Tilburg School of Economics and Management - TiSEM

Forecast Errors on the Implied Volatility of Equity Options

Bachelor Thesis

Student name: Joram Bakker
ANR: U176687
Date: May 24, 2024
Supervisor name: Hong Phuoc (Michael) Vo
Number of words: 6048

Abstract

This thesis investigates the rational pricing of S&P 500 index options from 1996 to 2022. This is done by first using the Stein (1989) method to investigate the dynamics of the mean reversion of implied variances. Secondly, we will inspect the predictability of ex-post forecast errors of implied variance using ex-ante forecast revisions, and the impact of uncertainty on this relationship. We find that long-term options tend to overreact to changes in the implied variances of short-term options. Furthermore, we find that forecast errors are predictable from forecast revisions, suggesting a degree of information rigidity among investors. This implies a systematic bias where upward revisions to forecasts predict higher future realizations, indicating that forecast adjustments are too conservative, and causes underreaction for short-term options. By adding the uncertainty variable to the model, we can see that uncertainty significantly influences forecast revision and that a higher level of uncertainty results in a larger underestimation of forecasts These findings are further tested through a trading strategy based on forecast revisions. The strategy produces high average monthly returns. However, the high volatility suggests that while the strategy is profitable, it comes with significant risk.

Table of contents

1. Introduction

Equity options are financial derivatives that give the holder the right to buy or sell a specified amount of a stock at a predetermined price at the option's expiration date. They are a way for investors to speculate on or hedge against the future price movements of a stock without directly buying or selling the stock itself. However, the price of an option is more difficult to determine. It depends on the volatility of the underlying, but this volatility is not observable and therefore the option market can be considered as a speculative market on volatility. Previous research have found that options are often irrationally priced due to forecast errors of investors.

This thesis will aim to answer the following research questions:
- *Are S&P 500 index options rationally priced in the period between 1996-2022?*
- *To what degree are ex-post forecast errors of implied volatility predictable with ex-ante forecast revision?*
- *Does the level of uncertainty influence the relationship between forecast errors and forecast revision in the option market?*

There are two main types of options: call options and put options. The former gives the holder the option to buy, while the latter gives the holder the right to sell the underlying asset. Since an option holder has the right, but not the obligation, to buy or sell the underlying asset, its payoff is at least zero or higher. Consequently, the buyer of an option pays a premium to the writer of the option. This premium is called the option price and can be calculated using various models, like the Black-Scholes model (Black & Scholes, 1973) and the Binomial model (Cox et al., 1979). These models provide a theoretical estimate of the price of options. Black and Scholes have shown that when the volatility of an asset is known, the price of an option is fully determined. However, because volatility is unknown and variable, options maintain a degree of independence from their associated stocks. Thus, options can be considered to represent a speculative market on volatility itself, which is indicated through the option's implied volatility. This implied volatility can be calculated by inverting the Black-Scholes formula and can be seen as the expected average volatility over the life of the option.

However, investors are not always rational. Stein (1989) found that investors overreact to the current information by ignoring the long-run mean reversion in implied volatility. Coibion and

Gorodnichenko (2015) have provided a framework that regresses ex-post mean forecast errors on ex-ante forecast revisions to test the degree of information rigidity of investors. In this thesis I will extend this framework by interacting forecast revision with the real uncertainty index data which uses the methodology described in Jurado et al. (2015) and Ludvigson et al. (2021).

First, I will research whether options on the S&P 500 are mispriced, using the framework of Stein (1989). After that I will look at the degree of information rigidity from investors using the framework of Coibion and Gorodnichenko (2015). Then, I will extend the framework of Coibion and Gorodnichenko with the variable *uncertainty* to research whether this variable has an interaction effect between the ex-post forecast errors and ex-ante forecast revision of investors. At last, I will test the performance of a trading strategy based on these findings, to check if it can create significant returns.

2. Literature overview

The pricing of options

The work of Black and Scholes (1973) and Cox et al. (1979) laid the foundation for understanding the pricing of options through their models. Black and Scholes (1973) introduced a model that revolutionized the field by linking the price of an option to the underlying asset's volatility, assuming it is known and constant. Cox et al. (1979) further contributed by offering a more flexible binomial model that accommodates various market conditions and time steps. However, the real-world application of these models faces the issue of volatility being unobservable and variable over time. Thus, implied volatility, which can be derived by inverting the Black-Scholes formula, represents the market's expectation of the future volatility of the underlying asset. Black and Scholes (1973) provide us with the following model, which is essential to understand the pricing of options.

