

## Article

# Revolutionizing Hedge Fund Risk Management: The Power of Deep Learning and LSTM in Hedging Illiquid Assets

Yige Wang<sup>1</sup>, Leyao Tong<sup>2</sup> and Yueshu Zhao<sup>3,\*</sup> <sup>1</sup> Numerix LLC, New York, NY 10017, USA; ywang@numerix.com<sup>2</sup> Financial Services Forum, Washington, DC 20005, USA; etong@fsforum.org<sup>3</sup> International Monetary Fund, Washington, DC 20431, USA

\* Correspondence: yzhao4@imf.org

**Abstract:** In the dynamic sphere of financial markets, hedge funds have emerged as a critical force, navigating through volatility with advanced risk management techniques yet grappling with the challenges posed by illiquid assets. This study aims to transcend traditional option pricing models, which struggle under the complexities of hedge fund investments, by exploring the applicability of machine learning in financial risk management. Leveraging Deep Neural Networks (DNNs) and Long Short-Term Memory (LSTM) cells, the research introduces a model-free, data-driven approach for discrete-time hedging problems. Through a comparative analysis of simulated data and the implementation of LSTM architectures, the paper elucidates the potential of these machine learning techniques to enhance the precision of risk assessments and decision-making processes in hedge fund investments. The findings reveal that DNNs and LSTMs offer significant advancements over conventional models, effectively capturing long-term dependencies and complex patterns within financial time series data. Consequently, the study underscores the transformative impact of machine learning on the methodologies employed in financial risk management, proposing a novel paradigm that promises to mitigate the intricacies of hedging illiquid assets. This research not only contributes to the academic discourse but also paves the way for the development of more adaptive and resilient investment strategies in the face of market uncertainties.

**Keywords:** financial risk management; artificial intelligence in finance; machine learning algorithms; illiquid asset hedging; LSTM; DNN; risk mitigation strategies; neural network forecasting



**Citation:** Wang, Yige, Leyao Tong, and Yueshu Zhao. 2024.

Revolutionizing Hedge Fund Risk Management: The Power of Deep Learning and LSTM in Hedging Illiquid Assets. *Journal of Risk and Financial Management* 17: 224. <https://doi.org/10.3390/jrfm17060224>

Academic Editor: Svetlozar (Zari) Rachev

Received: 10 April 2024

Revised: 22 May 2024

Accepted: 24 May 2024

Published: 26 May 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### 1.1. Literature Review

Hedge funds are investment funds that pool capital from accredited or institutional investors and invest in a variety of assets. They usually deliver higher Sharpe ratios than Buy-and-Hold strategies on traditional asset classes, benefiting from their complicated portfolio-construction and risk management techniques. The investors were especially attracted by the promising behaviors of hedge fund industry during the bear market between 2000 and 2003. However, investments in hedge funds are illiquid, since they often require investors to keep their money in the fund for at least one year, a time known as ‘the lock-up period’. Withdrawals may also only happen at certain frequencies, e.g., quarterly or bi-annually. In such cases, the Black-Scholes model on option pricing may suffer from restrictive assumptions when being used in the hedge fund index option pricing. Investors may also have to consider tail risk and hedge slippage in discrete-time hedging problems. In this report, we would like to develop a model-free approach to solve this illiquid option hedging problem, using multiple criteria to measure hedge errors.

In the realm of options trading, delta hedging plays a pivotal role in portfolio management. Delta, the most critical hedge parameter, can be easily adjusted through trades in the underlying asset. Since the advent of exchange-traded options markets in 1973, option

traders have frequently adjusted delta to near-zero levels by trading the underlying asset, highlighting its significance in risk mitigation strategies. In the literature, several stochastic volatility models have been proposed, including those by [Hull and White \(1987\)](#), [Hull \(1988\)](#), [Heston \(1993\)](#), and [Hagan et al. \(2002\)](#). A recent research by [Hull and White \(2017\)](#) has noted that the conventionally calculated delta doesn't minimize portfolio variance due to the correlation between asset price and volatility movements. The minimum variance delta considers both price fluctuations and volatility changes. And they empirically derived a model for this delta, demonstrating its superiority over stochastic volatility models using S&P 500 options data.

Instead of directly pricing the European style option on hedging non-traded assets, we investigate the techniques to value a payoff in an incomplete financial market. One of the traditional hedging methodologies studying optimal policies under such conditions is called 'the mean-variance hedging'. An initial research by [Duffie et al. \(1991\)](#) provided explicit optimal positions that minimize the quadratic objective, assuming that both tradable and non-tradable asset prices follow a geometric Brownian motion. Later, [Schweizer \(1995\)](#) provided a solution to one-dimensional, mean-variance hedging with a non-stochastic interest rate. An optimal hedging strategy in terms of parameters from a specific non-tradable asset payoff decomposition were thereafter derived by [Gourieroux et al. \(1998\)](#). With the help of stochastic dynamic programming, [Bertsimas et al. \(2001\)](#) solved the minimization of the mean-squared-error, and numerically computed the optimal replication strategy. Consecutive researches from [Černý and Kallsen \(2007, 2008\)](#) studied mean-variance hedging strategies in locally square-integrable semi-martingales context. Further, they also proposed solutions to the mean-variance hedging problem in Heston's model framework. A more recent study by [Rémillard and Rubenthaler \(2013\)](#) proposed the optimal solution for hedging portfolio in a discrete time context.

Another hedging methodology, CVaR (Conditional Value at Risk)-based hedging, is a risk management tool introduced by [Rockafellar and Uryasev \(2000\)](#). It measures the average loss of an asset or portfolio in the worst-case scenario within a given confidence level, typically 95% or 99%. Unlike VaR (Value at Risk), which focuses on the maximum potential loss within a specific confidence interval, CVaR considers the average loss beyond the VaR threshold, thus addressing the "tail risk" more comprehensively. Further researches including [Rockafellar and Uryasev \(2002\)](#) and [Krokhmal et al. \(2002\)](#) developed the potential and constraints of CVaR-based hedging. [Alexander et al. \(2003\)](#) discussed on derivative portfolio hedging utilizing CVaR.

As techniques in machine learning evolve rapidly within recent decades, attempts to solve financial problems with neural networks start to prosper. Promising results provided by [Hutchinson et al. \(1994\)](#) directly parameterized the pricing function of a derivative using a neural network, assuming relatively good liquidity and abundance in historical data of the underlying. [Moody and Wu \(1997\)](#) and [Jiang et al. \(2017\)](#) also apply machine learning techniques to deal with a non-linear objective functions setup of classic portfolio optimization. Solid outcomes given by [Du et al. \(2016\)](#) and [Lu \(2017\)](#) also confirm the problem-solving competence of neural network in algorithmic trading. Recent works from [Lütkebohmert et al. \(2022\)](#) and [Mikkilä and Kanninen \(2023\)](#) provide further insight on the potential of empirical deep hedging and robust deep hedging.

Deep feed forward networks, as an extension of the first and simplest type of artificial neural network devised, enjoy high reputation in its capability to satisfy universal approximation properties. Early in the 1990s [Hornik \(1991\)](#) revealed the effectiveness of deep feed forward networks in combining optimal approximation properties of all affine systems. Such efficiency to determine the optimal hedging strategy with corresponding input factors turns out to be an edge in the situation of particular hedging problems, where the availability of the price data of the derivative to be hedged is limited. More importantly, the deep hedging methodology provides the possibility to aggregate multiple hedging instruments and market frictions, which, in our case, is the transaction costs. A relevant research from [Föllmer and Schied \(2011\)](#) provided a general introduction focusing on such

incomplete markets. Modern reinforcement learning methods were applied by [Buehler et al. \(2019\)](#) to create a framework for hedging a portfolio of derivatives in the presence of market incompleteness. Several machine-learning-based algorithms were also generated by [Fecamp et al. \(2019\)](#) to solve hedging problems related to illiquidity, non-tradable risk factors, discrete hedging dates and proportional transaction costs. A flexible and accurate model based on reinforcement learning also appeared in recent research by [Kolm and Ritter \(2019\)](#) to resolve hedging problems where trading decisions are discrete and trading costs are nonlinear.

