϕ_i is the probability density function of the Gaussian distribution with mean μ_i and volatility σ_i. For all numerical experiments, the time 0 regime h_0 is sampled from the stationary distribution of the Markov Chain. Lastly, the benchmarking of equal risk prices to option prices obtained under conventional risk-neutral measures is also presented. Risk-neutral dynamics as well as the numerical scheme used to evaluate the risk-neutral price (including for alternative dynamics introduced subsequently) are presented in Section 5.10.

5.5.2.2 Numerical results sensitivity analysis to objective function

Table 5.1 presents equal risk prices obtained under the CVaR_α with $\alpha = 0.90, 0.95, 0.99$ as well as under the class of semi-$\mathbb{L}^p$ risk measures with $p = 2, 4, 6, 8, 10$. All equal risk prices are expressed relative to risk-neutral prices $C_0^{\mathbb{Q}}$. Also, hedging statistics obtained across the different objective functions are analyzed subsequently in Section 5.5.2.3.

Table 5.1: Sensitivity analysis of equal risk prices $C_0^\star$ for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 60/260$ under the regime-switching model.

		$C_0^\star$ under CVaR_α			$C_0^\star$ under semi-$\mathbb{L}^p$				
Moneyness	$C_0^{\mathbb{Q}}$	$\text{CVaR}_{0.90}$	$\text{CVaR}_{0.95}$	$\text{CVaR}_{0.99}$	$\mathbb{L}^2$	$\mathbb{L}^4$	$\mathbb{L}^6$	$\mathbb{L}^8$	$\mathbb{L}^{10}$
OTM	0.56	91%	119%	161%	50%	86%	116%	147%	184%
ATM	3.27	18%	24%	29%	10%	17%	22%	27%	32%
ITM	10.36	5%	7%	9%	2%	4%	6%	9%	11%

Notes: $C_0^\star$ results are computed based on 100,000 independent paths generated from the regime-switching model under $\mathbb{P}$ (see Section 5.5.2.1 for model definition and Section 5.9 for model parameters). Risk-neutral prices $C_0^{\mathbb{Q}}$ are computed under $\mathbb{Q}$-dynamics described in Section 5.10. The training of neural networks is performed as described in Section 5.4.2 with hyperparameters presented in Section 5.5.1. $C_0^\star$ are expressed relative to $C_0^{\mathbb{Q}}$ (% increase).

Values from Table 5.1 indicate that equal risk prices generated by the class of semi-$\mathbb{L}^p$ risk measures can span much more than the interval of prices obtained under the CVaR_α risk measures with the selected confidence level α values. The latter results holds across all moneyness levels for puts. For instance, the relative increase in the equal risk price $C_0^\star$ as compared to the risk-neutral price $C_0^{\mathbb{Q}}$ for OTM puts is $91\%, 119\%$ and 161% under $\text{CVaR}_{0.90}$, $\text{CVaR}_{0.95}$ and $\text{CVaR}_{0.99}$, and ranges between 50% to 184% from the semi-$\mathbb{L}^2$ to the semi-$\mathbb{L}^{10}$. Similar observations can be made for ATM and ITM moneyness levels. Furthermore, it is very interesting to observe that the use of the semi-$\mathbb{L}^2$ risk measure entails a significant reduction of $C_0^\star$ as compared to the price obtained under the $\text{CVaR}_{0.90}$. Indeed, the relative increase in the equal risk price $C_0^\star$ with $p = 2$ as compared to the risk-neutral price $C_0^{\mathbb{Q}}$ for OTM, ATM and ITM moneyness levels is respectively 50%, 10% and 2%, which is significantly smaller than the corresponding relative increase of $91\%, 18\%$ and 5% under the $\text{CVaR}_{0.90}$ measure. Moreover, as expected, equal risk prices $C_0^\star$ generated with the class of semi-$\mathbb{L}^p$ risk measures show a positive relation with the risk aversion parameter p. This observation can be explained by an analogous analysis provided in Carbonneau and Godin (2021b) under the CVaR_α measures: since the put option's payoff is bounded below at zero, the short position hedging error has a thicker right tail than the corresponding right tail of the long position hedging error. Consequently, an increase in the risk aversion parameter p entails more weights on extreme hedging losses, which result in a larger increase of perceived residual risk exposure for the short position than for the long position. The latter entails that $C_0^\star$ must be increased to equalize the residual hedging risk of both parties. In conclusion, all these results clearly demonstrate the benefit of using the class of semi-$\mathbb{L}^p$ risk measures from the standpoint of pricing derivatives by not only spanning wider ranges of prices than these generated by the CVaR with conventional confidence levels, but by also significantly alleviating the inflated option prices phenomenon observed under the CVaR_α. However, the question about whether or not the optimized global policies under the semi-$\mathbb{L}^p$ risk measures are effective from the standpoint of risk mitigation remains. This is examined in the following section.

5.5.2.3 Hedging performance benchmarking

This section conducts the benchmarking of the neural networks trading policies hedging performance under the CVaR_α and semi-$\mathbb{L}^p$ objective functions. For the sake of brevity, hedging metrics values considered to compare the different policies are only presented for the short position hedge of the ATM put with the usual market setup, i.e. time-to-maturity of $T = 60/260$ under the regime-switching model with daily stock hedges. Table 5.2 presents hedging statistics of the global hedging policies obtained with the CVaR_α and semi-$\mathbb{L}^p$ risk measures with the objective functions used to generate the $C_0^\star$ values in the previous section (i.e. $\alpha = 0.90, 0.95, 0.99$ and $p = 2, 4, 6, 8, 10$). To compare the trading policies on common grounds, the initial portfolio value is set as the risk-neutral price with $V_0 = 3.27$ for all examples.[9] Furthermore, hedging metrics used for the benchmarking consists of the VaR_α and CVaR_α statistics over various α's, the mean hedging error, the SMSE (i.e. semi-$\mathbb{L}^2$ metric) and the mean-squared-error (MSE). Note that all hedging statistics are estimated in an out-of-sample fashion on the test set of 100,000 independent additional simulated paths.

Hedging metrics values show that while the trading policy optimized with the $\text{CVaR}_{0.90}$ objective function entails the smallest values for $\text{CVaR}_{0.90}$, $\text{VaR}_{0.90}$ and $\text{VaR}_{0.95}$ statistics, it exhibits poor mitigation of tail risk as compared to the other policies. For instance, the relative reduction of the $\text{CVaR}_{0.99}$ statistic across all other penalties than the $\text{CVaR}_{0.90}$ ranges between 29.8% and 44.5% as compared to the $\text{CVaR}_{0.90}$ trading policy. Similar observations can be made for the $\text{CVaR}_{0.999}$ and $\text{VaR}_{0.999}$ statistics capturing extreme scenarios. Consequently, these results cast doubt on the practical effectiveness of the $\text{CVaR}_{0.90}$ hedging policy from a risk mitigation standpoint, and thus also of trading policies optimized with lower values for α, due to their poor mitigation of risk for quantiles above the CVaR confidence level. The latter conclusion has important implications in the context of the ERP framework. Indeed,

[9] Recall that optimal policies under the CVaR_α risk measures are independent of V_0 due to the translation invariance property. Furthermore, the optimal policies obtained under the semi-$\mathbb{L}^p$ risk measures can be used not only with a specific value for V_0, but with an interval of initial capital investments due to the proposed modified training algorithm in this paper.

Table 5.2: Hedging statistics for short position ATM put option of maturity $T = 60/260$ under the regime-switching model.

	CVaR_α			semi-$\mathbb{L}^p$				
Penalty	$\text{CVaR}_{0.90}$	$\text{CVaR}_{0.95}$	$\text{CVaR}_{0.99}$	$\mathbb{L}^2$	$\mathbb{L}^4$	$\mathbb{L}^6$	$\mathbb{L}^8$	$\mathbb{L}^{10}$
Statistics								
Mean	0.11	0.13	0.14	−0.03	0.04	0.08	0.10	0.14
$\text{CVaR}_{0.90}$	**2.64**	5.3%	22.6%	6.3%	6.4%	7.6%	11.2%	13.4%
$\text{CVaR}_{0.95}$	3.41	**−8.4%**	1.6%	0.5%	−6.2%	−6.1%	−4.5%	−2.6%
$\text{CVaR}_{0.99}$	6.86	−31.7%	**−44.5%**	−29.8%	−41.0%	−42.6%	−43.4%	−42.9%
$\text{CVaR}_{0.999}$	19.99	−48.5%	−76.1%	−65.4%	−73.2%	−74.4%	−75.6%	**−76.3%**
$\text{VaR}_{0.90}$	**1.75**	34.7%	59.9%	9.3%	26.9%	30.5%	37.3%	40.9%
$\text{VaR}_{0.95}$	**2.08**	21.9%	54.6%	22.3%	28.0%	30.4%	36.7%	39.5%
$\text{VaR}_{0.99}$	3.67	**−9.6%**	−2.9%	8.1%	−4.2%	−5.1%	−4.6%	−2.6%
$\text{VaR}_{0.999}$	11.00	−43.3%	**−62.5%**	−47.2%	−57.3%	−59.2%	−60.0%	−60.4%
SMSE	1.83	−7.0%	6.8%	**−34.5%**	−28.1%	−24.0%	−17.6%	−11.2%
MSE	2.93	−1.8%	12.2%	**−27.7%**	−20.9%	−18.0%	−11.5	−6.2%

Notes: Hedging statistics are computed based on 100,000 independent paths generated from the regime-switching model under $\mathbb{P}$ (see Section 5.5.2.1 for model definition and Section 5.9 for model parameters). The training of neural networks is performed as described in Section 5.4.2 with hyperparameters presented in Section 5.5.1. All hedging statistics except the mean hedging error are expressed relative to values obtained under the $\text{CVaR}_{0.90}$ penalty. **Bold** values are the lowest across all penalties.

as shown in Carbonneau and Godin (2021b), the equal risk price $C_0^\star$ has a positive relation to the risk aversion parameter α. Thus, the inflated equal risk price phenomenon observed under the class of CVaR_α measures cannot be effectively alleviated through the reduction of α as the resulting trading policies quickly exhibit poor hedging performance. On the other hand, hedging statistics obtained with the class of semi-$\mathbb{L}^p$ risk measures indicate that across all levels of risk aversion considered, optimized trading policies are effective for mitigating hedging risk with the parameter p controlling the relative weight associated to extreme hedging losses. From the combination of these hedging statistics values as well as equal risk price values presented in Table 5.1, we can conclude that the class of semi-$\mathbb{L}^p$ risk measures can simultaneously generate lower equal risk prices with trading policies exhibiting effective hedging risk mitigation.

