

Deep Learning for Options Trading: An End-To-End Approach

Wee Ling Tan

weeling@robots.ox.ac.uk

0000-0003-2664-8238

Oxford-Man Institute of Quantitative Finance University of Oxford Oxford United Kingdom

Stephen Roberts

sjrob@robots.ox.ac.uk

Oxford-Man Institute of Quantitative Finance University of Oxford Oxford United Kingdom

Stefan Zohren

stefan.zohren@eng.ox.ac.uk

Oxford-Man Institute of Quantitative Finance University of Oxford Oxford United Kingdom

Abstract.

We introduce a novel approach to options trading strategies using a highly scalable and data-driven machine learning algorithm. In contrast to traditional approaches that often require specifications of underlying market dynamics or assumptions on an option pricing model, our models depart fundamentally from the need for these prerequisites, directly learning non-trivial mappings from market data to optimal trading signals. Backtesting on more than a decade of option contracts for equities listed on the S&P 100, we demonstrate that deep learning models trained according to our end-to-end approach exhibit significant improvements in risk-adjusted performance over existing rules-based trading strategies. We find that incorporating turnover regularization into the models leads to further performance enhancements at prohibitively high levels of transaction costs.

Options, Derivatives, Trading Strategies, Machine Learning, Momentum, Mean-Reversion

1. Introduction

Options form a class of derivatives known as contingent claims, granting one of the counterparties the right to transact an underlying asset at a certain time in the future and at a specific price. For instance, the buyer of a call (put) stock option pays a premium to the seller for the right to buy (sell) the underlying stock by a specified expiration date and strike price. Unlike in cash markets that typically exhibit linear payoffs, the ability to generate synthetic and highly non-linear exposures has made options an in-

creasingly important tool among investors as trading instruments alongside other derivatives (Black, 1975).

The options market has continued to grow significantly over the last decade. According to the Options Clearing Corporation, the average daily combined volume of US equity and non-equity options has steadily risen and roughly tripled over the past decade from about 16.3 million in 2013 to 44.2 million contracts in 2023 (Corporation, 2024). Given the increasing popularity of options trading by market participants conducted in modern electronic exchanges, there exists an unprecedented opportunity to engage in large-scale data analysis of options trading from the perspective of an active investor.

Classical option pricing models and their parametric variants, stemming from the seminal works of Black and Scholes (Black and Scholes, 1973) and Merton (Merton, 1973), regard options as redundant assets. While these approaches offer a tractable no-arbitrage pricing framework, they necessitate thorough specifications of underlying dynamics and require assumptions such as a frictionless market and the ability of market makers to perfectly hedge their exposures. Practically, these models are often vulnerable to model misspecification and are usually too simplistic to sufficiently account for empirically observed variations in option returns (Büchner and Kelly, 2022). Recent works have challenged this framework, demonstrating that option prices are influenced by additional risks beyond the underlying asset's exposure (Coval and Shumway, 2001; Goyal and Saretto, 2009; Heston et al., 2023), with several studies documenting the existence of mispricing in the options market (Bali et al., 2023; Eisdorfer et al., 2022).

Despite the accelerating growth of the options market and the evidence of options mispricing, there is a surprising lack of research on a scalable machine learning-based strategy that is able to directly leverage the abundance of data to conduct options trading on behalf of an active investor. In this work, we address this gap by introducing a novel class of deep learning models capable of managing and trading a portfolio of options in a highly data-driven approach.

Previous research has conventionally approached the complexities of managing portfolios of options by developing methods for optimal hedging and accurately pricing options. Our model departs fundamentally from the need for these prerequisites, and are directly optimized to learn highly non-trivial mappings from observed data to optimal trading decisions based on risk-adjusted performance. The resulting end-to-end framework does not depend on specific market dynamics, and can be extended broadly across instruments where market data is available.

To illustrate our approach, we construct the framework by training end-to-end neural networks of different complexities and show that our models are able to outperform existing rules-based strategies in managing a portfolio of options, demonstrating strong risk-adjusted returns over a backtest period of over a decade. We subsequently extend our models to incorporate turnover regularization during optimization, leading to further performance enhancements in the presence of high transaction costs.

2. Related Work

To place our work within the context of existing research on options, we briefly review the broad literature in three areas: replication and hedging, pricing and valuation, and predicting option returns. We

subsequently demonstrate how our work contributes to existing systematic trading strategies within options markets.

2.1. Options Literature

Traditionally, options have been extensively studied due to their significance in both replication and hedging strategies. The works of (Buehler et al., 2019) and (Kolm and Ritter, 2019) apply reinforcement learning techniques towards approximating policies for optimal hedging in the presence of trading costs. While these works are pertinent to a product or liquidity provider needing to accurately hedge its exposures to books of options and other complex derivatives, our analysis fundamentally differs in its focus on options trading from the standpoint of an active investor seeking to profit from an options trading strategy and whose main objective is not predominantly about neutralizing exposures. Furthermore, hedging techniques which are based on reinforcement learning often require generating or simulating possible market paths to arrive at an optimal trading strategy. In contrast, the solution we offer does not require any assumption or simulation of market processes, and scales with available historical data and compute.

There has also been significant emphasis placed on the need to accurately price or value an option in order to facilitate optimal hedging and trading, with notable contributions from the classical Black-Scholes-Merton option pricing model for European-style options (Black and Scholes, 1973; Merton, 1973). Non-parametric models have also been developed, with (Hutchinson et al., 1994) and (Ivaşcu, 2021) using neural networks to approximate the market's option pricing function. Conversely, our framework focuses on automatically extracting features and making trading decisions directly from available data, effectively circumventing the need for engineering an option pricing or valuation model.

Several studies have also employed machine learning for predicting option returns, often framing the trading strategy as a standard regression problem and subsequently determining the direction of price movements based on forecasted returns. Bali et al. (2023) use both linear and nonlinear models and ex-ante option-based and stock-based characteristics to predict monthly returns of delta-hedged options, demonstrating high out-of-sample forecast accuracy. However in the case of a trend-following strategy, (Lim et al., 2019) shows that models that accurately predict returns do not necessarily guarantee a positive or superior strategy performance, given that the overall profitability of trading strategies is influenced by other factors including the distribution of returns, position sizes and the presence of risk adjustments like volatility targeting (Harvey et al., 2018). Taking this into consideration, our work effectively integrates both trend prediction and optimal position sizing simultaneously within a single end-to-end function, eliminating the need to forecast option returns in making trading decisions.