## 1.2. Problem Formulation

The main objective of this paper focuses on simultaneously determining the option prices  $V(t, S(t))$  and hedge ratios  $\Phi(t, S(t))$  throughout different time  $t$  of maturity  $T$  and corresponding underlying  $S(t)$ . Previous researches on delta hedging by [Hull and White \(1987\)](#) and [Hull \(1988\)](#) are referred to for the following problem formulation. Special attentions are paid to the initial endowment  $V(0, S(0))$  and hedging strategy  $\Phi(0, S(0))$ . Notice that the hedge ratio is considered as an independent entity to determine. It does *not* simply equal to some infinitesimal change in the option price relative to an infinitesimal change in the underlying asset price.

Consider the profit and loss (P&L) of an option seller over a time period  $(t, t + 1]$ . The wealth change of his/her delta hedged portfolio consists of two parts: the option part and the hedge part:

$$\Delta W(t, t + 1) = \Delta W_{\text{option}}(t, t + 1) + \Delta W_{\text{hedge}}(t, t + 1). \quad (1)$$

The option part can be written as

$$\begin{aligned} \Delta W_{\text{option}}(t, t + 1) &= W_{\text{option}}(t + 1) - W_{\text{option}}(t) \\ &= V(t, S(t))df(t, t + 1) - V(t, S(t)) + P(t)df(t, t + 1) \\ &= G(t) - V(t, S(t)) \end{aligned} \quad (2)$$

where

$$G(t) = V(t, S(t))df(t, t + 1) + P(t)df(t, t + 1). \quad (3)$$

Here  $V(t, S(t))$  stands for the option price at time  $t$ .  $P(t)$  denotes the payoff of the contract at time  $t$ . In the case of European options,  $P(t) = 0$  for all  $t \neq T$ .  $df(t, t + 1)$  represents the risk-free discount fact from time  $t$  to  $t + 1$ .

On the other hand, the option seller will also attempt to hedge the sold option position with hedge ratio of  $\Phi(t, S(t))$  over the time period  $(t, t + 1]$ . Here, the hedge part of the wealth change includes changes in underlying asset price, financing costs, dividends received or paid, and transaction costs. Therefore, the wealth change of the total hedged position without transaction costs is given by

$$\Delta W_{\text{hedge}}(t, t + 1) = \Phi(t, S(t))H(t) + Z(t, S(t)), \quad (4)$$

where

$$H(t) = \left( S(t + 1) - \frac{S(t)}{Df(t, t + 1)} \right) df(t, t + 1) + M(t + 1)df(t, t + 1). \quad (5)$$

Here  $Df(t, t + 1)$  denotes the financing cost discount factor based on the repurchase agreement (repo) rate of the underlying asset, which is essentially different from  $df(t, t + 1)$ .  $M(t)$  represents any discrete dividend paid by holding the underlying asset.  $Z(t, S(t))$  stands for the transaction costs for each hedging procedure.

To hedge the portfolio and figure out the initial endowment  $V(0, S(0))$  and hedging strategy  $\Phi(0, S(0))$ , the absolute value of total wealth change  $\Delta W$  is supposed to be minimized. Different criteria should be applied to the optimization for different purposes, and situations both with and without transaction costs should be investigated. We in-

tend to develop a model-free machine learning approach which could be applied to solve discrete-time hedging problems of illiquid assets.

The rest of this report is arranged as follows: Section 2 discusses the methodology used in this project. Hedging approach of two different loss function, the structure of LSTM cells and transaction costs will be included. Section 3 provides results and relevant discussion over numerical results and empirical results. Section 4 concludes our project and proposes possible topics for future researches.

## 2. Methodology

### 2.1. Hedging Approach

Similar to the binomial tree option pricing method, the simulation of all Monte Carlo paths for all time steps is required before further implementation. A specific stochastic process is chosen for the underlying asset for simulation. Such procedure resembles the creation of a full binomial tree first before starting the option pricing. With the simulated data in hand, we can work backwards from the maturity, solve for the option value  $V(t, S(t))$  and the hedge ratio  $\Phi(t, S(t))$  at each time step, and finally reach the initial endowment  $V(0, S(0))$  and hedging strategy  $\Phi(0, S(0))$ . Like the binomial tree method, the European option value at maturity is the payoff of that specific Monte Carlo simulation path. The method to solve for  $V(0, S(0))$  and  $\Phi(0, S(0))$  is based upon minimization of a loss function  $\mathcal{L}(\Delta W)$  of the average wealth change over all paths and all time steps. In this report, we discuss two different loss functions: mean variance and CVaR.

#### 2.1.1. Mean Variance

Mean-variance analysis is a popular method of weighing risk, expressed as variance, against expected return. Results of mean-variance analysis may help investors make decisions about which financial instruments to invest in, based on how much risk they are willing to take in exchange for different levels of reward. Mean-variance analysis allows investors to find the largest profit at a given level of risk or the least risk at a given level of return.

In our particular hedging problem, the total wealth of the portfolio is set to be the required mean (expected return) for the mean-variance analysis [Rémillard and Rubenthaler \(2013\)](#). We would like to find the initial endowment and the hedging strategy

$$V(0, S(0)) \text{ and } \Phi(0, S(0)), \quad (6)$$

such that the following loss function is minimized

$$\mathcal{L}(\Delta W) = \sigma_{\Delta W}^2 = \mathbb{E}[(\Delta W - \mathbb{E}[\Delta W])^2]. \quad (7)$$

Notice that here  $\Delta W$  denotes the total wealth change of all time steps from 0 to  $T$ ,

$$\Delta W = \Delta W(0, T) = \sum_{t=0}^{T-1} \Delta W(t, t+1) df(t, t+1). \quad (8)$$

The distribution of this total wealth change will give out an overall P&L distribution of the attempted option hedge through all time steps. This total wealth change distribution gives a complete picture of valuation and risk as compared to the methods producing one unique price with risk measures being simple calculated using sensitivities to infinitesimal changes of various input parameters (i.e., “delta”, “vega”, “rho”, etc).

#### 2.1.2. CVaR

Conditional value at risk (CVaR) is a risk measure evaluate the market risk or credit risk of a portfolio. It is also known as ‘the expected shortfall’: the “expected shortfall at  $q\%$  level” is the expected return on the portfolio in the worst  $q\%$  of cases. CVaR is an alternative to value at risk (VaR) because it is a coherent, and moreover a spectral, measure of financial

portfolio risk. It is calculated for a given quantile-level  $q$ , and is defined to be the mean loss of portfolio value given that a loss is occurring at or below the  $q$ -quantile.

CVaR values are derived from the calculation of VaR itself. Therefore, the assumptions that VaR is based on will all affect the value of CVaR, such as the shape of the distribution of returns, the cut-off level used, the periodicity of the data, and the assumptions about stochastic volatility. The value of CVaR equals to the average of the values that fall beyond the VaR:

$$\text{CVaR} = \frac{1}{1-a} \int_{-1}^{\text{VaR}} xp(x)dx, \quad (9)$$

where  $a$  represents the cut-off point (significance level) on the distribution,  $p(x)$  is the probability density, and VaR is the agreed-upon VaR level.

The illiquid asset hedging problem requires a discrete-time context. In our particular setting, the value of CVaR equals to the average of the smallest  $q\%$  of all possible  $\Delta W$  for all  $N$  simulations. The loss function we would like to minimize translates to:

$$\mathcal{L}(\Delta W) = \text{CVaR}(\Delta W) = \frac{1}{n} \sum_{k=1}^n \Delta W_k, \quad (10)$$

where  $n = q\% \times N$ . Unlike the mean-variance approach, the CVaR approach emphasizes the tail risk and tries to prevent extreme losses. Such characteristic coincides with real-life concerns and thus gains its popularity in risk management fields.