5.5.3 *Sensitivity analysis to dynamics of risky assets*

This section performs a sensitivity analysis of equal risk prices across different dynamics for the financial market. The motivation is to assess if the conclusion that the class of semi-$\mathbb{L}^p$ risk measures can dampen the inflated equal risk prices phenomenon as well as span wider price intervals than these obtained under the CVaR_α measures is robust to the presence of different equity risk features. For such a purpose, this paper considers the presence of jump risk with the Merton jump-diffusion model (MJD, Merton (1976)) and of volatility risk with the GJR-GARCH model (Glosten et al., 1993). The Black and Scholes (1973) and Merton (1973) (BSM) model is also considered due to its popularity and the fact that contrarily to the other dynamics, the BSM model does not exhibit fat tails. The assessment of the impact of the choice of risk measure controlling the weight associated to extreme scenarios is thus also of interest under the BSM dynamics since the optimal hedging strategies, and thus equal risk prices, should be less sensitive to the risk aversion parameter under a dynamics without fat tails.

The dynamics of all three models is now formally presented. All model parameters are estimated with the same time series of daily log-returns on the S&P 500 index for the period 1986-12-31 to 2010-04-01 (5863 log-returns).

5.5.3.1 *Black-Scholes model*

The Black-Scholes model assumes that log-returns are iid Gaussian random variables of yearly mean $\mu - \sigma^2/2$ and yearly volatility σ:

$$y_n = \left(\mu - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\epsilon_n, \quad n = 1, \ldots, N.$$

Stock prices have the Markov property under $\mathbb{P}$ with respect to the market filtration $\mathbb{F}$. The latter entails that no additional information should be added to the state variables of the neural networks, i.e. $\mathcal{I}_n = 0$ for all n.

5.5.3.2 GJR-GARCH model

The GJR-GARCH model relaxes the constant volatility assumption of the BSM model by assuming the presence of stochastic volatility which incorporates the leverage effect. Log-returns under this model have the representation

$$y_n = \mu + \sigma_n \epsilon_n,$$
$$\sigma_{n+1}^2 = \omega + \upsilon \sigma_n^2 (|\epsilon_n| - \gamma \epsilon_n)^2 + \beta \sigma_n^2,$$

where $\{\sigma_n^2\}_{n=1}^{N+1}$ are the daily variances of log-returns, $\{\mu, \omega, \upsilon, \gamma, \beta\}$ are the model parameters with $\{\omega, \upsilon, \beta\}$ being positive real values and $\{\mu, \gamma\}$ real values. Note that given σ_1^2, the sequence of variances $\sigma_2^2, \ldots, \sigma_{N+1}^2$ can be computed recursively with the observed path of log-returns. In this paper, the initial value σ_1^2 is set as the stationary variance of the process: $\sigma_1^2 \equiv \mathbb{E}[\sigma_n^2] = \frac{\omega}{1-\upsilon(1+\gamma^2)-\beta}$. Furthermore, it can be shown that $\{S_n^{(0,b)}, \sigma_{n+1}\}_{n=0}^{N}$ is an $(\mathbb{F}, \mathbb{P})$-Markov bivariate process. Consequently, the periodic volatility is added to the states variables of the neural networks at each time step: $\mathcal{I}_n = \sigma_{n+1}$ for $n = 0, \ldots, N-1$.

5.5.3.3 Merton jump-diffusion model

Contrarily to the GJR-GARCH model, the MJD dynamics assumes constant volatility, but deviates from the BSM assumptions by incorporating random Gaussian jumps to stock returns. Let $\{N_n\}_{n=0}^{N}$ be realizations of a Poisson process of parameter $\lambda > 0$, where N_n represents the cumulative number of jumps of the stock price from time 0 to time t_n. The Merton (1976) model assumes that jumps, denoted by $\{\zeta_j\}_{j=1}^{\infty}$, are iid Gaussian random variables of mean μ_J and variance σ_J^2:[10]

$$y_n = \left(\nu - \lambda(e^{\mu_J+\sigma_J^2/2} - 1) - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\epsilon_n + \sum_{j=N_{n-1}+1}^{N_n} \zeta_j,$$

where $\{\epsilon_n\}_{n=1}^N$, $\{N_n\}_{n=0}^N$ and $\{\zeta_j\}_{j=1}^\infty$ are independent. Model parameters consist of $\{\nu, \lambda, \sigma, \mu_J, \sigma_J\}$ where $\nu \in \mathbb{R}$ is the drift parameter and $\sigma > 0$ is the constant volatility term. Since stock prices are iid, this dynamics does not necessitate the addition of other state variables to the feature vectors, i.e. $\mathcal{I}_n = 0$ for all n.

5.5.3.4 Numerical results sensitivity analysis to dynamics

Table 5.3 presents the sensitivity analysis of equal risk prices with the same setup as in previous sections, i.e. for put options of maturity $T = 60/260$ with daily stock hedges, for the BSM, MJD and GJR-GARCH models. To save space, results are only presented for the OTM moneyness as the main conclusions are shared for both ATM and ITM moneyness levels. Furthermore, both the CVaR_α and semi-$\mathbb{L}^p$ classes of risk measures are considered with $\alpha = 0.90, 0.95, 0.99$ and $p = 2, 4, 6, 8, 10$.

Table 5.3: Sensitivity analysis of equal risk prices for OTM put options of maturity $T = 60/260$ under the BSM, MJD and GJR-GARCH models.

		$C_0^\star$ under CVaR_α			$C_0^\star$ under semi-$\mathbb{L}^p$				
Dynamics	$C_0^{\mathbb{Q}}$	$\text{CVaR}_{0.90}$	$\text{CVaR}_{0.95}$	$\text{CVaR}_{0.99}$	$\mathbb{L}^2$	$\mathbb{L}^4$	$\mathbb{L}^6$	$\mathbb{L}^8$	$\mathbb{L}^{10}$
BSM	0.53	5%	10%	17%	3%	13%	22%	31%	45%
MJD	0.46	23%	34%	129%	15%	43%	69%	95%	125%
GJR-GARCH	0.57	52%	71%	139%	27%	121%	155%	208%	237%

Notes: Equal risk prices $C_0^\star$ results are computed based on 100,000 independent paths generated from the BSM, MJD and GJR-GARCH model under $\mathbb{P}$ (see Section 5.5.3 for models definitions under $\mathbb{P}$ and Section 5.9 for model parameters). Risk-neutral prices $C_0^{\mathbb{Q}}$ are computed under $\mathbb{Q}$-dynamics described in Section 5.10. The training of feedforward neural networks is performed as described in Section 5.4.2 with hyperparameters presented in Section 5.5.1. $C_0^\star$ are expressed relative to $C_0^{\mathbb{Q}}$ (% increase).

[10] The convention that $\sum_{j=N_{n-1}+1}^{N_n} \zeta_j = 0$ if $N_{n-1} = N_n$ is adopted.

These results clearly demonstrate that the conclusion that equal risk prices generated by the class of semi-$\mathbb{L}^p$ risk measures can alleviate the price inflation phenomenon observed under the CVaR_α measures is robust to different dynamics. Indeed, values show that by using the semi-$\mathbb{L}^2$ risk measure, OTM equal risk prices $C_0^\star$ values exhibit a relative increase over risk-neutral prices $C_0^{\mathbb{Q}}$ of respectively of $3\%, 15\%$ and 27% under the BSM, MJD and GARCH models as compared to 5%, 23% and 52% under the $\text{CVaR}_{0.90}$ cost function. Furthermore, values presented in Table 5.3 demonstrate that the observation made in the previous section under the RS model with respect to the fact that equal risk prices generated by the class of semi-$\mathbb{L}^p$ risk measures can span a large interval of prices which encompasses values obtained with the CVaR_α measures is robust to different dynamics of the financial markets. Lastly, it is interesting to observe that the length of the price intervals generated by both classes of risk measures varies significantly with the dynamics of the financial market. Indeed, under the BSM model, the relative increase of $C_0^\star$ as compared to $C_0^{\mathbb{Q}}$ ranges between 5% to 17% under the CVaR_α and 3% to 45% under the semi-$\mathbb{L}^p$. On the other hand, with the GJR-GARCH dynamics, the relative increase in $C_0^\star$ under the CVaR_α ranges between 52% to 139%, while under the semi-$\mathbb{L}^p$, it ranges between 27% to 237%. Similar observations can be made under the MJD dynamics. This can be explained by the fact that contrarily to the other models, the BSM dynamics does not exhibits fat tails as the market incompleteness solely stems from discrete-time trading. Consequently, the trading policies are much less sensitive to the choice of risk aversion parameter p or α under the BSM model, which entails equal risk prices that are less sensitive to risk aversion parameters. From these results, we can conclude that the choice of both the risky asset dynamics and of the risk measure among the classes of CVaR_α and semi-$\mathbb{L}^p$ measures has a material impact on equal risk prices, and this impact becomes more important as the dynamics exhibits fatter tails for risky assets returns.