2.2. Systematic Strategies and Risk Premia in Options Markets

Our work contributes to a series of research documenting systematic strategies originating from studies demonstrating additional risk factors (Büchner and Kelly, 2022; Coval and Shumway, 2001; Goyal and Saretto, 2009; Heston et al., 2023; Vasquez, 2017), and the existence of mispricing (Bali et al., 2023; Eisdorfer et al., 2022) within the options market. Considering the broad range of possible risk factors and strategies, we concentrate our analysis on systematic trend-based strategies. Given their simplicity of being based primarily on historical price trends and returns, and unlike volatility-based strategies, trend-

based strategies often do not require additional assumptions on any underlying option pricing model or specifications of market dynamics.

2.3. Momentum, Mean-Reversion and Applications of Machine Learning

Trend-based strategies consists of primarily momentum and mean-reversion strategies. Momentum strategies operate on the principle that asset returns exhibit a tendency to persist in the same direction and aim to trade in the direction of the trend (Jegadeesh and Titman, 1993; Moskowitz et al., 2012). On the other hand, mean-reversion strategies adopt an opposite and contrarian view, taking on positions that bet on the eventual break down and correction of overextended trends (De Bondt and Thaler, 1985; Poterba and Summers, 1988). Recently, machine learning models have been increasingly utilized in trend-based trading strategies (Lim et al., 2019; Poh et al., 2021, 2022; Tan et al., 2023; Wood et al., 2023, 2022).

While momentum strategies have been extensively documented in a range of asset classes from cash equities (Rouwenhorst, 1998) to futures markets (Moskowitz et al., 2012), these strategies have received relatively less attention in options markets. Heston et al. (2023) document strong evidence of the presence of momentum in options markets and propose a series of rules-based momentum strategies. However, these rules-based strategies often require explicit specifications for the trading rule, often with insufficient evidence to justify selecting a certain rule over another. In light of these observations, we employ deep neural networks to automatically learn risk-adjusted trading rules in a data-driven approach, and provide a thorough comparison between these trend-based strategies.

3. Overview of Dataset

Our dataset consists of option contracts sourced from the OptionMetrics Ivy DB database, comprising end-of-day bid-ask prices, implied volatility and Greeks of individual options. We independently verify the accuracy of the provided implied volatility and Greeks using a binomial tree model of Cox et al. (1979) only for the purpose of initial delta hedging (refer to Section 4), but otherwise we do not assume any underlying option pricing model throughout this work. We focus our analysis on option contracts of equities listed on the S&P 100 Index, as these companies span a wide range of sectors, representing major large-cap optionable companies in the US market, and are associated with higher option liquidity. We obtain underlying stock prices from the Center for Research in Security Prices (CRSP), which we use for computing option moneyness and accounting for corporate actions such as stock splits. We perform backtesting with the most recent market data (as of this work) from 2010 to 2023, which notably includes the market selloff following the COVID-19 pandemic.

We impose a series of data filters established in the literature to ensure the consistency of our analysis. We consider only standard monthly option contracts that expire on the third Friday of the month and exclude any special settlements due to corporate actions. We then exclude options that contain price observations that breach American-style option bounds, having a bid price of zero, or where the ask price is smaller or equal to the bid price, and require options to have a positive open interest on the day of portfolio formation. In addition, we disregard options that contain one or more missing observations between the day of portfolio formation and expiration.

On the expiration day of each month and for each stock, we form portfolios of static delta-neutral straddle options by selecting a pair of call and put contracts with identical strike prices that are closest to at-the-money (ATM) and expire in the following month. Since it is usually not possible to select options with moneyness exactly equal to 1.0 (moneyness of call = S / K , put = K / S where K = strike price, S = stock price), we select the pair of options that are closest to ATM within a moneyness range of 0.95 to 1.05.

Our resulting universe consists of 29984 options that are traded, with a total of 603068 daily returns observations throughout the backtest period. The returns of straddle options have positive means (1.41% monthly), large standard deviations (90.85% monthly), and significant positive skewness as indicated by a low median (-15.61% monthly). We provide an in-depth explanation on portfolio formation and returns computation in Section 4.

4. Systematic Options Trading Strategies

Let $i = 1, 2, \dots, N_t$ denote individual underlying stocks. For a given portfolio of (1-month, ATM, static delta-neutral) straddle options of these stocks that is rebalanced daily, the overall returns of a strategy that equally diversifies over N_t straddles at day t can be expressed as follows:

$$(1) \quad r_{t,t+1}^{\text{STRATEGY}} = \frac{1}{N_t} \sum_{i=1}^{N_t} X_t^{(i, \text{straddle})} \left(\frac{\sigma_{\text{tgt}}}{\sigma_t^{(i, \text{straddle})}} \right) r_{t,t+1}^{(i, \text{straddle})}$$

whereby

$$(2) \quad r_{t,t+1}^{(i, \text{straddle})} = \frac{p_{t+1}^{(i, \text{straddle})} - p_t^{(i, \text{straddle})}}{p_t^{(i, \text{straddle})}}$$

$$p_t^{(i, \text{straddle})} = w_{\text{norm}}^{(i, \text{call})} p_t^{(i, \text{call})} + w_{\text{norm}}^{(i, \text{put})} p_t^{(i, \text{put})}$$

$$w_{\text{norm}}^{(i, \text{call})} = \frac{w^{(i, \text{call})}}{w^{(i, \text{call})} + w^{(i, \text{put})}}$$

$$w_{\text{norm}}^{(i, \text{put})} = 1 - w_{\text{norm}}^{(i, \text{call})}$$

$$w^{(i, \text{call})} = -\Delta_0^{(i, \text{put})}$$

$$w^{(i, \text{put})} = \Delta_0^{(i, \text{call})}$$

On the day of portfolio formation ($t = 0$), we construct static delta-neutral straddle options by holding the call and put options with weights $-\Delta_0^{(i, \text{put})}$ and $\Delta_0^{(i, \text{call})}$ respectively, with Δ_0 representing initial deltas. We subsequently normalize both weights to sum to one, resulting in a respective weightage of the call and put options that is generally close to 50-50. In evaluating the price of the straddle, we take $p_t^{(i, \text{call})}$ and $p_t^{(i, \text{put})}$ to be the bid-ask midpoints of the call and put options.