## 2.2. Machine Learning Approach

A deep neural network (DNN) is an artificial neural network (ANN) with multiple layers between the input and output layers, where each layer represents a unique mathematical manipulation. As an edge-cutting approximation technique, DNNs approximate the function with given input and output, by matching each layer with proper weight parameters. A prominent advantage of a DNN is that it approximates the target function effectively, no matter it is linear or non-linear, so we utilize such property to solve aforementioned hedging problems. [Hornik et al. \(1989\)](#) showed that the multi-layer feedforward architecture gives neural networks the competence for universal approximation.

**Theorem 1** (Universal Approximation Theorem, [Hornik et al. \(1989\)](#) Corollary 2.4). *For a given dimension  $I \in \mathbb{N}$ , let  $C^I$  be the set of all continuous Borel-measurable functions from  $\mathbb{R}^I$  to  $\mathbb{R}$ . For any monotonically increasing and bounded function  $\sigma(\cdot)$  (sigmoid activation function), there exists  $f(x) = v \cdot (\sigma(Wx + \theta))$ , where  $J \in \mathbb{N}$ ,  $v, \theta \in \mathbb{R}^J$ ,  $W \in \mathbb{R}^{J \times I}$ ,  $x \in \mathbb{R}^I$ , and  $g(x) \in C^I$ , such that*

$$|f(x) - g(x)| < \varepsilon, \quad (11)$$

for any  $\varepsilon > 0$ . The operator  $\cdot$  denotes the scalar product.

This theorem states that a feedforward neural network with one hidden layer (a three-layered feedforward neural network) has the capability to approximate any function in  $C^I$ . Corollary 1 further extends the theorem and shows that it holds for networks with multiple outputs.

**Corollary 1** ([Hornik et al. \(1989\)](#) Corollary 2.6). *Theorem 1 holds for the approximation of functions in  $C^{I,N}$  by extending the function  $f(x) = V(\sigma(Wx + \theta))$ , where  $V \in \mathbb{R}^{N \times J}$ ,  $W \in \mathbb{R}^{J \times I}$ ,  $\theta \in \mathbb{R}^J$ , and  $x \in \mathbb{R}^I$ .*

Consequently, three-layered multi-output feedforward neural networks are universal approximators for vector-valued functions. However, financial series data is time dependent. In this aspect, recurrent neural network (RNN) shows remarkable competency over regular DNNs for its capability in modeling sequence of time-dependent data. [Schäfer and Zimmermann \(2006\)](#) showed that RNN in state space model form are also



universal approximators and are able to approximate any open dynamical system with an arbitrary accuracy.

**Theorem 2** (Universal Approximation Theorem for RNN, Schäfer and Zimmermann (2006) Theorem 2). *For a measurable function  $g(\cdot) : \mathbb{R}^J \times \mathbb{R}^I \rightarrow \mathbb{R}^J$  and a continuous function  $h(\cdot) : \mathbb{R}^J \rightarrow \mathbb{R}^N$ , the external inputs  $x_t \in \mathbb{R}^I$ , the inner states  $s_t \in \mathbb{R}^J$ , and the outputs  $y_t \in \mathbb{R}^N$  ( $t = 1, \dots, T$ ), any open dynamical system of the form*

$$\begin{aligned} s_{t+1} &= g(s_t, x_t) \\ y_t &= h(s_t), \end{aligned} \quad (12)$$

*can be approximated with an arbitrary accuracy by a system of the following form*

$$\begin{aligned} s_{t+1} &= \sigma(U s_t + W x_t + \theta) \\ y_t &= C s_t, \end{aligned} \quad (13)$$

*where  $\sigma(\cdot)$  is a sigmoid activation function, the matrices  $U \in \mathbb{R}^{J \times J}$ ,  $W \in \mathbb{R}^{J \times I}$ , and  $C \in \mathbb{R}^{N \times J}$  and the bias  $\theta \in \mathbb{R}^J$ .*

Nevertheless, the deficiency of RNN emerges as the gradient vanishing effect becomes conspicuous. To avoid such effect that basic RNN suffers, Long Short-Term Memory (LSTM) cells were introduced by Hochreiter and Schmidhuber (1997) for their power to capture the long-range dependence of the data.

### 2.2.1. LSTM Cell

The architecture of a basic LSTM cell unrolled is illustrated in Figure 1. As input time series data is processed through the LSTM cell, structures named “gates” regulate the information by modifying its flow and produce two output vectors: a hidden state  $s_t$  (short term memory), and a cell state  $c_t$  (long term memory). The hidden state  $s_{t-1}$  from time  $t - 1$  is passed down to the current time step at time  $t$  and goes through a sigmoid function known as the “forget gate layer”, which determines the proportion of memory that is to be “remembered”. The “input gate layer” decides how much of the input  $x_t$  is used for the calculation of the memory state  $c_t$  at time  $t$ . The “output layer” determines the final output  $s_t$  and  $c_t$ . Meanwhile,  $c_t$  is adjusted by the previous cell state  $c_{t-1}$  and the outcome of the forget gate and the input gate.  $c_t$  together with  $s_t$  will flow to the next time step, while a copy of  $s_t$  is extracted as the output of the LSTM cell of current time step. As introduced in Hochreiter and Schmidhuber (1997), the compact forms of the equations for the forward pass of an LSTM unit with a forget gate are:

$$f_t = \sigma(U_f s_{t-1} + W_f x_t + \theta_f), \quad (14)$$

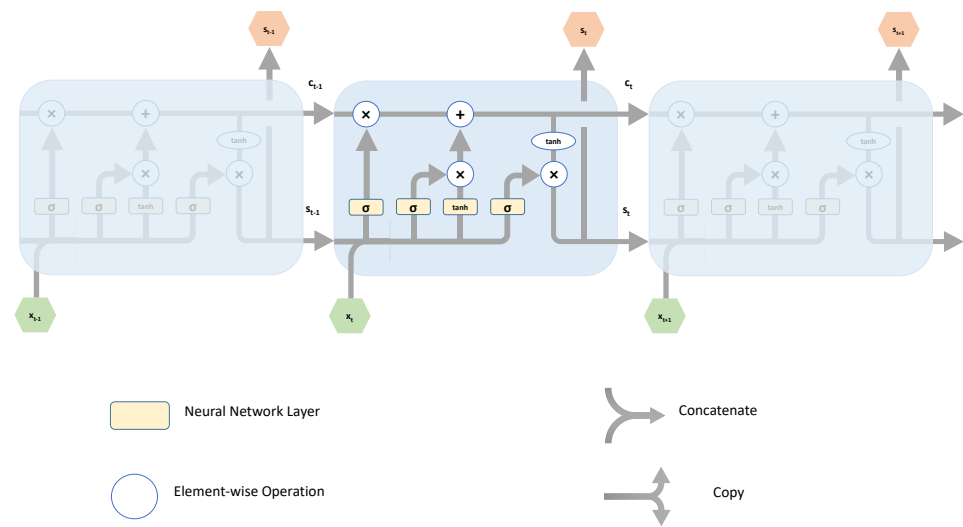
$$i_t = \sigma(U_i s_{t-1} + W_i x_t + \theta_i), \quad (15)$$

$$o_t = \sigma(U_o s_{t-1} + W_o x_t + \theta_o), \quad (16)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(U_c s_{t-1} + W_c x_t + \theta_c), \quad (17)$$

$$s_t = o_t \odot \tanh(c_t), \quad (18)$$

where the initial values are  $C_0 = 0$  and  $H_0 = 0$ , and the operator  $\odot$  denotes the Hadamard product (element-wise product).  $\sigma(\cdot)$  is the logistic sigmoid function, defined as  $\sigma(s) = 1/(1 + e^{-s})$ . The subscript  $t$  indexes the time step.  $x_t \in \mathbb{R}^I$  denotes the input vector to the LSTM unit.  $f_t \in \mathbb{R}^J$ ,  $i_t \in \mathbb{R}^J$ , and  $o_t \in \mathbb{R}^J$  represent the activation vectors of the forget gate, the input gate, and the output gate, respectively.  $s_t \in \mathbb{R}^J$  is the hidden state vector, also known as the output vector of the LSTM unit, and  $c_t \in \mathbb{R}^J$  is the cell state vector.  $W \in \mathbb{R}^{J \times I}$ ,  $U \in \mathbb{R}^{J \times J}$  and  $\theta \in \mathbb{R}^J$  stand for weight matrices and bias vector parameters to be trained, and the superscripts  $I$  and  $J$  refer to the number of input features and number of hidden units, respectively.



