Erasmus University Rotterdam
Erasmus School of Economics
Master Thesis Quantitative Finance

Modeling equity option Implied Volatility Surfaces using nonparametrically corrected parametric models

Name Student: Finn van der Meer
Student ID Number: 504905
Supervisor: dr. G. Freire
Second Assessor: dr. M. Grith
Date Final Version: August 11, 2024

Abstract

This paper explores the application of neural networks to correct implied volatility surface (IVS) models, focusing on equity options rather than the more commonly studied index options. Better predictions enable market participants to make more informed financial decisions while providing deeper insights into the dynamics of underlying stock prices. The study employs neural networks to model error terms in traditional parametric IVS models, integrating machine learning techniques with the economic principles underlying these models. The neural network correction method yields significantly lower Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) values across long-term forecasting and daily interpolation and prediction tasks. These results highlight the potential of this approach for both market practitioners seeking precise implied volatility estimates, as well as for researchers aiming to develop robust models to capture IVS dynamics. Furthermore, this approach is not limited to financial applications, as it can be applied to studies in other disciplines aiming to enhance forecasting accuracy.

Contents

1 Introduction

Option prices play an important role in financial markets, because they are influenced by many different factors such as expectations on underlying assets and the market as a whole. The price of an option is closely related to its Implied Volatility (IV). The IV of an option varies over time to maturity and moneyness, defined as the price of the underlying asset divided by the strike price, forming an Implied Volatility Surface (IVS). The predictability in IVS dynamics is well-documented in the literature (Bernales & Guidolin, 2014; Cont & Da Fonseca, 2002; Dumas et al., 1998). This predictability offers valuable insights into market expectations of future volatility, encompassing crucial information about market sentiment, risk, and the behavior of underlying assets. This allows traders and investors to make better informed decisions, leading to more effective speculative or hedging strategies. In addition to providing practical market insights, modeling IVS levels holds significant scientific importance, as evidenced by the ongoing research and interest in IV since its introduction by Black and Scholes (1973).

Despite extensive research efforts, conventional parametric IVS models can fall short due to their inability to capture the complex, non-linear IVS dynamics. This has led to the exploration of advanced computational techniques to model these variables. The recent surge in popularity of machine learning methods has greatly improved our understanding of these techniques, making them an excellent candidate to model non-linear IVS dynamics. While neural networks can be utilized to directly model variables of interest, they can also be used to model the error terms of traditional IVS models (Almeida et al., 2023). This corrective procedure aims to improve predictive power while keeping the original framework of the parametric models intact. Further investigation of such methods contributes to both theoretical and practical advancements in financial modeling.

In this paper, I investigate whether the nonparametric correction of parametric models lead to improved equity option IVS predictions. This analysis focuses on IVS levels of options with stocks as their underlying assets. Despite the robustness of parametric models, they can fail to capture complex, non-linear patterns present in financial data (Hutchinson et al., 1994). This research aims to determine to what extend the integration of neural networks as nonparametric corrections can address this issue. This results in the following research

question:

> *To what extend can neural network corrected models increase the forecasting accuracy of parametric equity option IVS models?.*

Equity option data from Wharton Research Data Services, n.d. is used to answer this research question. The data consists of approximately 34 million observations of options, spanning a five-year period from 2018 to 2023. Various methods are employed in this research. Initially, different parametric models are used to make IVS predictions. These models include the baseline model introduced by Black and Scholes (1973), as well as more sophisticated models presented in the works of Dumas et al. (1998) and Carr and Wu (2016). Each of these models has a neural network corrected (NNC) counterpart, where the errors terms are modeled using a neural network. Three distinct prediction exercises are conducted to assess the viability of these models for different applications. The first exercise involves daily interpolation, mimicking how a market practitioner would employ an IVS model. The second exercise focuses on predictions for one day, one week and one month ahead, evaluating the robustness and predictive accuracy of the different models. Lastly, the long-term predictions of the models are assessed to determine whether the NNC correction provides long-term stability. The models are evaluated using standard performance measures, alongside an assessment of their model-implied moments. This allows for a comprehensive assessment of the models' ability to capture the dynamics of the IVS and the options' underlying asset returns.

The majority of IVS research has focused on IV values of options with an index as their underlying asset. Consequently, IV models for options with equities as their underlying assets remain relatively underexplored. Additionally, there is reason to believe that equity option prices exhibit greater volatility compared to index option prices, given that individual stocks tend to be more volatile than broad market indices. As a result, the IV values of equity options are expected to reflect this high price volatility, enriching them with valuable information. In this paper, I aim to fill the gap in the literature by focusing on IVS models for equity options to provide deeper insights into their dynamics. Understanding these nuances can lead to more accurate pricing models for both options and the underlying equity assets.

I find that the NNC models produce predictions with significantly lower prediction errors compared to their original counterparts. In the daily interpolation exercise, the average decrease in RMSE and MAE across all different models is 35%. For the daily one day, one week and one month ahead predictions, the average prediction error decreases by 22%, with the largest improvements observed for the shorter horizons. For longer-term prediction extending up to two years, the NNC models also show a decrease of the average prediction error, with an average reduction of 22%. Diebold-Mariano test results confirm that all these decreases are statistically significant. Additionally, model performance is evaluated based on the options' underlying assets. The results indicate that the reduction in error metrics is consistent across options for all underlying assets, and not specifically linked to the liquidity of options.

This paper presents a method to estimate IV values more accurately, with these findings extending beyond the scope of option pricing. They demonstrate that incorporating a neural network can enhance predictive accuracy while maintaining the setup of the parametric models. This is particularly valuable for models based on theoretical assumptions or empirical observations. The hybrid approach preserves the theoretical foundations of parametric models while leveraging the flexibility and predictive power of neural networks. The random nature of errors in statistical models make them particularly suitable to model using neural networks, which can be criticized in certain applications for their lack of transparency. The broad applicability of this approach suggests potential improvements in predictability across other fields of research where similar prediction challenges exist. Consequently, this research not only advances the field of financial modeling but also offers valuable insights for other research areas interested in enhancing predictive models through the integration of machine learning techniques.

This remainder of this paper is structured as follows. Section 2 provides an overview of the relevant literature on IVS modeling. Section 3 details the data used in the study, after which section 4 describes the various models and evaluation methods used in this research. The results are presented in section 5. Finally, section 6 concludes the paper and offers suggestions for further research.

2 Literature Review

Implied volatility was first introduced through the option pricing formula of Black and Scholes (1973), where it serves as the volatility parameter that returns market observed option prices. The Black-Scholes (BS) formula establishes a direct relationship between option prices and IV values, which has encouraged extensive research into IVS models. Although this formula laid the foundation of the option pricing literature, some of its assumptions do not hold in practice. Most notably, the BS framework assumes that IV values are constant across strike prices and maturities. This assumption is often violated in practice, as IV's typically increase for moneyness levels further from one, leading to the well-documented phenomenon known as the "volatility smile" (Derman & Kani, 1994; Rubinstein, 1994). As a result, various models arose that incorporate the volatility smile into IVS modeling. Dumas et al. (1998) and Gonçalves and Guidolin (2006) introduce frameworks that model the IV of an option based on factors such as the option's moneyness and time to maturity. The direct modeling of IV values allows for non-constant IV values, dropping the assumption of constant volatility. More recent models, such as those by Carr and Wu (2016) and Ulrich et al. (2023), also propose parametric models to directly estimate IV values.

Recently, nonparametric methods have gained prominence in option pricing literature. Among these, machine learning techniques such as neural networks have emerged as effective tools to conduct complex research in the field of finance. Hutchinson et al. (1994) demonstrate that neural networks provide good results in option pricing, largely due to their ability to capture non-linear relationships. This notion is supported by the findings of Bali et al. (2021), who find that allowing for non-linearities in option pricing models significantly enhances performance. These promising results extend to the IVS modeling literature, aligning with the finding that IVS values also exhibit non-linear dynamics (Andersen et al., 2015). Cao et al. (2020) and Hamid and Iqbal (2004) show that neural networks are able to accurately predict IV values directly. These findings underscore the potential of neural networks to directly model IVS levels.

However, recent advancements in research have revealed that neural networks can also be used to enhance the performance of conventional models. Almeida et al. (2023) demonstrate that applying neural networks to model the error terms of existing models enhances the pre-

diction accuracy of index option IVS levels. This framework of nonparametrically correcting parametric models will be extended in this research to predict IVS levels for equity options rather than index options.

The majority of IVS research has concentrated on index IVS levels due to their substantial trading volumes. However, equity options, which also represent a significant segment of financial markets, remain relatively underexplored in the literature. Given that equity options are the second most exchange traded derivative product (World Federation of Exchanges (WFE) Full Year 2023 Market Highlights), expanding IVS research towards equity options can provide valuable insights. Equity IVS modeling is not completely uncharted territory, as evidenced by Ulrich et al. (2023) who successfully model IV values of options with stocks as underlying values. Moreover, the extensive index IVS literature can prove to be helpful in equity option IVS modeling, as Bernales and Guidolin (2014) demonstrate that IVS levels of options on the S&P 500 index can help predict equity option IVS levels.

The IVS models proposed in the literature can be employed to predict an option's IV value, which can subsequently be converted into option prices. Breeden and Litzenberger (1978) show how option prices can be utilized to derive probability distributions of the returns of the underlying assets, offering valuable insights into the movement of the underlying asset. Extending this research, Bakshi et al. (2003) demonstrate a method where option prices are used to construct moments of the risk-neutral distribution of the underlying asset's returns. More recent studies also focus on extracting moments from IVS models (François et al., 2022), highlighting the potential of using option-derived moments to gain more information about the return distributions of underlying assets. Therefore, models that more accurately predict IVS levels, and thereby option prices, are extremely valuable as they enhance the ability to derive a deeper understanding of asset behaviour, improving the ability to make informed financial decisions.

3 Data

To conduct the research on IVS models for equity options, I use equity option data from OptionMetrics via the Wharton Research Data Services database (OptionMetrics, 2024). The dataset spans a five-year period from February 28, 2018 to February 28, 2023. This time period includes high volatility periods like the COVID-19 pandemic, as well as more stable periods. This variety allows the models to be trained and tested under diverse market conditions. The dataset consists of approximately 34 million observations of options with different underlying assets. Every observation contains information on expiry dates, strike prices, best bid and ask prices, and IV values. Option prices are constructed as the average of the best bid and ask prices. Additionally, underlying asset prices are included to calculate the moneyness of the options.

The options in this dataset are written on different stocks. These stocks are ranked based on the number of options written on them, and the top fifty are selected to ensure that the most liquid options are used in this research. The fifty selected stocks and the exact amount of options written on them can be found in table 6.1 in the appendix. This selection strikes a balance between giving a broad overview of the market and ensuring that the quality of the options included is sufficient, as illiquid options can cause significant prediction challenges. The decision to focus on a subset of underlying assets is influenced by the research objective, which prioritizes gaining insights into equity option IVS dynamics over reducing dataset dimensions.

The option data for the fifty most traded underlying assets is filtered before being used in the models. Given that equity options are less liquid than index options (Bernales & Guidolin, 2014), using filters from existing literature on index IVS modeling would result in an insufficiently small sample. As a result, less restrictive filtering criteria are utilized to maintain a sufficiently large sample for meaningful model estimation and prediction. Options with zero trading volume are not excluded, as the bid and ask quotes on non-trading days still provide useful information that should be captured in the models (Bernales & Guidolin, 2014). However, options with zero open interest are excluded to eliminate options with no liquidity that offer little useful information (Goyal & Saretto, 2009). Options with a bid price of zero are also omitted, as they do not reflect realistic market conditions.

Table 3.1: Underlying Assets Statistics

Ticker	Amount	Mean IV	St. Dev.	Ticker	Amount	Mean IV	St. Dev.
AMZN	2,734,828	0.385	0.136	**AZO**	503,835	0.328	0.110
TSLA	1,739,849	0.703	0.222	**SQ**	501,233	0.607	0.198
GOOGL	1,639,177	0.330	0.102	**NOW**	498,366	0.446	0.133
GOOG	1,502,404	0.329	0.104	**REGN**	496,257	0.374	0.101
BKNG	1,434,118	0.373	0.143	**BIIB**	493,773	0.435	0.168
SHOP	1,026,193	0.609	0.190	**LMT**	490,139	0.285	0.100
NFLX	1,020,231	0.465	0.144	**CRM**	486,838	0.399	0.126
CMG	974,942	0.390	0.138	**V**	483,448	0.310	0.113
NVDA	912,362	0.498	0.132	**ULTA**	482,667	0.435	0.166
META	771,065	0.403	0.125	**PYPL**	470,100	0.428	0.147
BA	757,149	0.447	0.211	**TWLO**	469,647	0.617	0.200
AAPL	734,189	0.353	0.115	**AMD**	457,406	0.549	0.129
ADBE	728,721	0.392	0.134	**BLK**	447,588	0.320	0.115
MSFT	692,378	0.321	0.099	**LULU**	444,095	0.459	0.174
BABA	683,526	0.452	0.143	**RH**	436,787	0.600	0.196
MELI	656,538	0.563	0.166	**W**	417,768	0.738	0.279
LRCX	646,893	0.456	0.127	**PANW**	413,141	0.409	0.132
ROKU	593,691	0.722	0.210	**MDB**	399,738	0.653	0.193
ISRG	588,171	0.371	0.124	**CHTR**	396,929	0.341	0.112
COST	562,737	0.302	0.118	**NOC**	388,337	0.302	0.095
GS	561,009	0.344	0.130	**NTES**	367,503	0.446	0.117
AVGO	548,645	0.372	0.117	**MSTR**	367,227	0.958	0.309
MA	538,827	0.342	0.124	**ILMN**	366,757	0.432	0.137
TTD	522,596	0.663	0.194	**ALGN**	363,472	0.511	0.175
HD	504,415	0.311	0.119	**HUM**	352,310	0.328	0.106

