risks

MDPI

Article

Deep Equal Risk Pricing of Financial Derivatives with Non-Translation Invariant Risk Measures †

Alexandre Carbonneau and Frédéric Godin *

Department of Mathematics and Statistics, Concordia University, Montréal, QC H3G 1M8, Canada; alexandre.carbonneau@mail.concordia.ca
* Correspondence: frederic.godin@concordia.ca
† A GitHub repository with some samples of codes can be found at github.com/alexandrecarbonneau (accessed on 1 July 2023).

Abstract: The objective is to study the use of non-translation invariant risk measures within the equal risk pricing (ERP) methodology for the valuation of financial derivatives. 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 observed under the tail value at risk-based ERP as documented in previous work. 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 so as to enable obtaining a price through a single training run for the two neural networks associated with the respective 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: finance; option pricing; hedging; reinforcement learning; deep learning

check for updates

Citation: Carbonneau, Alexandre, and Frédéric Godin. 2023. Deep Equal Risk Pricing of Financial Derivatives with Non-Translation Invariant Risk Measures. *Risks* 11: 140. https://doi.org/10.3390/risks11080140

Academic Editor: Mogens Steffensen

Received: 26 June 2023
Revised: 14 July 2023
Accepted: 28 July 2023
Published: 1 August 2023

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 that 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 Carbonneau and Godin 2021b; Guo and Zhu 2017; Marzban et al. 2022), 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. (2022) 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 the improved interpretability of the 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 consider 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 hedging portfolio initial values.

The most natural convex risk measure to consider within the ERP approach is arguably 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 the worst-case risks occurring within an event of probability below or exactly at $1-\alpha$, 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 use the CVaR within the ERP framework: see Carbonneau and Godin (2021a, 2021b). It was observed in the former 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 can cast doubt on the applicability of the method in practice. An avenue that was explored in the aforementioned study 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 present 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 by Carbonneau and Godin (2021a) consists in incorporating other hedging instruments (e.g., short-term options) within dynamic hedging schemes. That approach is 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, alleviate the price inflation obtained with CVaR-based ERP. A straightforward route to explore so as to attempt obtaining a satisfactory ERP method while 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 lower option prices. However, such risk measures (e.g., the semi-variance and semi-root mean square error (SRMSE)) do not necessarily satisfy the 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 for 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 by Carbonneau and Godin (2021a, 2021b) 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 with simulated risky asset paths to the deep hedging algorithm from Buehler et al. (2019), 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, by Marzban et al. (2022).

The training algorithm modification is shown in the present 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 those 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 wider ranges of prices than those obtained under the CVaR$_\alpha$ risk measures with conventional confidence level α values. The latter phenomenon is shown to hold across all moneyness levels for puts and is robust for 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 with extreme hedging losses. This is in contrast with the CVaR$_\alpha$ objective function, where hedging policies optimized with relatively small confidence level α exhibit poor risk mitigation for loss quantiles 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 in reducing equal risk price values as compared to values obtained by trading shorter-term options with the CVaR$_\alpha$ 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 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 present work is presented in Section 3. Section 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 displays the results of numerical experiments associated with semi-$\mathbb{L}^p$ risk measures-based ERP. Section 6 concludes.

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 the literature considers pricing methods consistent with the 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 with 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 shown by 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 whether a long or short position in the derivative is considered. The ERP approach developed by Guo and Zhu (2017) identifying the derivative price equating the 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 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. (2022), Ma et al. (2022) and He and Zhu (2020). Marzban et al. (2022) 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. This setting is extended by Marzban et al. (2021) to rather consider dynamic risk measures within the objective function and obtain a time-consistent problem.

The computation of equal risk prices for derivatives is a highly non-trivial endeavor requiring advanced numerical schemes in most cases. Marzban et al. (2022) propose to use dynamic programming, which they apply on a robust optimization setting. Conversely, Carbonneau and Godin (2021a, 2021b) use the deep reinforcement learning approach of Buehler et al. (2019) 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), Horvath et al. (2021) and Lütkebohmert et al. (2022). Deep reinforcement learning is a very convenient 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 by Halperin (2020) and Cao et al. (2021), proximal policy optimization by Chong et al. (2021), least squares policy iteration and fitted Q-iteration for American option pricing by Li et al. (2009), or batch policy gradient by Buehler et al. (2019). Moreover, various other financial problems were tackled through reinforcement learning procedures in the literature, for instance, portfolio management by 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 by Bao and Liu (2019); or trading optimization by Hendricks and Wilcox (2014), Lu (2017) and Ning et al. (2021).

3. Mathematical Setup for the Financial Market

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

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 present 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 that $S_n^{(b)}$ 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 reach maturity and 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 being 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 that all positions in a given portfolio are liquidated at the end of any period and are repurchased at the beginning of the next period 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. Define a trading strategy $\delta \equiv \{\delta_n\}_{n=0}^N$ as an $\mathbb{F}$-predictable process[4], 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 dot product. 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.[5] 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 are 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, by Lamberton and Lapeyre (2007), 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 $n = 0, \ldots, N$ 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. The solutions of 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.[6] 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

$$\epsilon^{(\mathcal{L})}(V_0) \equiv \min_{\delta \in \Pi} \rho\left(-\Phi(S_N^{(0,b)}) - V_N^\delta\right), \quad \epsilon^{(\mathcal{S})}(V_0) \equiv \min_{\delta \in \Pi} \rho\left(\Phi(S_N^{(0,b)}) - V_N^\delta\right). \tag{1}$$

The optimal hedging strategies are the minimizing arguments of such optimization problems:

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

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.,

$$\epsilon^{(\mathcal{L})}(-C_0^*) = \epsilon^{(\mathcal{S})}(C_0^*). \tag{2}$$

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 (2), see Theorem 2.1 of the latter paper.

Remark 1. *Under a convex measure ρ, Marzban et al. (2022) and Carbonneau and Godin (2021b) also obtain the following characterization of the equal risk price*

$$C_0^* = 0.5 B_N \left(\epsilon^{(\mathcal{S})}(0) - \epsilon^{(\mathcal{L})}(0) \right). \tag{3}$$

Representation (3) is very convenient, as it only requires to 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 (2).

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

$$\rho(X) \equiv \mathbb{E}\left[X^p \mathbb{1}_{\{X>0\}} \right]^{1/p}, \quad p > 0. \tag{4}$$

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. One important advantageous property of this class of risk measures is in penalizing exclusively hedging losses, not gains. 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 present paper for benchmarking purposes, as it is used by Carbonneau and Godin (2021a, 2021b). 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 $\text{CVaR}_\alpha(X) = \mathbb{E}[X | X \geq \text{VaR}_\alpha(X)]$. The CVaR is a coherent risk measure as shown by Rockafellar and Uryasev (2002), which implies that it satisfies the monotonicity and translation invariance properties.