**Figure 1.** LSTM cell Hochreiter and Schmidhuber (1997).

### 2.2.2. Recurrent and LSTM Networks for Option Hedging

Our multi-layer LSTM network consists of several basic LSTM cells, where the output of each individual cell is used as the input of its following cell. The LSTM network is fed successively with  $S(t)$ ,  $t \in \{1 \dots N - 1\}$ . For each pair of  $(t, S(t))$ , the network provides the hedge ratio  $\Phi_t(S(t), \Theta)$  where  $\Theta$  includes the bias and weights to be estimated. In the case of  $t = 0$ , the initial option price  $V_0$  and hedge ratio  $\Phi_0$  are to be optimized without the LSTM network. For convenience of computation, we set discount factors to be zero, and exclude the payment of dividends. Since  $V_0$ ,  $\Phi_0$  and  $\Theta$  are trainable variables, the optimization problem is equivalent to minimize the loss function  $\mathcal{L}(\Delta W)$ , where

$$\Delta W = V_0 + \Phi_0 \Delta S(0, 1) + \sum_{j=1}^{T-1} \Phi_j(S(j), \Theta) \Delta S(j, j+1) - P. \quad (19)$$

We use TensorFlow to construct the LSTM neural network. The architecture of LSTM recurrent neural network is illustrated in Figure 2. Adaptive moment estimation (Adam) optimization algorithm is used to update network weights iterative based in training data. The parameters used in the optimization process are listed as follows: The number of simulations used for each iteration of the AdamOptimizer, namely the batch size, is 1000. The initial learning rate for the AdamOptimizer is 0.001 as default. The number of nodes of the LSTM cell neural network is [24,12,12,1]. The input data is normalized batch-wise before fed into the LSTM neural network. 10,000 simulations are generated for the mean and variance used for normalization.

### 2.3. Transaction Costs

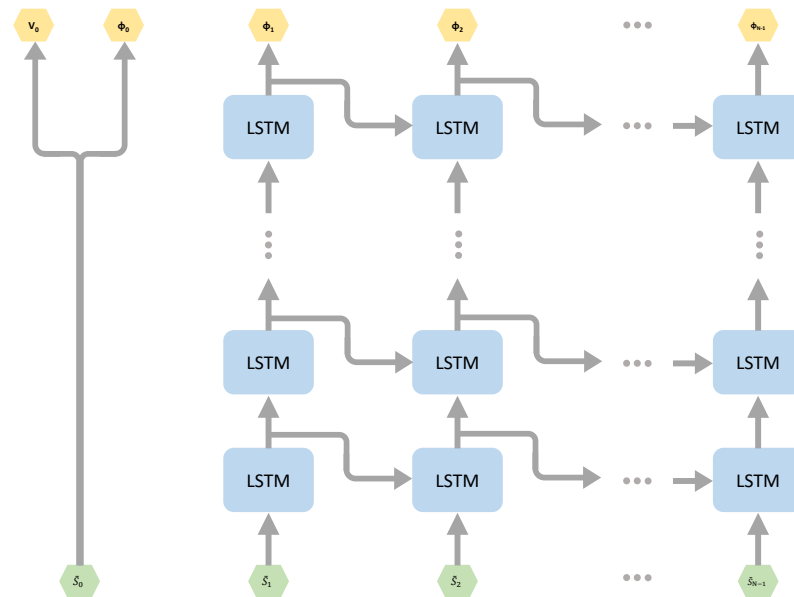
The transaction costs arise from changes in the hedge ratio during the dynamic hedging. Usually, the transaction costs include a  $\delta$  proportion of the value of the transaction and a flat rate (i.e.,  $c$  dollars per trade). Hence, the general fee structure is modeled as the following form:

$$Z = - \left[ c + \delta S(1) |\Phi_1(S(1), \Theta) - \Phi_0| + \sum_{j=1}^{T-2} (c + \delta S(j+1) |\Phi_{j+1}(S(j+1), \Theta) - \Phi_j(S(j+1), \Theta)|) \right]. \quad (20)$$

Taking transaction costs  $\chi$  into consideration, the total wealth change  $\Delta W$  described in (19) becomes:

$$\Delta W = V_0 + \Phi_0 \Delta S(0, 1) + \sum_{j=1}^{T-1} \Phi_j(S(j), \Theta) \Delta S(j, j+1) - P + Z \quad (21)$$

Since the impact of fixed transaction costs is generally overshadowed by that of the proportional part, we put our emphasis on the presence of  $\delta$  in our numerical examples and keep  $c = 0$ .



**Figure 2.** LSTM recurrent neural network architectures [Hochreiter and Schmidhuber \(1997\)](#).

### 3. Results and Discussion

In this section we provide both analytic and empirical results of our machine learning hedging model and discuss their properties. We first test the validity of our model using analytic solutions of the Heston-Nandi GARCH (HN-GARCH) model. The initial endowment  $V(0, S(0))$  and hedging strategy  $\Phi(0, S(0))$  will be justified and the distribution of total wealth change, i.e., the hedging error, will also be examined under different loss functions, and under conditions of with/without transaction costs. Similar analyses will be implemented onto empirical results generated from calibration of real-world data for particular illiquid assets using the Q-GARCH model.

#### 3.1. Analytic Results

Before importing actual data to our model, we would like to justify the validity of our LSTM architecture with simulated data through the comparison between results of analytic solution and our hedging model. Given that the illiquid asset hedging problem requires a discrete-time context, the hedging model we implemented should be model-free and data-driven. Therefore, its availability remains the same no matter the choice of particular calibration model to generate our simulated data. The model we choose is the Heston-Nandi GARCH model, namely the HN-GARCH model.

##### 3.1.1. HN-GARCH

We assume that we are equipped with a complete probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \{0,1,\dots,N\}}, \mathbb{P})$ , where  $\mathbb{P}$  is the physical measure. We denote by  $Y_t := \log(S_t/S_{t-1})$  the one-period log-return process, where  $S_t$  is the asset price at time  $t$ . The conditional variance  $h_t = \mathbb{V}ar[Y_t | \mathcal{F}_{t-1}]$  is an  $\mathcal{F}_t$ -predictable process. For the HN-GARCH model, one can derive the unconditional moment generating function of both  $\log S_{t-1}$  and  $h_t$  in an exponential affine form with coefficients satisfying some recursive relationships, which is the key ingredient in deriving closed-form solutions for variance-optimal hedging.



The dynamics of the log-return process are assumed to follow the Heston-Nandi GARCH(1,1) model under the physical measure  $\mathbb{P}$ , and are given by:

$$\begin{cases} Y_t = r + \lambda h_t + \sqrt{h_t} z_t, & z_t \sim \mathbf{N}(0, 1), \\ h_t = \omega + \alpha(z_{t-1} - \gamma\sqrt{h_{t-1}})^2 + \beta h_{t-1}. \end{cases} \quad (22)$$

In the above conditional mean equation,  $r$  denotes the one-period risk-free interest rate,  $\lambda$  is the equity risk-premium parameter and  $z_t$  is a sequence of i.i.d. standard Gaussian distributed random variables. The conditional variance process  $h_t$  as an affine GARCH(1,1) structure with the parameters  $\omega$ ,  $\alpha$ ,  $\beta$  and  $\gamma$  satisfying the standard positivity and stationarity constraints. The  $\gamma$  parameter captures asymmetry in the response of volatility to positive versus negative return shocks, and it reflects the leverage effect.