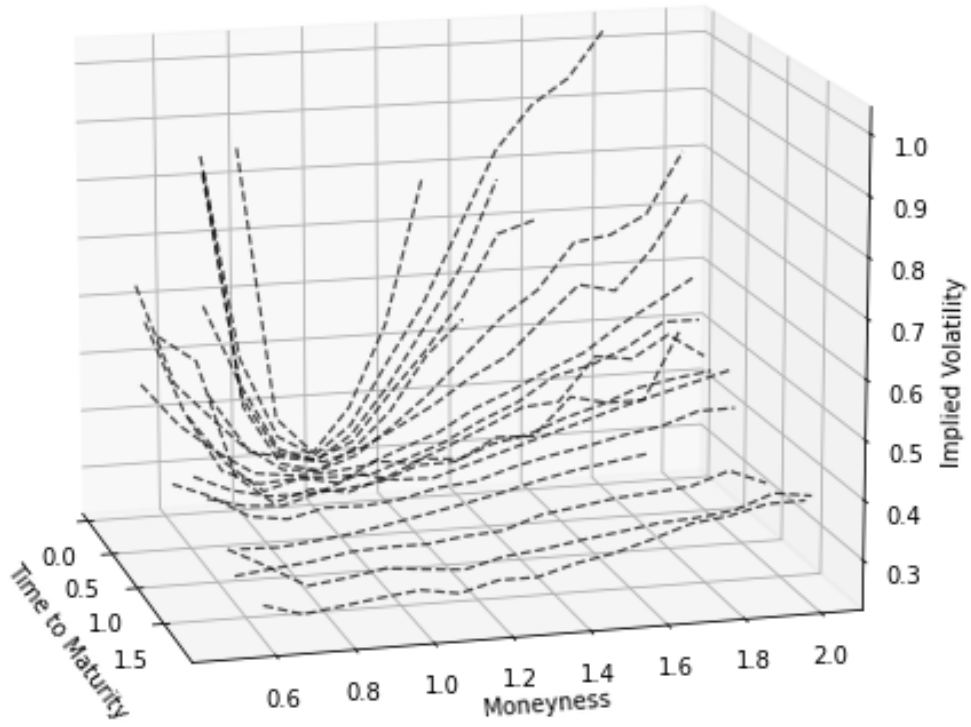
Notes: Statistics of the different underlying assets in the dataset after filtering. Per underlying asset (Ticker), the amount of observations, the mean Implied Volatility (IV) and the standard deviation (St. Dev.) is shown.

The final filters are based on the moneyness of the options. Only out-of-the-money (OTM) options are retained because the put-call parity ensures that the information of a call option is mirrored in a put option with the same strike price and time to maturity (Klemkosky & Resnick, 1979). Following the design of Almeida et al. (2023) and Andersen et al. (2015), only OTM options are used, as they contain more information because they are traded more often than in-the-money (ITM) options (Aït-Sahalia & Lo, 1998). Consequently, puts with moneyness greater than one and calls with moneyness less than one are included. Outliers with extreme moneyness values are also filtered out of the sample. Several studies on index option IVS only keep options with moneyness levels close to one (Dumas et al., 1998; Heston & Nandi, 2000). However, this study opts for more lenient moneyness bounds of $[0.5, 2]$ to maximize the trade off between maintaining a large sample size and ensuring moneyness levels remain reasonably close to one. These less strict bounds also acknowledge the higher volatility equity options possess compared to index options.

Figure 1 provides deeper insights into the filtered data by visualising the IVS of options on the first day of the sample. The graph is created by averaging the observed IV values across the fifty different underlying assets, offering a broad indication of the relation between moneyness levels and IV values. This average IV is plotted against moneyness levels for different maturities, with each dotted line representing a different time to maturity. This figure clearly illustrates the presence of the implied volatility smile in the data, where moneyness values near one exhibit lower IV values, while other moneyness values show a higher IV. This effect is particularly pronounced for shorter smaller maturities and diminishes for longer maturities.

As a result of the applied filters, the remaining dataset consists of 34.070.015 observations over a five-year period. Table 3.1 presents the summary statistics of options on the different underlying assets. Within this dataset, Amazon.com, Inc. (AMZN) is the stock with the highest number of options written on it, followed by Tesla, Inc. (TSLA) and Alphabet Inc. (GOOGL and GOOG). It is important to note the distinction between the tickers GOOGL and GOOG, as the former grants voting rights, whereas the latter does not, leading to a small price difference that justifies their separate inclusion in this study.

Figure 1: First Day Implied Volatility Surface



Notes: Implied Volatility values plotted against moneyness, taking the average implied volatility across the underlying assets. Every dotted line corresponds to a single time to maturity. The data is from the first day of the sample (2018-02-28)

The underlying assets of the options are diverse, spanning different sectors. AMZN is the stock on which the most options are written, and it falls within the consumer discretionary sector. In contrast, Apple Inc. (AAPL) and Microsoft Corporation (MSFT) are categorized under the information technology sector. The Boeing Company (BA) is an example of a stock in the industrial sector, while Biogen Inc. (BIIB) and The Goldman Sachs Group, Inc. (GS) originate from the healthcare and financial sectors, respectively. An examination of the properties of options on the different stocks, illustrated in table 3.1, reveals considerable variation in mean and standard deviation of the IV values across the different underlying assets. Among the most traded options, TSLA stands out with a high mean IV of 0.703 and a standard deviation of 0.222, underscoring its reputation as a volatile stock. Conversely, options on MSFT exhibit the lowest standard deviation of 0.099. These statistics highlight the diversity within the market, underlining the challenges and insights that can be gained from modeling equity options.

4 Methodology

This paper aims to model implied volatility surfaces (IVS), which are defined as implied volatility (IV) levels of options on a specific underlying asset across different strike prices (K) and times to maturity (τ). On day t, option i with corresponding K and τ has an IV of $\sigma_{i,t}$, and these values across K and τ form the IVS. The Black-Scholes (BS), ad-hoc Black-Scholes (AHBS) and Carr-Wu (CW) models are used as baseline parametric models, yielding IV predictions $\hat{\sigma}_{i,t}^{(M)}$ for option i with corresponding strike price K and time to maturity τ, using the different models $M \in \{BS, AHBS, CW\}$. These baseline models are then enhanced by incorporating a neural network to model the prediction errors, following the approach of Almeida et al. (2023), who show that this method improves prediction accuracy for index options. This section continues with an explanation of the baseline models. This is followed by a thorough explanation of the workings of neural networks, and how they are employed in this research. Finally, the used evaluation methods are introduced.

4.1 Parametric Models

4.1.1 Black-Scholes Model

The BS model, introduced by Black and Scholes (1973), is widely recognized as the foundation of option pricing literature. At the heart of this model is the assumption that a stock price S_t follows a geometric Brownian motion with constant drift and volatility. This is mathematically represented as:

$$\frac{dS_t}{S_t} = \mu d_t + \sigma dW_t, \tag{1}$$

where μ represents the expected rate of return on the stock and σ denotes the volatility that drives the stock's returns. Wiener process W_t introduces randomness into stock price movements, modeling the unpredictable nature of financial markets. The drift term μd_t accounts for the average expected return over time, while the volatility term σdW_t models the random fluctuations around this drift.

By dynamically hedging options with this underlying price process, Black and Scholes (1973) show that, at time t, the price of a European call option with strike K and time to maturity τ can be calculated as:

$$C(S_t, K, \tau, r, \sigma) = \Phi(d_1)S_t - \Phi(d_2)Ke^{(-r\tau)}, \tag{2}$$
$$d_1 = \frac{1}{\sigma\sqrt{\tau}}\left[\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\tau\right],$$
$$d_2 = d_1 - \sigma\sqrt{\tau},$$

where r is the risk-free rate and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. In the BS model, IV values are determined as the value of σ that equates the theoretical price derived from equation 2 to the market-observed option price. However, equation 2 illustrates that σ is independent of the strike price K, the time to maturity τ and the time t, indicating that the model assumes constant volatilities across strike prices and maturities. Consequently, in this research the BS IV predictions for the test set ($\hat{\sigma}_{i,t}^{BS}$), are computed as the average of the IV values observed in the model's training set.

However, this assumption of constant volatility is inconsistent with actual market behaviour, where IV values typically exhibit a 'volatility smile' (Derman & Kani, 1994). This phenomenon is characterized by lower IV values for at-the-money options compared to those that are in-the-money or out-of-the-money options. This caveat suggests that, while the BS model serves as a valuable benchmark, more advanced techniques have the potential to significantly enhance the accuracy of IVS predictions. In this research, the BS model is also used as a benchmark against which more sophisticated approaches are evaluated.

4.1.2 Ad-hoc Black-Scholes Model

The ad-hoc or practitioner Black-Scholes (AHBS) model, as proposed by Dumas et al. (1998), is an extension of the traditional BS model that drops the assumption of constant volatility, allowing IV values to vary across moneyness and time to maturity. The term 'ad-hoc' reflects the fact that the model is tailored to market conditions, rather than derived from a fundamental theoretical model. The title 'practitioner' stems from the assumption that practitioners employ the AHBS model to fit observed IV levels and subsequently use the BS framework to derive option prices.

The AHBS model distinguishes itself from the BS model by allowing IV values to vary with the option's moneyness and time to maturity. This variation is achieved by fitting IV values to these variables through a quadratic equation. Specifically, the IV of option i at time t, with time to maturity $\tau_{i,t}$ and moneyness $m_{i,t}$ is estimated through the following regression:

$$\sigma_{i,t}(\boldsymbol{a_t}, m_{i,t}, \tau_{i,t}) = a_{0,t} + a_{1,t}m_{i,t} + a_{2,t}m_{i,t}^2 + a_{3,t}\tau_{i,t} + a_{4,t}\tau_{i,t}^2 + a_{5,t}m_{i,t}\tau_{i,t} + \epsilon_{i,t}, \tag{3}$$

The parameter vector $\boldsymbol{a}_t$ captures the influence of the moneyness and time to maturity levels on the IV values. This vector is estimated through ordinary least squares, aiming to minimise the difference between the observed IV values and those predicted by the model:

$$\boldsymbol{a_t} = \underset{\boldsymbol{a_t}}{\arg\min} \left(\frac{1}{N} \sum_{i=1}^{N} \left[\sigma_{i,t} - \hat{\sigma}_{i,t}(\hat{\boldsymbol{a}}_{\boldsymbol{t}}, m_{i,t}, \tau_{i,t}) \right]^2 \right), \tag{4}$$

where N represents the number of observations in the training set. Equation 4 results in parameter vector $\hat{\boldsymbol{a}}_{\boldsymbol{t}}$, which is subsequently substituted into equation 3 to compute the IV values calculated using the AHBS model, denoted as $\hat{\sigma}_{i,t}^{AHBS}(\hat{\boldsymbol{a}}_{\boldsymbol{t}}, m_{i,t}, \tau_{i,t})$. The AHBS model diverges from the theoretical assumption of constant volatility inherent to the BS framework, offering a more accurate representation of IV values by incorporating the volatility smile observed in the market.

4.1.3 Carr-Wu Model

The CW model, introduced by Carr and Wu (2016), presents an alternative option pricing framework that differs from the BS model. Unlike the BS framework, which models the underlying price process, the CW model directly models IVS dynamics. This approach is consistent with market practices, where institutional investors manage their positions through IV values. The model derives IV values by applying no-arbitrage constraints to the shape of the IVS, enabling accurately modeling the volatility smile while ensuring that the information embedded within this smile is captured by the model.

Carr and Wu (2016) assume that at time t, an option with strike price K and time to maturity τ has the following dynamics for its underlying asset price S_t and implied volatility $\sigma_t(K,\tau)$:

$$\begin{aligned} \frac{dS_t}{S_t} &= \sqrt{v_t} dW_t, \\ \frac{d\sigma_t(K,\tau)}{\sigma_t(K,\tau)} &= e^{-\eta_t \tau}\left(m_t dt + w_t dZ_t\right), \end{aligned} \tag{5}$$

where v_t is the time t instantaneous variance rate of the underlying asset price, m_t and w_t are the respective drift and volatility of the implied volatility process and $e^{-\eta_t}$ is an exponential dampening factor to accommodate for the empirical observation that implied volatilities of options with long maturities tend to move less. The processes Z_t and W_t are modeled as Brownian Motions that are used as shocks for the stock price and volatility process. They have correlation ρ_t, which is a stochastic process that takes on values in the interval $[-1,1]$. The values of m_t, w_t and η_t are allowed to be stochastic processes, uncorrelated to K, τ and $\sigma_t(K,\tau)$.