5.5.4 Long-term maturity ERP with option hedges

This section examines the use of semi-$\mathbb{L}^p$ risk measures within the ERP framework for pricing long-term options with trades involving exclusively the underlying stock as compared to equal risk prices generated under the CVaR_α with trades involving shorter-term options. The motivation for this experiment is the following. The main finding of Carbonneau and Godin (2021a) is that under the CVaR_α measure, hedging long-term puts with shorter-term options in the presence of jump or volatility risks significantly reduces equal risk prices as compared to trading exclusively the underlying stock. However, the expected trading cost of setting up a trading strategy based solely on option hedges can be impractical in some cases in the face of highly illiquid options. In such context, the hedger could potentially be restricted to trading strategy relying exclusively on the underlying stock, which as shown in previous sections, can inflate equal risk prices under the CVaR_α measure. The objective of this last section is thus to assess if the use of the semi-$\mathbb{L}^p$ risk measure can achieve similar equal risk prices reduction when trading the underlying stock to the reduction obtained when trading options with the CVaR_α objective function. The setup to perform this experiment is the same as the one considered in Carbonneau and Godin (2021a), and numerical values for equal risk prices generated with trades involving exclusively options under the CVaR_α are taken directly from the latter work. This setup is now recalled below.

The derivative to price and hedge is a 1-year put with 252 days per year of moneyness levels OTM, ATM and ITM with strike prices of $90, 100$ and 110, respectively. As noted in Carbonneau and Godin (2021a), option trading strategies optimized with the confidence level α smaller than 0.95 when using the CVaR as the objective function often leads to hedging strategies exhibiting poor tail risk mitigation. Thus, the convex risk measure considered as the benchmark in the current study is the $\text{CVaR}_{0.95}$ measure with trades involving either exclusively the underlying stock on a daily or monthly basis (i.e. $N = 252$ or $N = 12$, respectively), or by trading solely with ATM 1-month and 3-months calls and puts (i.e. $N = 12$ or $N = 4$, respectively). Following the work of Carbonneau and Godin (2021a), the pricing of options used as hedging instruments is done through the modeling of the daily variations of the ATM logarithm implied volatility dynamics under $\mathbb{P}$ as an autoregressive (AR) model of order 1, named log-AR(1) hereafter. Furthermore, the model assumes for conveneice that the ATM 1-month and 3-months implied volatilities are the same.[11] It is worth highlighting that the implied volatility model is used exclusively for pricing options used as hedging instruments, not for the 1-year put Φ to be priced. Also, note that while the rebalancing frequency is either daily, monthly or quarterly, IV variations are always generated on a daily basis.

The log-AR(1) model is now formally defined. Denote by $\{IV_n\}_{n=0}^{252}$ the daily implied volatilities for the ATM 1-month and 3-months maturities which are used as hedging instruments. Also, let $\{Z_n\}_{n=1}^{252}$ be an additional sequence of iid standardized Gaussian random variables representing the random innovations of the log-IV dynamics. To capture the well-known leverage effect between asset returns and implied volatility variations (see for instance Cont and Da Fonseca (2002)), a correlation factor $\varrho \equiv corr(\epsilon_n, Z_n)$ set at -0.6 is assumed where $\{\epsilon_n\}_{n=1}^{252}$ are the daily random innovations associated to stock returns. The log-AR(1) model has the represensation

$$\log IV_{n+1} = \log IV_n + \kappa(\vartheta - \log IV_n) + \sigma_{IV} Z_{n+1}, \quad n = 0, \ldots, 251,$$

where $\{\kappa, \vartheta, \sigma_{IV}\}$ are the model parameters with κ and ϑ as real values and $\sigma_{IV} > 0$. The initial value of the process is set at the long-term parameter with $\log IV_0 \equiv \vartheta$. Moreover, the pricing of the calls and puts used as hedging instruments is performed with the well-known Black-Scholes formula with the annualized volatility set at the current implied volatility value. More precisely, denote by $C(IV, \Delta T, S, K)$ and $P(IV, \Delta T, S, K)$ the price of a call and put option respectively if the current implied volatility is IV, the time-to-maturity is ΔT, the

[11] Note that traded options with different maturities are never used simultaneously in the same hedging simulation.

underlying stock price is S and the strike price is K:

$$C(IV, \Delta T, S, K) \equiv S\mathcal{N}(d_1) - e^{-r\Delta T}K\mathcal{N}(d_2),$$
$$P(IV, \Delta T, S, K) \equiv e^{-r\Delta T}K\mathcal{N}(-d_2) - S\mathcal{N}(-d_1),$$

where $\mathcal{N}(\cdot)$ denotes the cumulative distribution function of a standardized Gaussian random variable with

$$d_1 \equiv \frac{\log(\frac{S}{K}) + (r + \frac{IV^2}{2})\Delta T}{IV\sqrt{\Delta T}}, \quad d_2 \equiv d_1 - IV\sqrt{\Delta T}.$$

Also, note that when option hedges are considered, the current implied volatility is added to the feature vectors of the neural networks. For instance, with 1-month calls and puts hedges, the nth trade at time $t_n = n/12$ uses as input vectors for the neural networks $X_n = [S^{(0,b)}_{21\times n}, IV_{21\times n}, T - t_n, \mathcal{I}_{21\times n}]$ for $n = 0, 1, \ldots, 11$.[12]

Moreover, the dynamics of the underlying asset returns considered for this last section is once again the MJD dynamics, but with different parameters than in previous sections since the ones considered in Carbonneau and Godin (2021a) are used for comparability purposes. The MJD as well as the log-AR(1) model parameters values are presented in Table 5.4 and Table 5.5, respectively, and are the same as in Carbonneau and Godin (2021a). These parameters were chosen in an ad hoc fashion so as to produce reasonable values for the dynamics of the financial market.

Table 5.4: Parameters of the 1-year Merton jump-diffusion model.

ν	σ	λ	μ_J	σ_J
0.1111	0.1323	0.25	−0.10	0.10

Notes: ν, σ and λ are on an annual basis.

[12] Note that with option hedges, the implied volatility of the options used as hedging instruments is added to feature vectors, not the price of each asset. This has the benefit of necessitating one less state variable with the implied volatility instead of adding two state variables with the price of the call and put used for hedging. Furthermore, this is a reasonable choice from a theoretical standpoint as implied volatilities are simply a nonlinear transformation of options prices due to the bijection relation between the two values.

Table 5.5: Parameters of the log-AR(1) model for the evolution of implied volatilities.

κ	ϑ	σ_{IV}	ϱ
0.15	log(0.15)	0.06	−0.6

5.5.4.1 Numerical results with option hedges

Table 5.6 presents equal risk prices $C_0^\star$ under $\text{CVaR}_{0.95}$ measure with daily or monthly stock trades as well as with 1-month or 3-months ATM calls and puts. Note that the latter values are taken directly from Table 3 of Carbonneau and Godin (2021a).[13] Furthermore, $C_0^\star$ values under the semi-$\mathbb{L}^2$ values with daily and monthly stock hedges are also presented.

Table 5.6: Sensitivity analysis of equal risk prices to jump risk for OTM ($K = 90$), ATM ($K = 100$) and ITM ($K = 110$) put options of maturity $T = 1$.

	$C_0^\star$ under $\text{CVaR}_{0.95}$				$C_0^\star$ under semi-$\mathbb{L}^2$	
Moneyness	Daily stock	Monthly stock	1-month opts	3-months opts	Daily stock	Monthly stock
OTM	2.58	2.60	2.24	2.08	2.18	2.23
ATM	6.01	5.77	5.36	5.12	5.38	5.22
ITM	11.68	11.44	10.86	10.51	10.42	10.54

Notes: These results are computed based on 100,000 independent paths generated from the MJD model under $\mathbb{P}$ (see Section 5.5.3.3 for model definition and Table 5.4 for model parameters). Options used as hedging instruments are priced with implied volatility modeled with a log-AR(1) dynamics (see Section 5.5.4 for model description and Table 5.5 for parameters values). Values for $C_0^\star$ under $\text{CVaR}_{0.95}$ are from Table 3 of Carbonneau and Godin (2021a). Values for $C_0^\star$ under semi-$\mathbb{L}^2$ are obtained with the training algorithm described in Section 5.4.2.2.