Under arbitrage-free condition, the price of any contingent claim can be expressed as the discounted expected value of its payoff at maturity under equivalent martingale measure. Here, we use the exponential affine pricing kernel first introduced for derivative valuation under GARCH models by Siu et al. (2004). Under this new pricing probability measure, denoted here by  $\mathbb{Q}$ , the risk-neutral returns dynamics coincide with those derived in Heston and Nandi (2000) and are given below:

$$\begin{cases} Y_t = r - \frac{1}{2}h_t + \sqrt{h_t} z_t^*, & z_t^* \sim \mathbf{N}(0, 1), \\ h_t = \omega + \alpha(z_{t-1}^* - \gamma^*\sqrt{h_{t-1}})^2 + \beta h_{t-1}. \end{cases} \quad (23)$$

Here, the innovation process  $z_t^*$  is standard Gaussian distributed under  $\mathbb{Q}$ . The risk-neutral leverage effect parameter  $\gamma^*$  is related to the physical counterpart by  $\gamma^* = \gamma + \lambda + \frac{1}{2}$ .

The risk-neutral parameters for the HN-GARCH risk-neutral dynamics used for our numerical exercises illustrated in Table 1 are taken from GARCH Options Toolbox<sup>1</sup>. Here we refer to Christoffersen et al. (2008) and Christoffersen et al. (2012) for the analytic solution of the mean-variance hedging approach for our model. We generate  $N = 10,000$  paths of  $T = 30$  time steps, with initial underlying price  $S_0 = 100$ , strike price  $K = 100$ , and risk-free rate  $r = 0$ . After obtaining the analytic results of the HN-GARCH model: the initial endowment  $V(0, S(0))$  and hedging strategy  $\Phi(0, S(0))$ , together with the distribution of hedging error  $\Delta W$ , we analyze how hedging results from our model coincide with these.

**Table 1.** Parameters for the HN-GARCH model.

| $\omega$                  | $\alpha$              | $\beta$ | $\gamma^*$ |
|---------------------------|-----------------------|---------|------------|
| $7.522908 \times 10^{-9}$ | $7.83 \times 10^{-7}$ | 0.881   | 378        |

### 3.1.2. Results

We first compare our LSTM results with the HN-GARCH model (Heston and Nandi 2000) analytic solutions with the mean-variance hedging approach to justify our model. The hedging results for LSTM and analytic solutions are listed in Table 2. Observations show that the discrepancy of the initial endowment  $V_0$  and the hedging strategy  $\Phi_0$  both remain less than 5%, which indicates the validity of our model.

**Table 2.** Hedging results for LSTM and analytic solutions.

|             | $V_0$  | $\Phi_0$ |
|-------------|--------|----------|
| LSTM        | 2.2833 | 0.4871   |
| Analytic    | 2.2877 | 0.4677   |
| Discrepancy | 0.19%  | 4.15%    |

After the validation of our model, we could move on and testify the difference between different loss functions. Similar comparison is applied here, and we notice that both initial

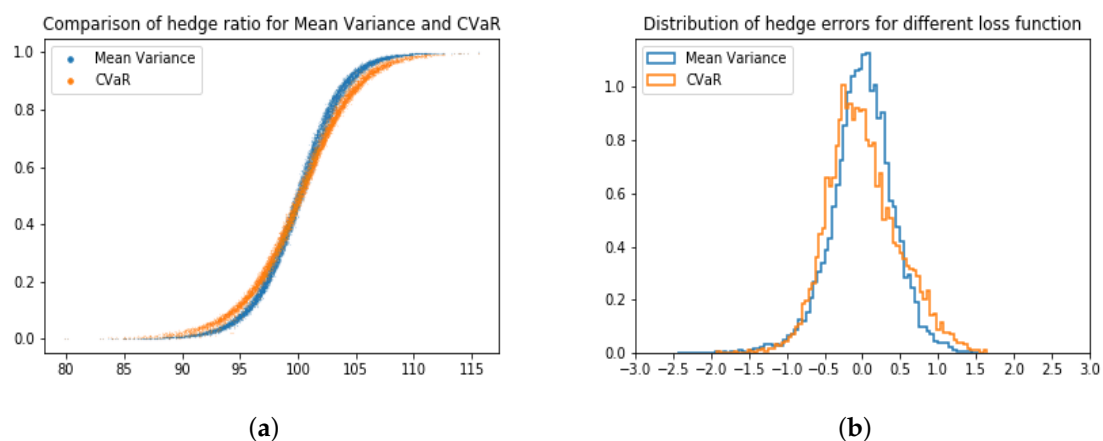
endowments are around \$2.28, and the discrepancy of hedging strategies is slightly above 4%, as shown in Table 3. Such results confirm that both loss functions are promising, and we would further investigate the difference between these two loss functions.

We compare the hedge ratios of two loss functions versus the change of underlying price at a snapshot of  $t = 20$  in Figure 3a. Both two lines converge to 1.0 at deep in-the-money range, and get close to 0.0 within out-of-the-money regime. However, the curve of Mean Variance shows a slightly larger slope than that of the CVaR curve. Such pattern results from the feature of the CVaR loss function, that it concerns more about the distribution of the hedge error  $\Delta W$ , so that it is less sensitive to the change of underlying price itself.

We move on to the investigation of the distribution of hedge error  $\Delta W$  as displayed in Figure 3b. More statistics are shown in Table 4, and it is obvious that the mean and variance for the mean variance loss function are smaller than their counterparts for the CVaR loss function. At the first glance, we might conclude from these statistics that the mean variance loss function outperforms CVaR. However, after a second thought, we realize that it is because the loss function is called “mean variance”, and the CVaR loss function should be reasonable or even more practical because it concerns more about the tail of the distribution. As shown in Figure 3b, the distribution for CVaR actually has a short tail than the mean variance loss function. In this case, we shall not say that one of the loss functions is better than the other. Instead, both of these loss functions have their advantages in relevant fields. The mean variance loss function might be utilized more by traders, because they care more about the expectation of their investments, while risk managers might prefer the CVaR loss function because their concern how much money they might lose in the worst case.

**Table 3.** Hedging results for different loss functions.

|               | $V_0$  | $\Phi_0$ |
|---------------|--------|----------|
| Mean Variance | 2.2833 | 0.4871   |
| CVaR          | 2.2792 | 0.5077   |
| Discrepancy   | 0.18%  | 4.06%    |