Black-Scholes model:

$$Call = S * N(d_1) - PV(K) * N(d_2),$$

$$Put = S * (N(d_1) - 1) + PV(K) * (1 - N(d_2).$$

Where,

$$d_1 = \frac{\ln(\frac{S}{PV(K)})}{\sigma\sqrt{T}}, \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{T}.$$

As we can see from this formula, volatility (σ) is a key determinant of option prices. Higher volatility increases the theoretical price of both call and put options. The logic behind this is that, due to the bigger price movements of the underlying asset, the holder of an option has a higher chance that the option will become in the money. Note that every other variable in this model is known. Because of this, the option market can be considered as a speculative market on volatility itself.

Mispricing of options

The method introduced by Stein (1989) focuses on the term structure of options' implied volatilities using data on S&P 100 index options. Implied volatility is strongly mean-reverting. This means that the implied volatility of longer-maturity options should move less in response to a given change in the implied volatility of shorter-maturity options. However, Stein discovered that the reaction of long-maturity options' implied volatilities to changes in short-maturity options' implied volatilities is larger than expected, indicating that long-maturity options tend to overreact.

Poteshman (2002) builds upon the work of Stein (1989) and his aim is to investigate whether options market investors overreact or underreact to the information contained in daily fluctuations in instantaneous variance over short horizons. He found that investors tend to underreact in the short horizon. He also checked whether the long-horizon overreaction found by Stein (1989) is also present in the S&P 500 index (SPX) options market. He found that the results of Stein (1989) of long-horizon overreaction also applies to the SPX options market over the period between 1988 and 1997.

These findings about underreaction in the short term and overreaction in the long term are also supported by Barberis et al. (1998). They suggest that investors are initially showing conservatism, which leads to underreaction to new information. Over time, this underreaction is replaced by overreaction, due to the representativeness heuristic bias.

Forecast errors and information rigidity

The concept of over- and under-reaction is linked to forecast errors, where the forecasted value deviates from the realized value. The model of information rigidity by Coibion and Gorodnichenko (2015) provides a theoretical framework to understand how forecast errors persist due to agents not fully incorporating new information. They based their model on two theoretical rational expectations models of information frictions. The first one is the sticky-information model of Mankiw and Reis (2002)**,** where agents update their information sets infrequently due to costs associated with acquiring new information. The degree of information rigidity in this model is then the probability of not acquiring new information each period. The second type of model they considered is a noisy-information model such as the models of Sims (2003) and Mackowiak and Wiederholt (2009). Here, agents continuously update their information sets but the information they receive is imperfect or noisy. This means that while agents receive new data regularly, the data is not entirely accurate, and they must filter out the noise to form their expectations. Both models predict a similar relationship between the mean ex post forecast errors among agents and the mean ex ante forecast revisions, with the coefficient on forecast revisions being just influenced by the level of information rigidity in each model. Coibion and Gorodnichenko (2015) combined these two models into a generalized model and their methodology allows for an empirical study into the dynamics of forecast revisions and their relationship with forecast errors.

The paper by Bordalo et al. (2020) examines the rationality of both individual and consensus forecasts about macroeconomic and financial variables. The study uses the methodology by Coibion and Gorodnichenko (2015) to explore the predictability of forecast errors from forecast revisions. They find that individual forecasters typically overreact to new information while consensus forecasts tend to underreact compared to full-information rational expectations.

Uncertainty

Caldara et al. (2016) found that uncertainty shocks negatively impact economic activity. These shocks result in reduced investment and hiring by firms due to increased risk and uncertainty about future economic conditions. The study found that uncertainty shocks can lead to significant declines in industrial production and employment. They also found that heightened uncertainty is associated with sharp declines in stock market valuations. This is because uncertainty affects investor confidence, leading to increased volatility and reduced stock prices. Jurado et al. (2015) introduces new methods for measuring uncertainty. It aims to provide more accurate estimates of uncertainty that are not overly dependent on specific theoretical models or a limited set of economic indicators. This will provide us with an uncertainty index which we can use in our analysis.

Relevance of this thesis

Understanding the accuracy of financial forecasts and the factors influencing forecast errors is crucial in the field of finance. This thesis explores the predictive power of uncertainty on the degree of ex-post forecast errors of investors. Firstly this research contributes to the literature on options markets by investigating the potential mispricing of S&P 500 index options using Stein's (1989) framework with recent data. Secondly, the inspection of information rigidity among investors using the framework of Coibion and Gorodnichenko (2015) adds to our understanding of behavioral finance. Particularly how investors process and react to new information. Finally, by extending the Coibion and Gorodnichenko framework to include the uncertainty variable, this study innovates in inspecting how uncertainty itself influences investor forecasts revision. Understanding whether and how the level of uncertainty can predict the ex-post forecast errors offers new insights into rational pricing .This could potentially lead to more robust forecasting techniques and thereby reduce the future forecast errors.