4. Methodology

The present 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. (2019) 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.

4.1. Neural Network Approximation of the Optimal Solution

The approach followed to obtain a numerical solution to the optimization problems (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 $\tilde{\delta}^{(\mathcal{S})}$:

$$\delta^{(\mathcal{L})}_{n+1}(V_0) = \tilde{\delta}^{(\mathcal{L})}\Big(T - t_n, S^{(b)}_n, V_n, \mathcal{I}_n\Big), \tag{5}$$

$$\delta^{(\mathcal{S})}_{n+1}(V_0) = \tilde{\delta}^{(\mathcal{S})}\Big(T - t_n, S^{(b)}_n, V_n, \mathcal{I}_n\Big), \quad n = 0, \ldots, N-1, \tag{6}$$

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 hedges for the long and short positions when time-0 capital investment is 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 asset has 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) and (6) enables approximating the optimal policies with 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 $\{T - t_n, S^{(b)}_n, V_n, \mathcal{I}_n\}$ into the respective (long or short) portfolio positions of risky assets $\delta^{(0:D)}_{n+1}$ for any $n = 0, \ldots, N-1$.[7] More precisely, denote by $F^{(\mathcal{L})}_\theta$ and $F^{(\mathcal{S})}_\theta$ 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.[8] For a given parameter set θ distinct for each neural network, the associated trading strategies are given by

$$\begin{aligned}
\delta^{(\mathcal{L},\theta)}_{n+1}(V_0) &\equiv F^{(\mathcal{L})}_\theta\Big(T - t_n, S^{(b)}_n, V_n, \mathcal{I}_n\Big), \\
\delta^{(\mathcal{S},\theta)}_{n+1}(V_0) &\equiv F^{(\mathcal{S})}_\theta\Big(T - t_n, S^{(b)}_n, V_n, \mathcal{I}_n\Big), \quad n = 0, \ldots, N-1.
\end{aligned}$$

The optimization of the trading strategy in problem (1) is thus replaced by the optimization of neural network parameters θ according to

$$\tilde{\epsilon}^{(\mathcal{L})}(V_0) \equiv \min_{\theta \in \mathbb{R}^q} \rho\Big(-\Phi(S^{(0,b)}_N) - V^{\delta^{(\mathcal{L},\theta)}}_N\Big), \quad \tilde{\epsilon}^{(\mathcal{S})}(V_0) \equiv \min_{\theta \in \mathbb{R}^q} \rho\Big(\Phi(S^{(0,b)}_N) - V^{\delta^{(\mathcal{S},\theta)}}_N\Big). \tag{7}$$

Note that the set of optimal parameters θ will be different for the long and the short trading strategies. Furthermore, problems (7) only lead to an approximate solution to the initial problems (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. (2019) show that there exist neural networks such that the solution $\tilde{\epsilon}^{(\mathcal{L})}, \tilde{\epsilon}^{(\mathcal{S})}$ from (7) can be made arbitrarily close to the solution $\epsilon^{(\mathcal{L})}, \epsilon^{(\mathcal{S})}$ from (1).

The mathematical definition of the 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:

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

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*, each of which successively applies an affine and a nonlinear transformation to the 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 $l = 1, \ldots, L+1$.

4.2. Calibration of Neural Networks through Reinforcement Learning

As shown by Buehler et al. (2019), 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 (7). Without loss of generality, the training algorithm is hereby only provided for the neural network $F_\theta^{(\mathcal{S})}$ associated with the short position, as steps for the long position are entirely analogous.

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 by Buehler et al. (2019). A slight modification to the algorithm will subsequently be presented in Section 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:

$$J(\theta, V_0) \equiv \rho\left(\Phi(S_N^{(0,b)}) - V_N^{\delta^{(\mathcal{S},\theta)}}\right), \quad \theta \in \mathbb{R}^q, V_0 \in \mathbb{R}. \tag{8}$$

The parameter 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 by 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 $N_{\text{batch}} = 1000$ of paths from the training set is used to estimate the cost function in (8). More precisely, for $\theta = \theta_j$, $F_\theta^{(\mathcal{S})}$ is used to compute the asset 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

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

Lastly, the computation of the gradient of the empirical cost function with respect to θ can be performed explicitly with modern deep learning libraries, such as Tensorflow (Abadi et al. 2016). Also, the Adam optimizer (Kingma and Ba 2014) can be used to dynamically determine 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.

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 $F_\theta^{(\mathcal{S})}$ and $F_\theta^{(\mathcal{L})}$ once, which thus circumvents the previously described computational burden.

The slight modification made to the training algorithm provided in Section 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 randomly sampled within the initial interval of values used for the bisection algorithm. For instance, in numerical experiments conducted in Section 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 $\mathbb{Q}$.[9] This approach 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, the Monte Carlo experiments provided in Appendix B show that incorporating this slight modification to the training algorithm does not materially impact the optimized neural networks performance.

Pseudo-codes of the training and bisection procedures are presented, respectively, in Algorithms A1 and A2 of Appendix A. An implementation in Python and Tensorflow to replicate numerical experiments presented in Section 5 can also be found online at github.com/alexandrecarbonneau (accessed on 1 July 2023).

Remark 2. *In numerical experiments of Section 5, the benchmarking of equal risk prices generated under the class of semi-$\mathbb{L}^p$ risk measures to these 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 (3), where $\tilde{\epsilon}^{(\mathcal{L})}(0)$ and $\tilde{\epsilon}^{(\mathcal{S})}(0)$ are computed with the steps of Section 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.

Remark 3. *For all numerical experiments under the semi-$\mathbb{L}^p$ risk measure conducted in this paper, a preprocessing of the feature vectors is applied, using $\{T-t_n, \log(S_n^{(b)}/K), V_n/\tilde{V}, \mathcal{I}_n\}$ instead of $\{T-t_n, S_n^{(b)}, V_n, \mathcal{I}_n\}$, where $\tilde{V}$ is 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 Carbonneau and Godin (2021a, 2021b) consider similar preprocessing for risky asset prices, while Carbonneau (2021) considers a similar preprocessing for portfolio values. Furthermore, under the CVaR$_\alpha$ 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 rather than $V_n/\tilde{V}$ is used in feature vectors.*

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$_\alpha$ measures as considered, for instance, by Carbonneau and Godin (2021a, 2021b) when relying on the neural network-based hedging scheme. Indeed, under the CVaR$_\alpha$ objective function, the use of minibatch stochastic gradient descent procedures to train neural networks restrains the use of extremely large quantiles for the CVaR$_\alpha$ (for instance, larger values than 0.99). The latter stems from the following observations. From a statistical standpoint, the estimation variance of CVaR$_\alpha$ increases with α. Furthermore, the empirical estimator of CVaR$_\alpha$ is biased in the finite sample size, whereas the empirical estimator of the semi-$\mathbb{L}^p$ risk measure is unbiased for any sample size. 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 by Goodfellow et al. (2016), the amount of memory required by hardware setups can be a limiting factor to increasing the minibatch 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 size on SGD procedures. This computational pitfall of pairing stochastic gradient descent with extreme values of α under the CVaR$_\alpha$ measure is not present under the semi-$\mathbb{L}^p$, which further motivates its use in the context of equal risk pricing and optimal hedging.