Consistent with other works (Goyal and Saretto, 2009; Heston et al., 2023), we focus on static delta-neutral straddles as these instruments are on average invariant to movements in the underlying. (Tian and Wu, 2023) demonstrate that performing one-time delta hedging at initiation neutralizes the overall directional risks associated with an option by about 70%. As such, we opt for delta hedging at initiation

and following similar works, we disregard the early exercise premium embedded in the American-style options and therefore ignore the possibility of early assignment of short options.

Here, $X_t^{(i, \text{straddle})} \in [-1, 1]$ denotes the trading signal or position for the straddle option of stock i at day t , and $r_{t, t+1}^{(i, \text{straddle})}$ is the realized returns of the straddle from t to $t + 1$. We find evidence that straddle options in the cross-section of individual constituents of the S&P 100 exhibit different levels of volatility using a Levene's test at a significance level of 1%. Given these differences, we include volatility targeting at the level of individual straddle options to scale the realized returns $r_{t, t+1}^{(i, \text{straddle})}$ by their volatility in order to target equal assignments of risk. We set the annualized volatility target σ_{tgt} to be 15% and estimate the ex-ante volatility $\sigma_t^{(i, \text{straddle})}$ with a 20-day exponentially weighted moving standard deviation of daily straddle returns.

All of the following trend-based benchmarks which we incorporate in our work adhere to this general framework and are concerned with constructing an accurate trading signal $X_t^{(i, \text{straddle})}$.

Long Only (Short Only)

This strategy takes a maximum long (short) position $X_t^{(i, \text{straddle})} = 1$ or -1 for all straddle options in the portfolio. Since the performance of a short only strategy is the exact opposite of a long only strategy, we focus only on the long only strategy.

TSMOM (TSMR)

Following the time-series momentum strategy (TSMOM) of Moskowitz et al. (2012), we adopt the strategy with a monthly horizon. The position taken for a straddle option is based on the sign of the option's returns over the past 20 days: $X_t^{(i, \text{straddle})} = \text{sgn}(r_{t-20, t}^{(i, \text{straddle})})$. In the time-series mean reversion strategy (TSMR), we modify TSMOM to take on a negative loading on past returns, where $X_t^{(i, \text{straddle})} = -\text{sgn}(r_{t-20, t}^{(i, \text{straddle})})$. This strategy takes a contrarian approach by taking on a long (short) position for straddles with negative (positive) returns over the past 20 days.

MACD (MACDMR)

We use volatility normalised moving average convergence divergence (MACD) indicators based on Baz et al. (2015) in place of the sign of returns for estimating the trading signal:

$$\begin{aligned} \text{MACD}(i, t, S, L) &= m(i, t, S) - m(i, t, L) \\ \text{MACD}_{\text{norm}}(i, t, S, L) &= \frac{\text{MACD}(i, t, S, L)}{\text{std}(p_{t-5:t})} \\ Y_t^{(i, \text{straddle})} &= \frac{\text{MACD}_{\text{norm}}(i, t, S, L)}{\text{std}(\text{MACD}_{\text{norm}}(i, t-20:t, S, L))} \end{aligned}$$

$$(3) \quad X_t^{(i, \text{straddle})} = \frac{1}{3} \sum_{k=1}^3 \phi(Y_t^{(i, \text{straddle})})(S_k, L_k)$$

where $\text{MACD}(i, t, S, L)$ denotes the MACD value of the straddle option of stock i at day t with a short time scale S and long time scale L . $m(i, t, j)$ is defined as the exponentially weighted moving average of the straddle's prices at day t , with a time scale j corresponding to a half-life of $HL = \log(0.5) / \log(1 - 1/j)$. We combine MACD signals in an equally weighted sum over multiple short and long time scales $S_k \in \{2, 4, 8\}$,

$L_k \in \{8, 16, 32\}$ to construct a position $X_t^{(i, \text{straddle})}$ where $\phi(y) = \frac{y \exp(-\frac{y^2}{4})}{0.89}$. Likewise, we consider both the momentum $X_t^{(i, \text{straddle})}$ (MACD) and mean reversion $-X_t^{(i, \text{straddle})}$ (MACDMR) strategies. We refer the reader to (Baz et al., 2015) for further details on the strategy implementation.

TSHestonMOM (TSHestonMR)

Following Heston et al. (2023), we consider a strategy that constructs the trading signal for the straddle option of stock i on the day of portfolio formation ($t = 0$), holding the position unchanged to expiry:

$$(4) \quad Y_{0:t}^{(i, \text{straddle})} = r_{0-nM, 0}^{(i, \text{straddle})}$$

$$(5) \quad X_{0:t}^{(i, \text{straddle})} = \text{sgn}(Y_{0:t}^{(i, \text{straddle})})$$

where $r_{0-nM, 0}^{(i, \text{straddle})}$ is the average returns for the series of (1-month, ATM, static delta-neutral) straddle options for stock i over a lookback period of n months from the day of portfolio formation. A key distinction between TSMOM and TSHestonMOM is that the momentum signals identified in the latter strategy are more accurately referred to as cross-serial correlations, in that the options used to compute the n -period average returns $r_{0-nM, 0}^{(i, \text{straddle})}$ are different from those used to construct the trading signal for the options to be traded following $t = 0$. In other words for the TSHestonMOM, all types of momentum signals involve past returns of a set of options predicting the future returns on an entirely new set. Similar to (Heston et al., 2023), we consider multiple strategies corresponding to average returns of monthly straddle returns over various lookback periods ranging from $n = 1, 3, 6, 12$ corresponding to monthly, quarterly, semiannual and annual returns. We again consider both the momentum $X_{0:t}^{(i, \text{straddle})}$ (TSHestonMOM) and mean reversion $-X_{0:t}^{(i, \text{straddle})}$ (TSHestonMR) strategies.