3. Methodology

The methodology used in this thesis is mostly based on the papers of Stein (1989) and Coibion and Gorodnichenko (2015). I will extend the framework of Coibion and Gorodnichenko to research the predicting power of the variable *uncertainty*.

3.1 Proving mispricing by the method of Stein

In this section I will use the methodology of Stein to research whether there is found mispricing in the given period. The main idea is to regress the spread between the implied variance of the distant option (i_t^d) and the average historical variance $(\overline{\sigma})$ to the spread between the implied variance of the near option (i_t^n) and the average historical variance $(\overline{\sigma})$ (see equation (1)).

$$\left(i_t^d - \overline{\sigma}\right) = \beta_i(i_t^n - \overline{\sigma}) + c. \tag{1}$$

To explain the concept behind this test we will consider the following. Assume that volatility typically averages around 20%, but it tends to fluctuate up and down rather quickly, following a strong mean-reverting process. If a one-month option now has an implied volatility of 25%, the implied volatility for a two-month option should be a bit lower, depending on the rate of mean reversion. On the other hand, if a one-month option has an implied volatility of 10%, the two-month option's implied volatility should be higher. The expectation is that the implied volatility should mean-revert. This would mean that changes in the implied volatility of near options would lead to smaller changes in the implied volatility of distant options.

β_i is the elasticity of implied volatility between the distant option and the nearby option and if it exceeds the theoretically β_{th}, mispricing is found. β_{th} can be calculated using a linear endpoint approximation (see equation (2))

$$\beta \simeq \frac{(1+\rho^4)}{2}. \tag{2}$$

Where ρ is the mean reversion parameter, which measures the autocorrelation on the implied volatility of the nearby options. The rate of mean reversion refers to how quickly the volatility reverts to its long-term average. This ρ can be determined by producing an autocorrelogram of the near implied volatility with 4 different lag lengths. The lag length will be ranging between 1 and 4 weeks. The autocorrelation of these different lag lengths *n* will be calculated using an AR(1) model:

$$i_t = \rho(i_{t-n}). \tag{3}$$

To compare the 4 different ρ's we must raise it to the $1/n$ power so that we get a weekly ρ. We will take the average of the 4 values and then use this value of ρ in equation (2) to calculate the theoretical value of β_{th}. So, if β_i in the regression of equation (1) turns out to be higher than β_{th} it would mean that we have found overreaction.

3.2 Forecasting errors and information rigidity of implied variance

Assuming that mispricing is proven in the previous section, we will look in this section why mispricing is there. Coibion and Gorodnichenko (2015) have provided a framework to test the full-information rational expectations (FIRE) hypothesis. This hypothesis assumes that economic agents make forecasts about future economic variables using all available information without any biases, and their predictions, on average, match actual outcomes. This hypothesis assumes agents process all relevant data perfectly to form their expectations, leading to efficient markets and optimal decision-making based on these accurate forecasts.

Coibion and Gorodnichenko (2015) tested this by examining the relationship between forecast errors and forecast revisions. If forecast errors are predictable by forecast revisions, it indicates that information rigidity exists, and agents don't fully process all new information. Since the implied variance of an option can be seen as a forecast of the average variance over the life of the option, we can implement it in the model of CG (2015). The degree of information rigidity is tested using the following equation:

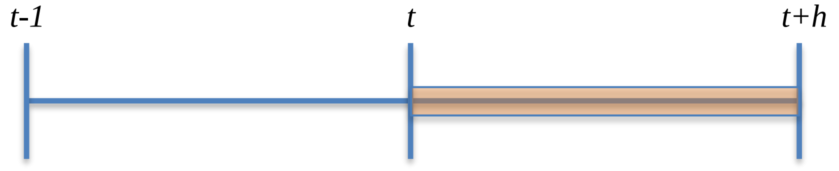
$$x_{t+h} - F_t x_{t+h} = c + \beta_1 (F_t x_{t+h} - F_{t-1} x_{t+h}) \quad (4)$$

Where "$x_{t+h} - F_t x_{t+h}$" is the forecast error and "$(F_t x_{t+h} - F_{t-1} x_{t+h})$" is the forecast revision. Underneath I will provide an explanation of the variables used in equation (4) in the context of implied variance:

- x_{t+h}: The realized variance over period *t+h*.
- $F_t x_{t+h}$: Forecast of variance over period *t+h* made at time *t*

  This is the implied variance of the most near (30 days) option made at time *t*.
- $F_{t-1} x_{t+h}$: Forecast of variance over period *t+h* made at time *t-1*

  This variable is however harder to calculate. Figure 1 is used as an illustration.