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 5.2 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$_\alpha$ 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 5.3 so as to test the impact of the inclusion of jump or volatility risk. Lastly, Section 5.4 presents the benchmarking of equal risk prices for long maturity options obtained under the semi-$\mathbb{L}^p$ risk measures with trades involving exclusively the underlying stock against these generated with option hedges under the CVaR$_\alpha$ objective function.

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 $T = 60/260$ 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 in Section 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 $r = 0.02$. Different moneyness levels are considered with $K = 90, 100$ and 110, respectively, for out-of-the-money (OTM), at-the-money (ATM), and in-the-money (ITM) puts.

Moreover, as described in Section 4, two distinct feedforward neural networks are considered for 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 100 epochs[10] is used with a minibatch size of 1000 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.

5.2. 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) observe 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 extent that can cast doubt on the applicability of ERP in practice. Furthermore, while the latter paper observes a positive relation between the risk aversion level α and equal risk prices $C_0^\star$ as shown in subsequent sections of this present 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 present section is to assess if the use of the semi-$\mathbb{L}^p$ class of risk measures helps in alleviating this price inflation phenomenon while simultaneously resulting in optimized trading policies which provide 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 objective functions but also the assessment of the effectiveness of the resulting global trading policies.

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 suitable dynamics for the underlying stock. Indeed, the model should incorporate salient stylized facts of financial markets with a specific focus on 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 performed 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 covering the period 31 December 1986 to 1 April 2010 (5863 observations). Parameter estimates are presented in Appendix C.

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 period of time, 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 $h_n \in \{1, \ldots, H\}$ 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

$$\mathbb{P}(h_{n+1} = j | \mathcal{F}_n, h_n, \ldots, h_0) = \gamma_{h_n,j}, \quad j = 1, \ldots, H, \tag{9}$$

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

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

where $\{\mu_i, \sigma_i\}_{i=1}^{H}$ are model parameters representing the means and volatilities on a yearly basis in 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 $\xi_n \equiv [\xi_{n,1}, \ldots, \xi_{n,H}]$ and $\xi_{n,j} \equiv \mathbb{P}(h_n = j|\mathcal{F}_n)$. Under the RS model, $\mathcal{I}_n = \xi_n$ for $n = 0, \ldots, N-1$. Following the work of François et al. (2014), the predictive probabilities can be computed recursively for $n = 0, \ldots, N-1$ as

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

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 Appendix D.

5.2.2. Numerical Results of the Sensitivity Analysis to the Objective Function

Table 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 $C_0^{\mathbb{Q}}$. Hedging statistics obtained across the different objective functions are analyzed subsequently in Section 5.2.3.

Table 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$_\alpha$			$C_0^\star$ under Semi-$\mathbb{L}^p$				
Moneyness	$C_0^{\mathbb{Q}}$	**CVaR$_{0.90}$**	**CVaR$_{0.95}$**	**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%	88%	111%	140%	175%
ATM	3.27	18%	24%	29%	10%	17%	22%	28%	35%
ITM	10.36	5%	7%	9%	2%	5%	7%	8%	9%

Notes: $C_0^\star$ results are computed based on 100,000 independent paths generated from the regime-switching model under $\mathbb{P}$ (see Section 5.2.1 for model definition and Appendix C for model parameters). Risk-neutral prices $C_0^{\mathbb{Q}}$ are computed under $\mathbb{Q}$-dynamics described in Appendix D. The training of neural networks is performed as described in Section 4.2 with hyperparameters presented in Section 5.1. $C_0^\star$ are expressed relative to $C_0^{\mathbb{Q}}$ (% increase).

Values from Table 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$_\alpha$ risk measures with the selected values for the confidence level α. The latter observation holds across all moneyness levels for puts. For instance, the relative increases in the equal risk price $C_0^\star$ as compared to the risk-neutral price $C_0^{\mathbb{Q}}$ for OTM puts are $91\%, 119\%$ and 161% under CVaR$_{0.90}$, CVaR$_{0.95}$ and CVaR$_{0.99}$, and range between 50% and 175% using the semi-$\mathbb{L}^p$, with p going from 2 to 10. Similar observations can be made for ATM and ITM moneyness levels. Furthermore, the use of the semi-$\mathbb{L}^2$ risk measure entails a significant reduction in $C_0^\star$ as compared to the price obtained under the CVaR$_{0.90}$. Indeed, the relative increases 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 are, respectively, 50%, 10% and 2%, which are significantly smaller than the corresponding relative increases of $91\%, 18\%$ and 5% under the 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 a rationale analogous to that mentioned by Carbonneau and Godin (2021b) under the CVaR$_\alpha$ risk measure case: since the put option 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 placing more weight on extreme hedging losses, which results in a larger increase in the 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 those generated by the CVaR with conventional confidence levels but by also significantly alleviating the inflated option prices phenomenon observed under the CVaR$_\alpha$. 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.

5.2.3. Hedging Performance Benchmarking

This section conducts the benchmarking of the neural network trading policies hedging performance under the CVaR$_\alpha$ 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 2 presents hedging statistics of the global hedging policies obtained with the CVaR$_\alpha$ and semi-$\mathbb{L}^p$ risk measures with the same 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.[11] Furthermore, the hedging metrics used for the benchmarking consist of the VaR$_\alpha$ and CVaR$_\alpha$ statistics over various α, 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 additional independent simulated paths.

Hedging metrics values show that while the trading policy optimized with the $\text{CVaR}_{0.90}$ objective function lead to the smallest values for the $\text{CVaR}_{0.90}$, $\text{VaR}_{0.90}$ and $\text{VaR}_{0.95}$ statistics, it exhibits poor mitigation of tail risk as compared to the other policies. For instance, the relative reduction in the $\text{CVaR}_{0.99}$ statistic achieved with all other penalties than the $\text{CVaR}_{0.90}$ ranges between 31.4% 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. The latter 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 CVaR_α with lower values for α, due to their poor mitigation of risk for quantiles above the CVaR confidence level. This conclusion has important implications in the context of the ERP framework. Indeed, as shown by Carbonneau and Godin (2021b), the equal risk price $C_0^\star$ obtained with the CVaR_α exhibits a positive relationship to α values. Consequently, the inflated equal risk price phenomenon observed under the class of CVaR_α measures cannot be effectively alleviated through the reduction in 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 p considered, optimized trading policies are effective for mitigating hedging risk. Recall that p controls the weight associated with extreme hedging losses. From the combination of these hedging statistics values as well as equal risk price values presented in Table 1, we can conclude that the class of semi-$\mathbb{L}^p$ risk measures is a successful choice within the ERP framework by simultaneously generating lower and more reasonable equal risk prices than those obtained with the CVaR_α and by resulting in effective trading policies.