Numerical results indicate that the use of the semi-$\mathbb{L}^2$ objective function is successful at reducing significantly equal risk prices when relying on trades involving exclusively the underlying stock. Indeed, the relative reduction in $C_0^\star$ obtained by using the semi-$\mathbb{L}^2$ risk measure as compared to the $\text{CVaR}_{0.95}$ for OTM, ATM and ITM moneyness levels is respectively

[13] The type of neural networks considered in Carbonneau and Godin (2021a) is the long-short term memory (LSTM). The current paper found that FFNN trading policies performed significantly better for the numerical experiments conducted under the semi-$\mathbb{L}^p$ risk measure which motivated their use over LSTMs. The reader is referred to Section 3 of Carbonneau and Godin (2021a) for the formal description of the LSTM architecture.

$15\%, 11\%$ and 11% with daily stock and $14\%, 10\%$ and 8% with monthly stock rebalancing.[14] Furthermore, equal risk price values under the semi-$\mathbb{L}^2$ risk measure with daily or monthly stock hedges are relatively close to those obtained with 1-month or 3-months option hedges under the $\text{CVaR}_{0.95}$. These results have important implications for ERP procedures. Indeed, this demonstrates that in the face of highly illiquid options, the use of the semi-$\mathbb{L}^p$ class of risk measures with stock hedges can effectively reduce equal risk prices to similar levels as the ones obtained with option hedges with the CVaR_α measures. This avenue is thus successful to alleviate the price inflation phenomenon of ERP procedures for the pricing of long-term options. It is worth highlighting that in the presence of jump risk, the use of options as hedging instruments is much more effective for risk mitigation as compared to hedging strategies involving exclusively the underlying stock (see for instance Coleman et al. (2007) and Carbonneau (2021)). Nevertheless, $C_0^\star$ values presented in Table 5.6 indicate that when setting up trading strategies with options is impractical due to high expected trading costs, the use of stock hedges coupled with semi-$\mathbb{L}^p$ risk measures can effectively reduce option prices.

5.6 Conclusion

This paper studies the class of semi-$\mathbb{L}^p$ risk measures in the context of equal risk pricing (ERP) for the valuation of European financial derivatives. The ERP framework prices contingent claims as the initial hedging portfolio value which equates the residual hedging risk of the long and short positions under optimal hedging strategies. Despite lacking the translation invariance property which complexify the numerical evaluation of equal risk prices, the use of semi-$\mathbb{L}^p$ risk measures as the objective function measuring residual hedging risk is shown to have several preferable properties over the CVaR_α which is explored for instance in Carbonneau and Godin (2021b) and Carbonneau and Godin (2021a) in the context of ERP. The optimal hedging problems underlying the ERP framework are solved with deep

[14] For instance, if $C_0^\star(\text{CVaR}_{0.95})$ and $C_0^\star(\mathbb{L}^2)$ are respectively equal risk prices under the $\text{CVaR}_{0.95}$ and semi-$\mathbb{L}^2$ objective functions, the relative reduction is computed as $1 - \frac{C_0(\mathbb{L}^2)}{C_0^\star(\text{CVaR}_{0.95})}$.

reinforcement learning procedures by representing trading policies with neural networks as proposed in the work of Buehler et al. (2019b). A modification to the training algorithm for neural networks is presented in this current paper to tackle the additional complexity of using the semi-$\mathbb{L}^p$ risk measures with the ERP framework. This modification consists in training the neural networks to learn the optimal mappings for an interval of initial capital investments instead of a unique fixed value. The latter is shown not to lead to material deterioration in the hedging accuracy of the neural networks trading policies.

Several numerical experiments are performed to examine option prices generated by the ERP framework under the class of semi-$\mathbb{L}^p$ risk measures. First, a sensitivity analysis of equal risk price values with respect to the choice of objective function is conducted by comparing prices obtained with the CVaR_α and semi-$\mathbb{L}^p$ objectives across different values of α and p controlling the risk aversion of the hedger. Numerical results demonstrate that equal risk prices under the semi-$\mathbb{L}^p$ risk measures are spanning a larger interval of values than the one obtained with the CVaR_α by alleviating the price inflation phenomenon observed under the CVaR_α documented in previous studies. Furthermore, the trading policies parameterized as neural networks are shown to be highly effective for risk mitigation under the semi-$\mathbb{L}^p$ objective functions across all values of p considered, with the risk aversion parameter controlling the relative weight associated with extreme scenarios. Moreover, additional numerical experiments show that the use of the semi-$\mathbb{L}^2$ objective function for the pricing of long-term puts with hedges exclusively relying on the underlying asset is successful at reducing equal risk prices roughly to the level of prices produced with option hedges under the CVaR_α objective function. The latter conclusion is highly important in the context of ERP as it demonstrates that in the case where options are not or cannot be used within the hedging strategy, the ERP methodology used in conjunction with the semi-$\mathbb{L}^p$ class of risk measures can produce reasonable option prices.

5.7 Pseudo-code

This section presents the pseudo-codes for the training of neural networks and of the bisection method. Algorithm 5.1 describes the pseudo-code to carry out a single SGD step, i.e. given θ_j and the initial portfolio value V_0, the steps to perform an update of the set of parameters to θ_{j+1} are presented. Without loss of generality, the training pseudo-code is presented only for the short neural network $F_\theta^{(\mathcal{S})}$ and for trades involving only the underlying stock. Note that for all numerical experiments under the semi-$\mathbb{L}^p$ risk measure conducted in this paper, a preprocessing of the feature vectors is used with $\{T - t_n, \log(S_n^{(b)}/K), V_n/\tilde{V}, \mathcal{I}_n\}$ instead of $\{T - t_n, S_n^{(b)}, V_n, \mathcal{I}_n\}$ for $\tilde{V}$ defined as the midpoint value of the initial search interval of the bisection algorithm $[V_A, V_B]$, i.e. $\tilde{V} \equiv 0.5(V_A + V_B)$. Note that Buehler et al. (2019b), Carbonneau and Godin (2021b) and Carbonneau and Godin (2021a) consider similar preprocessing for risky asset prices, while Carbonneau (2021) considers a similar preprocessing of portfolio values. Furthermore, under the CVaR_α objective function, the same preprocessing for risky asset prices is used, but portfolio values are not preprocessed as the bisection algorithm is not required to be used in this case, i.e. V_n is used in feature vectors. Moreover, the update rule for portfolio values in step (6) of Algorithm 5.1 can be obtained directly from the self-financing representation of V_n^δ as shown below

$$\begin{aligned}
V_n^\delta &= B_n(V_0^\delta + G_n^\delta) \\
&= B_n\left(V_0^\delta + G_{n-1}^\delta + \delta_n^{(0:D)} \bullet (B_n^{-1}S_{n-1}^{(e)} - B_{n-1}^{-1}S_{n-1}^{(b)})\right) \\
&= \frac{B_n}{B_{n-1}}V_{n-1}^\delta + \delta_n^{(0:D)} \bullet (S_{n-1}^{(e)} - \frac{B_n}{B_{n-1}}S_{n-1}^{(b)}) \\
&= e^{r\Delta}V_{n-1}^\delta + \delta_n^{(0:D)} \bullet (S_{n-1}^{(e)} - e^{r\Delta}S_{n-1}^{(b)}).
\end{aligned} \tag{5.9}$$

Algorithm 5.2 presents the pseudo-code for the bisection algorithm taking as inputs the two trained neural networks $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$, the test set of 100,000 paths as well as the initial search range $[V_A, V_B]$ so as to output the equal risk price.

Algorithm 5.1 Pseudo-code training neural networks $F_\theta^{(\mathcal{S})}$ with underlying stock hedges
Input: θ_j, V_0^δ
Output: θ_{j+1}

for $i = 1, \ldots, N_{\text{batch}}$ **do** ▷ Loop over each path of minibatch
    $X_{0,i} = [T, \log(S_{0,i}^{(0,b)}/K), V_{0,i}^\delta/\tilde{V}, \mathcal{I}_{0,i}]$ ▷ Time 0 feature vector of $F_\theta^{(\mathcal{S})}$
    for $n = 0, \ldots, N-1$ **do**
        $\delta_{n+1,i}^{(0)} \leftarrow$ time t_n output of FFNN $F_\theta^{(\mathcal{S})}$ with $\theta = \theta_j$
        $S_{n+1,i}^{(0,b)} = S_{n,i}^{(0,b)} e^{y_{n+1,i}}$ ▷ Sample next stock price
        $V_{n+1,i}^\delta = e^{r\Delta} V_{n,i}^\delta + \delta_{n+1,i}^{(0)} (S_{n+1,i}^{(0,b)} - e^{r\Delta} S_{n,i}^{(0,b)})$ ▷ See (5.9) for details
        $\mathcal{I}_{n+1,i} \leftarrow$ update additional state variables
        $X_{n+1,i} = [T - t_n, \log(S_{n+1,i}^{(0,b)}/K), V_{n+1,i}^\delta/\tilde{V}, \mathcal{I}_{n+1,i}]$ ▷ Time t_{n+1} feature vector of $F_\theta^{(\mathcal{S})}$
    end for
    $\Phi(S_{N,i}^{(0,b)}) = \max(K - S_{N,i}^{(0,b)}, 0)$
    $\pi_{i,j} = \Phi(S_{N,i}^{(0,b)}) - V_{N,i}^\delta$
end for
$\widehat{J} = \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \pi_{i,j}^p \mathbb{1}_{\{\pi_{i,j} > 0\}}$
$\eta_j \leftarrow$ Adam algorithm
$\theta_{j+1} = \theta_j - \eta_j \nabla_\theta \widehat{J}$ ▷ $\nabla_\theta \widehat{J}$ computed with Tensorflow

Notes: Subscript i represents the ith simulated path among the minibatch of size N_{batch}. Also, the time 0 feature vector is fixed for all paths, i.e. $S_{0,i}^{(0,b)} = S_0^{(0,b)}$, $V_{0,i}^\delta = V_0^\delta$ and $\mathcal{I}_{0,i} = \mathcal{I}_0$.