(Figure 1)

We want to calculate the forecast of the orange period from figure 1 made at time *t-1*. Because there is no such option that only covers this orange period at time *t-1*, we must do some calculations. At time t-1 our forecast of the variance of the orange area is a weighted average between the implied variance of the near option and the implied variance of the distant option. We will use equation (5) for this calculation:

$$F_{t-1}x_{t+h} = \frac{B}{B-A} * (i^{d}_{t-1}) - \frac{A}{B-A} * (i^{n}_{t-1}) \tag{5}$$

Where A = days till maturity of nearby option and B = days till maturity of distant option. For example: at time t-1 the implied variance of the near option (30 days till maturity) with expiration date *t* is equal to 15%, and the implied variance of the distant option (60 days till maturity) with expiration date *t+h* is equal to 20%. In this case the implied variance of the orange area should be: $\frac{60}{60-30} * 0.2 - \frac{30}{60-30} * 0.15 = 25\%$.

Now we can test the relationship between forecast errors and forecast revision by regressing equation (4). Under full-information rational expectations (FIRE) the forecast error is unpredictable, and this β_1 should be 0. When this β_1 is positive, upward revisions will predict higher realizations compared to the forecasts, meaning that the forecast underreacts to information relative to FIRE. This implies that investors do not fully incorporate new information into their forecasts promptly, indicating the presence of information rigidities. When this beta is negative, upward forecast revisions will predict lower realizations compared to the forecasts, meaning that the forecast overreacts relative to FIRE. This suggests that agents may be overestimating the impact of new information or reacting too strongly to new data

After that I will extend equation (4) to test whether the variable $\mathrm{Uncertainty_t}$ has a moderating effect on forecast revisions. This will result in regression equation (6)

$$x_{t+h} - F_t x_{t+h} = c + \beta_1(F_t x_{t+h} - F_{t-1}x_{t+h}) + \beta_2(F_t x_{t+h} - F_{t-1}x_{t+h})\, Uncertainty_t + \beta_3 Uncertainty_t \tag{6}$$

Where β_1 still has the same interpretation as in equation (4) if *Uncertainty* is zero, however as shown in Table 2 (in section 4), Uncertainty is always positive. Thus β_1 alone does not fully describe the relationship between forecast revisions and forecast errors. Therefore, we should combine $(\beta_1 + \beta_2 * Uncertainty)$ to get the final coefficient for forecast revision.. If β_2 is significant different from 0, $\mathrm{Uncertainty_t}$ has an interaction effect between forecast errors and forecast revision. If β_2 is positive, it tells us that when uncertainty is increasing, upward forecast revision predicts an increasing underestimation of the forecast. Conversely, if this β_2 is negative, it means that when uncertainty is increasing, upward forecast revision predicts an increasing overestimation of the forecast. The coefficient β_3 shows us the standalone effect of uncertainty on forecast errors. To quantify the total effect of forecast revisions and uncertainty on forecast errors, we should take all the betas into consideration. However, we are particularly interested in β_2, because it describes how *Uncertainty* influences the forecast revision.

4. Data

In this thesis I will use the following data:

1) Volatility surface data of the S&P 500 from January 1996 to December 2022 with maturities of 30 and 60 days. This data contains the interpolated implied volatility of the S&P 500. I will use the most at-the-money options because these options are the most liquid and therefore probably not contain illiquidity premium. In this thesis I will work with variance instead of volatility because it will simplify calculations. The implied variance will be calculated using the average of the squared volatility of the most ATM put- and call-option. This data will be gathered from WRDS - OptionMetrics Ivy DB US.
2) S&P 500 daily closing prices from January 1996 to December 2022. This data will be gathered from WRDS - OptionMetrics Ivy DB US. This data is gathered because we need to calculate the moneyness of each option to filter the dataset to only include the most ATM put- and call-option at each date
3) Historical realized volatility data gathered from WRDS - OptionMetrics Ivy DB US.
4) *Uncertainty* Data gathered from the website of Sydney Ludvigson from January 1996 to December 2022

Data used in section: Proving mispricing by the method of Stein

In this part we need the implied variance of both the nearby and distant option series. As stated before, those implied variances are calculated using the average of the squared volatility of the most ATM put- and call-option. Therefore, we also need to have the closing price of the S&P 500 for each date to calculate the moneyness. The moneyness of an option is calculated by dividing the price of the S&P 500 by the strike price of the option, as shown by the formula: $moneyness = \frac{P}{K}$. We will filter the dataset to include only the options with moneyness closest to 1. We also need the historical variance, so that we can calculate the historical average.

Combining this data will give me a dataset containing the following variables:

- $Date(t)$: Date of the information.
- i_t^n: Implied variance at *Date* of the near ATM option with a maturity of 30 days.
- i_t^d: Implied variance at *Date* of the distant ATM option with a maturity of 60 days.
- $\sigma_{Realized}$: The realized variance from *Date* till Date + 30 days.