Using no-arbitrage constraints on the shape of the IVS, Carr and Wu (2016) show that the following quadratic equation must hold for every option i with strike K, time to maturity τ, implied volatility $\sigma_{i,t}$ and relative strike $k = ln(K/S_t)$ on day t:

$$\begin{aligned} &\frac{1}{4} e^{-2n_t\tau} w_t^2 \tau^2 \sigma_{i,t}^4 + \left(1 - 2e^{-2n_t\tau} m_t \tau - e^{-n_t\tau} w_t \rho_t \sqrt{v_t}\tau\right)\sigma_{i,t}^2 \\ &\quad - \left(v_t + 2e^{-n_t\tau} w_t \rho_t \sqrt{v_t} k + e^{-2n_t\tau} w_t^2 k^2\right) = 0. \end{aligned} \tag{6}$$

It is crucial to note that the solution of this equation only depends on the current values of the five processes $(v_t, m_t, w_t, \eta_t, \rho_t)$. Consequently, parameter vector $\boldsymbol{\theta_t} = (v_t, m_t, w_t, \eta_t, \rho_t)$, must be estimated daily through non-linear least squares. Using option data from day t, where the IV of option i is $\sigma_{i,t} = \sigma_t(k_{i,t}, \tau_{i,t})$, $\boldsymbol{\theta_{t+1}}$ can be estimated as follows:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{\boldsymbol{t+1}} = \arg\min_{\boldsymbol{\theta_t}} \sum_{i=1}^{n} \Bigg[&\frac{1}{4} e^{-2\eta_t\tau_{i,t}} w_t^2 \tau_{i,t}^2 \sigma_{i,t}^4 + \left(1 - 2e^{-2n_t\tau_{i,t}} m_t \tau_{i,t} - e^{-n_t\tau_{i,t}} w_t \rho_t \sqrt{v_t}\tau_{i,t}\right)\sigma_{i,t}^2 \\ &- \left(v_t + 2e^{-n_t\tau_{i,t}} w_t \rho_t \sqrt{v_t} k_{i,t} + e^{-2n_t\tau_{i,t}} w_t^2 k_{i,t}^2\right) \Bigg]^2, \end{aligned} \tag{7}$$

where n represents the number of observations in the training set. After obtaining $\hat{\boldsymbol{\theta}}_{\boldsymbol{t}}$, IV estimates $\hat{\sigma}_{i,t}^{CW}$ can be obtained for the test set by solving equation 6, using $\hat{\boldsymbol{\theta}}_{\boldsymbol{t}}$ and the relative strike k and time to maturity τ of the option as inputs.

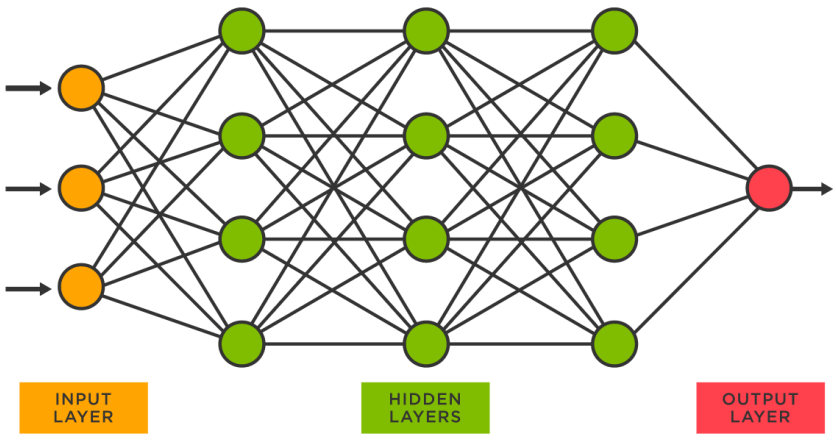
4.2 Neural Networks

The models introduced in section 4.1 generate IV predictions $\hat{\sigma}_{i,t}^{(M)}$, where $M \in \{BS, AHBS, CW\}$. To improve the accuracy of these predictions, they are refined using neural networks, resulting in enhanced neural network-corrected (NNC) predictions for each model: $\hat{\sigma}_{i,t}^{(M,NNC)}$. This subsection begins with an explanation of the fundamental mechanisms underlying neural networks, followed by an explanation of their application within the context of this research.

4.2.1 Neural Network Framework

The feed-forward neural networks used in this research map input vectors $\boldsymbol{x}$ to predicted outputs $\boldsymbol{y}$, effectively learning the function $f(\boldsymbol{x}) = \boldsymbol{y}$. This is done through a series of interconnected layers of neurons, consisting of an input layer, one or more hidden layers and an output layer. The neurons in each layer are connected to every neuron in the previous layer, forming a network capable of approximating complex functions. Figure 2 provides a schematic overview of such a network, where a neural network maps $\boldsymbol{x} \in \mathbb{R}^3$ onto $\boldsymbol{y} \in \mathbb{R}^1$ through three hidden layers, each containing four neurons. The number of hidden layers and the amount of neurons within those hidden layers, known as the respective depth and width of the neural network, are hyperparameters that can be adjusted to influence model performance.

Figure 2: Neural Network Framework Example



Notes: Schematic overview of a Neural Network with 3 hidden layers, 4 neurons per hidden layer, 3 input variables and one output variable.

In a neural network, each neuron receives input, which is subsequently transformed into output. The input of a neuron is the weighted sum of the outputs from the neurons in the previous layer. Denoting $n_l^{(i)}$ as the output of the i-th neuron in layer l, its input is equal to:

$$w_{i,l}^{(1)} n_{l-1}^{(1)} + w_{i,l}^{(2)} n_{l-1}^{(2)} + \cdots + w_{i,l}^{(K)} n_{l-1}^{(K)}, \tag{8}$$

where $w_{i,l}^{(k)}$ represents the weight of the connection between $n_l^{(i)}$ and $n_{l-1}^{(k)}$, and K denotes the number of neurons in layer $l-1$. To obtain output $n_l^{(i)}$, a bias term is added to the sum of inputs, and the resulting value is passed through an activation function $a(\cdot)$:

$$n_l^{(i)} = a\left(b_{i,l} + \sum_{k=1}^{K} w_{i,l}^{(k)} n_{l-1}^{(k)}\right). \tag{9}$$

The output value of the neuron reflects its level of activity and the importance within the network. Neurons with higher values are more active and have a bigger influence on the output of the network. The same principle applies to weights, which play a crucial part in determining the significance of connections between neurons. Higher weights indicate stronger connections, implying that the value of the previous neuron substantially influences the value of the current neuron. At its core, the weights determine the relative importance of specific inputs on the prediction process.

Extending equation 9 to matrix algebra, the outputs of all the neurons in layer l can be represented as:

$$N_l = a\left(W_l N_{l-1} + B_l\right), \tag{10}$$

where W_l denotes the weights corresponding to the connections between layer $l-1$ and layer l, N_{l-1} represents the outputs of the neurons in layer $l-1$, and B_l denotes the biases for layer l. The activation function $a(\cdot)$ is a critical component, introducing non-linearity into the model. It determines whether a neuron should be activated and its level of importance. Popular choices include the sigmoid function $\left(a(x) = \frac{e^x}{1+e^x}\right)$ and the ReLu function $\Big(a(x) = max(0, x)\Big)$, both of which are simple non-linear functions that map the input onto a nonnegative output space. In this context, the biases B_l play a significant role in the activation of the neurons. Adjustments in bias shift the activation threshold, directly influencing whether neurons remain active or become dormant. These adjustments are essential for a neural network's ability to learn and adapt to different data.

Equation 10 underscores the influence that weights and biases have on neuron outputs, highlighting the importance of accurately estimating these parameters. The process of estimating these variables and training the neural network is known as backpropagation. Training data is used to make N predictions, and the Root Mean Squared Error (RMSE) between the predictions and the target variables is calculated as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)}, \tag{11}$$

where $\hat{y}_i$ represents the output of the model, specifically the value of the neuron in the final layer L. Due to the recursive nature of equation 10. The model output can be expressed in terms of the model input $N_0 = \boldsymbol{x}$ as follows:

$$N_L = a\Big(W_L a\big(W_{L-1}\cdots a(W1\boldsymbol{x} + B_1) + \cdots + B_{L-1}\big) + B_L\Big) \tag{12}$$

The value of N_L can be optimized by selecting appropriate values for the weights and biases. This optimization is achieved using algorithms like gradient descent, which iteratively adjust the weights and biases to reduce the RMSE. This process involves computing the gradient of the error surface, resulting in a vector that indicates the direction in which the weights and biases should be adjusted to achieve the biggest loss in RMSE. Each observation in the training set provides a small update to the parameters, nudging them towards the values that yield the lowest possible error term.

This section provided a broad overview of neural networks, highlighting key components and mechanisms. However, it only offers a brief introduction to the intricate workings and potential of these systems. For a more in-depth understanding of neural networks and their applications in finance, readers are referred to the work of Gu et al. (2020). Additionally, Schmidhuber (2015) offer a comprehensive overview of the existing neural network literature.

4.2.2 Implementation of Neural Network

This research employs neural networks to correct parametric models, following an approach similar to that of Almeida et al. (2023) for index option IVS models. The parametric models generate IV predictions $\hat{\sigma}_{i,t}^{(M)}$, where $M \in \{BS, AHBS, CW\}$, for option i with strike price K and time to maturity τ. A neural network is then applied to correct these predictions by modeling their error terms. The error terms for model M are defined as the difference

between the model's predictions and the observed values:

$$\epsilon_{i,t}^{(M)} = \sigma_{i,t} - \hat{\sigma}_{i,t}^{(M)}. \tag{13}$$

Subsequently, a neural network is employed to model the function $f_t^{(M)}(\cdot)$, which aims to capture the error the IV prediction of option i made by model M, based on its strike price K and time to maturity τ as $f_t^{(M)}(K,\tau) = \epsilon_{i,t}^{(M)}$. The NNC predictions are then constructed by combining the original model predictions with the errors modeled by the neural networks:

$$\hat{\sigma}_{i,t}^{(M,NNC)} = \hat{\sigma}_{i,t}^{(M)} + f_t^{(M)}(K,\tau). \tag{14}$$

This method leverages the structure provided by the parametric model while improving predictive accuracy through a neural network. This approach is advantageous as it combines the parametric model assumptions with the flexibility of neural networks.

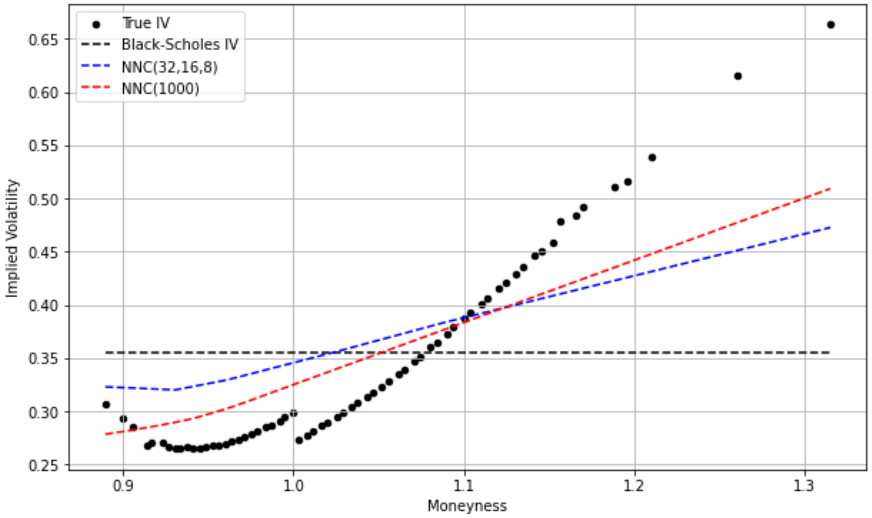
4.2.3 Neural Network Configuration

The initialization of a neural network requires several critical decisions, including the choice of activation function, the size of the hidden layer and the choice of solver, among other considerations. In this study, the ReLu activation function $\big(a(x) = max(0,x)\big)$ is employed due to its computational efficiency and its ability to deactivate neurons by assigning them a value of zero when their influence on the model is minimal. The neural network structure utilized in this research comprises a single hidden layer containing 1,000 neurons. This configuration is selected after experimentation, which demonstrated that a single layer with a large number of neurons outperformed multi-layered structures.

This finding is illustrated in figure 3, which visualizes the IV predictions generated by two NNC models with different neural network structures. The NNC model with a single hidden layer of 1,000 neurons, denoted in red, is compared against an NNC model employing three hidden layers, consisting of 32, 16 and 8 neurons, denoted in blue. This latter configuration is the set-up that yielded the best results in the study by Almeida et al. (2023). Figure 3 demonstrates that the IV predictions of the NNC model with a single hidden layer more closely resemble the true IV values for AMZN options on the first day of the sample. This pattern is consistent across the sample and the different models. One possible explanation for this outcome is that neural networks with multiple hidden layers are prone to overfitting,

particularly when trained on a small sample size (Li et al., 2019). In this study, the neural networks are trained on daily equity option data, which can have limited sample sizes compared to index options.

Figure 3: Neural Network Correction Black-Scholes Implied Volatility Surface



Notes: This graph compares true IV values against Black-Scholes predictions, and these predictions corrected with two different Neural Networks, one with a single hidden layer of 1,000 nodes, and one with three hidden layers with respective sizes 32, 16 and 8. The options in question have AMZN as their underlying asset and day t = 2018-02-28.

Additionally, the neural network framework provided by scikit-learn library in Python (Pedregosa et al., 2011) is employed. Specifically, the MLPRegressor function is utilized with the default stochastic gradient descent based solver 'adam', the `random_state` variable set to 1 and the `max_iter` variable set to 500. All other settings are left at their default. It is important to emphasize that the primary objective of this paper is to explore the application of neural networks to enhance parametric models. Consequently, this study does not focus on hyperparameter optimization to achieve the lowest possible error metrics. As a result, the findings presented in this paper should be interpreted as indicative of the potential capabilities of neural networks, rather than as definitive benchmarks.

4.3 Implementation Details of Models

Following the approach outlined by Almeida et al. (2023), the IV predictions from models are evaluated from both a practitioner's and an academic's perspective. Practitioners in financial markets typically use IVS models to interpolate observed IVS levels to estimate target IV

values due to the unavailability of many exact IV values on the market. Additionally, market practitioners use these models to make one day, one week (five trading days) and one month (twenty-one trading days) ahead IV predictions to help make informed trading decisions. Given the volatile nature of the options market, predictions beyond these horizons tend to be unreliable (Brownlees et al., 2011). However, the long-term behaviour of these models holds significant interest to researchers who seek to understand whether the neural networks help in capturing the complexities of IVS fluctuations over extended periods. Consequently, this study evaluates both the short-term interpolation and predictive performance, as well as the long-term performance of the models. This subsection now briefly explains how these predictions are constructed.