Algorithm 5.2 Pseudo-code bisection algorithm
Input: $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$ trained neural networks, initial search range $[V_A, V_B]$ and test set paths
Output: $C_0^\star$

nbs_iter $= 0$, $\Delta(V) = \infty$
while $|\Delta(V)| > \zeta$ and nbs_iter $<$ max_iter **do**
    $V = 0.5(V_A + V_B)$
    Compute $\tilde{\epsilon}^{(\mathcal{L})}(-V)$ and $\tilde{\epsilon}^{(\mathcal{S})}(V)$ on the test set with $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$
    $\Delta(V) = \tilde{\epsilon}^{(S)}(V) - \tilde{\epsilon}^{(L)}(-V)$
    if $\Delta(V) > 0$ **then**
        $V_A \leftarrow V$
    else
        $V_B \leftarrow V$
    end if
    nbs_iter $\leftarrow$ nbs_iter $+ 1$
end while
$C_0^\star = V$.

Notes: ζ and max_iter represent respectively the admissible level of pricing error and the maximum number of iterations for the bisection algorithm. For all numerical experiments conducted in Section 5.5, ζ is set to 0.01 and max_iter to 100.

5.8 Validation of modified training algorithm

The goal of this section is to demonstrate that the proposed modification to the training algorithm described in Section 5.4.2.2 to tackle the non-translation invariant risk measures case of the ERP framework does not materially impact the optimized neural networks hedging performance. Denote by F_θ the neural network trained with the additional step of sampling $V_0 \in [V_A, V_B]$ on top of the minibatch of paths at the beginning of each stochastic gradient descent step. One conclusive test to validate that the proposed modification does not deteriorate the neural networks accuracy is to compare the hedging performance of F_θ assuming $V_0 = V^\star$ to another neural network denoted as F_θ^{fixed} trained exclusively with a fixed initial capital investment set at $V^\star$. If F_θ exhibits similar hedging performance to F_θ^{fixed} over multiple iterations of $V^\star$, this demonstrates that F_θ learned the optimal trading policy over a range of possible initial capital investments.

The experiment conducted to perform the latter test is now formally presented. The setup considered is similar to the one presented in Section 5.5.1 with the hedging of an ATM put option of maturity $T = 60/260$ with daily stock hedges under the regime-switching model. The objective function is the semi-$\mathbb{L}^p$ for $p \in \{2, 4, 6, 8\}$:

1) Train F_θ with the procedure described Section 5.4.2.2 where V_0 is sampled in the interval $[0.75C_0^{\mathbb{Q}}, 1.50C_0^{\mathbb{Q}}]$ at the beginning of each SGD step with $C_0^{\mathbb{Q}}$ being the risk-neutral price. A total of 200 epochs is used on the train set.

2) For a fixed randomly sampled value $V^* \in [0.75C_0^{\mathbb{Q}}, 1.50C_0^{\mathbb{Q}}]$, set $V_0 = V^*$ and train F_θ^{fixed} with the methodology described in Section 5.4.2.1. A total of three iterations of this step is performed.

3) For the three values of $V^\star$ sampled, compute the SMSE (i.e. semi-$\mathbb{L}^2$) statistics on the test set with F_θ and F_θ^{fixed}.

4) Repeat steps 1) - 3) for $p \in \{2, 4, 6, 8\}$.

Table 5.7 presents the SMSE statistics for the three values of $V_0 = V^\star$ across the semi-$\mathbb{L}^p$ objective functions with $p = 2, 4, 6, 8$. These results clearly demonstrate that the modified

Table 5.7: Semi-mean-square-error (SMSE) statistics of the modified training algorithm for ATM ($K = 100$) put options of maturity $T = 60/260$ under the regime-switching model.

		$\mathbb{L}^2$		$\mathbb{L}^4$		$\mathbb{L}^6$		$\mathbb{L}^8$	
Iteration	V_0	F_θ	F_θ^{fixed}	F_θ	F_θ^{fixed}	F_θ	F_θ^{fixed}	F_θ	F_θ^{fixed}
1	4.343	0.3899	0.3835	0.4017	0.4107	0.4442	0.4543	0.4637	0.4755
2	2.503	2.3892	2.3470	2.6245	2.4532	2.7520	2.6084	2.8259	2.7272
3	4.005	0.5770	0.5750	0.6101	0.6157	0.6666	0.6728	0.6920	0.7239

Notes: SMSE statistics results are computed based on 100,000 independent paths generated with the regime-switching model under $\mathbb{P}$ (see Section 5.5.2.1 for model definition and Section 5.9 for model parameters). F_θ is the neural network trained with the modified algorithm described in Section 5.4.2.2. F_θ^{fixed} is the neural network trained with fixed initial capital investment of V_0 as described in Section 5.4.2.1.

training algorithm does not materially impact the accuracy of the neural network as the difference in SMSE statistics between the FFNNs F_θ and F_θ^{fixed} is most often marginal.

5.9 Maximum likelihood estimates results

This section presents maximum likelihood model parameters estimates for the different risky asset dynamics considered in numerical experiments of Section 5.5.2 and Section 5.5.3. All parameters are estimated with the same time-series of daily log-returns on the S&P 500 index for the period 1986-12-31 to 2010-04-01 (5863 log-returns).

Table 5.8: Maximum likelihood parameter estimates of the Black-Scholes model.

μ	σ
0.0892	0.1952

Notes: Both μ and σ are on an annual basis.

Table 5.9: Maximum likelihood parameter estimates of the GJR-GARCH model.

μ	ω	υ	γ	β
2.871e-04	1.795e-06	0.0540	0.6028	0.9105

Table 5.10: Maximum likelihood parameter estimates of the regime-switching model.

	Regime	
Parameter	1	2
μ	0.1804	-0.2682
σ	0.1193	0.3328
ν	0.7543	0.2457
Γ	0.9886	0.0114
	0.0355	0.9645

Notes: Parameters were estimated with the EM algorithm of Dempster et al. (1977). μ and σ are on an annual basis.

Table 5.11: Maximum likelihood parameter estimates of the Merton jump-diffusion model.

ν	σ	λ	μ_J	σ_J
0.0875	0.1036	92.3862	-0.0015	0.0160

Notes: ν, σ and λ are on an annual basis.

5.10 Risk-neutral dynamics

This section presents the risk-neutral dynamics for the RS, BSM, GARCH and MJD models. The absence of arbitrage opportunities implied by each model entails by the first fundamental theorem of asset pricing that there exist a probability measure $\mathbb{Q}$ such that $\{S_n^{(0,b)} e^{-rt_n}\}_{n=0}^N$ is an $(\mathbb{F}, \mathbb{Q})$-martingale (Delbaen and Schachermayer, 1994). Denote by $\{\epsilon_n^{\mathbb{Q}}\}_{n=1}^N$ a sequence of iid standardized Gaussian random variables under $\mathbb{Q}$. Hereby are the $\mathbb{Q}$-dynamics for the four models described in Section 5.5.2.1 and Section 5.5.3 as well as the corresponding method to compute the risk-neutral price $C_0^{\mathbb{Q}}$ of European puts.

5.10.1 Regime-switching

The change of measure used in this study is the popular choice of shifting the periodic drift to obtain risk-neutrality and model invariance as considered for instance in Hardy (2001). Under this change of measure $\mathbb{Q}$, the drift of each regime $\mu_i \Delta$ is shifted to $(r - \sigma_i^2/2)\Delta$, and the transition probabilities are left unchanged. The risk-neutral dynamics has the representation

$$y_{n+1} = \left(r - \frac{\sigma_{h_n}^2}{2} \right) \Delta + \sigma_{h_n} \sqrt{\Delta} \epsilon_{n+1}^{\mathbb{Q}}, \quad n = 0, \ldots, N-1.$$

To compute the risk-neutral price of Φ, the approach used follows the work of Godin et al. (2019) (see Section 5.3 of the latter paper). Let $\mathbb{H} \equiv \{\mathcal{H}_n\}_{n=0}^N$ be the filtration generated by the regimes and $\mathbb{G}$ be the filtration containing all latent factors and all market information available to financial participants, i.e. $\mathbb{G} \equiv \mathbb{F} \vee \mathbb{G}$. Using the law of iterative expectations, the risk-neutral price of Φ allows for the representation

$$\begin{aligned} C_0^{\mathbb{Q}} &\equiv e^{-rT} \mathbb{E}^{\mathbb{Q}}[\Phi(S_N^{(0,b)}) | \mathcal{F}_0] \\ &= e^{-rT} \mathbb{E}^{\mathbb{Q}} \left[\mathbb{E}^{\mathbb{Q}}[\Phi(S_N^{(0,b)}) | \mathcal{G}_0] | \mathcal{F}_0 \right] \\ &= e^{-rT} \sum_{i=1}^{H} \xi_{0,i}^{\mathbb{Q}} \mathbb{E}^{\mathbb{Q}}[\Phi(S_N^{(0,b)}) | h_0 = i, S_0^{(0,b)}], \end{aligned} \tag{5.10}$$

where $\xi^{\mathbb{Q}}_{0,i}$ is assumed to be equal to $\xi^{\mathbb{P}}_{0,i}$ for all regimes, i.e. to the stationary distribution of the Markov chain under $\mathbb{P}$. The computation of the conditional expectations in (5.10) can be done for instance with Monte Carlo simulations or with the closed-form solution of Hardy (2001) when $H = 2$.