Table 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	**−0.04**	0.03	0.11	0.13	0.15
$\text{CVaR}_{0.90}$	**2.64**	5.3%	22.6%	5.4%	5.6%	7.4%	11.6%	16.9%
$\text{CVaR}_{0.95}$	3.41	**−8.4%**	1.6%	−1.6%	−5.1%	−5.8%	−3.6%	0.2%
$\text{CVaR}_{0.99}$	6.86	−31.7%	**−44.5%**	−31.4%	−39.1%	−41.8%	−43.1%	−42.2%
$\text{CVaR}_{0.999}$	19.99	−48.5%	−76.1%	−65.6%	−72.4%	−74.2%	−75.8%	**−76.4%**
$\text{VaR}_{0.90}$	**1.75**	34.7%	59.9%	12.8%	21.3%	30.2%	36.7%	45.1%
$\text{VaR}_{0.95}$	**2.08**	21.9%	54.6%	21.4%	25.9%	29.6%	37.6%	44.9%
$\text{VaR}_{0.99}$	3.67	**−9.6%**	−2.9%	5.1%	−1.8%	−4.1%	−3.9%	−0.6%
$\text{VaR}_{0.999}$	11.00	−43.3%	**−62.5%**	−47.6%	−55.4%	−57.8%	−60.3%	−60.4%
SMSE	1.83	−7.0%	6.8%	**−33.5%**	−30.5%	−22.2%	−15.4%	−5.7%
MSE	2.93	−1.8%	12.2%	**−26.4%**	−24.2%	−15.6%	−9.7%	−0.2%

Notes: Hedging statistics are computed based on 100,000 independent paths generated from the regime switching model under $\mathbb{P}$ (see Section 5.2.1 for model definition and Appendix C for model parameters). The training of neural networks is performed as described in Section 4.2 with hyperparameters presented in Section 5.1. All hedging statistics except the mean hedging error are expressed relative to values obtained under the $\text{CVaR}_{0.90}$ penalty (% increase). **Bold** values are the lowest across all penalties.

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 those 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, contrary 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 with 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 covering the period 31 December 1986 to 1 April 2010 (5863 log-returns). Parameter estimates are presented in Appendix C.

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 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.

5.3.2. 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

$$\begin{aligned} y_n &= \mu + \sigma_n \epsilon_n, \\ \sigma^2_{n+1} &= \omega + \upsilon\sigma^2_n(|\epsilon_n| - \gamma\epsilon_n)^2 + \beta\sigma^2_n, \end{aligned}$$

where $\{\sigma^2_n\}_{n=1}^{N+1}$ are the daily variances of log-returns and $\{\mu, \omega, \upsilon, \gamma, \beta\}$ are the model parameters, with $\{\omega, \upsilon, \beta\}$ being positive real values and $\{\mu, \gamma\}$ real values. Note that given σ^2_1, the sequence of variances $\sigma^2_2, \ldots, \sigma^2_{N+1}$ can be computed recursively with the observed path of log-returns. In this paper, the initial value σ^2_1 is set as the stationary variance of the process: $\sigma^2_1 \equiv \mathbb{E}[\sigma^2_n] = \frac{\omega}{1-\upsilon(1+\gamma^2)-\beta}$. Furthermore, it can be shown that $\{S^{(0,b)}_n, \sigma_{n+1}\}_{n=0}^N$ 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 $n = 0, \ldots, N-1$.

5.3.3. Merton Jump Diffusion Model

Contrary 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^2_J under the physical measure[12]

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

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 returns are iid, the dynamics does not necessitate the addition of other state variables to the feature vectors, i.e., $\mathcal{I}_n = 0$ for all n.

5.3.4. Numerical Results of the Sensitivity Analysis to Underlying Asset Dynamics

Table 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 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%	10%	22%	31%	43%
MJD	0.46	23%	34%	129%	15%	41%	71%	102%	125%
GJR-GARCH	0.57	52%	71%	139%	29%	96%	156%	219%	265%

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 5.3 for models definitions under $\mathbb{P}$ and Appendix C for model parameters). Risk-neutral prices $C_0^{\mathbb{Q}}$ are computed under $\mathbb{Q}$-dynamics described in Appendix D. The training of feedforward neural networks is performed as described in Section 4.2 with hyperparameters presented in Section 5.1. $C_0^\star$ are expressed relative to $C_0^{\mathbb{Q}}$ (% increase).

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, by using the semi-$\mathbb{L}^2$ risk measure, OTM equal risk prices $C_0^\star$ exhibit relative increases over risk-neutral prices $C_0^{\mathbb{Q}}$ of, respectively, 3%, 15% and 29% under the BSM, MJD and GARCH models as compared to 5%, 23% and 52% under the $\text{CVaR}_{0.90}$ objective function. Furthermore, values presented in Table 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 in $C_0^\star$ as compared to $C_0^{\mathbb{Q}}$ ranges between 5% and 17% under the CVaR_α and between 3% and 43% 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% and 139%, while under the semi-$\mathbb{L}^p$, it ranges between 29% and 265%. Similar observations can be made under the MJD dynamics. This can be explained by the fact that contrary 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 results in equal risk price values 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 asset returns.

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$_\alpha$ 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$_\alpha$ 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 dynamic trading strategy based solely on option hedges can be impractical in some cases in the face of highly illiquid options. In such a context, the hedger could potentially be restricted to a trading strategy relying exclusively on the underlying stock, which as shown in previous sections can inflate equal risk prices under the CVaR$_\alpha$ 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 price reduction when trading exclusively the underlying stock to that obtained when trading options with the CVaR$_\alpha$ objective function. The setup to perform this experiment is the same as the one considered by Carbonneau and Godin (2021a), and numerical values for equal risk prices generated with trades involving exclusively options under the CVaR$_\alpha$ are taken directly from the latter work. This setup is now recalled.