In the interpolation exercise, the models are used to make same-day predictions. Following the setup from Almeida et al. (2023), approximately 60% of the daily data is reserved for training. This training dataset is created by sorting daily option data on strike price and time to maturity, and then selecting the first, third and fifth observation from every set of five. This sampling technique ensures a diverse and representative training dataset. The remaining 40% of the data serves as the test set, where IV values are estimated to evaluate model performance. For the daily prediction task, the training set comprises all option data from day t, while the different test sets consist of data for day $t + h$, where h represents 1, 5 or 20 trading days. Due to the high volatility inherent to the options market, only data from day t is used to estimate the models.

For long-term predictions, the training set includes the first three years of the sample period (28 February 2018 to 28 February 2021), while the test set comprises the last two years (1 March 2021 to 28 February 2023). Notably, the CW model is excluded from long-term prediction tasks due to its dependence on time-specific values of five dynamic processes $\boldsymbol{\theta_t} = (v_t, m_t, w_t, \eta_t, \rho_t)$. Although daily CW predictions can be generated by estimating $\boldsymbol{\theta_t}$ at day t, this approach is not feasible for extending predictions up to two years. This limitation arises because the values of $\boldsymbol{\theta_t}$ are time-dynamic processes which are fundamental for the model. Consequently, using the CW-model for long-term forecasts would not result in reasonable results, justifying its exclusion from this exercise.

The IV predictions of the remaining BS and AHBS models over this longer time period are corrected using a neural network similar to those employed for the other NNC models, where the moneyness and time to maturity of the option serve as input variables. However, an additional time-varying neural network correction procedure (denoted as NNC*) is implemented to further enhance the accuracy of long-term IV predictions. This method utilizes a neural network that incorporates time-varying input variables, specifically the VIX level and the return of the option's underlying asset. These variable are selected based on findings by Cao et al. (2020), who found that a neural network using these inputs could accurately predict IV values. It is important to note that a direct comparison between the parametric models and their NNC counterparts is not entirely fair, given that the neural networks receive daily updates. Nevertheless, this approach can be interpreted as a modeling technique aimed at delivering optimal long-term results by correcting parametric models with neural networks, aligning with the overarching methodology of this research.

The previously outlined training and testing sets are applied consistently across both parametric models and neural networks. When neural networks are employed, they are trained on the pricing errors generated by the parametric models within the training set. These fitting errors provide an effective training method, thereby facilitating a fair comparison between NNC models and their parametric counterparts.

4.4 Evaluation Methods

This subsection outlines the evaluation methods utilized in this research to assess the performance of the different IVS models.

4.4.1 Implied Volatility Root Mean Squared Error

The first evaluation metric used is the Implied Volatility Root Mean Squared Error (IVRMSE), as used by Almeida et al. (2023). This metric is defined in terms of implied volatilities and is therefore easily interpretable. A lower IVRMSE indicates smaller errors in the model, correlating to a good performance.

For model M that generates IV predictions $\hat{\sigma}_{i,t}^{(M)}$, its IVRMSE is computed by comparing these predictions against actual IV values $\sigma_{i,t}$ for the test set:

$$\text{IVRMSE}^{(M)} = \sqrt{\frac{1}{N}\sum_{i,t}\left(\sigma_{i,t} - \hat{\sigma}_{i,t}^{(M)}\right)^2}, \tag{15}$$

where N represents the number of observations in the test set. The RMSE framework is widely used in research due to its ability to compare different models. Its quadratic loss function assigns greater weight to larger errors, making it more sensitive to outliers. This sensitivity can prove to be valuable in the context of IV prediction, where significant miscalculations can lead to incorrect option price predictions and substantial financial losses.

4.4.2 Mean Absolute Error

In addition to the IVRMSE, the Mean Absolute Error (MAE) metric provides an alternative perspective by treating all errors equally. The MAE for a given model M is computed as:

$$\text{MAE}^{(M)} = \frac{1}{N}\sum_{i,t}\left|\sigma_{i,t} - \hat{\sigma}_{i,t}^{(M)}\right| \tag{16}$$

where N denotes the number of observations in the test set. Unlike the IVRMSE, the MAE does not square the errors but takes their absolute values. Consequently, the MAE penalizes large errors less than the IVRMSE, focusing on the average error size and being less influenced by extreme values. Employing both error metrics allows for a thorough evaluation of model performance, enabling the identification of robust models capable of capturing different types of errors.

4.4.3 Diebold-Mariano test

Given that this research explores the potential of neural networks to enhance parametric models, it is crucial to compare the performance of a model with its NNC counterpart. To formally assess this, the Diebold-Mariano (DM) test is used. Introduced by Diebold and Mariano (1995), this test evaluates whether the difference in forecasting performance between two models is statistically significant.

The test operates by analyzing the errors of the different models and determining which model exhibits larger errors. Let $e_{1,t}$ and $e_{2,t}$ represent the forecast errors of model 1 and model 2 at time t, respectively. The loss differential, computed as the difference between the squared errors, is defined as:

$$d_t = e_{1,t}^2 - e_{1,t}^2. \tag{17}$$

The mean of this loss differential series, $\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t$, is then used to calculate the DM statistic as:

$$DM = \frac{\bar{d}}{\sqrt{\sigma_t^2/T}}, \tag{18}$$

where T is the number of observations and σ_t^2 is the variance of the loss differential series. This statistic tests the null hypothesis that the two models have equal forecasting accuracy ($i.e., E[d_t] = 0$). A significantly positive (negative) DM statistic indicates that model 1 (2) has larger forecast errors, implying lower forecast accuracy. In this research, the DM test is employed to compare the parametric models against their NNC counterparts, where model 1 represents the original parametric model and model 2 corresponds to the NNC model. A significant positive test statistic implies that the NNC models have lower forecast errors and thus outperform the parametric models.

4.4.4 Moment Evaluation

The final evaluation technique determines the effectiveness of the different models in capturing real-world stock dynamics. Bakshi et al. (2003) introduce a method to calculate the moments of the risk-neutral return distribution of the underlying asset based on option prices, which can be derived from the different IVS models. This approach yields model implied moments for each model, which can then be compared against the moments of the observed asset return distribution. This allows for an evaluation of which model most accurately captures the dynamics of actual asset returns.

Theorem 1 of Bakshi et al. (2003) states that the time-t, τ-period risk-neutral return skewness can be computed as:

$$SKEW(t,\tau) = \frac{e^{r\tau}W(t,\tau) - 3\mu(t,\tau)e^{r\tau}V(t,\tau) + 2\mu(t,\tau)^3}{\left(e^{r\tau}V(t,\tau) - \mu(t,\tau)^2\right)^{3/2}}, \tag{19}$$

and the time-t, τ-period risk-neutral return kurtosis as:

$$KURT(t,\tau) = \frac{e^{r\tau}X(t,\tau) - 4\mu(t,\tau)e^{r\tau}W(t,\tau) + 6e^{r\tau}\mu(t,\tau)^2V(t,\tau) - 3\mu(t,\tau)^4}{\left(e^{r\tau}V(t,\tau) - \mu(t,\tau)^2\right)^2}, \tag{20}$$

where the values of $V(t,\tau)$, $W(t,\tau)$ and $X(t,\tau)$ are calculated as:

$$\begin{aligned} V(t,\tau) = &\int_{S(t)}^{\infty} \frac{2\left(1 - \ln\left[\frac{K}{S(t)}\right]\right)}{K^2} C(t,\tau;K)\,dK \\ &+ \int_{0}^{S(t)} \frac{2\left(1 + \ln\left[\frac{S(t)}{K}\right]\right)}{K^2} P(t,\tau;K)\,dK, \end{aligned} \tag{21}$$

$$\begin{aligned} W(t,\tau) = &\int_{S(t)}^{\infty} \frac{6\ln\left[\frac{K}{S(t)}\right] - 3\left(\ln\left[\frac{K}{S(t)}\right]\right)^2}{K^2} C(t,\tau;K)\,dK \\ &- \int_{0}^{S(t)} \frac{6\ln\left[\frac{S(t)}{K}\right] + 3\left(\ln\left[\frac{S(t)}{K}\right]\right)^2}{K^2} P(t,\tau;K)\,dK, \end{aligned} \tag{22}$$

$$\begin{aligned} X(t,\tau) = &\int_{S(t)}^{\infty} \frac{12\left(\ln\left[\frac{K}{S(t)}\right]\right)^2 - 4\left(\ln\left[\frac{K}{S(t)}\right]\right)^3}{K^2} C(t,\tau;K)\,dK \\ &+ \int_{0}^{S(t)} \frac{12\left(\ln\left[\frac{S(t)}{K}\right]\right)^2 + 4\left(\ln\left[\frac{S(t)}{K}\right]\right)^3}{K^2} P(t,\tau;K)\,dK, \end{aligned} \tag{23}$$

which can be combined to create $\mu(t,\tau)$ as follows:

$$\mu(t,\tau) = e^{r\tau} - 1 - \frac{e^{r\tau}}{2}V(t,\tau) - \frac{e^{r\tau}}{6}W(t,\tau) - \frac{e^{r\tau}}{24}X(t,\tau). \tag{24}$$

The terms $C(t,\tau;K)$ and $P(t,\tau;K)$ represent the respective call and put price at time t with strike price K and time to maturity τ. These prices are obtained by converting IV values $\hat{\sigma}_{i,t}^{(M)}$ into option prices using BS equation (1), using the 3-month Treasury bill rate from the St. Louis Federal Reserve Economic Data (FRED) database as a proxy for the risk-free rate.

To evaluate the performance of the parametric models and their NNC counterparts, their model-implied one-year ahead skewness and kurtosis are calculated. The integrals in equations 21, 22 and 23 are approximated by summing over a linear grid of 1,000 strike prices, spanning from the lowest observed strike price K^{min} to the highest K^{max} on day t for the given underlying asset. Using the values on this grid strike prices, in combination with $\tau = 1$

and current day t, IV values $\hat{\sigma}_{i,t}^{(M)}$ are calculated for the different models M. These IV values are then transformed into option prices through the BS formula, yielding synthetic put prices $P(t,\tau;K)$ and call prices $C(t,\tau;K)$, depending on whether the moneyness of the option is above or below one. Finally, these option prices are multiplied with the value $\Delta K = \frac{k^{max}-k^{min}}{1000}$ to accurately transform the integral into a sum.

Following the computation of the model-implied yearly skewness and kurtosis levels, denoted as $SKEW(t,1)$ and $KURT(t,1)$, these values are compared to the actual skewness and kurtosis levels of the underlying asset returns over the following year. This comparison is conducted for each of the five years covered by the dataset, ensuring that the observed skewness and kurtosis are well defined based on a full year of data. This evaluation technique measures how accurately the different models capture the return distribution of the underlying asset, highlighting the potential that options have in predicting market behaviour.

5 Results

This section presents an evaluation of the performance of the methods introduced in section 4. It starts with an analysis of the daily predictions, which includes both daily interpolation and prediction exercises. These methods demonstrate how the models can be utilized by market participants for short-term forecasting or for obtaining specific IV values through interpolation. Following this, the evaluation shifts to assessing the performance of long-term predictions. This analysis focuses on IV prediction and the effectiveness of the NNC procedure over extended time periods, thereby aligning with a scientific perspective aimed at investigating the robustness and stability of the models over time.

5.1 Daily Predictions

Beginning with the results of the interpolation exercise, table 5.1 presents the average RMSE and MAE values for the specified models and their NNC counterparts, evaluated across the different underlying assets. The table indicates that the application of the NNC procedure result in lower RMSE and MAE values for all three models. Specifically, the RMSE of the BS model is reduced by approximately 50% after the application of the neural network. Similarly, the AHBS and CW models show respective decreases of 24% and 13%.

Table 5.1: Average Error Evaluation Daily Interpolation

	BS	**BS-NNC**	**AHBS**	**AHBS-NNC**	**CW**	**CW-NNC**
RMSE	0.110	0.055	0.072	0.052	0.100	0.087
MAE	0.071	0.030	0.041	0.031	0.062	0.038

Notes: Root Mean Squared Errors (RMSE) and Mean Absolute Errors (MAE) for the Black-Scholes (BS) model, the ad-hoc Black-Scholes (AHBS) model, the Carr-Wu (CW) model and their Neural Network Corrected (NNC) counterparts for a daily interpolation exercise. The reported values are averages over fifty different underlying assets.

Table 5.2 presents the results of the DM test, which compares the average performance of the standard models to their NNC counterparts, across all underlying assets. The results further reinforce the superiority of the NNC models, as the DM statistics are positive across all three models. Recalling equation 18, where the parametric models are defined as model 1 and the NNC models as model 2, a positive DM statistic implies that the errors of the standard models are larger, indicating better performance by the NNC models. The results are statistically significant, with DM statistics exhibiting p-values effectively equal to zero. This high level of significance is likely due to the large sample size used for making predictions, which increase the confidence in rejecting the null hypothesis of equal performance between the standard and NNC models. These findings underscore the potential advantages of using neural networks to correct parametric models, particularly within an interpolation framework.