The summary statistics of this dataset are shown in Table 1

Table 1: Summary statistics of dataset for Stein (1989) part.

	N	Mean	SD	Min	Max
Date	6796			1996-01-04	2022-12-30
$\boldsymbol{i_t^n}$	6796	0.0452	0.0452	0.0071	0.5947
$\boldsymbol{i_t^d}$	6796	0.0459	0.0392	0.0010	0.4959
$\boldsymbol{\sigma_{Realized}}$	6796	0.0384	0.0727	0.0012	0.8833

Data used in section: Forecasting errors and information rigidity of implied variance

In this part we need the forecast of variance over period t+h made at time t. Note that this is the same as the implied variance of the nearby option. We also need the forecast of variance over period t+h made at time t-1. This is calculated by using equation (5) provided in section 3.. We also need the (30 days) realized variance at each date. We will also need the value of the uncertainty index.

Combining this data will give me a dataset containing the following variables:

- $Date(t)$*:* Date of the information.
- x_{t+h}: The realized variance over period *t+h*.
- $F_t x_{t+h}$: Forecast of variance over period *t+h* made at time *t*
- $F_{t-1} x_{t+h}$: Forecast of variance over period *t+h* made at time *t-1*
- $Uncertainty_t$: Real uncertainty index at *Date*

The summary statistics of this dataset are shown in Table 2.

Table 2: Summary statistics of dataset for CG (2015) part.

	N	Mean	SD	Min	Max
Date	6796			1996-01-04	2022-12-30
$\boldsymbol{F_t x_{t+h}}$	6796	0.0452	0.0452	0.0071	0.5947
$\boldsymbol{F_{t-1} x_{t+h}}$	6775	0.0466	0.0340	0.0121	0.4077
$\boldsymbol{x_{t+h}}$	6796	0.0384	0.0727	0.0012	0.8833
$\boldsymbol{Uncertainty_t}$	6796	0.6399	0.1305	0.5271	1.3776
Forecast error	6796	-0.0068	0.0567	-0.2731	0.8214
Forecast revision	6775	-0.0013	0.0344	-0.3703	0.5645

Note that some variables have 6775 observations instead of 6796. This is because these variables contain a (21 trading-days $\approx$ 30 calendar-days) lagged value. Since there is no data available before January 1996, the first 21 observations do not have a lagged value and are therefore omitted.
Recall that Forecast error is ($x_{t+h} - F_t x_{t+h}$), and Forecast revision is $(F_t x_{t+h} - F_{t-1} x_{t+h})$.

5. Results

5.1 Results of Stein (1989) method

First, we want to find the mean reversion parameter ρ. Table 3 presents the autocorrelogram for the near implied variance series for different lag lengths. The table indicates that the implied weekly ρ is almost equal to 0.9 for every lag length. This result is consistent to the mean reversion parameter ρ found in the paper of Stein (1989).

Table 3: Autocorrelogram for i_t^n series

Lag Length (weeks)	Autocorrelation*	Implied Weekly ρ
1	0.8864 (0.0056)	0.8864
2	0.8139 (0.0070)	0.9022
3	0.7430 (0.0081)	0.9057
4	0.6772 (0.0089)	0.9071

* Standard errors are in parentheses.
The implied weekly ρ is the autocorrelation raised to the 1/n power

The following step is to use this mean reversion parameter ρ to calculate the theoretical upper bound of beta. If we fit this ρ = 0.9 in equation (2) we will get a β_{th} that is equal to 0.8281. This means that the distant option should be at most 0.8281 percent above its mean, when the near option is 1 percent above its mean. We will check whether this is the case, by regressing equation (1). Table 4 presents the results of this regression. β_i is the elasticity of implied volatility between the distant option and the nearby option. We find that this β_i is equal to 0.8635. This is higher than the theoretical β_{th}, which means that we found overreaction for long-term options. For example: if the implied variance of the nearby option is currently 10%, and the average historical variance is 5%, then the implied variance of the distant option should theoretically be 9.14% (= 5 + 0.8281*(10-5)), but it tends to be too high at 9.32% (= 5 + 0.8635*(10-5)). So mispricing is found. This finding is consistent with the one of Stein (1989), who also found overreaction of long-maturity options.

Table 4: Regression results of equation (1): $\left(i_t^d - \overline{\sigma}\right) = \beta_i(i_t^n - \overline{\sigma}) + c.$

Variable	Coefficient	Standard Error	t-value	p-value
(Constant)	0.0006	(0.0001)	9.145	< 0.0001
$(\boldsymbol{i_t^n} - \boldsymbol{\overline{\sigma}})$	0.8635	(0.0015)	562.707	< 0.0001
Adj. $\boldsymbol{R^2}$	97.9%			
N	6796			

5.2 Results of CG (2015) method

Now that we have proven mispricing, we will look at why mispricing is present, by examining the relationship between forecast errors and forecast revisions. The summary statistics of the important variables used in this section are provided in Table 2. Under full-information rational expectations (FIRE) the forecast error should be unpredictable. To test whether this holds in the case of implied variance forecasts, we will regress equation (4).