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. The annualized continuous risk-free rate is $r = 0.03$. Also, as noted by 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 lead to hedging strategies exhibiting poor tail risk mitigation. Thus, the convex risk measure considered the benchmark in the present study is the 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-month 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 performed 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 convenience that the ATM 1-month and 3-month implied volatilities are the same.[13] 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 calls and puts of 1-month and 3-month 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 considered, where $\{\epsilon_n\}_{n=1}^{252}$ are the daily random innovations associated with stock returns. The log-AR(1) model has the representation

$$\log IV_{n+1} = \log IV_n + \kappa(\vartheta - \log IV_n) + \sigma_{IV} Z_{n+1}, \quad n = 0, \ldots, 251, \tag{10}$$

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

$$C(IV, \Delta T, S, K) \equiv S\mathcal{N}(d_1) - e^{-r\Delta T} K\mathcal{N}(d_2), \tag{11}$$
$$P(IV, \Delta T, S, K) \equiv e^{-r\Delta T} K\mathcal{N}(-d_2) - S\mathcal{N}(-d_1), \tag{12}$$

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 vector of the neural networks. For instance, with 1-month calls and puts hedges, the n-th 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 $n = 0, 1, \ldots, 11$, where 21 represents the number of days in a given month.[14]

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 by Carbonneau and Godin (2021a) are used for comparability purposes. The MJD as well as the log-AR(1) model parameters values are presented in Tables 4 and 5. These parameters were chosen in an ad hoc fashion so as to produce reasonable values for the dynamics of the financial market.

Table 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.

Table 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

Numerical Results with Option Hedges

Table 6 presents equal risk prices $C_0^\star$ under the $\text{CVaR}_{0.95}$ measure with daily or monthly stock trades as well as with 1-month or 3-month ATM calls and puts trades. Note that the latter values are from Table 3 of Carbonneau and Godin (2021a).[15] Furthermore, $C_0^\star$ values under the semi-$\mathbb{L}^2$ objective function with daily and monthly stock hedges are also presented.

Numerical results indicate that the use of the semi-$\mathbb{L}^2$ objective function is successful at significantly reducing equal risk prices when relying on trades involving exclusively the underlying stock. Indeed, the relative reductions 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 are, respectively, 15%, 11% and 11% with daily stock and 14%, 10% and 8% with monthly stock rebalancing.[16] Furthermore, equal risk prices 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-month 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 levels similar to those obtained with option hedges under the CVaR_α measures. This avenue is thus successful in alleviating the price inflation phenomenon when using 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 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.

Table 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 5.3.3 for model definition and Table 4 for model parameters). Options used as hedging instruments are priced with implied volatility modeled with log-AR(1) dynamics (see Section 5.4 for model description and Table 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 4.2.2.

6. Conclusions

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 complexifies the numerical evaluation of equal risk prices, the use of semi-$\mathbb{L}^p$ risk measures as the objective functions measuring residual hedging risk is shown to have several preferable properties over the use of CVaR_α, the latter being explored, for instance, by Carbonneau and Godin (2021a, 2021b) in the context of ERP. The optimal hedging problems underlying the ERP framework are solved with deep reinforcement learning procedures by representing trading policies with neural networks as proposed in the work of Buehler et al. (2019). A modification to the training algorithm for neural networks is presented in this current paper to tackle the additional complexity of using semi-$\mathbb{L}^p$ risk measures within 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 span a larger interval of values than the one obtained with the CVaR_α, thereby allowing to alleviate 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.

A potential avenue for further research would consist in considering objective functions reflecting a risk–reward trade-off rather than simply considering hedging risk minimization. This could, for instance, be achieved by using rank-dependent expected utility functions as suggested by Wu and Jaimungal (2023).

Author Contributions: Conceptualization, A.C. and F.G. Programming, A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Natural Sciences and Engineering Research Council grant number RGPIN-2017-06837, and Fonds de recherche du Québec—Nature et technologies grant number 205683.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Pseudo-Code

This section presents the pseudo-codes for the training of neural networks and the bisection method. Algorithm A1 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}. 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. The update rule for portfolio values in step (6) of Algorithm A1 can be obtained directly from the self-financing representation of V_n^δ as shown below:

$$\begin{aligned}
V_n^\delta &= B_n(V_0^\delta + G_n^\delta) \\
&= B_n\Big(V_0^\delta + G_{n-1}^\delta + \delta_n^{(0:D)} \bullet (B_n^{-1} S_{n-1}^{(e)} - B_{n-1}^{-1} S_{n-1}^{(b)})\Big) \\
&= \frac{B_n}{B_{n-1}} V_{n-1}^\delta + \delta_n^{(0:D)} \bullet (S_{n-1}^{(e)} - \frac{B_n}{B_{n-1}} S_{n-1}^{(b)}) \\
&= e^{r\Delta} V_{n-1}^\delta + \delta_n^{(0:D)} \bullet (S_{n-1}^{(e)} - e^{r\Delta} S_{n-1}^{(b)}).
\end{aligned} \tag{A1}$$

Algorithm A1 Pseudo-code training neural networks $F_\theta^{(\mathcal{S})}$ with underlying stock hedges.

Input: θ_j, V_0^δ
Output: θ_{j+1}
1: **for** $i = 1, \ldots, N_{\text{batch}}$ **do** ▷ Loop over each path of minibatch
2: $X_{0,i} = [T, \log(S_{0,i}^{(0,b)}/K), V_{0,i}^\delta/\tilde{V}, \mathcal{I}_{0,i}]$ ▷ Time-0 feature vector of $F_\theta^{(\mathcal{S})}$
3: **for** $n = 0, \ldots, N-1$ **do**
4: $\delta_{n+1,i}^{(0)} \leftarrow$ time-t_n output of FFNN $F_\theta^{(\mathcal{S})}$ with $\theta = \theta_j$
5: $S_{n+1,i}^{(0,b)} = S_{n,i}^{(0,b)} e^{y_{n+1,i}}$ ▷ Sample next stock price
6: $V_{n+1,i}^\delta = e^{r\Delta} V_{n,i}^\delta + \delta_{n+1,i}^{(0)} (S_{n+1,i}^{(0,b)} - e^{r\Delta} S_{n,i}^{(0,b)})$ ▷ See (A1) for details
7: $\mathcal{I}_{n+1,i} \leftarrow$ update additional state variables
8: $X_{n+1,i} = [T - t_n, \log(S_{n+1,i}^{(0,b)}/K), V_{n+1,i}^\delta/\tilde{V}, \mathcal{I}_{n+1,i}]$ ▷ Time t_{n+1} feature vector of $F_\theta^{(\mathcal{S})}$
9: **end for**
10: $\Phi(S_{N,i}^{(0,b)}) = \max(K - S_{N,i}^{(0,b)}, 0)$
11: $\pi_{i,j} = \Phi(S_{N,i}^{(0,b)}) - V_{N,i}^\delta$
12: **end for**
13: $\widehat{J} = \left(\frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \pi_{i,j}^p \mathbb{1}_{\{\pi_{i,j}>0\}}\right)^{1/p}$
14: $\eta_j \leftarrow$ Adam algorithm
15: $\theta_{j+1} = \theta_j - \eta_j \nabla_\theta \widehat{J}$ ▷ $\nabla_\theta \widehat{J}$ 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., $S_{0,i}^{(0,b)} = S_0^{(0,b)}$, $V_{0,i}^\delta = V_0^\delta$ and $\mathcal{I}_{0,i} = \mathcal{I}_0$.