CSHestonMOM (CSHestonMR)

Based on the long-short approach of (Heston et al., 2023) and (Jegadeesh and Titman, 1993), we implement a cross-sectional momentum strategy that scores and ranks a stock based on its average straddle returns computed as per TSHestonMOM. On the day of portfolio formation, the strategy utilizes a high-minus-low decile portfolio, taking a maximum long and short position for the top and bottom 10% of ranked stocks respectively and holding the positions to expiry. We first calculate raw momentum scores $Y_{0:t}^{(i, \text{straddle})} = r_{0-nM, 0}^{(i, \text{straddle})}$ according to Equation (4). Then, we rank stocks by their raw momentum scores:

$$(6) \quad X_{0:t}^{(i, \text{straddle})} = \{+1, -1, 0\}$$

where $X_{0:t}^{(i, \text{straddle})} = +1$ for stocks ranked in the top 10% and -1 for the bottom 10% ranked stocks, and 0 otherwise. We examine both the momentum $X_{0:t}^{(i, \text{straddle})}$ (CSHestonMOM) and mean reversion $-X_{0:t}^{(i, \text{straddle})}$ (CSHestonMR) strategies.

5. Deep Learning for Options Trading

5.1. General End-To-End Framework

We frame the problem of generating optimal trading decisions $X_t^{(i, \text{straddle})}$ for a portfolio of options with a model f as an end-to-end framework. Given time t and a straddle option of stock i , we have an input

$\mathbf{u}_t^{(i, \text{straddle})} \in \mathbb{R}^d$ of option features. We learn a model f with trainable parameters θ :

$$(7) \quad X_t^{(i, \text{straddle})} = f(\mathbf{u}_t^{(i, \text{straddle})}; \theta)$$

In this framework, trading signals $X_t^{(i, \text{straddle})}$ are directly computed using the point-in-time snapshot of the option's features, integrating both trend prediction and optimal position sizing within a single function f . Unlike standard supervised learning paradigms, a distinctive characteristic of our end-to-end framework is the unavailability of ground-truth labels for the optimal trading signal $X_t^{(i, \text{straddle})}$ of a given option at any point in time. This necessitates the learning of a non-trivial mapping from an option's features to optimal trading signals. We discuss the choice of architectures and optimization of these models in the following sections.

5.2. Network Architectures

Given the problem of learning a mapping from option features to trading signals, it is not immediately clear which choice of architecture would best suit an end-to-end model, warranting the need to consider multiple options. Taking this into account, we examine various choices of neural networks for f .

Linear

Beginning with the elementary case of a neural network comprising of a single fully connected layer:

$$(8) \quad X_t^{(i, \text{straddle})} = g(\mathbf{W}^\top \mathbf{u}_t^{(i, \text{straddle})} + b)$$

where $\mathbf{W} \in \mathbb{R}^m$, $\mathbf{u}_t^{(i, \text{straddle})} := \mathbf{u}_{t-\tau+1:t}^{(i, \text{straddle})} \in \mathbb{R}^m$ with $m = \tau d$, $b \in \mathbb{R}$ and $g = \tanh$ is the activation function. The model computes a linear combination of the input features prior to the tanh activation function that outputs the trading signal to be within the finite range of $[-1, 1]$. To factor in the immediate temporal history of observations for making predictions, we concatenate features from the past $\tau = 5$ days from time t into a single input vector.

Multilayer Perceptron (MLP)

We include an additional hidden layer to the Linear model, enhancing the model's depth:

$$(9) \quad X_t^{(i, \text{straddle})} = g[\mathbf{W}^{[2] \top} \sigma(\mathbf{W}^{[1] \top} \mathbf{u}_t^{(i, \text{straddle})} + b^{[1]}) + b^{[2]}]$$

with $g = \sigma = \tanh$ corresponding to the activation functions of each layer.

Convolutional Neural Networks (CNN)

Modified for time series data, CNNs have been designed to incorporate causal convolutions that utilize only past information for forecasting (Binkowski et al., 2018), maintaining the autoregressive ordering of temporal features. We consider a 1-D autoregressive CNN:

$$(10) \quad \begin{aligned} \mathbf{h}_t^{(i, \text{straddle})} &= P\sigma[\mathbf{W}_c^{[2]} * \sigma(\mathbf{W}_c^{[1]} * \mathbf{u}_t^{(i, \text{straddle})} + \mathbf{b}_c^{[1]}) + \mathbf{b}_c^{[2]}] \\ X_t^{(i, \text{straddle})} &= g[\mathbf{W}^{[2] \top} \sigma(\mathbf{W}^{[1] \top} \mathbf{h}_t^{(i, \text{straddle})} + b^{[1]}) + b^{[2]}] \end{aligned}$$

where $\mathbf{W}_c^{[l]}$ represent convolutional kernels with bias terms $\mathbf{b}_c^{[l]}$ and activation functions $g = \sigma = \tanh$, and $*$ represents the causal convolution operator. Subsequently, we perform average pooling with P prior to passing the activations \mathbf{h}_t to a fully connected neural network.

Long Short-term Memory (LSTM)

Recurrent neural networks (RNNs) have traditionally found applications in sequence modelling and time series forecasting (Lim and Zohren, 2021). Given the sequential nature of our prediction task, it is natural to consider recurrent architectures. We implement a single layer LSTM model (Hochreiter and Schmidhuber, 1997) that takes in an input sequence of option features $\mathbf{u}_t^{(i, \text{straddle})} := \mathbf{u}_{t-\tau+1:t}^{(i, \text{straddle})} \in \mathbb{R}^m$ with $m = \tau d$ where τ represents the length of a trajectory, and subdivide the time series into trajectories of $\tau = 20$ during backpropagation. We omit the technical equations for the LSTM architecture for brevity.