Table 5 describes the results of this regression. We find that β_1 = 0.291 and it is significantly different from zero. This result suggests that investors do not adjust their forecasts accordingly in response to new information, leading to predictable forecast errors. In this case, upward forecast revisions will predict higher realizations compared to the forecasts, meaning that the forecast revision underreacts relative to FIRE. In practical terms, if there is an upward forecast revision of 1 percentage point, it will on average lead to an ex-post forecast error of 0.29 percentage points. Hence, the correct upward forecast revision should have been 1.29 percentage points in this case, instead of 1 percentage point. This finding is consistent with Lochstoer and Muir (2022) who also found underreaction in index volatility due to slow-moving beliefs of agents.

Table 5: Regression results of equation (4): $x_{t+h} - F_t x_{t+h} = c + \beta_1(F_t x_{t+h} - F_{t-1} x_{t+h})$

Variable	Coefficient	Standard Error	t-value	p-value
(Constant)	-0.0065	0.0007	-9.509	< 0.0001
Forecast revision	0.2910	0.0197	14.745	< 0.0001
Adj. $\boldsymbol{R^2}$	3.10%			
N	6774			

Note that $(F_t x_{t+h} - F_{t-1} x_{t+h})$ is the forecast revision.

Now we will test whether adding the variable $Uncertainty_t$ can improve the model. This is done by regressing equation (4). Table 6 presents the results of this regression.
As we can see β_2 and β_3 are both significantly different from zero. Now β_1 is statistically insignificant, which means that forecast revision alone does not predict forecast errors anymore when we include Uncertainty in the model.
However, β_2 = 0.3031, which means that when uncertainty is increasing, upward forecast revision predicts towards an underestimation of the forecast. As we can see in the summary statistics (Table 2), the average value of *Uncertainty* is around 0.64. So, in practical terms, if there is an upward forecast revision of 1 percentage point, it will on average lead to an ex-post forecast error of 0.1866 percentage points. But if *Uncertainty* is very high (1.37), then an upward forecast revision of 1 percentage point will predict a forecast error of 0.4496 percentage points.

The adjusted R^2 in this extended model is also significantly higher, which means that this model explains more of the variation in the forecast errors. So, adding the variable *Uncertainty* to the model not only improves the overall fit of the model, but also provides important insights into how uncertainty affects forecast revisions. The significance of β_2 and β_3 supports the hypothesis that uncertainty plays a critical role in forecasting, and thus, its inclusion is beneficial for enhancing the accuracy and reliability of the model.

Table 6: Regression results of equation (6):

$$x_{t+h} - F_t x_{t+h} = c + \beta_1 (F_t x_{t+h} - F_{t-1} x_{t+h}) + \beta_2 (F_t x_{t+h} - F_{t-1} x_{t+h}) \, Uncertainty_t + \beta_3 Uncertainty_t$$

Variable	Coefficient	Standard Error	t-value	p-value
(Constant)	-0.0434	0.0034	-12.855	< 0.0001
Forecast revision	-0.0007	0.0737	-0.010	0.9923
Interaction term	0.3031	0.0788	3.845	0.0001
Uncertainty	0.0573	0.0052	11.099	< 0.0001
Adj. R^2	5.00%			
N	6774			

Note that the forecast revision is: $(F_t x_{t+h} - F_{t-1} x_{t+h})$.
The interaction term is: Forecast revision * $Uncertainty$.

5.3 Testing a trading strategy

In this section I will check the performance of a trading strategy that is designed to take advantage of the findings from section 5.2. In the previous section we found a significant positive β_1 for forecast revision. This means that agents tend to underreact compared to FIRE. If the most recent forecast revision is upward, this means that variance expectations have increased recently. Yet with underreaction, it is not fully incorporated into prices, so the rational expectations should be higher. We can take advantage from this by buying a straddle because this straddle should be underpriced. If the recent forecast revision is downward, we should sell a straddle because this straddle should be overpriced. A straddle position is formed by either buying or selling one call and one put with the same strike price and maturity. We will take a position in a straddle because our interest is in studying option returns based on their volatility characteristics and not the direction of the price of the underlying. Figure 2 illustrates the payoff of buying a straddle.