Algorithm A2 presents the pseudo-code for the bisection algorithm, taking as inputs the two trained neural networks $F_\theta^{(\mathcal{L})}$ and $F_\theta^{(\mathcal{S})}$ as well as the initial search range $[V_A, V_B]$ so as to output the equal risk price.

Algorithm A2 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, Δ(V) = ∞
while |Δ(V)| > ζ and nbs_iter < max_iter do
    V = 0.5(V_A + V_B)
    Compute ε̃^(L)(−V) and ε̃^(S)(V) on the test set with F_θ^(L) and F_θ^(S)
    Δ(V) = ε̃^(S)(V) − ε̃^(L)(−V)
    if Δ(V) > 0 then
        V_A ← V
    else
        V_B ← V
    end if
    nbs_iter ← nbs_iter + 1
end while
C_0^⋆ = 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, ζ is set to 0.01 and max_iter to 100.

Appendix B. Validation of Modified Training Algorithm

The goal of this section is to demonstrate that the proposed modification to the training algorithm described in Section 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 accurately 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 5.1 with the hedging of an ATM put option of maturity $T = 60/260$ with daily stock hedges under the regime switching model. The steps are the following for all semi-$\mathbb{L}^p$ objective functions, $p \in \{2, 4, 6, 8, 10\}$:

(1) Train F_θ with the procedure described Section 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 100 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 4.2.1. A total of three iterations of this step is performed (i.e., three different values of $V^\star$ are considered).
(3) For the three sampled values of $V^\star$, compute the semi-$\mathbb{L}^p$ statistics on the test set with F_θ and F_θ^{fixed}.

Table A1 presents the semi-$\mathbb{L}^p$ statistics for the three values of $V_0 = V^\star$ with p = 2, 4, 6, 8, 10. These results clearly demonstrate that the modified training algorithm does not materially impact the accuracy of the neural network, as the difference in semi-$\mathbb{L}^p$ statistics between the FFNNs F_θ and F_θ^{fixed} is most often marginal.

Table A1. Semi-$\mathbb{L}^p$ statistics of the modified training algorithm for ATM ($K = 100$) put options of maturity $T = 60/260$ under the regime-switching model.

	$\mathbb{L}^2$		$\mathbb{L}^4$		$\mathbb{L}^6$		$\mathbb{L}^8$		$\mathbb{L}^{10}$	
V_0	F_θ	F_θ^{fixed}	F_θ	F_θ^{fixed}	F_θ	F_θ^{fixed}	F_θ	F_θ^{fixed}	F_θ	F_θ^{fixed}
4.343	0.6236	0.6209	1.1523	1.1483	1.5291	1.5466	1.8981	1.8628	2.2747	2.2069
2.503	1.5457	1.5355	2.2485	2.2466	2.6865	2.6843	3.0126	3.0277	3.3121	3.3311
4.005	0.7596	0.7578	1.3288	1.3317	1.716	1.7292	2.0703	2.0542	2.4308	2.3866

Notes: Semi-$\mathbb{L}^p$ statistics results are computed based on 100,000 independent paths generated with the regime-switching model under $\mathbb{P}$ (see Section 5.2.1 for model definition and Appendix C for model parameters). F_θ is the neural network trained with the modified algorithm described in Section 4.2.2. F_θ^{fixed} is the neural network trained with fixed initial capital investment of V_0 as described in Section 4.2.1.

Appendix C. Maximum Likelihood Estimates Results

This section presents the maximum likelihood model parameter estimates for the different risky asset dynamics considered in the numerical experiments of Sections 5.2 and 5.3. All parameters are estimated with the same time series of daily log-returns on the S&P 500 index for the period 31 December 1986 to 1 April 2010 (5863 log-returns). Except for the regime-switching model, all parameters are identical to these from Carbonneau and Godin (2021b), see their Tables A1, A2 and A4. The estimated parameters are presented in Tables A2–A5.

Table A2. 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 A3. Maximum likelihood parameter estimates of the GJR-GARCH model.

μ	ω	υ	γ	β
2.871×10^{-4}	1.795×10^{-6}	0.0540	0.6028	0.9105

Table A4. 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). ν represent probabilities associated with the stationary distribution of the Markov chain. μ and σ are on an annual basis.

Table A5. 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.

Appendix D. 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 exists 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 $\mathbb{Q}$. The $\mathbb{Q}$-dynamics for the four models are described in Sections 5.2.1 and 5.3 as well as the corresponding methods to compute the risk-neutral price $C_0^{\mathbb{Q}}$ of European puts.

Appendix D.1. Regime-Switching

The change of measure used in this study is the popular choice of shifting the drift to obtain risk neutrality and model invariance as considered, for instance, by Hardy (2001). Under this change of measure $\mathbb{Q}$, the drift $\mu_i \Delta$ in each regime 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 has the representation

$$\begin{aligned} C_0^{\mathbb{Q}} &\equiv e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi(S_N^{(0,b)})|\mathcal{F}_0] \\ &= e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}[\Phi(S_N^{(0,b)})|\mathcal{G}_0]|\mathcal{F}_0\right] \\ &= e^{-rT}\sum_{i=1}^{H}\xi_{0,i}^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}[\Phi(S_N^{(0,b)})|h_0 = i, S_0^{(0,b)}], \end{aligned} \tag{A2}$$

where $\xi_{0,i}^{\mathbb{Q}}$ is assumed to be equal to $\xi_{0,i}^{\mathbb{P}}$ for all regimes i, i.e., to the stationary distribution of the Markov chain under $\mathbb{P}$. The computation of the conditional expectations in (A2) can be performed, for instance, with Monte Carlo simulations or with the closed-form solution of Hardy (2001) when $H = 2$.

Appendix D.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.

Appendix D.3. GARCH

The change of measure from the $\mathbb{P}$ 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 $\mathbb{F}$. The one-period expected conditional gross return under $\mathbb{Q}$ must be equal to the one-period risk-free rate accrual factor 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

$$\psi_n \equiv \frac{\mu - r\Delta + \sigma_n^2/2}{\sigma_n}, \quad n = 1, \ldots, N. \tag{A3}$$

With (A3), the GARCH risky asset dynamics under $\mathbb{Q}$ is

$$y_n = r\Delta - \sigma_n^2/2 + \sigma_n \epsilon_n^{\mathbb{Q}},$$
$$\sigma_{n+1}^2 = \omega + \upsilon\sigma_n^2(|\epsilon_n^{\mathbb{Q}} - \psi_n| - \gamma(\epsilon_n^{\mathbb{Q}} - \psi_n))^2 + \beta\sigma_n^2.$$

The computation of the risk-neutral price $C_0^{\mathbb{Q}}$ can be performed with Monte Carlo simulations.