5.3. Training Details

5.3.1. Loss Function

To facilitate the learning of a non-trivial mapping from option features to optimal trading signals that effectively balances both risk and reward, we directly calibrate the models using the Sharpe ratio (Sharpe, 1994), a risk-adjusted performance metric. Given a set of contemporaneous option features and their respective trading signals $\mathcal{D}_\Omega = \{(\mathbf{u}_t^{(i, \text{straddle})}, X_t^{(i, \text{straddle})} = f(\mathbf{u}_t^{(i, \text{straddle})}; \boldsymbol{\theta}))\}$ with $\Omega = \{(i, t) | i = 1, \dots, N^t, t = 1, \dots, T\}$ denoting all straddle-time pairs, we define the loss $\mathcal{L}_{\text{sharpe}}(\boldsymbol{\theta})$ over \mathcal{D}_Ω as the annualized Sharpe ratio:

$$(11) \quad \mathcal{L}_{\text{sharpe}}(\boldsymbol{\theta}) = - \frac{\frac{1}{|\Omega|} \sum_{\Omega} R_i(t) \times 252}{\left[\frac{1}{|\Omega|} \sum_{\Omega} R_i(t)^2 - \left[\frac{1}{|\Omega|} \sum_{\Omega} R_i(t) \right]^2 \right)^{1/2}}$$

$$(12) \quad R_i(t) = X_t^{(i, \text{straddle})} \left(\frac{\sigma_{\text{tgt}}}{\sigma_t^{(i, \text{straddle})}} \right) r_{t, t+1}^{(i, \text{straddle})}$$

5.3.2. Optimization

Within each in-sample window, we perform a train-validation split with the earlier 90% of data used for calibrating the models and the most recent 10% reserved for validation. To calibrate the models, we perform backpropagation using minibatch stochastic gradient descent with Adam (Kingma and Ba, 2014), and trigger early stopping with a patience of 25 epochs based on the validation loss. In order to select optimal candidates for each machine learning model, we conduct hyperparameter optimization with 100 iterations of random search. We refer the reader to Appendix A for the detailed description of hyperparameter search ranges. Model calibration was performed on a server equipped with an AMD EPYC7713 CPU and multiple NVIDIA L40 GPUs.

6. Performance Evaluation

6.1. Backtest Details

Following an expanding window approach, we train all models with every block of 5 additional years. In each block, we fix the weights and hyperparameters of the trained models and evaluate the models out-of-sample in the following 5-year window. We perform model calibration over multiple seeded runs and present the aggregated out-of-sample results in Section 6.3.

6.2. Option Features

To construct an input of option features $\mathbf{u}_t^{(i, \text{straddle})} \in \mathbb{R}^d$ as described in Equation (7), we include a combination of the predictors used in the strategies outlined in Section 4:

- I. **Normalized Returns** – we use $r_{t-k, t}^{(i, \text{straddle})} / (\sigma_t^{(i, \text{straddle})} k)$, representing straddle returns normalized by daily volatility estimates scaled to a time scale $k \in \{1, 5, 10, 15, 20\}$, which corresponds to daily, weekly, biweekly, triweekly and monthly returns.
- II. **MACD Indicators** – we take volatility normalised MACD signals $Y_t^{(i, \text{straddle})}(S_k, L_k)$ from Equation (3) with short and long time scales $S_k \in \{2, 4, 8\}$ and $L_k \in \{8, 16, 32\}$.
- III. **Option Momentum Features** – we expand our set of predictors to include the momentum features as defined in Equation (4), taking the average returns of straddle options for the stock over look-back periods of $n = 1, 3, 6, 12$ months.
- IV. **Core Option Features** – to facilitate comparability with the trend-based benchmarks outlined in Section 4, and to focus on the predictive power of the features above, we maintain a parsimonious set of core option features which includes the log-moneyness (of both call and put options forming the straddle) and days to expiry (DTE, in years). Given that the moneyness and DTE of a straddle option changes over time, these core contract features are necessary in identifying the option at particular stages of its lifespan. Distinctively, we exclude other features such as option implied volatility or sensitivity measures such as Greeks which would require making an assumption of an underlying option pricing model such as the Black-Scholes or the binomial model. Given our focus on delta-neutral straddles, we also exclude underlying stock characteristics and stock returns from our core set of features.

6.3. Results and Discussion

Table 1. Performance Metrics – Raw Signal Outputs

	E[Return]	Vol.	Downside Deviation	MDD	Sharpe	Sortino	Calmar	Hit Rate	Ave. P Ave. L
Benchmarks									
Long Only	0.055	0.103	0.058	0.188	0.534	0.940	0.291	0.451	1.341
TSMOM	-0.076	0.091	0.077	0.504	-0.827	-0.982	-0.150	0.553	0.693
TSMR	0.057*	0.088	0.049	0.168	0.646	1.160	0.340	0.445	1.406*
MACD	-0.038	0.056	0.047	0.308	-0.671	-0.800	-0.123	0.559	0.697
MACDMR	0.029	0.054	0.030	0.106	0.529	0.945	0.271	0.443	1.388
TSHestonMOM	0.002	0.053	0.035	0.150	0.043	0.065	0.015	0.495	1.027
TSHestonMR	0.023	0.048	0.031	0.085	0.476	0.741	0.266	0.484	1.163
CSHestonMOM	-0.000	0.015	0.010	0.031	-0.005	-0.007	-0.002	0.507	0.973
CSHestonMR	0.009	0.015*	0.010*	0.028*	0.565	0.862	0.301	0.500	1.109
Deep Learning Models									
Linear	0.021	0.018	0.014	0.036	1.258	1.672	0.754	0.561	1.073
MLP	0.016	0.017	0.013	0.034	0.954	1.262	0.504	0.580*	0.939
CNN	0.009	0.016	0.012	0.035	0.534	0.717	0.240	0.507	1.199
LSTM	0.026	0.019	0.014	0.031	1.399*	1.917*	0.850*	0.557	1.231

(bold and * denotes best performing strategy for each column)

Table 2. Performance Metrics – Rescaled to Target Volatility

	E[Return]	Vol.	Downside Deviation	MDD	Sharpe	Sortino	Calmar	Hit Rate	Ave. P Ave. L
Benchmarks									
Long Only	0.112	0.160	0.088	0.292	0.697	1.275	0.383	0.451	1.384
TSMR	0.123	0.161	0.087*	0.254	0.762	1.414	0.484	0.445	1.441*
MACDMR	0.106	0.162	0.088	0.264	0.655	1.207	0.402	0.443	1.424
TSHestonMR	0.098	0.157*	0.100	0.191*	0.626	0.984	0.514	0.484	1.193
CSHestonMR	0.091	0.159	0.104	0.318	0.573	0.878	0.286	0.500	1.112
Deep Learning Models									
Linear	0.245	0.190	0.146	0.276	1.290	1.690	0.917	0.561	1.094
MLP	0.167	0.186	0.141	0.344	0.895	1.199	0.523	0.580*	0.929
CNN	0.119	0.217	0.163	0.435	0.551	0.736	0.284	0.507	1.216
LSTM	0.266*	0.200	0.144	0.274	1.329*	1.862*	0.974*	0.557	1.212