Table 5.2: Diebold-Mariano Test Statistics Daily Interpolation

	BS	**AHBS**	**CW**
DM	650***	315***	98.54***

*Notes: Diebold-Mariano (DM) test statistics for the comparison of the Black-Scholes (BS), ad-hoc Black-Scholes (AHBS) and Carr-Wu (CW) models with their Neural Network Corrected (NNC) counterparts for a daily interpolation exercise. *** indicates a P-value below 0.001*

Table 5.6 presents the RMSE and MAE metrics for the models per underlying asset. The results indicate that the NNC models consistently outperform the parametric models across a diverse range of stocks. This effect is particularly pronounced for liquid options with underlying assets such as Amazon.com, Inc. (AMZN), Alphabet Inc. (GOOGL) and Tesla, Inc. (TSLA), where the NNC models demonstrate substantial improvements in predictive

accuracy. Interestingly, this improved predictive performance extends beyond the most liquid options, as options on stocks like Align Technology, Inc. (ALGN), Charter Communications, Inc. (CHTR) and Visa Inc. (V) also exhibit more accurate IV predictions following the NNC procedure.

However, there are instances where the NNC models yield higher RMSE values, notably for options on Wayfair Inc. (W), MongoDB, Inc. (MDB) and MicroStrategy Incorporated (MSTR). In these cases, the CW model in particular shows limited benefit from the NNC procedure. It is important to note that options on these assets are not among the most actively traded, suggesting that the lower performance of the NNC models is potentially caused by the limited number of available observations. Nevertheless, even for models with higher RMSE values, the MAE values for the NNC models remain below those of the non-NNC counterparts, indicating that despite the presence of some significant outliers, the NNC models generally provide more accurate predictions compared to traditional parametric models.

Table 5.3: Average Error Evaluation Daily Prediction

		BS	**BS-NNC**	**AHBS**	**AHBS-NNC**	**CW**	**CW-NNC**
t+1	***RMSE***	0.117	0.066	0.082	0.063	0.110	0.098
	MAE	0.074	0.034	0.045	0.029	0.065	0.041
t+5	***RMSE***	0.128	0.091	0.102	0.093	0.126	0.122
	MAE	0.080	0.048	0.056	0.044	0.074	0.054
t+21	***RMSE***	0.148	0.130	0.137	0.130	0.149	0.144
	MAE	0.093	0.071	0.078	0.072	0.090	0.076

Notes: Root Mean Squared Errors (RMSE) and Mean Absolute Errors (MAE) for the Black-Scholes (BS) model, the ad-hoc Black-Scholes (AHBS) model, the Carr-Wu (CW) model and their Neural Network Corrected (NNC) counterparts for a daily prediction exercise. The reported values are averages over fifty different underlying assets.

Proceeding to the results of the daily prediction exercise, table 5.3 provides the RMSE and MAE values of the one day, one week and one month ahead forecasts across different models. These values are averages computed over all underlying assets. The table shows that, consistent with the findings from the interpolation exercise, the NNC models consistently outperform their standard counterparts across all time horizons. Notably, the largest re-

ductions are observed in the one day ahead predictions of the BS model, with both RMSE and MAE values decreasing by approximately 50%. It should be emphasized that prediction accuracy decreases as the forecast horizon increases, a trend observed across all models. This decline in accuracy reflects the increasing uncertainty associated with modeling further into the future.

Table 5.4 provides further evidence of the improvements achieved by the NNC models, presenting the results of the DM test which compares the predictive performance of the parametric models with their NNC counterparts. Consistent with the DM statistics observed in the interpolation exercise, the significant positive DM statistics indicate that the NNC models achieve lower errors terms, thereby significantly improving predictive accuracy. The positive and significant DM statistics across all models and time horizons highlight the robustness of the NNC procedure within this framework. These results suggest that integrating neural networks to model error terms of traditional models can substantially improve forecast accuracy, while preserving the economic foundations underlying the models.

Table 5.4: Diebold-Mariano Test Statistics Daily Prediction

	Horizon	**BS**	**AHBS**	**CW**
DM	*t+1*	1015***	514***	104***
DM	*t+5*	883***	360***	78***
DM	*t+21*	611***	347***	54***

*Notes: Diebold-Mariano (DM) test statistics for the comparison of the Black-Scholes (BS), ad-hoc Black-Scholes (AHBS) and Carr-Wu (CW) models with their Neural Network Corrected (NNC) counterparts for a daily prediction exercise. *** indicates that the P-value is below 0.001*

Table 5.7 and table 5.8 present the RMSE and MAE values for the model predictions across individual underlying assets. The results indicate that the prediction accuracy diminishes over longer time horizons, a consequence of the increased uncertainty associated with predicting further into the future. Notably, for nearly all underlying assets the NNC procedure increases predictive performance. The largest improvements are observed in short-term predictions for the BS and AHBS models, consistent with the findings in table 5.3.

Similar to the interpolation exercise, the prediction accuracy of liquid options, such as those

on stocks of Amazon.com, Inc. (AMZN) or Alphabet Inc. (GOOGL) as their underlying assets, is notably improved by the NNC procedure. Furthermore, even options with fewer observations in the sample, like those on Charter Communications, Inc. (CHTR) and Visa Inc. (V), predictive accuracy improves after the NNC procedure. Nonetheless, certain options do not experience significant benefits from the NNC procedure. Consistent with the interpolation exercise, IV predictions for options on stocks as Wayfair Inc. (W), MongoDB, Inc. (MDB) and MicroStrategy Incorporated (MSTR) become less accurate when using the CW-NNC model compared to the CW model. This pattern is reflected primarily in RMSE values, where the NNC models occasionally underperform. However, MAE values generally remain lower for NNC models compared to their parametric counterparts, indicating that despite some significant errors, the overall predictions of NNC models are generally closer to the true values.

As the final evaluation metric, the model-implied moments are evaluated. Table 5.5 presents the average RMSE and MAE values for the annual skewness and kurtosis implied by the models, compared to the moments of the observed return distributions. Several notable results arise from this analysis. Both error metrics are consistently lower for skewness than for kurtosis across all models. This suggests that the models more accurately capture the skewness of the underlying assets return distributions. Given that skewness measures asymmetry and kurtosis measures the tailedness of a distribution, these results imply that the model-implied moments are better at capturing the direction in which the underlying asset returns move, rather than the presence of outliers in the data. In terms of model improvement between the parametric and NNC models, the results are mixed. For the AHBS and CW models, the application of neural networks improves performance. Conversely, the error metrics of the BS model increase when a correcting neural network is applied.

Remarkably, the unmodified BS model records the lowest error metrics for both skewness and kurtosis among all evaluated models. The application of the NNC procedure to the BS model does not improve moment estimation accuracy, despite previous findings indicating that it increases the accuracy of IV predictions. In contrast, for the AHBS and CW models, a correcting neural network does increase the estimation accuracy of both the skewness and kurtosis. However, it is important to recognize that a higher risk perception can lead to

unusual patterns in skewness and kurtosis values. Following the 1987 market crash, return distributions have become more negatively skewed (Bates, 1991), even when the physical process is symmetrical (Bakshi et al., 2003). This shift can introduce additional errors into moment estimation, and the variation in skewness and kurtosis across different models may reflect underlying market sentiment and perceived risk of extreme negative outcomes. Therefore, while the NNC model does not enhance moment prediction accuracy for the BS model, its performance for the AHBS and CW model suggests that it can be a valuable tool for improving moment estimation under certain conditions. Accurate moment estimation can be very important as these metrics provide insights into stock price behaviour. Given the large amount of factors influencing option prices, using IVS models to refine stock movement predictions could offer substantial benefits.

Table 5.5: Error Evaluation Model Implied Moments

		BS	**BS-NNC**	**AHBS**	**AHBS-NNC**	**CW**	**CW-NNC**
Skewness	*RMSE*	0.502	0.597	0.762	0.575	0.531	0.528
	MAE	0.379	0.481	0.626	0.465	0.417	0.413
Kurtosis	*RMSE*	1.806	2.062	2.321	1.986	1.907	1.889
	MAE	1.265	1.628	1.874	1.554	1.449	1.410

Notes: Root Mean Squared Errors (RMSE) and Mean Absolute Errors (MAE) for the comparison between the Skewness and Kurtosis of the observed return series and the Skewness and Kurtosis values implied by the Black-Scholes (BS) model, the ad-hoc Black-Scholes (AHBS) model, the Carr-Wu (CW) model and their Neural Network Corrected (NNC) counterparts.

In conclusion, the NNC models demonstrate significant improvement in both daily interpolation and prediction exercises. Although the comparison between NNC models does not identify a clear winner, all NNC models consistently outperform their parametric counterparts. This underscores the overall effectiveness of neural networks in reducing error terms and suggests that the enhancement achieved through the NNC procedure are not heavily dependent on the choice of the underlying parametric model.

Table 5.6: Error Evaluation Daily Interpolation per Underlying Asset

(*) = NNC	**BS**		**AHBS**		**CW**		*(*) = NNC*	**BS**		**AHBS**		**CW**	
Ticker	RMSE	MAE	RMSE	MAE	RMSE	MAE	**Ticker**	RMSE	MAE	RMSE	MAE	RMSE	MAE
AAPL	0.088	0.061	0.045	0.029	0.039	0.023	LULU	0.133	0.082	0.095	0.052	0.129	0.083
AAPL*	0.038	0.023	0.036	0.022	0.033	0.019	LULU*	0.072	0.042	0.075	0.043	0.108	0.051
ADBE	0.110	0.074	0.071	0.042	0.109	0.075	MA	0.104	0.071	0.060	0.036	0.088	0.060
ADBE*	0.059	0.033	0.059	0.033	0.068	0.037	MA*	0.056	0.032	0.054	0.031	0.056	0.035
ALGN	0.137	0.089	0.093	0.055	0.109	0.068	MDB	0.141	0.087	0.093	0.052	0.136	0.079
ALGN*	0.077	0.048	0.075	0.046	0.091	0.049	MDB*	0.073	0.041	0.073	0.042	0.140	0.063
AMD	0.086	0.056	0.057	0.034	0.087	0.057	MELI	0.130	0.084	0.084	0.046	0.130	0.085
AMD*	0.051	0.030	0.051	0.030	0.066	0.035	MELI*	0.065	0.034	0.064	0.034	0.109	0.048
AMZN	0.112	0.074	0.079	0.048	0.113	0.079	META	0.091	0.059	0.055	0.032	0.049	0.026
AMZN*	0.046	0.026	0.048	0.027	0.069	0.034	META*	0.041	0.025	0.04	0.025	0.049	0.022
AVGO	0.097	0.066	0.057	0.033	0.097	0.068	MSFT	0.081	0.058	0.038	0.024	0.044	0.020
AVGO*	0.050	0.029	0.049	0.028	0.058	0.034	MSFT*	0.033	0.021	0.032	0.020	0.041	0.017
AZO	0.079	0.052	0.048	0.027	0.078	0.052	MSTR	0.225	0.150	0.141	0.089	0.224	0.153
AZO*	0.046	0.027	0.043	0.025	0.058	0.031	MSTR*	0.104	0.058	0.109	0.063	0.337	0.185
BA	0.106	0.067	0.056	0.033	0.054	0.027	NFLX	0.110	0.072	0.073	0.044	0.056	0.036
BA*	0.045	0.026	0.044	0.027	0.049	0.022	NFLX*	0.052	0.032	0.052	0.033	0.057	0.028
BABA	0.094	0.062	0.055	0.033	0.091	0.06	NOC	0.075	0.054	0.042	0.025	0.053	0.035
BABA*	0.046	0.027	0.046	0.027	0.053	0.03	NOC*	0.043	0.027	0.042	0.026	0.041	0.026
BIIB	0.122	0.071	0.087	0.047	0.124	0.073	NOW	0.105	0.069	0.068	0.039	0.099	0.064
BIIB*	0.073	0.039	0.073	0.040	0.113	0.048	NOW*	0.060	0.034	0.059	0.033	0.067	0.035
BKNG	0.102	0.067	0.059	0.035	0.103	0.071	NTES	0.084	0.053	0.058	0.032	0.076	0.046
BKNG*	0.044	0.025	0.044	0.025	0.072	0.034	NTES*	0.056	0.033	0.056	0.032	0.057	0.033
BLK	0.090	0.062	0.050	0.030	0.068	0.041	NVDA	0.104	0.068	0.065	0.037	0.102	0.069
BLK*	0.050	0.030	0.047	0.027	0.053	0.027	NVDA*	0.047	0.026	0.048	0.027	0.065	0.031
CHTR	0.090	0.058	0.062	0.033	0.091	0.061	PANW	0.106	0.067	0.073	0.040	0.108	0.071
CHTR*	0.056	0.032	0.056	0.031	0.062	0.037	PANW*	0.063	0.036	0.063	0.035	0.079	0.040
CMG	0.102	0.065	0.066	0.038	0.103	0.067	PYPL	0.098	0.064	0.063	0.037	0.095	0.061
CMG*	0.052	0.030	0.051	0.030	0.081	0.037	PYPL*	0.052	0.032	0.051	0.032	0.073	0.035
COST	0.099	0.066	0.064	0.036	0.093	0.061	REGN	0.079	0.049	0.056	0.029	0.078	0.05
COST*	0.056	0.032	0.054	0.030	0.067	0.038	REGN*	0.052	0.027	0.052	0.027	0.054	0.030
CRM	0.101	0.068	0.065	0.037	0.091	0.06	RH	0.142	0.089	0.099	0.056	0.131	0.084
CRM*	0.059	0.033	0.059	0.033	0.058	0.034	RH*	0.077	0.047	0.079	0.047	0.135	0.063
GOOG	0.084	0.059	0.049	0.031	0.070	0.047	ROKU	0.155	0.095	0.114	0.067	0.157	0.100
GOOG*	0.039	0.023	0.039	0.023	0.043	0.026	ROKU*	0.078	0.047	0.079	0.049	0.185	0.084
GOOGL	0.083	0.059	0.049	0.031	0.083	0.060	SHOP	0.133	0.084	0.089	0.053	0.115	0.073
GOOGL*	0.038	0.022	0.038	0.022	0.045	0.029	SHOP*	0.065	0.035	0.067	0.038	0.100	0.048
GS	0.102	0.069	0.054	0.031	0.100	0.065	SQ	0.132	0.081	0.084	0.047	0.133	0.084
GS*	0.048	0.028	0.047	0.026	0.070	0.035	SQ*	0.065	0.037	0.065	0.037	0.126	0.058
HD	0.094	0.066	0.048	0.028	0.057	0.029	TSLA	0.175	0.113	0.109	0.064	0.080	0.050
HD*	0.045	0.026	0.042	0.025	0.052	0.023	TSLA*	0.058	0.031	0.059	0.032	0.066	0.035
HUM	0.081	0.056	0.044	0.025	0.063	0.044	TTD	0.148	0.089	0.109	0.060	0.141	0.088
HUM*	0.047	0.029	0.044	0.026	0.045	0.029	TTD*	0.077	0.045	0.078	0.046	0.156	0.070
ILMN	0.107	0.066	0.077	0.041	0.104	0.065	TWLO	0.149	0.090	0.110	0.063	0.153	0.096
ILMN*	0.071	0.040	0.069	0.038	0.078	0.042	TWLO*	0.087	0.049	0.088	0.051	0.147	0.070
ISRG	0.102	0.065	0.064	0.035	0.090	0.056	ULTA	0.121	0.073	0.084	0.045	0.101	0.060
ISRG*	0.056	0.029	0.054	0.028	0.063	0.031	ULTA*	0.064	0.037	0.065	0.037	0.083	0.040
LMT	0.079	0.057	0.044	0.027	0.037	0.024	V	0.095	0.066	0.049	0.028	0.096	0.069
LMT*	0.043	0.027	0.041	0.026	0.034	0.022	V*	0.047	0.027	0.044	0.025	0.057	0.036
LRCX	0.104	0.068	0.072	0.038	0.094	0.060	W	0.179	0.111	0.123	0.072	0.159	0.098
LRCX*	0.063	0.032	0.064	0.032	0.060	0.032	W*	0.093	0.058	0.095	0.058	0.205	0.097