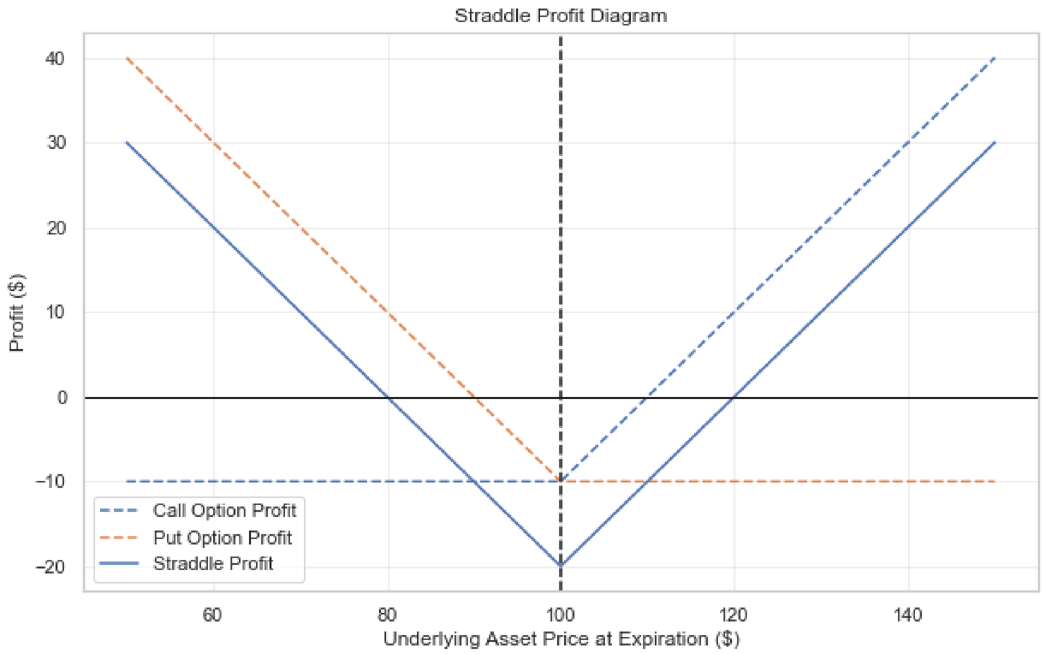


Figure 2.

Profit diagram of going long a straddle with strike price K = $100, and a premium of $10 for both the put and the call.

The profit of a call option can be calculated with the following formula: Profit = max(0, S-K) – premium.

The profit of a put option can be calculated with the following formula: Profit = max(0, K-S) – premium.

For a trading strategy to be realistic and applicable in real-world scenarios, it must be implementable in real-time. So, the trading strategy should only be based on information available up to that point in time, without incorporating data about the future. Therefore, we will first run regression equation (4) based on the data from the period 1996-2008. This results in a β_1 of 0.36 at the end of 2008. Starting from 2009, we will implement the strategy and the model will be updated continuously each date to check whether the strategy remains the same. Figure 3 provides the value of β_1 over the period 2009-2022.

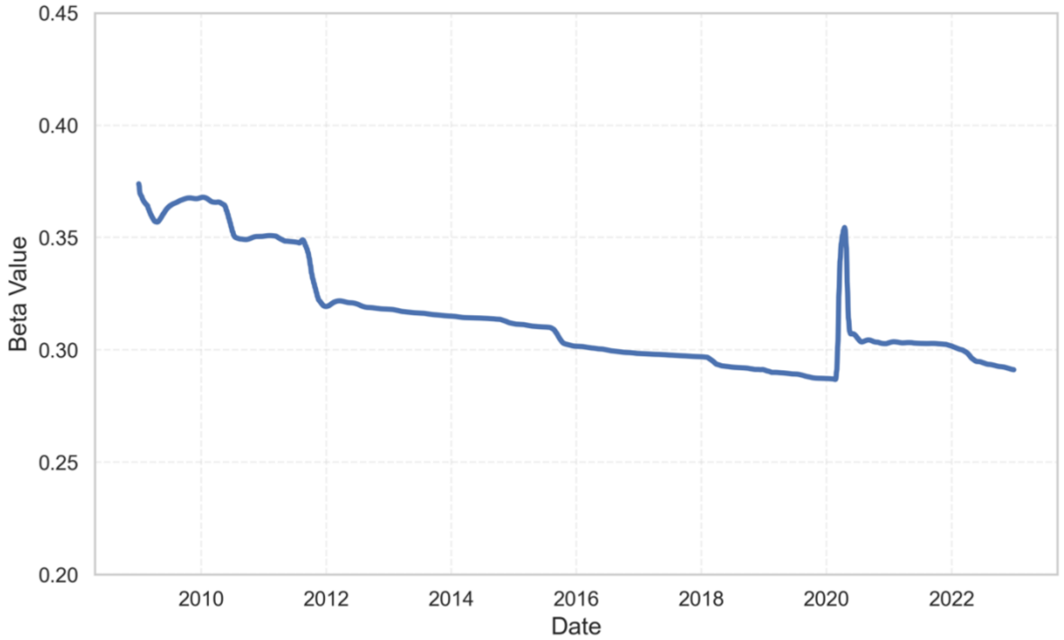


Figure 3.
Rolling window beta over the period 2009-2022. The value of beta is calculated using data from 1996 till Date.