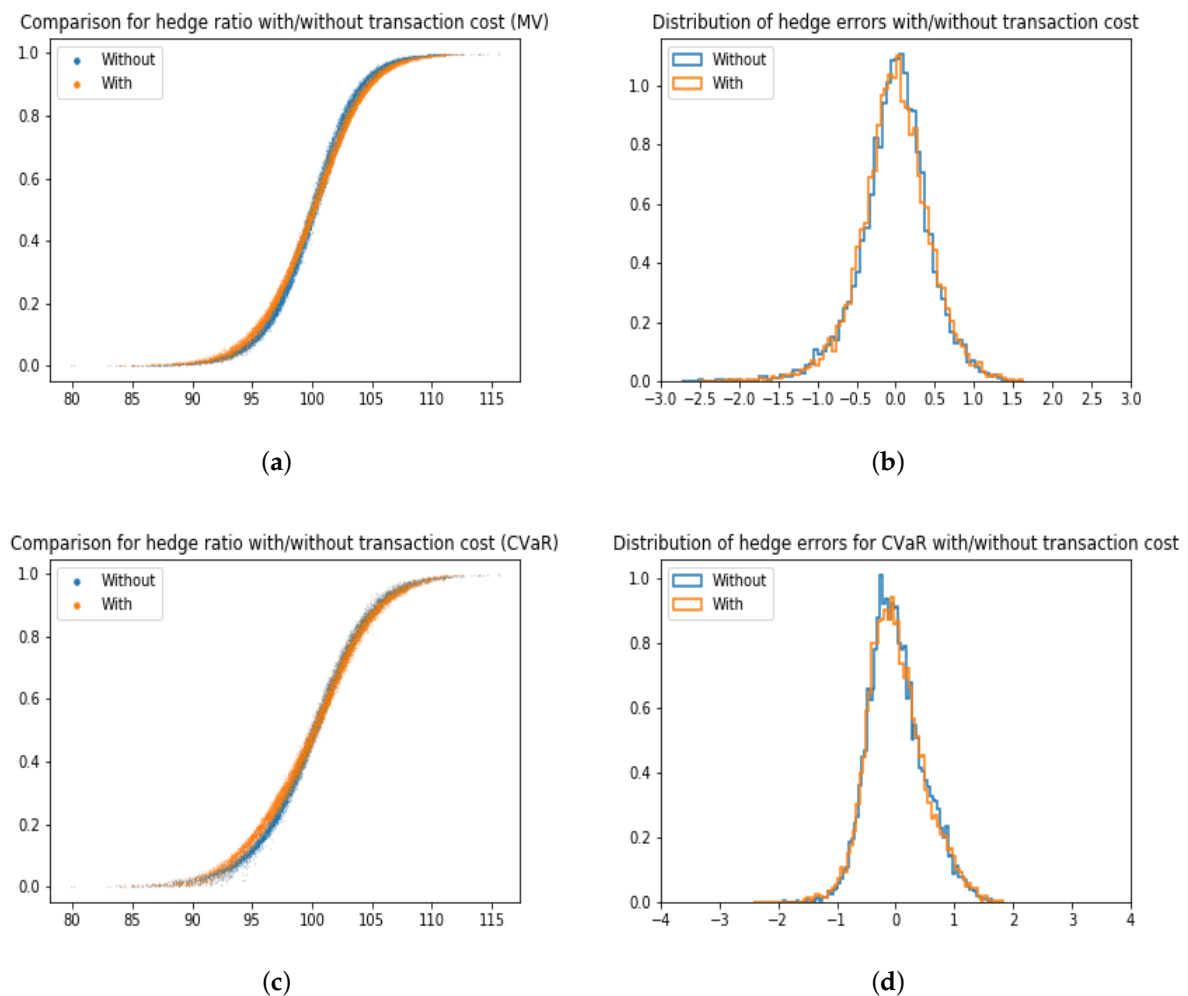
**Figure 3.** Comparison of difference between loss functions. (a) Comparison of hedge ratio for different loss functions. (b) Distribution of hedge errors for different loss functions.

After the comparison between two loss functions, we move on to testify the impact of transaction costs. In our project, we implied a proportional transaction cost of 0.2%, and investigate its influence on the initial endowment  $V_0$ , hedging strategy  $\Phi_0$ , and the distribution of hedging error  $\Delta W$ .

**Table 4.** Statistics of hedge error distribution for different loss functions.

|          | Mean Variance          | CVaR   |
|----------|------------------------|--------|
| Mean     | $-3.98 \times 10^{-7}$ | 0.012  |
| Variance | 0.1676                 | 0.2179 |

Similar comparison is shown in Table 5. We notice that the presence of transaction costs causes almost no change in hedge ratio  $\Phi_0$ , but increases the option price  $V_0$ . This is because for every hedge step, the investor has to pay a small amount of money. These small amounts of money accumulate and are reflected in the final option price. The curves of hedge ratio versus underlying price are also presented in Figure 4a. Observation shows that no conspicuous impacts are done by the presence of transaction costs in terms of hedge ratio. Furthermore, the distribution of hedging error described in Figure 4b and Table 6 indicates that the distribution of  $\Delta W$  remains unchanged when a proportional transaction cost is implied. Since all aforementioned analyses are studied using the mean variance loss function, similar investigation are implemented under CVaR loss function, and the results displayed in Figure 4c,d confirm our conclusion.



**Figure 4.** Comparison of difference with/without transaction costs. (a) Comparison of hedge ratio with/without transaction costs (mean variance). (b) Distribution of hedge errors with/without transaction costs (mean variance). (c) Comparison of hedge ratio with/without transaction costs (CVaR). (d) Distribution of hedge errors with/without transaction costs (CVaR).

**Table 5.** Hedging results with/without transaction costs.

|         | $V_0$  | $\Phi_0$ |
|---------|--------|----------|
| Without | 2.2833 | 0.4871   |
| With    | 2.5875 | 0.4862   |

**Table 6.** Statistics of hedge error distribution with/without transaction costs.

|          | Without | With    |
|----------|---------|---------|
| Mean     | −0.0109 | −0.0089 |
| Variance | 0.1963  | 0.1942  |

### 3.2. Empirical Results

Empirical properties of asset returns are mainly characterized by volatility clustering, high kurtosis, and slow decay of the auto-correlations in squared returns. GARCH models are commonly employed in modelling financial series that exhibit these properties. However, standard GARCH models assume that positive and negative error terms have a symmetric effect on the volatility. In practice, this assumption is frequently violated due to leverage effect, i.e., the asymmetric response of volatility to positive and negative returns. Therefore, QGARCH(1,1) is introduced as a realistic objective measure model for equity index returns, to help us discover how the hedging performance affected by leverage effect. Such properties of the QGARCH model fit the asymmetric characteristics of illiquid assets.

#### 3.2.1. Q-GARCH

The QGARCH model was proposed by [Sentana \(1995\)](#) to overcome the weakness of the GARCH model. Under QGARCH(1,1) framework, the asset and its volatility evolves as follows:

$$\begin{cases} S_{t+1} - S_t = S_t(\mu\Delta t + \sigma_t\sqrt{\Delta t}z_t), \\ \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma\varepsilon_{t-1}, \end{cases} \quad (24)$$

where the innovation process is define as  $\varepsilon_t = \sigma_t z_t$ .  $\{z_t\}$  is a sequence of i.i.d random variables and assumed to follow the standard normal distribution  $N(0,1)$ . The auto-regressive parameter  $\beta$  partly determines the persistence of the variance in the model, and the innovation parameter  $\alpha$  determines the volatility of volatility. When  $\alpha$  is not zero, the kurtosis of the spot return increases and consequently the distribution of returns exhibit fat-tail phenomenon. This characteristic renders the model consistent with stylized facts that financial time series have positive excess kurtosis and heavy-tailed distributions. The parameter  $\gamma$  captures the asymmetry in the response of volatility to positive versus negative return shocks, and it also captures the leverage effect. If the parameter  $\gamma$  is zero, the distribution is symmetric, while a value of  $\gamma$  different from zero results in asymmetric influences of the shocks, e.g., a large negative shock  $z_t$  raises the variance more than a large positive shock does.

#### 3.2.2. Data

The analysis in this chapter was based on the HFRI Fund of Funds Index (HFRIFO). Our main tests use monthly data from 31 December 2005 to 30 June 2016. The datasets are obtained from Bloomberg and the sample spans 126 trading months. The sample data are used to calibrate the Q-GARCH(1,1) model by using maximum likelihood method.

The HFRX Global Index (HFRX) is an investable index with daily liquidity, including a subset of managers from the HFR database (approximately 6800 funds) that are open for investment and will accept managed account investments from HFR, along with other restrictions. In terms of Hedge Fund Index, survivor-ship bias commonly occurs. An upward bias is created when obsolete funds cease to report to a database. In addition, hedge funds may also choose to stop reporting funds that will result in a downward bias.