Notes: RMSE and MAE values for the BS, AHBS, CW models and their NNC counterparts for a daily interpolation exercise for different underlying assets, denoted by their ticker. The shaded rows marked with an asterisk (*) correspond to the NNC models.

Table 5.7: Error Evaluation Daily Prediction per Underlying Asset (part 1)

	BS						AHBS						CW					
(*) = NNC	*t+1*		*t+5*		*t+21*		*t+1*		*t+5*		*t+21*		*t+1*		*t+5*		*t+21*	
Ticker	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
AAPL	0.090	0.062	0.094	0.064	0.109	0.073	0.051	0.032	0.063	0.039	0.091	0.055	0.048	0.027	0.060	0.035	0.085	0.050
AAPL*	0.043	0.024	0.057	0.033	0.087	0.050	0.042	0.024	0.057	0.033	0.087	0.051	0.043	0.022	0.057	0.031	0.085	0.048
ADBE	0.112	0.076	0.118	0.091	0.135	0.065	0.078	0.045	0.090	0.053	0.119	0.073	0.111	0.071	0.113	0.076	0.137	0.089
ADBE*	0.064	0.034	0.080	0.044	0.110	0.036	0.065	0.036	0.082	0.046	0.112	0.066	0.073	0.038	0.082	0.047	0.117	0.066
ALGN	0.143	0.093	0.161	0.103	0.185	0.123	0.106	0.060	0.138	0.077	0.177	0.109	0.138	0.079	0.159	0.093	0.195	0.121
ALGN*	0.088	0.049	0.132	0.069	0.174	0.103	0.088	0.049	0.132	0.070	0.173	0.104	0.122	0.053	0.155	0.071	0.197	0.107
AMD	0.092	0.059	0.104	0.068	0.128	0.086	0.069	0.040	0.089	0.055	0.125	0.080	0.090	0.059	0.104	0.069	0.130	0.088
AMD*	0.062	0.035	0.086	0.051	0.123	0.077	0.063	0.036	0.087	0.051	0.123	0.078	0.078	0.038	0.100	0.054	0.137	0.081
AMZN	0.116	0.076	0.121	0.080	0.133	0.089	0.087	0.051	0.099	0.059	0.120	0.075	0.120	0.081	0.124	0.085	0.138	0.093
AMZN*	0.055	0.028	0.078	0.040	0.104	0.060	0.055	0.029	0.079	0.041	0.105	0.061	0.064	0.035	0.091	0.046	0.111	0.063
AVGO	0.099	0.067	0.104	0.071	0.121	0.080	0.064	0.036	0.076	0.044	0.106	0.062	0.086	0.055	0.095	0.061	0.125	0.075
AVGO*	0.057	0.030	0.072	0.040	0.101	0.057	0.056	0.030	0.071	0.039	0.102	0.057	0.061	0.033	0.079	0.042	0.113	0.059
AZO	0.082	0.054	0.090	0.058	0.110	0.070	0.054	0.030	0.073	0.040	0.104	0.059	0.082	0.055	0.092	0.060	0.113	0.072
AZO*	0.052	0.029	0.071	0.038	0.102	0.057	0.051	0.028	0.071	0.038	0.102	0.057	0.070	0.034	0.083	0.042	0.111	0.060
BA	0.110	0.069	0.127	0.075	0.177	0.096	0.066	0.038	0.098	0.051	0.168	0.081	0.077	0.032	0.099	0.045	0.161	0.073
BA*	0.054	0.030	0.093	0.045	0.167	0.076	0.055	0.031	0.093	0.046	0.166	0.077	0.075	0.028	0.101	0.043	0.170	0.073
BABA	0.096	0.064	0.104	0.069	0.114	0.077	0.062	0.037	0.079	0.047	0.097	0.062	0.055	0.032	0.073	0.043	0.092	0.057
BABA*	0.052	0.029	0.073	0.041	0.092	0.056	0.052	0.030	0.073	0.042	0.092	0.057	0.053	0.026	0.072	0.038	0.091	0.054
BIIB	0.127	0.073	0.141	0.082	0.170	0.099	0.095	0.051	0.120	0.063	0.169	0.091	0.129	0.076	0.146	0.085	0.176	0.104
BIIB*	0.082	0.041	0.114	0.055	0.167	0.085	0.082	0.042	0.114	0.056	0.168	0.087	0.117	0.049	0.144	0.059	0.194	0.090
BKNG	0.105	0.069	0.113	0.073	0.134	0.085	0.068	0.039	0.085	0.049	0.121	0.071	0.111	0.071	0.117	0.076	0.154	0.089
BKNG*	0.055	0.029	0.078	0.042	0.116	0.065	0.055	0.029	0.079	0.042	0.117	0.065	0.089	0.037	0.104	0.048	0.148	0.068
BLK	0.091	0.063	0.096	0.066	0.113	0.074	0.057	0.033	0.068	0.040	0.099	0.056	0.075	0.033	0.083	0.040	0.097	0.054
BLK*	0.055	0.030	0.066	0.038	0.097	0.054	0.052	0.029	0.065	0.037	0.096	0.053	0.073	0.028	0.080	0.035	0.095	0.051
CHTR	0.092	0.059	0.096	0.062	0.107	0.070	0.067	0.035	0.078	0.041	0.097	0.057	0.093	0.062	0.099	0.065	0.110	0.074
CHTR*	0.060	0.031	0.076	0.038	0.095	0.054	0.060	0.031	0.075	0.038	0.095	0.054	0.062	0.036	0.084	0.044	0.099	0.057
CMG	0.106	0.067	0.118	0.074	0.146	0.091	0.075	0.043	0.096	0.055	0.138	0.079	0.108	0.070	0.121	0.077	0.151	0.094
CMG*	0.062	0.034	0.090	0.048	0.134	0.074	0.062	0.034	0.090	0.048	0.133	0.074	0.089	0.040	0.117	0.054	0.150	0.077
COST	0.101	0.068	0.107	0.070	0.120	0.078	0.069	0.038	0.080	0.044	0.103	0.059	0.098	0.063	0.103	0.067	0.120	0.077
COST*	0.058	0.031	0.073	0.038	0.098	0.053	0.058	0.031	0.073	0.038	0.098	0.054	0.065	0.038	0.077	0.044	0.102	0.057
CRM	0.104	0.070	0.111	0.073	0.124	0.083	0.073	0.042	0.086	0.050	0.112	0.067	0.101	0.059	0.118	0.065	0.125	0.076
CRM*	0.064	0.034	0.080	0.044	0.105	0.062	0.065	0.035	0.081	0.045	0.107	0.063	0.079	0.036	0.105	0.045	0.114	0.061
GOOG	0.086	0.060	0.091	0.064	0.102	0.071	0.056	0.034	0.067	0.042	0.089	0.055	0.084	0.056	0.090	0.060	0.113	0.069
GOOG*	0.045	0.026	0.061	0.035	0.084	0.050	0.045	0.026	0.061	0.035	0.085	0.050	0.055	0.030	0.067	0.038	0.101	0.051
GOOGL	0.085	0.060	0.090	0.063	0.101	0.069	0.056	0.034	0.067	0.041	0.087	0.054	0.086	0.062	0.091	0.065	0.103	0.072
GOOGL*	0.044	0.025	0.059	0.034	0.083	0.049	0.045	0.025	0.060	0.034	0.083	0.049	0.052	0.032	0.066	0.039	0.087	0.052
GS	0.104	0.070	0.109	0.073	0.129	0.081	0.060	0.034	0.073	0.041	0.110	0.058	0.115	0.067	0.105	0.070	0.143	0.081
GS*	0.052	0.028	0.067	0.036	0.106	0.053	0.052	0.028	0.067	0.036	0.106	0.054	0.084	0.035	0.081	0.042	0.132	0.058
HD	0.095	0.067	0.101	0.069	0.118	0.077	0.053	0.030	0.067	0.037	0.097	0.052	0.079	0.050	0.098	0.055	0.111	0.065
HD*	0.048	0.026	0.063	0.034	0.094	0.048	0.047	0.025	0.063	0.033	0.094	0.048	0.057	0.030	0.084	0.037	0.100	0.050
HUM	0.083	0.057	0.087	0.060	0.101	0.066	0.050	0.029	0.062	0.036	0.089	0.051	0.074	0.047	0.076	0.052	0.102	0.062
HUM*	0.050	0.029	0.061	0.036	0.087	0.050	0.048	0.027	0.060	0.035	0.088	0.050	0.060	0.030	0.063	0.037	0.095	0.051
ILMN	0.110	0.068	0.118	0.074	0.131	0.083	0.085	0.045	0.102	0.055	0.122	0.072	0.114	0.072	0.122	0.078	0.135	0.088
ILMN*	0.076	0.039	0.099	0.050	0.120	0.068	0.075	0.039	0.099	0.051	0.120	0.068	0.083	0.042	0.114	0.054	0.131	0.070
ISRG	0.104	0.066	0.109	0.069	0.126	0.078	0.070	0.038	0.082	0.046	0.114	0.065	0.104	0.066	0.108	0.070	0.129	0.081
ISRG*	0.062	0.031	0.076	0.040	0.111	0.060	0.061	0.031	0.076	0.040	0.110	0.060	0.074	0.037	0.093	0.045	0.122	0.062
LMT	0.081	0.058	0.084	0.060	0.098	0.067	0.048	0.030	0.057	0.035	0.081	0.048	0.043	0.027	0.052	0.033	0.077	0.045
LMT*	0.045	0.027	0.054	0.033	0.079	0.046	0.044	0.027	0.054	0.032	0.079	0.045	0.039	0.024	0.049	0.030	0.075	0.043
LRCX	0.107	0.070	0.112	0.073	0.123	0.082	0.079	0.043	0.089	0.051	0.109	0.067	0.108	0.072	0.113	0.076	0.125	0.085
LRCX*	0.069	0.034	0.082	0.044	0.103	0.061	0.070	0.035	0.083	0.045	0.104	0.062	0.075	0.037	0.092	0.047	0.117	0.064

Notes: Notes: RMSE and MAE values for the BS, AHBS, CW models and their NNC counterparts for a daily prediction exercise for different underlying assets, denoted by their ticker. The shaded rows marked with an asterisk (*) correspond to the NNC models.