5.10.2 BSM

The change of measure from $\mathbb{P}$ to $\mathbb{Q}$ under the BSM dynamics is the one obtained with the discrete-time version of the Girsanov theorem: there exists a market price of risk process denoted as $\psi \equiv \{\psi_n\}_{n=1}^{N}$ such that $\epsilon_n^{\mathbb{Q}} = \epsilon_n + \psi_n$. By setting $\psi_n \equiv \sqrt{\Delta}(\frac{\mu - r}{\sigma})$, it is easy to show that $\{S_n^{(0,b)} e^{-rt_n}\}_{n=0}^{N}$ is an $(\mathbb{F}, \mathbb{Q})$-martingale and that the $\mathbb{Q}$-dynamics of log-returns is

$$y_n = \left(r - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\epsilon_n^{\mathbb{Q}}.$$

Risk-neutral put option prices presented in this paper are computed with the well-known Black-Scholes closed-form solution.

5.10.3 GARCH

The change of measure from $\mathbb{P}$ to $\mathbb{Q}$ considered is the one from Duan (1995) where the one-period conditional expected log-return is shifted, but the one-period conditional variance is unchanged when going from the physical to the risk-neutral measure. More precisely, let $\epsilon_n^{\mathbb{Q}} = \epsilon_n + \psi_n$ where $\psi \equiv \{\psi_n\}_{n=1}^{N}$ is predictable with respect to the filtration $\mathbb{F}$. In order to respect the condition of invariance of the one-period conditional variance from $\mathbb{P}$ to $\mathbb{Q}$, the one-period expected conditional return under $\mathbb{Q}$ must be equal to the periodic risk-free rate for $n = 1, \ldots, N$:

$$\mathbb{E}^{\mathbb{Q}}[e^{y_n}|\mathcal{F}_{n-1}] = \mathbb{E}^{\mathbb{Q}}[e^{\mu - \psi_n\sigma_n + \sigma_n\epsilon_n^{\mathbb{Q}}}|\mathcal{F}_{n-1}] = e^{\mu - \psi_n\sigma_n + \sigma_n^2/2} = e^{r\Delta}.$$

Thus, ψ_n has the representation

$$\psi_n \equiv \frac{\mu - r\Delta + \sigma_n^2/2}{\sigma_n}, \quad n = 1, \ldots, N. \tag{5.11}$$

With (5.11), the GARCH risky asset dynamics under $\mathbb{Q}$ is

$$\begin{aligned} y_n &= r\Delta - \sigma_n^2/2 + \sigma_n \epsilon_n^{\mathbb{Q}}, \\ \sigma_{n+1}^2 &= \omega + \upsilon \sigma_n^2 (|\epsilon_n^{\mathbb{Q}} - \psi_n| - \gamma(\epsilon_n^{\mathbb{Q}} - \psi_n))^2 + \beta \sigma_n^2. \end{aligned}$$

The computation of the risk-neutral price $C_0^{\mathbb{Q}}$ can be performed with Monte-Carlo simulations.

5.10.4 Merton jump-diffusion

For this model, the change of measure used is the one originally proposed in Merton (1976) which assumes no risk premia for jump risk: parameters $\{\mu_J, \sigma_J, \lambda, \sigma\}$ are left unchanged, and the drift parameter υ is shifted to the annualized continuously compounded risk-free rate r. The $\mathbb{Q}$-dynamics is

$$y_n = \left(r - \lambda(e^{\mu_J + \sigma_J^2/2} - 1) - \frac{\sigma^2}{2} \right) \Delta + \sigma\sqrt{\Delta}\epsilon_n^{\mathbb{Q}} + \sum_{j=N_{n-1}^{\mathbb{Q}}+1}^{N_n^{\mathbb{Q}}} \zeta_j^{\mathbb{Q}},$$

where $\{N_n^{\mathbb{Q}}\}_{n=0}^{N}$ and $\{\zeta_j^{\mathbb{Q}}\}_{j\geq 1}$ have the same distributions under $\mathbb{Q}$ than under $\mathbb{P}$. The risk-neutral price of put options $C_0^{\mathbb{Q}}$ can be computed with the well-known closed-form solution.

Conclusion

This thesis studies the use of deep reinforcement learning methods for pricing and hedging financial derivatives in incomplete markets. Throughout all four papers, global hedging problems are solved by using the deep hedging algorithm of Buehler et al. (2019b). This approach is based on a parametric approximation of trading policies with neural networks trained with reinforcement learning.

The first paper presents a universal and tractable methodology based on deep reinforcement learning for implementing the equal risk pricing framework of Marzban et al. (2020) under convex risk measures. The methodology consists in representing the long and short global trading policies with two distinct neural networks. This paper also has theoretical contributions by providing a rigorous proof that equal risk prices under the setup considered in the paper are arbitrage-free. Furthermore, this paper introduces ϵ-completeness metrics based on non-quadratic risk measures quantifying the level of market incompleteness.

The second paper studies the problem of global hedging long-term financial derivatives. The setup considered is similar to the work of Coleman et al. (2007) with the risk management of long-term lookback options embedded in variable annuities guarantees with ratchet features. The main contribution of this paper is in conducting extensive benchmarking of global policies under different penalties (quadratic and non-quadratic) with the use of multiple hedging instruments (standard options and underlying stock) in the presence of different dynamics for the financial market (presence of jump risk). These experiments demonstrate the significant benefit of using the neural network-based hedging scheme with non-quadratic penalties by successfully reducing downside risk as compared to typical hedging schemes (global quadratic hedging, local risk minimization and greek-based hedging) as well as earning significant positive average returns.

The main theme of the third paper is the examination of the impact of including short-term options as hedging instruments for pricing longer-term derivatives with the equal risk pricing framework. Broad analysis of the interrelation between different stylized features of jump and volatility risks and the choice of hedging instruments (short-term options or underlying stock) on the equal risk pricing framework are provided.

The fourth paper of the thesis investigates the use of non-translation invariant risk measures within the equal risk pricing framework. A modification to the conventional deep hedging training algorithm is proposed to tackle the additional difficulty of using non-translation invariant risk measures. The use of the class of semi-$\mathbb{L}^p$ objective functions within the equal risk pricing framework is shown to be highly effective from the standpoint of pricing derivatives by spanning larger interval of prices than under CVaR measures considered in previous papers as well as from a risk mitigation standpoint with global trading policies exhibiting highly effective hedging risk mitigation.

Future research on deep hedging methods for pricing and hedging derivatives would prove worthwhile. Robustness analysis of the empirical performance of the neural network trading policies in the presence of model uncertainty when dynamics differ from the ones used for training would be of interest. For instance, an assessment of the hedging performance of the neural networks in extreme market scenarios (e.g. financial crisis) could be of interest. Furthermore, in the same vein, a comprehensive empirical study of the equal risk pricing framework with methods developed in this thesis would prove worthwhile.

Moreover, in the context of mitigating the risk exposure associated with options embedded in guarantees of variable annuities with deep hedging methods, the inclusion of additional risk factors to have a more realistic depiction of the hedging problem such as stochastic interest rate, mortality risk, basis risk and lapse risk would be of interest. Another potential avenue to explore to improve the effectiveness of deep hedging methods is the development of more realistic financial market simulators to train the neural networks. Some studies are considering the use of deep learning methods such as Generative Adversarial Networks and Variational Autoencoders for model-free market simulation: Wiese et al. (2019), Buehler et al.

(2020) and Wiese et al. (2020).

Appendix

The objective of this appendix is twofold. First, Section 7.1 presents the proof that if Φ is attainable, than the equal risk price under convex risk measures coincides with the initial capital investment of the replicating strategy. Some parts of the proof are inspired by the proof of Lemma 3.2 of Buehler et al. (2019b). Second, Section 7.2 presents several papers which identified sufficient conditions for the existence of optimal trading strategies for global hedging problems under similar setups than the ones studied in this thesis.