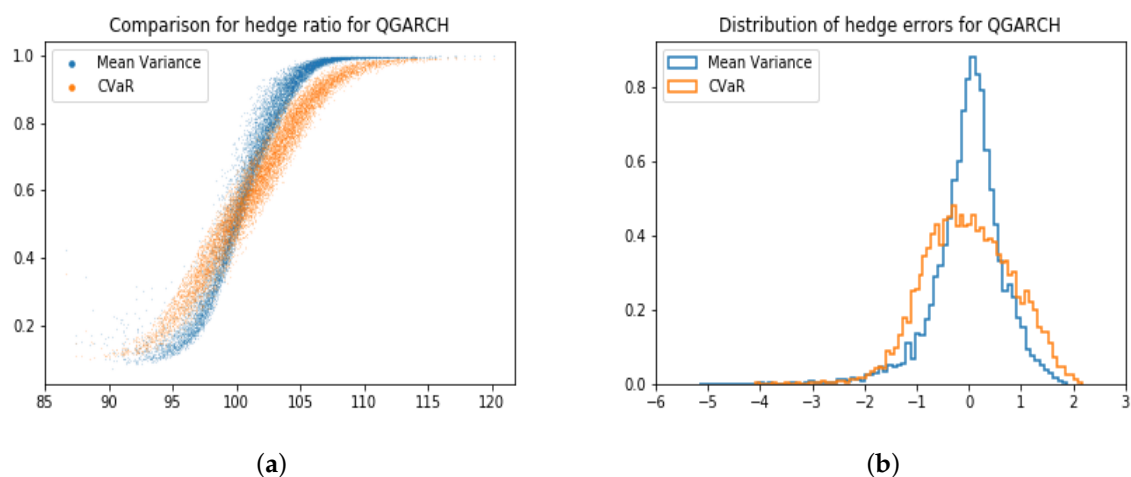
### 3.2.3. Results

The analysis on empirical results will resemble the one implemented on numerical results. We first investigate the difference between two loss functions. As shown in Table 7, the discrepancies of option price  $V_0$  and hedge ratio  $\Phi_0$  are all restricted within a safe range of 5%, which indicates both loss functions are valid under real-life scenarios. Moving on, a snapshot of hedge ratio versus the underlying price change is plotted in Figure 5a. We notice that both curves converge to 1 in the money and goes down to 0 out of the money. And the slope for mean variance is larger than CVaR, just as described in numerical results. However, a significant difference here is that the points in the plot is rather scattered, compared to clear curves we obtained for numerical results. Such phenomenon comes from the use of the Q-GARCH model. The Q-GARCH model captures the asymmetric property of the illiquid asset, while leaves each single dot rather path-dependent. Moreover, the distribution of hedging error  $\Delta W$  is investigated in Figure 5b and Table 8. The mean variance loss function shows promising capability in minimizing total mean and variance, while the CVaR loss function indeed secure a shorter tail for the distribution.

Similarly, we try to figure out how transaction costs influence our hedging procedure for empirical results. Table 9 demonstrates that we obtain an increase option price with the presence of transaction costs, and unchanged hedge ratio. Table 10 and Figure 6 confirms that the distribution of the hedging error  $\Delta W$  is not influenced by transactions costs, no matter which loss function is applied.

**Table 7.** Hedging results for different loss functions.

|               | $V_0$  | $\Phi_0$ |
|---------------|--------|----------|
| Mean Variance | 2.0030 | 0.5895   |
| CVaR          | 2.1064 | 0.5993   |
| Discrepancy   | 4.91%  | 1.64%    |



**Figure 5.** Comparison of difference between loss functions (Q-GARCH). (a) Comparison of hedge ratio for different loss functions (Q-GARCH). (b) Distribution of hedge errors for different loss functions (Q-GARCH).

**Table 8.** Statistics of hedge error distribution for different loss functions.

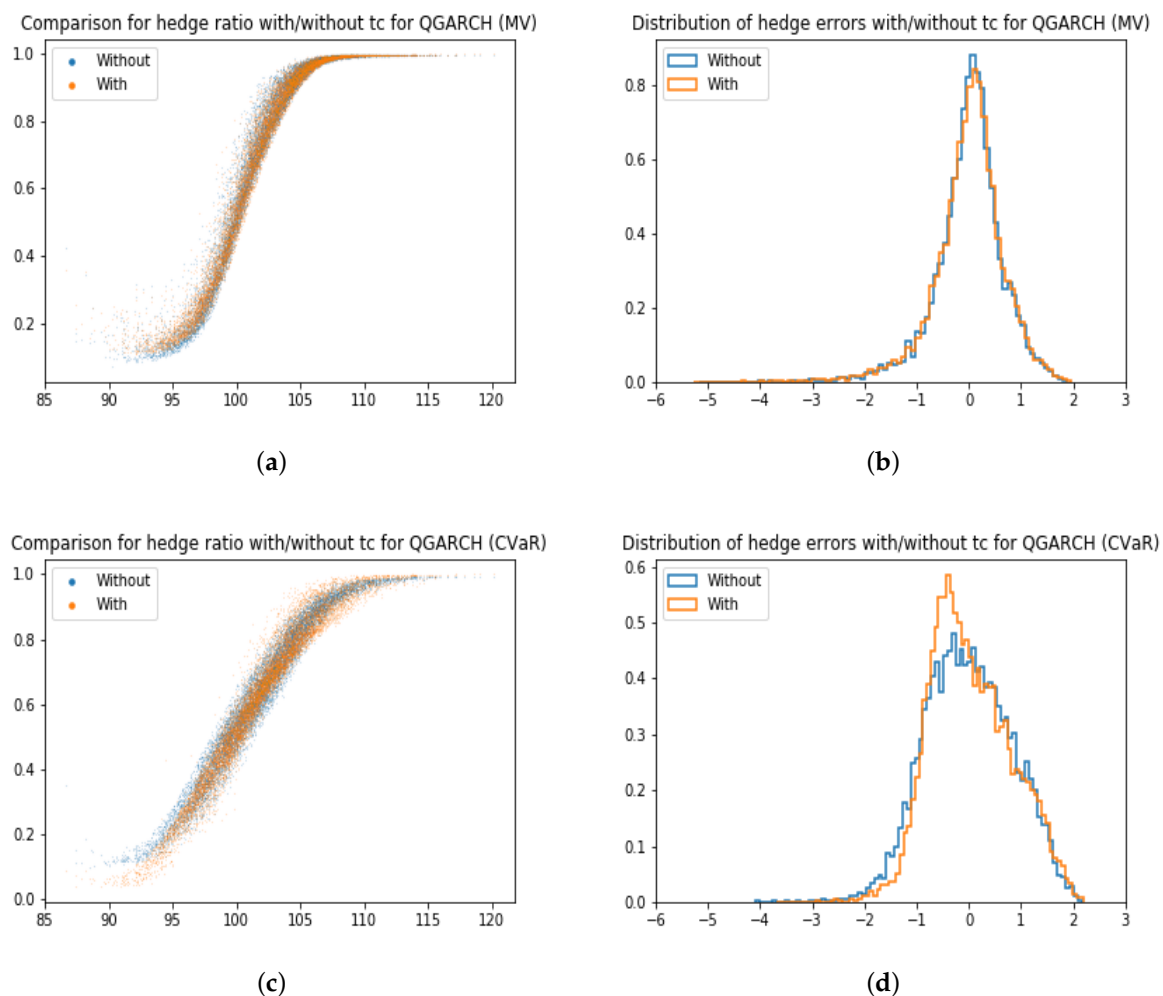
|          | Mean Variance           | CVaR      |
|----------|-------------------------|-----------|
| Mean     | $-1.092 \times 10^{-5}$ | $-0.0086$ |
| Variance | 0.4317                  | 0.6966    |

**Table 9.** Hedging results with/without transaction costs.

|         | $V_0$  | $\Phi_0$ |
|---------|--------|----------|
| Without | 2.0030 | 0.5895   |
| With    | 2.1319 | 0.5865   |

**Table 10.** Statistics of hedge error distribution with/without transaction costs.

|          | Without                 | With                    |
|----------|-------------------------|-------------------------|
| Mean     | $-1.092 \times 10^{-5}$ | $-1.014 \times 10^{-6}$ |
| Variance | 0.4317                  | 0.4530                  |

**Figure 6.** Comparison of with/without transaction costs (Q-GARCH). (a) Comparison of hedge ratio with/without transaction costs (mean variance). (b) Distribution of hedge errors for different loss functions (mean variance). (c) Comparison of hedge ratio for different loss functions (CVaR). (d) Distribution of hedge errors for different loss functions (CVaR).