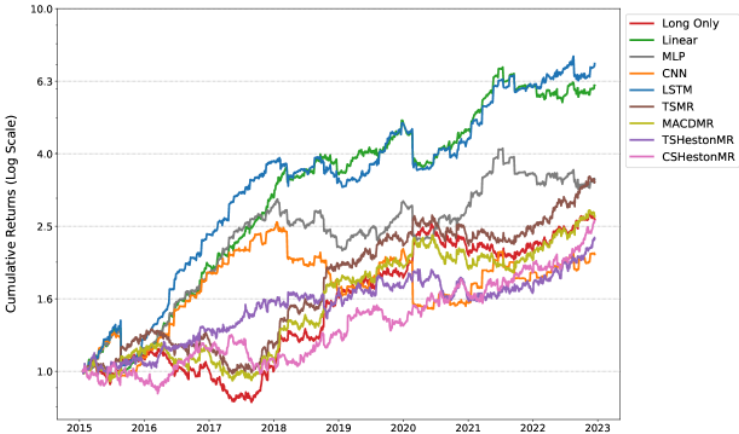


Figure 1. Cumulative Returns - Rescaled to Target Volatility

\Description

[Plot showing Cumulative Returns - Rescaled to Target Volatility]Plot showing Cumulative Returns - Rescaled to Target Volatility, LSTM and Linear models exhibit highest cumulative returns

Table 3. Impact of Transactions Costs on Sharpe Ratio – Rescaled to Target Volatility

Transaction Costs (bps)	0.0	1.0	2.0	3.0	5.0	10.0	20.0	50.0
Benchmarks								
Long Only	0.697	0.695	0.692	0.690	0.685	0.673	0.650	0.578
TSMR	0.762	0.757	0.752	0.747	0.737	0.713	0.663	0.514
MACDMR	0.655	0.646	0.638	0.629	0.612	0.568	0.481	0.219
TSHestonMR	0.626	0.621	0.615	0.610	0.600	0.574	0.522	0.367
CSHestonMR	0.573	0.570	0.567	0.563	0.557	0.541	0.508	0.410
Deep Learning Models								
LSTM	1.329*	1.310*	1.291*	1.272*	1.235*	1.140	0.952	0.388
LSTM + TC Reg.	1.282	1.270	1.259	1.247	1.223	1.164*	1.045*	0.689*

We use the following annualized metrics to evaluate the out-of-sample performance of all strategies:

- I. Profitability Measures – Expected Returns ($\mathbb{E}[\text{Returns}]$), Hit Rate
- II. Risk Measures – Volatility (Vol.), Downside Deviation, Maximum Drawdown (MDD)
- III. Performance Ratios – Sharpe, Sortino and Calmar Ratios, Average Profit over Loss $\left(\frac{\text{Ave. P}}{\text{Ave. L}}\right)$

We present the aggregated out-of-sample performance metrics of all strategies computed using the over-all returns according to Equation (1). Firstly, we present the performance of all strategies from their raw signal outputs in Table 1. We then apply to all strategies (excluding unprofitable strategies) an additional

layer of volatility scaling to target an annualized volatility of 15% at the portfolio level and report the performance in Table 2 and plot their cumulative returns in Figure 1. This adjustment at the portfolio level facilitates comparison between individual strategy returns in line with our 15% volatility target. For each Heston portfolio (TSHestonMOM, TSHestonMR, CSHestonMOM, CSHestonMR), we report results for the best performing lookback period for the sake of brevity. In this section, we report the performance of all strategies without factoring in transaction costs to evaluate their raw predictive ability. In Section 6.4, we include an analysis of the impact of transaction costs and the effect of turnover regularization.

From Table 1, we find that the Long Only straddle portfolio was a profitable strategy over the backtest period. We find this observation interesting, running contrary to (Hu et al., 2023) who show that retail and institutional investors executing short volatility strategies tend to perform well. Furthermore, a Long Only options portfolio would typically benefit from limited downside exposures as opposed to Short Only. Turning our attention to trend-based strategies, we observe that mean-reversion portfolios (TSMR, MACDMR, TSHestonMR, CSHestonMR) exhibit positive performances compared to their opposite momentum counterparts (TSMOM, MACD, TSHestonMOM, CSHestonMOM), which were generally unprofitable over the backtest period. We note that both TSMR and CSHestonMR exhibited only slight performance improvements over Long Only. We obtain similar results for the Heston portfolios as reported in (Heston et al., 2023), who document significant reversals in option returns at short-term horizons, and our best performing Heston models correspond mostly with strategies adopting the shortest $n = 1$ month lookback period. Moving on to our deep learning models, apart from the CNN, we observe that the Linear, MLP and LSTM exhibit a clear disparity in performance above all other strategies as seen from their higher performance ratios.

Referring to Table 2, with the implementation of volatility targeting at the portfolio level, we observe modest improvements in performance across all profitable benchmarks. Portfolio volatility targeting resulted in minimal changes to the deep learning models, allowing them to retain a large gap in their performance ratios over the benchmarks. In particular, we see that the Linear and LSTM models exhibited the best performances, outperforming the benchmarks by roughly twice in their Sharpe ratios of 1.290 and 1.329 respectively. We note that the simplest Linear model was able to perform the MLP, most likely due to our introduction of an L1 regularization penalty only for the Linear model during training. Examining the period during the COVID-19 market selloff, we see clearly from Figure 1 that Long Only exhibited sharp gains during the COVID-19 market selloff at the start of 2020, demonstrating the profitability of long straddles during periods of high market volatility. While we observe that the deep learning models experienced brief drawdowns during the selloff, performance of the models swiftly recovered following the market rebound.