7.1 Attainable case

Suppose that a European-type contingent claim $\Phi(S_N)$ is attainable, i.e. there exist $\hat{\delta} \in \Pi$ and $C_0 \in \mathbb{R}$ such that $-\Phi(S_N) + B_N(C_0 + G_N^{\hat{\delta}}) = 0$. Furthermore, let ρ be a convex risk measure. For any $\delta \in \Pi$, we have

$$\begin{aligned}\rho(\Phi(S_N) - B_N G_N^{\delta}) &= \rho(\Phi(S_N) + B_N C_0 - B_N C_0 + B_N G_N^{\hat{\delta}} - B_N G_N^{\hat{\delta}} - B_N G_N^{\delta}) \\ &= B_N C_0 + \rho(-B_N G_N^{\delta - \hat{\delta}}) \qquad (7.1)\end{aligned}$$

where (7.1) stems from the translation invariance property of ρ and from the fact that $\Phi(S_N) - B_N(C_0 + G_N^{\hat{\delta}}) = 0$. Taking the minimum over all admissible trading strategies, we obtain

$$\epsilon^{(S)}(0) = \min_{\delta \in \Pi} \rho\left(\Phi(S_N) - B_N G_N^{\delta}\right) = B_N C_0 + \min_{\delta \in \Pi} \rho(-B_N G_N^{\delta - \hat{\delta}}). \qquad (7.2)$$

Similarly, for any $\delta \in \Pi$, we have

$$\begin{aligned}\rho(-\Phi(S_N) - B_N G_N^{\delta}) &= \rho(-\Phi(S_N) + B_N C_0 - B_N C_0 + B_N G_N^{\hat{\delta}} - B_N G_N^{\hat{\delta}} - B_N G_N^{\delta}) \\ &= -B_N C_0 + \rho(-B_N G_N^{\delta + \hat{\delta}}) \qquad (7.3)\end{aligned}$$

where (7.3) stems from the translation invariance property of ρ and from the fact that $-\Phi(S_N) + B_N(C_0 + G_N^{\hat{\delta}}) = 0$. Taking the minimum over all admissible trading strategies, we obtain

$$\epsilon^{(L)}(0) = \min_{\delta \in \Pi} \rho\left(-\Phi(S_N) - B_N G_N^{\delta}\right) = -B_N C_0 + \min_{\delta \in \Pi} \rho(-B_N G_N^{\delta + \hat{\delta}}). \tag{7.4}$$

Lastly, using the characterization (2.11) of equal risk prices under convex risk measures, we have

$$C_0^\star = \frac{\epsilon^{(S)}(0) - \epsilon^{(L)}(0)}{2B_N} = \frac{B_N C_0 + \min_{\delta \in \Pi} \rho(-B_N G_N^{\delta - \hat{\delta}}) + B_N C_0 - \min_{\delta \in \Pi} \rho(-B_N G_N^{\delta + \hat{\delta}})}{2B_N} = C_0. \tag{7.5}$$

This concludes the demonstration that if a European-type contingent claim $\Phi(S_N)$ is attainable, than the equal risk price under convex risk measures coincides with the initial capital investment of the replicating portfolio.

7.2 Existence of optimal trading strategy

Throughout the thesis, for each example of global hedging problems considered, the identification of sufficient conditions leading to the existence of an optimal trading strategy is left out-of-scope. Such conditions were investigated in other literature works. Here are some papers providing sufficient conditions for the existence of optimal global trading strategy for optimization problems that are related to the ones considered in this thesis.

- Schweizer (1995) considers the case of a quadratic penalty in a discrete-time frictionless market with a single traded asset being the underlying stock. The latter work also considers the case where both the trading strategy as well as the initial capital investment are jointly optimized. The optimization problems (3.6) with $\mathcal{L}(x) = x^2$ and (4.30) are both related to the global hedging problems considered by Schweizer (1995). Theorem 2.2 of the latter paper demonstrates the existence of an admissible hedging strategy under the assumption of a so-called nondegeneracy condition for the stock price process.

- François et al. (2014) considers the case of expected penalties as risk measures in a discrete-time frictionless market with trades involving the underlying stock. The optimization problem (3.6) with $\mathcal{L}(x) = x^2 \mathbb{1}_{\{x>0\}}$ as well as the global hedging problem (5.1) with the objective function (5.4) are both related to the global hedging problem considered in François et al. (2014). Theorem 3.3 of the latter paper provides sufficient conditions for the existence of an optimal admissible hedging strategy in such context.

- Godin (2016) considers the case of the CVaR risk measure in discrete-time in the presence of transaction costs for trades in the underlying stock. Throughout this thesis, the CVaR risk measure is extensively studied in various numerical experimentations of the ERP framework. Lemma 2.1 of Godin (2016) provides sufficient conditions for an optimal admissible hedging strategy to exist under the CVaR objective function.

- Xu (2006) considers the case of convex risk measures as the objective function in continuous-time. Theorem 2.6 of the latter paper provides sufficient conditions for an optimal admissible hedging strategy to exist in this setup.

Bibliography

Abadi, M. et al. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*.

Alexander, S., Coleman, T. F., and Li, Y. (2003). Derivative portfolio hedging based on CVaR. *New Risk Measures in Investment and Regulation: Wiley*.

Alfeus, M., He, X.-J., and Zhu, S.-P. (2019). An empirical study of the option pricing formula with the underlying banned from short sell. *Available at SSRN 3478355*.

Almahdi, S. and Yang, S. Y. (2017). An adaptive portfolio trading system: A risk-return portfolio optimization using recurrent reinforcement learning with expected maximum drawdown. *Expert Systems with Applications*, 87:267–279.

Ang, A. and Timmermann, A. (2012). Regime changes and financial markets. *Annual Review of Financial Economics*, 4(1):313–337.

Ankirchner, S., Schneider, J. C., and Schweizer, N. (2014). Cross-hedging minimum return guarantees: Basis and liquidity risks. *Journal of Economic Dynamics and Control*, 41:93–109.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Augustyniak, M. and Boudreault, M. (2017). Mitigating interest rate risk in variable annuities: An analysis of hedging effectiveness under model risk. *North American Actuarial Journal*, 21(4):502–525.

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

Bao, W. and Liu, X.-Y. (2019). Multi-agent deep reinforcement learning for liquidation strategy analysis. *arXiv preprint arXiv:1906.11046*.

Barigou, K., Bignozzi, V., and Tsanakas, A. (2020). Insurance valuation: A two-step generalised regression approach. *arXiv preprint arXiv:2012.04364*.

Basak, S. and Chabakauri, G. (2012). Dynamic hedging in incomplete markets: a simple solution. *The Review of Financial Studies*, 25(6):1845–1896.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107.

Bauer, D., Kling, A., and Russ, J. (2008). A universal pricing framework for guaranteed minimum benefits in variable annuities. *ASTIN Bulletin: The Journal of the IAA*, 38(2):621–651.

BCBS (2016). *Standards. Minimum Capital Requirements for Market Risk. January 2016*. Basel Committee on Banking Supervision. Basel: Bank for International Settlements.

Becker, S., Cheridito, P., and Jentzen, A. (2019). Deep optimal stopping. *Journal of Machine Learning Research*, 20:1–25.

Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166.

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

Betancourt, C. and Chen, W.-H. (2021). Deep reinforcement learning for portfolio management of markets with a dynamic number of assets. *Expert Systems with Applications*, 164:114002.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Bollen, N. P. (1998). Valuing options in regime-switching models. *Journal of Derivatives*, 6:38–50.

Boyle, P. P. and Hardy, M. R. (1997). Reserving for maturity guarantees: Two approaches.

Insurance: Mathematics and Economics, 21(2):113–127.

Boyle, P. P. and Schwartz, E. S. (1977). Equilibrium prices of guarantees under equity-linked contracts. *Journal of Risk and Insurance*, 44:639–660.

Brennan, M. J. and Schwartz, E. S. (1976). The pricing of equity-linked life insurance policies with an asset value guarantee. *Journal of Financial Economics*, 3(3):195–213.

Breton, M. and Godin, F. (2017). Global hedging through post-decision state variables. *Journal of Risk and Financial Management*, 10(3):16.

Buehler, H. et al. (2019a). Deep hedging: hedging derivatives under generic market frictions using reinforcement learning. Technical Report 19-80.

Buehler, H., Gonon, L., Teichmann, J., and Wood, B. (2019b). Deep hedging. *Quantitative Finance*, 19(8):1271–1291.

Buehler, H., Horvath, B., Lyons, T., Perez Arribas, I., and Wood, B. (2020). A data-driven market simulator for small data environments. *Available at SSRN 3632431*.

Cao, H., Cui, Z., and Liu, Y. (2020). Discrete-time variance-optimal deep hedging in affine GARCH models. *Available at SSRN 3659275*.

Cao, J., Chen, J., Hull, J., and Poulos, Z. (2021). Deep hedging of derivatives using reinforcement learning. *The Journal of Financial Data Science*, 3(1):10–27.

Carbonneau, A. (2021). Deep hedging of long-term financial derivatives. *Insurance: Mathematics and Economics*, 99:327–340.

Carbonneau, A. and Godin, F. (2021a). Deep equal risk pricing of financial derivatives with multiple hedging instruments. *arXiv preprint arXiv:2102.12694*.

Carbonneau, A. and Godin, F. (2021b). Equal risk pricing of derivatives with deep hedging. *Quantitative Finance*, 21(4):593–608.

Carr, P. and Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal* *of Computational Finance*, 2(4):61–73.

Carr, P. and Wu, L. (2014). Static hedging of standard options. *Journal of Financial Econometrics*, 12(1):3–46.

Christoffersen, P., Elkamhi, R., Feunou, B., and Jacobs, K. (2010). Option valuation with conditional heteroskedasticity and nonnormality. *The Review of Financial Studies*, 23(5):2139–2183.

Coleman, T., Kim, Y., Li, Y., and Patron, M. (2007). Robustly hedging variable annuities with guarantees under jump and volatility risks. *Journal of Risk and Insurance*, 74(2):347–376.