Appendix D.4. Merton Jump-Diffusion

For this model, the change of measure used is the one originally proposed by Merton (1976), which assumes no risk premia for the 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

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

where $\{N_n\}_{n=0}^N$ and $\{\zeta_j\}_{j\geq 1}$ have the same distribution than under the physical measure. The risk-neutral price of put options $C_0^{\mathbb{Q}}$ can be computed with the well-known closed-form solution.

Notes

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.
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.
4. This means δ_0 is $\mathcal{F}_0$-measurable and δ_n is $\mathcal{F}_{n-1}$-measurable for $n = 1, \ldots, N$.
5. Details characterizing well-behavedness in the context of the present study are omitted to avoid lengthy discussions straying us away from the main research objectives of this work.
6. A risk measure is a mapping taking a random variable representing a random loss as input, and returning a real number representing its perceived risk as an output.
7. Recall that since the trading strategy is self-financing, $\delta_{n+1}^{(B)}$ is characterized by $\delta_{n+1}^{(0:D)}$ and V_n.
8. 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 q of parameters to fit 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 difficulty.
9. 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.
10. 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 1000, one epoch consists of 400 updates of the set of trainable parameters θ.
11. Recall that optimal policies under the CVaR$_\alpha$ 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 that include the risk-neutral price due to the proposed modified training algorithm in this paper.
12. The convention that $\sum_{j=N_{n-1}+1}^{N_n} \zeta_j = 0$ if $N_{n-1} = N_n$ is adopted.

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

14 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.

15 The type of neural networks considered by 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.

16 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 $1 - \frac{C_0(\mathbb{L}^2)}{C_0^\star(\text{CVaR}_{0.95})}$.

References

Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, and et al. 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv* arXiv:1603.04467.

Alfeus, Mesias, Xin-Jiang He, and Song-Ping Zhu. 2022. An empirical analysis of option pricing with short sell bans. *International Journal of Theoretical and Applied Finance* 25: 2250012. [CrossRef]

Ang, Andrew, and Allan Timmermann. 2012. Regime changes and financial markets. *Annual Review of Financial Economics* 4: 313–37. [CrossRef]

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk. *Mathematical Finance* 9: 203–28. [CrossRef]

Augustyniak, Maciej, Frédéric Godin, and Clarence Simard. 2017. Assessing the effectiveness of local and global quadratic hedging under GARCH models. *Quantitative Finance* 17: 1305–18. [CrossRef]

Bao, Wenhang, and Xiao-Yang Liu. 2019. Multi-agent deep reinforcement learning for liquidation strategy analysis. *arXiv* arXiv:1906.11046.

BCBS. 2016. *Standards. Minimum Capital Requirements for Market Risk. January 2016*. *Basel Committee on Banking Supervision*. Basel: Bank for International Settlements.

Bertsimas, Dimitris, Leonid Kogan, and Andrew W. Lo. 2001. Hedging derivative securities and incomplete markets: An ϵ-arbitrage approach. *Operations Research* 49: 372–97. [CrossRef]

Betancourt, Carlos, and Wen-Hui Chen. 2021. Deep reinforcement learning for portfolio management of markets with a dynamic number of assets. *Expert Systems with Applications* 164: 114002. [CrossRef]

Black, Fischer, and Myron Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81: 637–54. [CrossRef]

Breton, Michèle, and Frédéric Godin. 2017. Global hedging through post-decision state variables. *Journal of Risk and Financial Management* 10: 16. [CrossRef]

Buehler, Hans, Lukas Gonon, Josef Teichmann, and Ben Wood. 2019. Deep hedging. *Quantitative Finance* 19: 1271–91. [CrossRef]

Cao, Hongkai, Zhenyu Cui, and Yanchu Liu. 2020. Discrete-time variance-optimal deep hedging in affine GARCH models. *SSRN*, 3659275. [CrossRef]

Cao, Jay, Jacky Chen, John Hull, and Zissis Poulos. 2021. Deep hedging of derivatives using reinforcement learning. *The Journal of Financial Data Science* 3: 10–27. [CrossRef]

Carbonneau, Alexandre. 2021. Deep hedging of long-term financial derivatives. *Insurance: Mathematics and Economics* 99: 327–40. [CrossRef]

Carbonneau, Alexandre, and Frédéric Godin. 2021a. Deep equal risk pricing of financial derivatives with multiple hedging instruments. *arXiv* arXiv:2102.12694.

Carbonneau, Alexandre, and Frédéric Godin. 2021b. Equal risk pricing of derivatives with deep hedging. *Quantitative Finance* 21: 593–608. [CrossRef]

Chong, Wing Fung, Haoen Cui, and Yuxuan Li. 2021. Pseudo-model-free hedging for variable annuities via deep reinforcement learning. *arXiv* arXiv:2107.03340.

Christoffersen, Peter, Redouane Elkamhi, Bruno Feunou, and Kris Jacobs. 2010. Option valuation with conditional heteroskedasticity and nonnormality. *The Review of Financial Studies* 23: 2139–83. [CrossRef]

Coleman, Thomas F., Yohan Kim, Yuying Li, and M. Patron. 2007. Robustly hedging variable annuities with guarantees under jump and volatility risks. *Journal of Risk and Insurance* 74: 347–76. [CrossRef]

Cont, Rama, and José Da Fonseca. 2002. Dynamics of implied volatility surfaces. *Quantitative Finance* 2: 45–60. [CrossRef]

Delbaen, Freddy, and Walter Schachermayer. 1994. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* 300: 463–520. [CrossRef]

Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39: 1–22.