6.4. Transaction Costs and Turnover Regularization

Some of the key challenges of rebalancing an options portfolio include market microstructure considerations arising from market liquidity and bid-ask spreads, which can result in high transaction costs. To examine the impact of transaction costs on the profitability of all strategies, we first compute Sharpe ratios adjusted for transaction costs by taking into account turnover-adjusted returns:

$$(13) \quad \tilde{r}_{t,t+1}^{\text{STRATEGY}} = \frac{1}{N_t} \sum_{i=1}^{N_t} \left(R_i(t) - c \cdot \sigma_{\text{tgt}} \left| \frac{X_t^{(i, \text{straddle})}}{\sigma_t^{(i, \text{straddle})}} - \frac{X_{t-1}^{(i, \text{straddle})}}{\sigma_{t-1}^{(i, \text{straddle})}} \right| \right)$$

where c represents a measure of average transaction costs in basis points. Focusing on the LSTM model, we observe from Table 3 that the model maintains superior risk-adjusted performance over the best performing benchmark - TSMR up to transaction costs of $c = 20$ bps, deteriorating at higher transaction costs of $c = 50$ bps.

Following the methodology detailed in (Lim et al., 2019; Tan et al., 2023), we modify the training loss to utilize turnover-adjusted returns as defined in Equation (13), which in effect results in optimizing the Sharpe ratio while regularizing for the turnover generated by the trading signals of the LSTM. Based on Table 3, we see that turnover regularization further enhances the performance of the LSTM for high transaction costs of $c = 10$ to 50 bps, allowing the regularized model to outperform other strategies at prohibitively high levels of transaction costs.

7. Conclusions

We present a general end-to-end framework for trading options using a highly data-driven machine learning algorithm, adopting the point of view of an active investor seeking to profit from options trading. Departing from conventional approaches that typically rely on specific market dynamics or option pricing models, we train end-to-end neural networks to directly learn mappings from options data to optimal trading positions, removing the need to simulate market processes, price options or predict option returns. Backtesting our approach on portfolios of delta-neutral equity options, our models demonstrate significant improvements in risk-adjusted performance when calibrated using the Sharpe ratio. Crucially, our framework is agnostic to specific underlying market assumptions, potentially allowing for further extensions to a broader set of derivatives or complex instruments where data is available.

Acknowledgements.

We would like to thank the Oxford-Man Institute of Quantitative Finance for providing compute resources. Wee Ling Tan thanks Bryan Lim, Martin Luk, Guillaume Andrieux and Hans Buehler for their insightful comments.

References

- (1) Bali et al. (2023) Turan G Bali, Heiner Beckmeyer, Mathis Moerke, and Florian Weigert. 2023. Option Return Predictability with Machine Learning and Big Data.

The Review of Financial Studies 36, 9 (2023), 3548–3602.

<https://doi.org/10.1093/rfs/hhad017>

Baz et al. (2015)

Jamil Baz, Nicolas Granger, Campbell R Harvey, Nicolas Le Roux, and Sandy Rattray. 2015. Dissecting Investment Strategies in the Cross Section and Time Series. *SSRN 2695101* (2015).

Binkowski et al. (2018)

Mikolaj Binkowski, Gautier Marti, and Philippe Donnat. 2018. Autoregressive Convolutional Neural Networks for Asynchronous Time Series. In *International Conference on Machine Learning*. PMLR, 580–589.

Black (1975)

Fischer Black. 1975. Fact and Fantasy in the Use of Options. *Financial Analysts Journal* 31, 4 (1975), 36–41. <https://doi.org/10.2469/faj.v31.n4.36>

Black and Scholes (1973)

Fischer Black and Myron Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81, 3 (1973), 637–654. <http://www.jstor.org/stable/1831029>

Büchner and Kelly (2022)

Matthias Büchner and Bryan Kelly. 2022. A Factor Model for Option Returns. *Journal of Financial Economics* 143, 3 (2022), 1140–1161. <https://doi.org/10.1016/j.jfineco.2021.11>

Buehler et al. (2019)

Hans Buehler, Lukas Gonon, Josef Teichmann, and Ben Wood. 2019. Deep Hedging. *Quantitative Finance* 19, 8 (2019), 1271–1291. <https://doi.org/10.1080/14697688.2019.15>

Corporation (2024)

Options Clearing Corporation. 2024. OCC - Historical Volume Statistics. <https://www.theocc.com/market-data/market-data-reports/volume-and-open-interest/historical-volume-statistics>.

- Coval and Shumway (2001) Joshua D Coval and Tyler Shumway. 2001. Expected Option Returns. *The Journal of Finance* 56, 3 (2001), 983–1009. <https://doi.org/10.1111/0022-1082.00352>
- Cox et al. (1979) John C Cox, Stephen A Ross, and Mark Rubinstein. 1979. Option Pricing: A Simplified Approach. *Journal of Financial Economics* 7, 3 (1979), 229–263. [https://doi.org/10.1016/0304-405X\(79\)90015-1](https://doi.org/10.1016/0304-405X(79)90015-1)
- De Bondt and Thaler (1985) Werner FM De Bondt and Richard Thaler. 1985. Does the Stock Market Overreact? *The Journal of Finance* 40, 3 (1985), 793–805. <https://doi.org/10.2307/2327804>
- Eisdorfer et al. (2022) Assaf Eisdorfer, Ronnie Sadka, and Alexei Zhdanov. 2022. Maturity Driven Mispricing of Options. *Journal of Financial and Quantitative Analysis* 57, 2 (2022), 514–542. <https://doi.org/10.1017/S002210902100003>
- Goyal and Saretto (2009) Amit Goyal and Alessio Saretto. 2009. Cross-section of Option Returns and Volatility. *Journal of Financial Economics* 94, 2 (2009), 310–326. <https://doi.org/10.1016/j.jfineco.2009.01.001>
- Harvey et al. (2018) Campbell R Harvey, Edward Hoyle, Russell Korgaonkar, Sandy Rattray, Matthew Sargaison, and Otto Van Hemert. 2018. The Impact of Volatility Targeting. *The Journal of Portfolio Management* 45, 1 (2018), 14–33. <https://doi.org/10.3905/jpm.2018.45.1.014>
- Heston et al. (2023) Steven L Heston, Christopher S Jones, Mehdi Khorram, Shuaiqi Li, and Haitao Mo. 2023. Option Momentum. *The Journal of Finance* 78, 6 (2023), 3141–3192. <https://doi.org/10.1111/jofi.13279>
- Hochreiter and Schmidhuber (1997) Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory.