Table 5.8: Error Evaluation Daily Prediction per Underlying Asset (part 2)

	BS						AHBS						CW					
(*) = NNC	*t+1*		*t+5*		*t+21*		*t+1*		*t+5*		*t+21*		*t+1*		*t+5*		*t+21*	
Ticker	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
LULU	0.139	0.085	0.156	0.096	0.191	0.118	0.107	0.057	0.135	0.073	0.192	0.107	0.147	0.092	0.165	0.103	0.218	0.127
LULU*	0.086	0.045	0.127	0.064	0.181	0.099	0.087	0.047	0.128	0.065	0.185	0.101	0.117	0.053	0.160	0.069	0.218	0.103
MA	0.106	0.072	0.109	0.075	0.125	0.082	0.066	0.039	0.076	0.046	0.105	0.062	0.094	0.063	0.098	0.068	0.119	0.079
MA*	0.059	0.033	0.071	0.040	0.101	0.057	0.059	0.033	0.071	0.041	0.102	0.057	0.064	0.038	0.073	0.044	0.102	0.059
MDB	0.146	0.091	0.163	0.101	0.192	0.121	0.108	0.059	0.142	0.077	0.208	0.111	0.148	0.086	0.167	0.100	0.214	0.124
MDB*	0.091	0.048	0.134	0.069	0.182	0.103	0.092	0.049	0.136	0.070	0.204	0.106	0.160	0.068	0.183	0.085	0.239	0.118
MELI	0.134	0.086	0.143	0.092	0.162	0.106	0.095	0.052	0.114	0.064	0.150	0.091	0.138	0.091	0.149	0.097	0.167	0.111
MELI*	0.077	0.040	0.106	0.056	0.146	0.085	0.077	0.040	0.107	0.057	0.146	0.085	0.111	0.051	0.152	0.067	0.180	0.094
META	0.094	0.061	0.102	0.065	0.116	0.074	0.063	0.036	0.080	0.044	0.106	0.061	0.094	0.060	0.104	0.066	0.120	0.076
META*	0.050	0.027	0.074	0.038	0.102	0.056	0.050	0.027	0.074	0.038	0.102	0.057	0.067	0.031	0.087	0.040	0.112	0.058
MSFT	0.083	0.059	0.086	0.061	0.098	0.067	0.044	0.027	0.054	0.033	0.077	0.045	0.056	0.024	0.060	0.030	0.077	0.043
MSFT*	0.039	0.022	0.051	0.029	0.075	0.043	0.038	0.021	0.050	0.029	0.075	0.042	0.054	0.020	0.058	0.027	0.077	0.041
MSTR	0.227	0.149	0.244	0.162	0.277	0.189	0.156	0.097	0.193	0.123	0.244	0.163	0.256	0.161	0.255	0.173	0.296	0.201
MSTR*	0.121	0.068	0.169	0.100	0.227	0.146	0.122	0.069	0.171	0.101	0.229	0.147	0.371	0.194	0.435	0.224	0.462	0.249
NFLX	0.115	0.075	0.125	0.082	0.146	0.099	0.082	0.048	0.102	0.060	0.134	0.085	0.081	0.042	0.099	0.054	0.134	0.080
NFLX*	0.062	0.034	0.092	0.049	0.128	0.076	0.063	0.036	0.093	0.050	0.128	0.077	0.076	0.033	0.097	0.047	0.132	0.074
NOC	0.077	0.055	0.081	0.058	0.096	0.065	0.047	0.028	0.056	0.034	0.082	0.047	0.061	0.037	0.065	0.042	0.092	0.054
NOC*	0.046	0.027	0.055	0.033	0.081	0.046	0.044	0.026	0.054	0.032	0.081	0.046	0.050	0.028	0.056	0.033	0.087	0.047
NOW	0.108	0.071	0.116	0.077	0.132	0.088	0.076	0.043	0.094	0.054	0.121	0.074	0.105	0.068	0.118	0.075	0.134	0.089
NOW*	0.068	0.037	0.091	0.049	0.119	0.070	0.068	0.037	0.090	0.049	0.118	0.071	0.079	0.038	0.106	0.051	0.129	0.072
NTES	0.087	0.055	0.097	0.062	0.111	0.073	0.066	0.037	0.083	0.048	0.105	0.065	0.089	0.056	0.098	0.063	0.114	0.075
NTES*	0.062	0.034	0.082	0.046	0.104	0.063	0.062	0.034	0.081	0.046	0.103	0.063	0.071	0.036	0.091	0.048	0.111	0.065
NVDA	0.108	0.070	0.115	0.075	0.130	0.086	0.074	0.042	0.090	0.053	0.118	0.072	0.108	0.070	0.119	0.076	0.134	0.088
NVDA*	0.057	0.030	0.081	0.043	0.113	0.065	0.058	0.031	0.082	0.044	0.113	0.065	0.077	0.034	0.102	0.047	0.127	0.068
PANW	0.110	0.070	0.119	0.076	0.137	0.090	0.082	0.044	0.100	0.055	0.132	0.078	0.108	0.067	0.123	0.075	0.144	0.092
PANW*	0.070	0.037	0.095	0.049	0.127	0.073	0.070	0.037	0.095	0.050	0.129	0.074	0.077	0.039	0.113	0.051	0.144	0.075
PYPL	0.102	0.066	0.112	0.073	0.130	0.086	0.074	0.041	0.093	0.053	0.124	0.075	0.100	0.065	0.115	0.073	0.132	0.088
PYPL*	0.063	0.034	0.089	0.048	0.121	0.072	0.063	0.035	0.089	0.048	0.121	0.072	0.075	0.037	0.108	0.050	0.134	0.074
REGN	0.080	0.050	0.084	0.053	0.097	0.060	0.059	0.031	0.068	0.037	0.086	0.049	0.078	0.041	0.074	0.045	0.109	0.056
REGN*	0.054	0.028	0.065	0.035	0.084	0.047	0.054	0.028	0.064	0.035	0.084	0.047	0.068	0.029	0.065	0.035	0.103	0.048
RH	0.150	0.094	0.174	0.109	0.214	0.142	0.116	0.063	0.156	0.084	0.221	0.133	0.153	0.096	0.180	0.112	0.224	0.148
RH*	0.097	0.053	0.169	0.078	0.227	0.126	0.098	0.054	0.171	0.078	0.229	0.128	0.161	0.069	0.202	0.089	0.264	0.140
ROKU	0.164	0.100	0.189	0.117	0.218	0.142	0.131	0.075	0.172	0.098	0.214	0.132	0.169	0.106	0.197	0.124	0.227	0.149
ROKU*	0.097	0.054	0.161	0.083	0.207	0.121	0.098	0.057	0.162	0.085	0.208	0.123	0.184	0.085	0.247	0.107	0.275	0.147
SHOP	0.137	0.087	0.146	0.095	0.167	0.109	0.098	0.059	0.116	0.072	0.150	0.095	0.121	0.077	0.134	0.086	0.164	0.104
SHOP*	0.075	0.040	0.100	0.057	0.138	0.084	0.076	0.042	0.101	0.059	0.139	0.085	0.108	0.051	0.125	0.066	0.159	0.091
SQ	0.138	0.086	0.151	0.095	0.178	0.116	0.099	0.055	0.126	0.072	0.167	0.102	0.141	0.089	0.156	0.099	0.185	0.121
SQ*	0.081	0.043	0.119	0.064	0.163	0.096	0.081	0.044	0.120	0.064	0.162	0.096	0.131	0.062	0.178	0.080	0.218	0.111
TSLA	0.179	0.116	0.192	0.125	0.219	0.146	0.119	0.071	0.146	0.088	0.190	0.123	0.107	0.061	0.133	0.079	0.193	0.114
TSLA*	0.077	0.040	0.117	0.066	0.172	0.108	0.076	0.040	0.117	0.066	0.171	0.108	0.098	0.047	0.130	0.070	0.199	0.112
TTD	0.156	0.094	0.179	0.109	0.210	0.133	0.124	0.068	0.163	0.090	0.210	0.126	0.159	0.095	0.186	0.112	0.218	0.138
TTD*	0.097	0.053	0.156	0.080	0.205	0.118	0.098	0.055	0.156	0.081	0.206	0.120	0.163	0.074	0.217	0.095	0.255	0.135
TWLO	0.156	0.095	0.174	0.107	0.199	0.126	0.123	0.070	0.154	0.088	0.191	0.116	0.162	0.101	0.181	0.115	0.206	0.134
TWLO*	0.101	0.054	0.144	0.076	0.185	0.106	0.101	0.056	0.145	0.078	0.186	0.108	0.162	0.071	0.192	0.088	0.223	0.119
ULTA	0.126	0.075	0.138	0.083	0.168	0.104	0.094	0.049	0.117	0.061	0.164	0.092	0.124	0.068	0.138	0.078	0.174	0.103
ULTA*	0.075	0.040	0.112	0.055	0.158	0.086	0.076	0.040	0.112	0.056	0.159	0.087	0.099	0.043	0.132	0.057	0.169	0.087
V	0.097	0.067	0.100	0.068	0.113	0.075	0.055	0.031	0.063	0.036	0.091	0.050	0.098	0.070	0.101	0.072	0.115	0.078
V*	0.050	0.027	0.060	0.033	0.089	0.047	0.049	0.026	0.058	0.032	0.088	0.047	0.061	0.036	0.070	0.041	0.098	0.053
W	0.188	0.117	0.214	0.133	0.258	0.166	0.140	0.081	0.185	0.104	0.253	0.154	0.189	0.115	0.217	0.133	0.271	0.171
W*	0.114	0.065	0.177	0.094	0.248	0.145	0.114	0.066	0.178	0.095	0.249	0.147	0.251	0.111	0.293	0.132	0.370	0.186

Notes: Notes: RMSE and MAE values for the BS, AHBS, CW models and their NNC counterparts for a daily prediction exercise for different underlying assets, denoted by their ticker. The shaded rows marked with an asterisk (*) correspond to the NNC models.

5.2 Long-term Predictions

Table 5.9 presents the RMSE and MAE values for long-term predictions for the last two years of the sample, based on the first three years of the sample. The NNC models incorporating time-varying variables are marked with an asterisk (NNC*). The results reveal that the NNC models exhibit lower RMSE and MAE values compared to the original parametric models. It is important to note that the differences in error metrics are slightly reduced, and the errors are generally higher than those observed in the daily interpolation and prediction exercises. This observation is expected, as the models are required to forecast further into the future, increasing uncertainty.

Furthermore, the NNC* models, which include time-varying parameters, achieve lower error metrics than both the parametric models and normal NNC models. However, the difference in RMSE and MAE values between the NNC and NNC* models are not large and not statistically significant. This implies that including the VIX level and the return of the underlying asset in the neural network only slightly enhances performance compared to the classic NNC model in this specific research setup.

Table 5.9: Average Error Evaluation Long-Term Prediction

	BS	**BS-NNC**	**BS-NNC***	**AHBS**	**AHBS-NNC**	**AHBS-NNC***
RMSE	0.156	0.114	0.111	0.140	0.114	0.109
MAE	0.102	0.075	0.070	0.090	0.075	0.070

Notes: Root Mean Squared Errors (RMSE) and Mean Absolute Errors (MAE) for the Black-Scholes (BS) model, the ad-hoc Black-Scholes (AHBS) model and their Neural Network Corrected (NNC) counterparts for a long-term prediction exercise. The asterisk (*) for the NNC models indicates that time-varying parameters are incorporated in the neural network. The reported values are averages over fifty different underlying assets.

Table 5.10 displays the DM statistics comparing the parametric models with their NNC counterparts. The positive and significant DM statistics indicate that, even over extended periods, the NNC models significantly outperform the parametric models. This holds true for both the standard NNC and NNC* models.

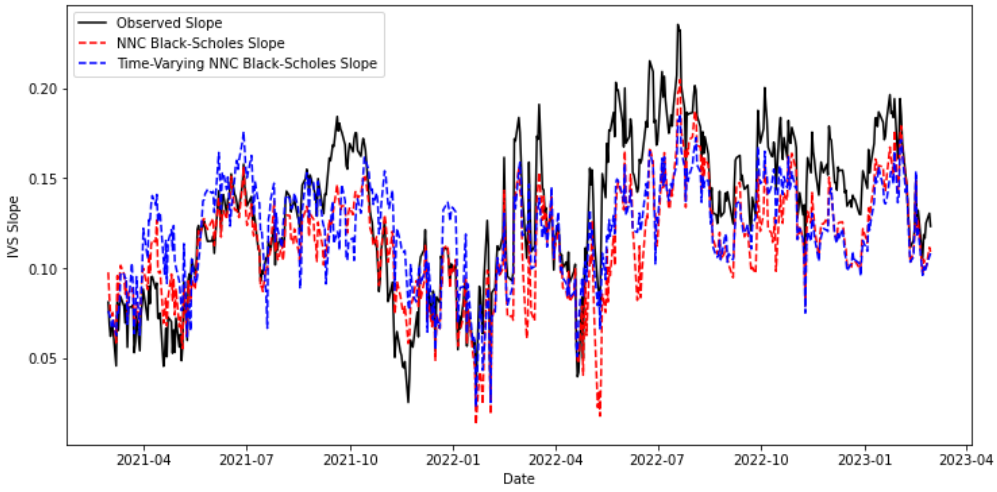
Table 5.10: Diebold-Mariano Test Statistic Long-Term Prediction

	BS	**AHBS**
NNC	507***	344***
NNC*	534***	359***

*Notes: Diebold-Mariano (DM) test statistics for the comparison of the Black-Scholes (BS), ad-hoc Black-Scholes (AHBS) and Carr-Wu (CW) models with their Neural Network Corrected (NNC) counterparts for a long-term prediction exercise. The asterisk (*) for the NNC models indicates that time-varying parameters are incorporated in the neural network. *** indicates that the P-value is below 0.001*

Figure 4 offers a more detailed comparison between the NNC and NNC* enhancements for a baseline BS model, while figure 5, located in the appendix, illustrates this comparison for the AHBS model. This figure is included to offer additional insights into why the time-varying NNC* model does not significantly outperform the standard NNC model. It presents the slopes of the volatility surfaces estimated by the different NNC extensions, compared to the slope observed in the data. In both figures, the blue line represents the standard NNC model, whereas the red line denotes the time-varying NNC* model.

Figure 4: Long-Term Prediction Daily IVS slope



Notes: This graph shows the daily moneyness slope of implied volatility surfaces with respect to moneyness. The observed slope is compared to the slopes constructed by twice correcting the Black-Scholes model, once with a Neural Network that incorporates time-varying inputs and once with a Neural Network without time-varying inputs.

The slopes are closely aligned, suggesting that the standard NNC model effectively captures the majority of the underlying market dynamics. However, there are moments where the

slopes diverge, with certain outliers present in the NNC model but not in the time-varying NNC* model. These deviations imply that the inclusion of time-varying inputs offers some additional flexibility in capturing subtle changes in market dynamics, although the overall impact remains relatively small. This finding is consistent with the results presented in table 5.9, which show that the time-varying NNC* model yields only small reductions in RMSE and MAE values compared to the NNC versions for both the BS and AHBS models.