Duan, Jin-Chuan. 1995. The GARCH option pricing model. *Mathematical Finance* 5: 13–32. [CrossRef]

Elliott, Robert J., and Dilip B. Madan. 1998. A discrete time equivalent martingale measure. *Mathematical Finance* 8: 127–52. [CrossRef]

Föllmer, Hans, and Martin Schweizer. 1988. Hedging by sequential regression: An introduction to the mathematics of option trading. *ASTIN Bulletin: The Journal of the IAA* 18: 147–60. [CrossRef]

Föllmer, Hans, and Peter Leukert. 1999. Quantile hedging. *Finance and Stochastics* 3: 251–73. [CrossRef]

Föllmer, Hans, and Peter Leukert. 2000. Efficient hedging: Cost versus shortfall risk. *Finance and Stochastics* 4: 117–46. [CrossRef]

François, Pascal, Geneviève Gauthier, and Frédéric Godin. 2014. Optimal hedging when the underlying asset follows a regime-switching Markov process. *European Journal of Operational Research* 237: 312–22. [CrossRef]

Frittelli, Marco. 2000. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance* 10: 39–52. [CrossRef]

García-Galicia, Mauricio, Alin A. Carsteanu, and Julio B. Clempner. 2019. Continuous-time reinforcement learning approach for portfolio management with time penalization. *Expert Systems with Applications* 129: 27–36. [CrossRef]

Gerber, Hans U., and Elias S. W. Shiu. 1994. Option pricing by Esscher transforms. *Transactions of the Society of Actuaries* 46: 99–191.

Glorot, Xavier, and Yoshua Bengio. 2010. Understanding the difficulty of training deep feedforward neural networks. Paper presented at Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, May 13–15. JMLR Workshop and Conference Proceedings, pp. 249–56.

Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle. 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance* 48: 1779–801. [CrossRef]

Godin, Frédéric. 2016. Minimizing CVaR in global dynamic hedging with transaction costs. *Quantitative Finance* 16: 461–75. [CrossRef]

Godin, Frédéric, and Denis-Alexandre Trottier. 2019. Option pricing under regime-switching models: Novel approaches removing path-dependence. *Insurance: Mathematics and Economics* 87: 130–42. [CrossRef]

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. Cambridge: MIT Press.

Guo, Ivan, and Song-Ping Zhu. 2017. Equal risk pricing under convex trading constraints. *Journal of Economic Dynamics and Control* 76: 136–51. [CrossRef]

Halperin, Igor. 2020. Qlbs: Q-learner in the Black-Scholes (-Merton) worlds. *The Journal of Derivatives* 28: 99–122. [CrossRef]

Hamilton, James D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society* 57: 357–84. [CrossRef]

Hardy, Mary R. 2001. A regime-switching model of long-term stock returns. *North American Actuarial Journal* 5: 41–53. [CrossRef]

He, Xin-Jiang, and Song-Ping Zhu. 2020. A revised option pricing formula with the underlying being banned from short selling. *Quantitative Finance* 20: 935–48. [CrossRef]

Hendricks, Dieter, and Diane Wilcox. 2014. A reinforcement learning extension to the Almgren-Chriss framework for optimal trade execution. Paper presented at 2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr), London, UK, March 27–28. pp. 457–64.

Hodges, Stewart, and Anthony Neuberger. 1989. Optimal replication of contingent claims under transactions costs. *Review of Futures Markets* 8: 222–39.

Hornik, Kurt. 1991. Approximation capabilities of multilayer feedforward networks. *Neural Networks* 4: 251–57. [CrossRef]

Horvath, Blanka, Josef Teichmann, and Zan Zuric. 2021. Deep hedging under rough volatility. *Risks* 9: 138. [CrossRef]

Jiang, Zhengyao, Dixing Xu, and Jinjun Liang. 2017. A deep reinforcement learning framework for the financial portfolio management problem. *arXiv* arXiv:1706.10059.

Kingma, Diederik P., and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv* arXiv:1412.6980.

Lamberton, Damien, and Bernard Lapeyre. 2007. *Introduction to Stochastic Calculus Applied to Finance*. Boca Raton: CRC Press.

Li, Yuxi, Csaba Szepesvari, and Dale Schuurmans. 2009. Learning exercise policies for American options. In *Artificial Intelligence and Statistics*. PMLR, Clearwater Beach, Florida, USA, pp. 352–59.

Lu, David W. 2017. Agent inspired trading using recurrent reinforcement learning and LSTM neural networks. *arXiv* arXiv:1707.07338.

Lütkebohmert, Eva, Thorsten Schmidt, and Julian Sester. 2022. Robust deep hedging. *Quantitative Finance* 22: 1465–80. [CrossRef]

Ma, Guiyuan, Song-Ping Zhu, and Ivan Guo. 2022. Valuation of general contingent claims with short selling bans: An equal-risk pricing approach. *International Journal of Theoretical and Applied Finance* 25: 2250022. [CrossRef]

Marzban, Saeed, Erick Delage, and Jonathan Yumeng Li. 2021. Deep reinforcement learning for equal risk pricing and hedging under dynamic expectile risk measures. *arXiv*, arXiv:2109.04001.

Marzban, Saeed, Erick Delage, and Jonathan Yu-Meng Li. 2022. Equal risk pricing and hedging of financial derivatives with convex risk measures. *Quantitative Finance* 22: 47–73. [CrossRef]

Melnikov, Alexander, and Ivan Smirnov. 2012. Dynamic hedging of conditional value-at-risk. *Insurance: Mathematics and Economics* 51: 182–90. [CrossRef]

Merton, Robert C. 1973. Theory of rational option pricing. *The Bell Journal of Economics and Management Science* 4: 141–83. [CrossRef]

Merton, Robert C. 1976. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3: 125–44. [CrossRef]

Moody, John, and Lizhong Wu. 1997. Optimization of trading systems and portfolios. Paper presented at IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr), New York City, NY, USA, March 24–25. pp. 300–7.

Ning, Brian, Franco Ho Ting Lin, and Sebastian Jaimungal. 2021. Double deep Q-learning for optimal execution. *Applied Mathematical Finance* 28: 361–80. [CrossRef]

Pendharkar, Parag C., and Patrick Cusatis. 2018. Trading financial indices with reinforcement learning agents. *Expert Systems with Applications* 103: 1–13. [CrossRef]

Rockafellar, R. Tyrrell, and Stanislav Uryasev. 2002. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* 26: 1443–71.

Schweizer, Martin. 1995. Variance-optimal hedging in discrete time. *Mathematics of Operations Research* 20: 1–32. [CrossRef]

Wang, Haoran, and Xun Yu Zhou. 2020. Continuous-time mean-variance portfolio selection: A reinforcement learning framework. *Mathematical Finance* 30: 1273–308. [CrossRef]

Wu, David, and Sebastian Jaimungal. 2023. Robust risk-aware option hedging. *arXiv*, arXiv:2303.15216.

Xu, Mingxin. 2006. Risk measure pricing and hedging in incomplete markets. *Annals of Finance* 2: 51–71. [CrossRef]

Ye, Yunan, Hengzhi Pei, Boxin Wang, Pin-Yu Chen, Yada Zhu, Ju Xiao, and Bo Li. 2020. Reinforcement-learning based portfolio management with augmented asset movement prediction states. Paper presented at AAAI Conference on Artificial Intelligence, New York, NY, USA, February 7–12; Vol. 34, pp. 1112–19.