Coleman, T., Li, Y., and Patron, M. (2006). Hedging guarantees in variable annuities under both equity and interest rate risks. *Insurance: Mathematics and Economics*, 38(2):215–228.

Cont, R. and Da Fonseca, J. (2002). Dynamics of implied volatility surfaces. *Quantitative Finance*, 2(1):45–60.

Cont, R. and Tankov, P. (2003). *Financial modelling with jump processes.* CRC press.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):463–520.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22.

Deng, Y. et al. (2016). Deep direct reinforcement learning for financial signal representation and trading. *IEEE Transactions on Neural Networks and Learning Systems*, 28(3):653–664.

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

Dupuis, D., Gauthier, G., and Godin, F. (2016). Short-term hedging for an electricity retailer. *The Energy Journal*, 37(2):31–59.

Eberlein, E. and Jacod, J. (1997). On the range of options prices. *Finance and Stochastics*, 1(2):131–140.

El Karoui, N. and Quenez, M.-C. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal on Control and Optimization*, 33(1):29–66.

Elliott, R. J. and Madan, D. B. (1998). A discrete time equivalent martingale measure. *Mathematical Finance*, 8(2):127–152.

Föllmer, H. and Leukert, P. (1999). Quantile hedging. *Finance and Stochastics*, 3(3):251–273.

Föllmer, H. and Leukert, P. (2000). Efficient hedging: cost versus shortfall risk. *Finance and Stochastics*, 4(2):117–146.

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447.

Föllmer, H. and Schweizer, M. (1988). Hedging by sequential regression: An introduction to the mathematics of option trading. *ASTIN Bulletin: The Journal of the IAA*, 18(2):147–160.

Föllmer, H. and Schweizer, M. (1991). Hedging of contingent claims under incomplete information. *Applied Stochastic Analysis*, 5:389–414.

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

Frittelli, M. (2000). The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance*, 10(1):39–52.

Gan, G. (2013). Application of data clustering and machine learning in variable annuity valuation. *Insurance: Mathematics and Economics*, 53(3):795–801.

García-Galicia, M., Carsteanu, A. A., and Clempner, J. B. (2019). Continuous-time reinforcement learning approach for portfolio management with time penalization. *Expert Systems with Applications*, 129:27–36.

Gerber, H. U. and Shiu, E. S. W. (1994). Option pricing by Esscher transforms. *Transactions of the Society of Actuaries*, 46:99–191.

Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth International Conference on Artificial Intelligence and Statistics*, pages 249–256.

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

Godin, F. (2016). Minimizing CVaR in global dynamic hedging with transaction costs. *Quantitative Finance*, 16(3):461–475.

Godin, F., Lai, V. S., and Trottier, D.-A. (2019). Option pricing under regime-switching models: Novel approaches removing path-dependence. *Insurance: Mathematics and Economics*, 87:130–142.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.

Guo, I. and Zhu, S.-P. (2017). Equal risk pricing under convex trading constraints. *Journal of Economic Dynamics and Control*, 76:136–151.

Gushchin, A. A. and Mordecki, E. (2002). Bounds of option prices for semimartingale market models. *Proceedings of the Steklov Institute of Mathematics*, 237:73–113.

Halperin, I. (2020). Qlbs: Q-learner in the black-scholes (-merton) worlds. *The Journal of Derivatives*.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384.

Han, J. and E, W. (2016). Deep learning approximation for stochastic control problems. *arXiv preprint arXiv:1611.07422*.

Hardy, M. R. (2000). Hedging and reserving for single-premium segregated fund contracts. *North American Actuarial Journal*, 4(2):63–74.

Hardy, M. R. (2001). A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 5(2):41–53.

Hardy, M. R. (2003). *Investment guarantees: modeling and risk management for equity-linked life insurance*, volume 215. John Wiley & Sons.

Harrison, J. M. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3):215–260.

He, X.-J. and Zhu, S.-P. (2020). A revised option pricing formula with the underlying being banned from short selling. *Quantitative Finance*, 20(6):935–948.

Hendricks, D. and Wilcox, D. (2014). A reinforcement learning extension to the Almgren-Chriss framework for optimal trade execution. In *2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pages 457–464. IEEE.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Hodges, S. and Neuberger, A. (1989). Optimal replication of contingent claims under transactions costs. *Review of Futures Markets*, 8:222–239.

Hong, L. J., Hu, Z., and Liu, G. (2014). Monte Carlo methods for value-at-risk and conditional value-at-risk: a review. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 24(4):1–37.

Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2):251–257.

Horvath, B., Teichmann, J., and Zuric, Z. (2021). Deep hedging under rough volatility. *arXiv preprint arXiv:2102.01962*.

Hull, J. C. (2003). *Options, futures, and other derivatives.* Pearson Education India.

IASB (2017). *IFRS 17 Insurance Contracts.* International Accounting Standards Board.

Jiang, Z., Xu, D., and Liang, J. (2017). A deep reinforcement learning framework for the financial portfolio management problem. *arXiv preprint arXiv:1706.10059.*

Kélani, A. and Quittard-Pinon, F. (2017). Pricing and hedging variable annuities in a Lévy market: a risk management perspective. *Journal of Risk and Insurance*, 84(1):209–238.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980.*

Kolm, P. N. and Ritter, G. (2019). Dynamic replication and hedging: A reinforcement learning approach. *The Journal of Financial Data Science*, 1(1):159–171.

Lamberton, D. and Lapeyre, B. (2011). *Introduction to Stochastic Calculus Applied to Finance.* Chapman and Hall/CRC.

Li, Y., Szepesvari, C., and Schuurmans, D. (2009). Learning exercise policies for American options. In *Artificial Intelligence and Statistics*, pages 352–359. PMLR.

Lu, D. W. (2017). Agent inspired trading using recurrent reinforcement learning and lstm neural networks. *arXiv preprint arXiv:1707.07338.*

Ma, G., Zhu, S.-P., and Guo, I. (2019). Pricing contingent claims with short selling bans. *arXiv preprint arXiv:1910.04960.*

Marzban, S., Delage, E., and Li, J. Y. (2020). Equal risk pricing and hedging of financial derivatives with convex risk measures. *arXiv preprint arXiv:2002.02876.*

Melnikov, A. and Smirnov, I. (2012). Dynamic hedging of conditional value-at-risk. *Insurance: Mathematics and Economics*, 51(1):182–190.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, pages 141–183.

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

Moody, J. and Saffell, M. (2001). Learning to trade via direct reinforcement. *IEEE Transactions on Neural Networks*, 12(4):875–889.

Moody, J. and Wu, L. (1997). Optimization of trading systems and portfolios. In *Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr)*, pages 300–307. IEEE.

Ng, A. C.-Y. and Li, J. S.-H. (2011). Valuing variable annuity guarantees with the multivariate Esscher transform. *Insurance: Mathematics and Economics*, 49(3):393–400.

Ning, B., Lin, F. H. T., and Jaimungal, S. (2018). Double deep Q-learning for optimal execution. *arXiv preprint arXiv:1812.06600*.

Pelsser, A. (2003). Pricing and hedging guaranteed annuity options via static option replication. *Insurance: Mathematics and Economics*, 33(2):283–296.

Pendharkar, P. C. and Cusatis, P. (2018). Trading financial indices with reinforcement learning agents. *Expert Systems with Applications*, 103:1–13.

Powell, W. B. (2009). What you should know about approximate dynamic programming. *Naval Research Logistics (NRL)*, 56(3):239–249.

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

Rockafellar, R. T. and Uryasev, S. (2002). Conditional Value-at-Risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533–536.

Schweizer, M. (1991). Option hedging for semimartingales. *Stochastic processes and their*

Applications, 37(2):339–363.

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

Schweizer, M. (1996). Approximation pricing and the variance-optimal martingale measure. *The Annals of Probability*, 24(1):206–236.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Trindade, A. A., Uryasev, S., Shapiro, A., and Zrazhevsky, G. (2007). Financial prediction with constrained tail risk. *Journal of Banking & Finance*, 31(11):3524–3538.

Trottier, D.-A., Godin, F., and Hamel, E. (2018a). Local hedging of variable annuities in the presence of basis risk. *ASTIN Bulletin: The Journal of the IAA*, 48(2):611–646.

Trottier, D.-A., Godin, F., and Hamel, E. (2018b). On fund mapping regressions applied to segregated funds hedging under regime-switching dynamics. *Risks*, 6(3):78.

Wang, H. and Zhou, X. Y. (2020). Continuous-time mean-variance portfolio selection: A reinforcement learning framework. *Mathematical Finance*, 30(4):1273–1308.

Wiese, M., Bai, L., Wood, B., and Buehler, H. (2019). Deep hedging: learning to simulate equity option markets. *Available at SSRN 3470756*.

Wiese, M., Knobloch, R., Korn, R., and Kretschmer, P. (2020). Quant GANs: Deep generation of financial time series. *Quantitative Finance*, 20(9):1419–1440.

Xu, M. (2006). Risk measure pricing and hedging in incomplete markets. *Annals of Finance*, 2(1):51–71.

Ye, Y., Pei, H., Wang, B., Chen, P.-Y., Zhu, Y., Xiao, J., and Li, B. (2020). Reinforcement-learning based portfolio management with augmented asset movement prediction states. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 1112–1119.

Zhang, F. (2010). Integrating robust risk management into pricing: New thinking for VA writers. *Risk and Rewards*, 55:34–36.