To conclude, these results confirm that the neural networks consistently outperform traditional parametric models, even when applied to longer time horizons. The smaller differences and higher error metric values compared to the daily interpolation and prediction exercises reflect the increased difficulty and uncertainty that is known in long-term forecasting. Nonetheless, the significant DM statistics underscore the robustness of the NNC models, demonstrating their superior performance over the standard parametric models across extended time horizons. The incorporation of time-varying parameters yields only modest improvements in predictive performance, indicating that while such parameters can enhance prediction accuracy, their impact is limited in this specific setup.

6 Conclusion

In this thesis, I explored the application of neural networks to enhance the predictive accuracy of IV values of equity options. The main objective was to determine the extent to which NNC models improve the forecasting accuracy of IVS levels compared to their uncorrected parametric counterparts. This study aimed to address the limitations of parametric models, particularly their inability to capture non-linear relationships, by using neural networks.

The findings of this research indicate that NNC models significantly outperform their parametric counterparts in terms of predictive accuracy. Specifically, a daily interpolation exercise showed an average decrease in RMSE and MAE of 35% across all models. For one day, one week and one month ahead IV predictions, the RMSE and MAE values dropped 22% on average. Long-term predictions, designed to assess the stability of the NNC models over time, also showed an average decrease of 22% for RMSE and MAE values. These results are consistent across options with various underlying assets, regardless of their liquidity levels.

Diebold-Mariano test statistics confirmed the significance of these results, further underlining the robustness of the NNC approach.

The practical implications of these results are significant for traders and investors in financial markets. With more accurate IVS models, market participants can make better-informed decisions, allowing for better speculative and hedging techniques. The improvements to these models increase the accuracy of option pricing and risk management, providing a competitive edge in market analysis. Furthermore, the application of neural networks to model error terms is not limited to financial markets. This approach shows potential to increase predictive accuracy across a wide range of disciplines, highlighting the potential and versatility of machine learning techniques.

This paper provides significant contributions to the existing literature on IVS modeling, particularly by focusing on equity option IVS modeling rather than the more commonly studied IVS models for index options. Additionally, it advances our understanding of non-linear IVS dynamics by effectively modeling them using neural networks. The innovative approach of using neural networks to model errors preserves the theoretical foundations of the parametric models while improving their predictive accuracy. This methodology not only improves the performance of IVS models, but it also has the potential to be applied to studies in other disciplines.

Despite the promising results of this research, there are some limitations that can be addressed in further investigations. One limitation is related to the lower liquidity of options on certain underlying assets, resulting in fewer observations of these options. This issue arose due to the inclusion of options on a wide range of underlying assets, a deliberate choice to evaluate whether the NNC procedure improved predictions across a wide range of options. However, this decision led to the use of less restrictive filters to maintain a sufficiently large sample size. While this research demonstrated that NNC models improve IVS predictions of options on different underlying assets, a more selective panel could be beneficial for a deeper analysis. A focused selection would allow for a more thorough investigation into the reasons why the NNC procedure sometimes produces better results, an aspect not fully explored in this thesis due to its broad scope and time constraints.

Another observation from this study is the similar performance across NNC models with different parametric model baselines. While this paper successfully demonstrated how neural networks improved prediction accuracy of parametric models, it did not thoroughly examine the differences between the NNC models themselves. This raises the interesting question of how much the choice of the uncorrected parametric model impacts the accuracy of IVS predictions. This finding suggests a potential direction for further research: examining which models serve as the most effective baselines for NNC enhancements. Additionally, exploring the conditions under which the NNC procedure yields the most significant improvements could provide valuable insights. Identifying the types of parametric models that benefit the most from NNC enhancements could be valuable information for the application of this methodology in various other fields of research.

Overall, this research highlights the significant impact neural networks can have in the field of IVS modeling. The combination of traditional parametric models with state-of-the-art machine learning techniques presents a unique way to enhance predictive accuracy while preserving underlying model assumptions. In the ever-involving landscape of financial markets, the ability to leverage such hybrid methodologies is crucial for developing robust predictive models. These findings can be extended beyond the realm of option pricing and financial modeling to any discipline that relies on statistical modeling. The promising results encourage further exploration of these techniques, suggesting that the combination of traditional and modern approaches can lead to significant improvements in predictive analytics.

References

Aït-Sahalia, Y., & Lo, A. (1998). Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *Journal of Finance*, *53*(2), 499–547. https://doi.org/10.1111/0022-1082.215228

Almeida, C., Fan, J., Freire, G., & Tang, F. (2023). Can a Machine Correct Option Pricing Models? *Journal of Business and Economic Statistics*, *41*(3), 995–1009. https://doi.org/10.1080/07350015.2022.2099871

Andersen, T. G., Fusari, N., & Todorov, V. (2015). The risk premia embedded in index options. *Journal of Financial Economics*, *117*(3), 558–584. https://doi.org/https://doi.org/10.1016/j.jfineco.2015.06.005

Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *The Review of Financial Studies*, *16*(1), 101–143. http://www.jstor.org/stable/1262727

Bali, T. G., Beckmeyer, H., Moerke, M., & Weigert, F. (2021). Option return predictability with machine learning and big data. https://ideas.repec.org/p/zbw/cfrwps/2108.html

Bates, D. S. (1991). The Crash of '87: Was It Expected? The Evidence from Options Markets. *Journal of Finance*, *46*(3), 1009–44. https://EconPapers.repec.org/RePEc:bla:jfinan:v:46:y:1991:i:3:p:1009-44

Bernales, A., & Guidolin, M. (2014). Can we forecast the implied volatility surface dynamics of equity options? Predictability and economic value tests. *Journal of Banking Finance*, *46*, 326–342. https://doi.org/https://doi.org/10.1016/j.jbankfin.2014.06.002

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

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business*, *51*(4), 621–651. http://www.jstor.org/stable/2352653

Brownlees, C. T., Engle, R. F., & Kelly, B. T. (2011). A Practical Guide to Volatility Forecasting through Calm and Storm. https://doi.org/10.2139/ssrn.1502915

Cao, J., Chen, J., & Hull, J. (2020). A Neural Network Approach to Understanding Implied Volatility Movements. *Quantitative Finance*, *20*(9), 1405–1413. https://doi.org/10.1080/14697688.2020.1750679

Carr, P., & Wu, L. (2016). Analyzing volatility risk and risk premium in option contracts: A new theory. *Journal of Financial Economics*, *120*(1), 1–20. https://doi.org/https://doi.org/10.1016/j.jfineco.2016.01.004

Cont, R., & Da Fonseca, J. (2002). Dynamics of implied volatility surfaces. *Quantitative Finance*, *2*(1), 45–60. https://EconPapers.repec.org/RePEc:taf:quantf:v:2:y:2002:i:1:p:45-60

Derman, E., & Kani, I. (1994). The Volatility Smile and its Implied Tree. *Goldman Sachs Quantitative Strategies Research Notes*, *2*, 45–60.

Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business Economic Statistics*, *13*(3), 253–263.

Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied Volatility Functions: Empirical Tests. *The Journal of Finance*, *53*(6), 2059–2106. http://www.jstor.org/stable/117461

François, P., Galarneau-Vincent, R., Gauthier, G., & Godin, F. (2022). Venturing into uncharted territory: An extensible implied volatility surface model. *Journal of Futures Markets*, *42*(10), 1912–1940.

Gonçalves, S., & Guidolin, M. (2006). Predictable Dynamics in the SP 500 Index Options Implied Volatility Surface. *The Journal of Business*, *79*(3), 1591–1635. http://www.jstor.org/stable/10.1086/500686

Goyal, A., & Saretto, A. (2009). Cross-section of option returns and volatility. *Journal of Financial Economics*, *94*(2), 310–326. https://doi.org/https://doi.org/10.1016/j.jfineco.2009.01.001

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Hamid, S. A., & Iqbal, Z. (2004). Using neural networks for forecasting volatility of S&P 500 Index futures prices. *Journal of Business Research*, *57*(10), 1116–1125. https://doi.org/https://doi.org/10.1016/S0148-2963(03)00043-2

Heston, S. L., & Nandi, S. (2000). A Closed-Form GARCH Option Valuation Model. *The Review of Financial Studies*, *13*(3), 585–625. http://www.jstor.org/stable/2645997

Hutchinson, J. M., Lo, A., & Poggio, T. (1994). A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks. (4718). https://EconPapers.repec.org/RePEc:nbr:nberwo:4718

Klemkosky, R. C., & Resnick, B. G. (1979). Put-Call Parity and Market Efficiency. *The Journal of Finance*, *34*(5), 1141–1155. http://www.jstor.org/stable/2327240

Li, H., Li, J., Guan, X., Liang, B., Lai, Y., & Luo, X. (2019). Research on Overfitting of Deep Learning. *2019 15th International Conference on Computational Intelligence and Security (CIS)*, 78–81. https://doi.org/10.1109/CIS.2019.00025

OptionMetrics. (2024). Ivy DB US [data set] [Retrieved from Wharton Research Data Services]. https://wrds-web.wharton.upenn.edu/wrds

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Rubinstein, M. (1994). Implied Binomial Trees. *Journal of Finance*, *49*(3), 771–818. https://EconPapers.repec.org/RePEc:bla:jfinan:v:49:y:1994:i:3:p:771-818

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117. https://doi.org/https://doi.org/10.1016/j.neunet.2014.09.003

Ulrich, M., Zimmer, L., & Merbecks, C. (2023). Implied volatility surfaces: a comprehensive analysis using half a billion option prices. *Review of Derivatives Research*, *26*(2), 135–169.

Wharton Research Data Services. (n.d.). *"WRDS"*. Retrieved February 8, 2024, from wrds.wharton.upenn.edu

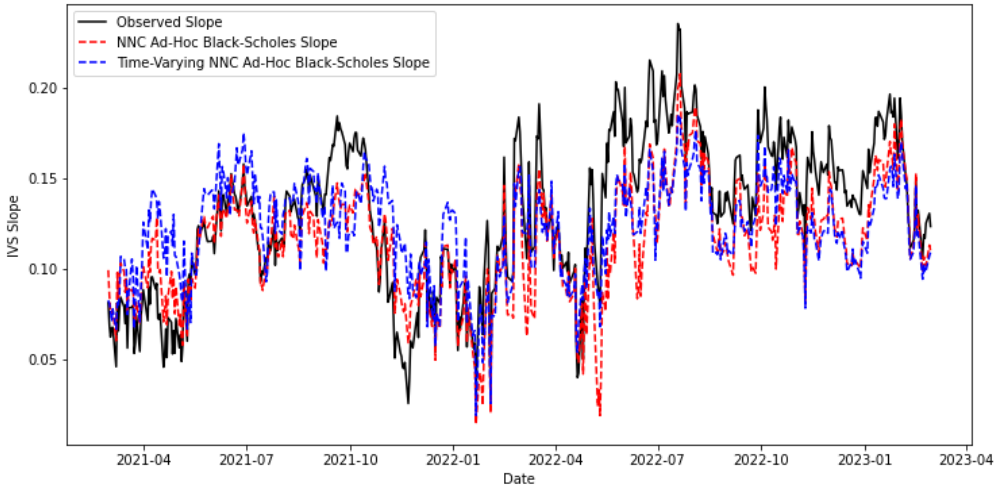
Appendix

Table 6.1: Most Traded Equity Options, Unfiltered

Rank	Ticker	Observations	Ranking	Ticker	Observations
1	AMZN	6,641,431	26	AVGO	1,611,456
2	TSLA	5,443,713	27	BLK	1,604,305
3	BKNG	5,005,834	28	COST	1,600,691
4	GOOGL	4,626,783	29	GS	1,596,795
5	GOOG	4,323,198	30	RH	1,529,967
6	SHOP	3,127,637	31	MA	1,522,766
7	CMG	3,089,316	32	ULTA	1,510,315
8	NFLX	2,840,092	33	LMT	1,490,225
9	NVDA	2,536,050	34	TWLO	1,477,142
10	MELI	2,326,765	35	SQ	1,475,303
11	META	2,253,784	36	HD	1,462,912
12	AZO	2,136,767	37	ALGN	1,450,013
13	AAPL	2,096,494	38	HUM	1,445,553
14	ISRG	2,059,543	39	W	1,412,516
15	BA	1,995,370	40	MSTR	1,404,711
16	ADBE	1,985,606	41	NTES	1,386,760
17	LRCX	1,967,019	42	MDB	1,371,363
18	BABA	1,914,413	43	ILMN	1,370,818
19	MSFT	1,817,530	44	PYPL	1,362,979
20	REGN	1,807,625	45	CRM	1,361,847
21	ROKU	1,705,308	46	V	1,356,189
22	TTD	1,696,684	47	AMD	1,354,024
23	CHTR	1,644,907	48	NOC	1,353,486
24	BIIB	1,639,627	49	PANW	1,339,581
25	NOW	1,638,221	50	LULU	1,329,337

Notes: This table presents the number of options on the different underlying assets present in the Wharton Research Data Services (n.d.) dataset before any filters. The assets are denoted by their ticker and are sorted based on the number of observations.

Figure 5: Long-Term Prediction Daily IVS slope - AHBS



Notes: Notes: This graph shows the daily moneyness slope of implied volatility surfaces with respect to moneyness. The observed slope is compared to the slopes constructed by twice correcting the ad-hoc Black-Scholes model, once with a Neural Network that incorporates time-varying inputs and once with a Neural Network without time-varying inputs.