#### 4. Conclusions

This study has successfully introduced and validated a model-free, data-driven approach leveraging Long Short-Term Memory (LSTM) neural networks to navigate the hedging challenges associated with illiquid assets. Through meticulous analysis, we examined the effects of initial endowment, hedging strategies, and the distribution of hedging errors across two distinct loss functions, while also assessing the impact of transaction costs on these elements. Our findings affirm the efficacy and relevance of both loss functions,



each displaying unique advantages and applicability within specific contexts. Importantly, our research highlights that transaction costs contribute to an escalation in the final option pricing, albeit without altering the hedging strategy or the extent of hedging error. A critical strength of the proposed LSTM-based model lies in its flexibility and lack of reliance on pre-defined models or assumptions, offering robust adaptability to a wide array of data within a discrete-time framework. The implementation of LSTM neural networks allows for a sophisticated examination of data patterns that traditional models might possibly overlook, significantly shortens the calculation runtime compared with analytical approaches, and offers a more nuanced understanding of risk in hedge fund portfolios. Looking forward, potential avenues for further investigation include exploring the effects of varied hedging frequencies and the implications of incorporating diverse option types, such as binary options or those with target volatility. This work not only broadens the horizons of financial risk management strategies but also lays the groundwork for future innovations in the field.

**Author Contributions:** Writing—original draft preparation, Y.W. and Y.Z.; writing—review and editing, L.T. and Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Data available upon request.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Note

- <sup>1</sup> [https://www.dropbox.com/s/og8iiatgh28d7fq/GARCHOptionsToolbox18Nov2014.zip?dl=0&file\\_subpath=%2FGARCHOptionsToolbox%2FGARCHOptionsToolboxDoc.pdf](https://www.dropbox.com/s/og8iiatgh28d7fq/GARCHOptionsToolbox18Nov2014.zip?dl=0&file_subpath=%2FGARCHOptionsToolbox%2FGARCHOptionsToolboxDoc.pdf) on 27 April 2023.

## References

- Alexander, Siddharth, Thomas F. Coleman, and Yuying Li. 2003. Derivative portfolio hedging based on cvar. In *New Risk Measures in Investment and Regulation*. Hoboken: Wiley.
- Bertsimas, Dimitris, Leonid Kogan, and Andrew W. Lo. 2001. Hedging derivative securities and incomplete markets: An  $\epsilon$ -arbitrage approach. *Operations Research* 49: 372–97. [CrossRef]
- Buehler, Hans, Lukas Gonon, Josef Teichmann, and Ben Wood. 2019. Deep hedging. *Quantitative Finance* 1–21. [CrossRef]
- Černý, Aleš, and Jan Kallsen. 2007. On the structure of general mean-variance hedging strategies. *The Annals of Probability* 35: 1479–531. [CrossRef]
- Černý, Aleš, and Jan Kallsen. 2008. Mean-variance hedging and optimal investment in heston's model with correlation. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* 18: 473–92. [CrossRef]
- Christoffersen, Peter, Kris Jacobs, and Chayawat Ornathanalai. 2012. Dynamic jump intensities and risk premiums: Evidence from s&p500 returns and options. *Journal of Financial Economics* 106: 447–72.
- Christoffersen, Peter, Kris Jacobs, Chayawat Ornathanalai, and Yintian Wang. 2008. Option valuation with long-run and short-run volatility components. *Journal of Financial Economics* 90: 272–97. [CrossRef]
- Du, Xin, Jinjian Zhai, and Koupin Lv. 2016. Algorithm trading using q-learning and recurrent reinforcement learning. *Positions* 1: 1.
- Duffie, Darrell, and Henry R. Richardson. 1991. Mean-variance hedging in continuous time. *The Annals of Applied Probability* 1: 1–15. [CrossRef]
- Fecamp, Simon, Joseph Mikael, and Xavier Warin. 2019. Risk management with machine-learning-based algorithms. *arXiv* arXiv:1902.05287.
- Föllmer, Hans, and Alexander Schied. 2011. *Stochastic Finance: An Introduction in Discrete Time*. Berlin: Walter de Gruyter.
- Gourieroux, Christian, Jean Paul Laurent, and Huyên Pham. 1998. Mean-variance hedging and numéraire. *Mathematical Finance* 8: 179–200. [CrossRef]
- Hagan, Patrick S., Deep Kumar, Andrew S. Lesniewski, and Diana E. Woodward. 2002. Managing smile risk. *The Best of Wilmott* 1: 249–96.
- Heston, Steven L. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* 6: 327–43. [CrossRef]
- Heston, Steven L., and Saikat Nandi. 2000. A closed-form garch option valuation model. *The Review of Financial Studies* 13: 585–625. [CrossRef]
- Hochreiter, Sepp, and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation* 9: 1735–80. [CrossRef] [PubMed]
- Hornik, Kurt. 1991. Approximation capabilities of multilayer feedforward networks. *Neural Networks* 4: 251–57. [CrossRef]
- Hornik, Kurt, Maxwell Stinchcombe, and Halbert White. 1989. Multilayer feedforward networks are universal approximators. *Neural Networks* 2: 359–66. [CrossRef]

- Hull, John. 1988. An analysis of the bias in option pricing caused by a stochastic volatility. *Advances in Futures and Options Research* 3: 29–61.
- Hull, John, and Alan White. 1987. The pricing of options on assets with stochastic volatilities. *The Journal of Finance* 42: 281–300. [\[CrossRef\]](#)
- Hull, John, and Alan White. 2017. Optimal delta hedging for options. *Journal of Banking & Finance* 82: 180–90.
- Hutchinson, James M., Andrew W. Lo, and Tomaso Poggio. 1994. A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance* 49: 851–89. [\[CrossRef\]](#)
- Jiang, Zhengyao, Dixing Xu, and Jinjun Liang. 2017. A deep reinforcement learning framework for the financial portfolio management problem. *arXiv* arXiv:1706.10059.
- Kolm, Petter N., and Gordon Ritter. 2019. Dynamic replication and hedging: A reinforcement learning approach. *The Journal of Financial Data Science* 1: 159–71. [\[CrossRef\]](#)
- Krokhmal, Pavlo, Jonas Palmquist, and Stanislav Uryasev. 2002. Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk* 4: 43–68. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Mikkilä, Oskari, and Juho Kannianen. 2023. Empirical deep hedging. *Quantitative Finance* 23: 111–22. [\[CrossRef\]](#)
- 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.
- Rémillard, Bruno, and Sylvain Rubenthaler. 2013. Optimal hedging in discrete time. *Quantitative Finance* 13: 819–25. [\[CrossRef\]](#)
- Rockafellar, R. Tyrrell, and Stanislav Uryasev. 2002. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* 26: 1443–71.
- Rockafellar, R. Tyrrell, and Stanislav Uryasev. 2000. Optimization of conditional value-at-risk. *Journal of Risk* 2: 21–42. [\[CrossRef\]](#)
- Schäfer, Anton Maximilian, and Hans Georg Zimmermann. 2006. Recurrent neural networks are universal approximators. In *International Conference on Artificial Neural Networks*. Berlin: Springer, pp. 632–40.
- Schweizer, Martin. 1995. Variance-optimal hedging in discrete time. *Mathematics of Operations Research* 20: 1–32. [\[CrossRef\]](#)
- Sentana, Enrique. 1995. Quadratic arch models. *The Review of Economic Studies* 62: 639–61. [\[CrossRef\]](#)
- Siu, Tak Kuen, Howell Tong, and Hailiang Yang. 2004. On pricing derivatives under garch models: A dynamic gerber-shiu approach. *North American Actuarial Journal* 8: 17–31. [\[CrossRef\]](#)

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.