Neural Computation 9, 8 (1997), 1735–1780.
<https://doi.org/10.1162/neco.1997.9.8.1735>

Hu et al. (2023)

Jianfeng Hu, Antonia Kirilova, Seongkyu Park, and Doojin Ryu. 2023.
 Who Profits from Trading Options?
Management Science 70, 7 (2023), 4742–4761.
<https://doi.org/10.1287/mnsc.2023.4916>

Hutchinson et al. (1994)

James M Hutchinson, Andrew W Lo, and Tomaso Poggio. 1994.
 A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks.
The Journal of Finance 49, 3 (1994), 851–889.
<https://doi.org/10.2307/2329209>

Ivaşcu (2021)

Codruţ-Florin Ivaşcu. 2021.
 Option Pricing using Machine Learning.
Expert Systems with Applications 163 (2021), 113799.
<https://doi.org/10.1016/j.eswa.2020.113799>

Jegadeesh and Titman (1993)

Narasimhan Jegadeesh and Sheridan Titman. 1993.
 Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency.
The Journal of Finance 48, 1 (1993), 65–91.
<https://doi.org/10.1111/j.1540-6261.1993.tb04702.x>

Kingma and Ba (2014)

Diederik P Kingma and Jimmy Ba. 2014.
 Adam: A Method for Stochastic Optimization.
arXiv:1412.6980 (2014).

Kolm and Ritter (2019)

Petter N Kolm and Gordon Ritter. 2019.
 Dynamic Replication and Hedging: A Reinforcement Learning Approach.
The Journal of Financial Data Science 1, 1 (2019), 159–171.
<https://doi.org/10.3905/jfds.2019.1.1.159>

Lim and Zohren (2021)

Bryan Lim and Stefan Zohren. 2021.
 Time Series Forecasting With Deep Learning: A Survey.
Philosophical Transactions of the Royal Society A 379, 2194 (2021), 20200209.
<https://doi.org/10.1098/rsta.2020.0209>


- Lim et al. (2019)
Bryan Lim, Stefan Zohren, and Stephen Roberts. 2019.
Enhancing Time-Series Momentum Strategies Using Deep Neural Networks.
The Journal of Financial Data Science 1, 4 (2019), 19–38.
<https://doi.org/10.3905/jfds.2019.1.015>
- Merton (1973)
Robert C Merton. 1973.
Theory of Rational Option Pricing.
The Bell Journal of Economics and Management Science 4, 1 (1973), 141–183.
<https://doi.org/10.2307/3003143>
- Moskowitz et al. (2012)
Tobias J Moskowitz, Yao Hua Ooi, and Lasse Heje Pedersen. 2012.
Time Series Momentum.
Journal of Financial Economics 104, 2 (2012), 228–250.
<https://doi.org/10.1016/j.jfineco.2011.1>
- Poh et al. (2021)
Daniel Poh, Bryan Lim, Stefan Zohren, and Stephen Roberts. 2021.
Building Cross-Sectional Systematic Strategies by Learning to Rank.
The Journal of Financial Data Science 3, 2 (2021), 70–86.
<https://doi.org/10.3905/jfds.2021.1.060>
- Poh et al. (2022)
Daniel Poh, Bryan Lim, Stefan Zohren, and Stephen Roberts. 2022.
Enhancing Cross-Sectional Currency Strategies by Context-Aware Learning to Rank with Self-Attention.
The Journal of Financial Data Science 4, 3 (2022), 89–107.
<https://doi.org/10.3905/jfds.2022.1.099>
- Poterba and Summers (1988)
James M Poterba and Lawrence H Summers. 1988.
Mean Reversion in Stock Prices: Evidence and Implications.
Journal of Financial Economics 22, 1 (1988), 27–59.
[https://doi.org/10.1016/0304-405X\(88\)90021-9](https://doi.org/10.1016/0304-405X(88)90021-9)

- Rouwenhorst (1998) K Geert Rouwenhorst. 1998. International Momentum Strategies. *The Journal of Finance* 53, 1 (1998), 267–284.
- Sharpe (1994) William F Sharpe. 1994. The Sharpe Ratio. *The Journal of Portfolio Management* 21, 1 (1994), 49–58.
<https://doi.org/10.3905/jpm.1994.409501>
- Tan et al. (2023) Wee Ling Tan, Stephen Roberts, and Stefan Zohren. 2023. Spatio-Temporal Momentum: Jointly Learning Time-Series and Cross-Sectional Strategies. *The Journal of Financial Data Science* 5, 3 (2023), 107–129.
<https://doi.org/10.3905/jfds.2023.1.130>
- Tian and Wu (2023) Meng Tian and Liuren Wu. 2023. Limits of Arbitrage and Primary Risk-Taking in Derivative Securities. *The Review of Asset Pricing Studies* 13, 3 (2023), 405–439.
<https://doi.org/10.1093/rapstu/raad003>
- Vasquez (2017) Aurelio Vasquez. 2017. Equity Volatility Term Structures and the Cross Section of Option Returns. *Journal of Financial and Quantitative Analysis* 52, 6 (2017), 2727–2754.
<https://doi.org/10.1017/S002210901700076>
- Wood et al. (2023) Kieran Wood, Sven Giegerich, Stephen Roberts, and Stefan Zohren. 2023. Trading with the Momentum Transformer: An Intelligent and Interpretable Architecture. *arXiv:2112.08534, Risk* (2023).
- Wood et al. (2022) Kieran Wood, Stephen Roberts, and Stefan Zohren. 2022. Slow Momentum with Fast Reversion: A Trading Strategy Using Deep Learning and Changepoint Detection. *The Journal of Financial Data Science* 4, 1 (2022), 111–129.
<https://doi.org/10.3905/jfds.2021.1.081>

Appendix A Hyperparameter Optimization

Table 4. Hyperparameter Search Range

Hyperparameters	Search Grid
Minibatch Size	32, 64, 128, 256
Dropout Rate	0.1, 0.2, 0.3, 0.4, 0.5
Hidden Layer Size	5, 10, 20, 40, 80, 160
Learning Rate	10^{-5} , 10^{-4} , 10^{-3} , 10^{-2} , 10^{-1} , 10^0
Max Gradient Norm	10^{-4} , 10^{-3} , 10^{-2} , 10^{-1} , 10^0 , 10^1



Feeling lucky?

Conversion report (E)

Report an issue

View original on arXiv