As we can see in this figure, the value of β_1 remains always positive. This means that our strategy over the entire period will be the same: buying a straddle if there is upward revision, and selling a straddle if there is downward revision. We will take a position in a trade at the first trading day after the expiration day of the month (all the options expire immediately after the third Friday of the month). We will only consider options that will expire in the following month, and we will choose the options which are closest of at-the-money. The straddle returns are constructed as follows: The beginning price (P_1) is the sum of the average of the closing bid and ask prices of the call and put. The closing price (P_2) is the terminal payoff of the options. The return for buying (going long) a straddle is $R_l = \frac{P_2}{P_1} - 1$. The return for selling (going short) a straddle is simply $R_s = -R_l$.

In summary, we will perform an out-of-sample test about the trading strategy where at every first trading day after the expiration Friday of the month a straddle is either bought or sold. The trading signal depends on the sign of beta and on the sign of forecast revision. Given that beta remains positive over the whole testing period (see Figure 3), the strategy is simplified to the following:

- If forecast revision > 0: buy an at-the-money straddle,
- If forecast revision < 0: sell an at-the money straddle.

Table 7 reports summary statistics of the return of the trading strategy if we perform it on the period between 2009-2022.

Table 7: Summary statistics of trading strategy between 2009-2022

	N	**Mean**	**SD**	**Min**	**Max**
$\boldsymbol{R_{strategy}}$	167	0.0832	0.8331	-2.3176	5.2484
$\boldsymbol{R_{S\&P\ 500}}$	167	0.0097	0.0527	-0.3703	0.2209

The summary statistics of the return on the S&P 500 are also provided into this table.
The returns are divined as monthly returns.

Table 7 shows that the trading strategy would yield an average monthly return of 8.32%. If we convert this to a yearly return (including compounding) we get a return of about 161%. To put it into perspective, the average return of investing in the S&P 500 over this period is only 0.97% per month (12.28% per year). However, is this large return systematic compensation for risk or is this return abnormal? To test this we will first look at the Sharpe Ratio (Sharpe, 1994). This ratio is used to evaluate the risk-adjusted return of an investment. We could calculate it using the following equation:

$$Sharpe\ Ratio\ (SR) = \frac{R_i - r_f}{\sigma_i} \tag{7}$$

By using equation (7) we get a monthly SR of 9.93% for the strategy. Comparing this to the monthly SR of just investing in the S&P 500 (17.46%), the SR of our strategy is not quite high. This is mainly driven by the fact that the volatility of the strategy is much higher than the volatility of the S&P 500.

Additionally, we will check whether our strategy generates positive alpha in the Capital Asset Pricing Model (CAPM). This will be done by regressing this equation:

$$(R_i - r_f) = \alpha + \beta(R_{mkt} - r_f) \tag{8}$$

A positive alpha indicates that the trading strategy has delivered returns that are higher than what would be expected based on the CAPM model. The results of this regression are reported in Table 8.

Table 8: CAPM regression output

Variable	Coefficient	Standard Error	t-value	p-value
Intercept (α)	0.1120	0.066	1.707	0.090
$R_{mkt} - r_f$	-2.3480	1.262	-1.861	0.065
R^2	0.021			
N	167			

The intercept in the regression output represents the alpha of the Capital Asset Pricing Model (CAPM). In practical terms, alpha measures the strategy's excess performance relative to what would be expected based on its exposure to market risk.

The regression output shows that the CAPM alpha is positive and statistically significant only at the 10% level. This means that there is (weak) evidence that the strategy is delivering returns that are abnormal. Since the statistical significance is weak, investors should be cautious and not overly rely on this strategy's past performance as an indicator of future results. Other factors, such as transaction costs might also affect the strategy's performance.

However, the strategy's returns are not strongly explained by the CAPM model, as indicated by the low value of R^2. This calls for further investigation into other factors that might influence the strategy's performance and whether these results hold over different time periods or market conditions. However, this further investigation is not addressed within the scope of this thesis and is recommended for future research. Future studies could explore additional variables to gain a deeper understanding of the determinants of the strategy's returns.

6. Conclusion

This thesis aimed to investigate the rational pricing of S&P 500 index options, the predictability of ex-post forecast errors on implied volatility with ex-ante forecast revisions, and the impact of uncertainty on this relationship. The research builds on the foundational work of Stein (1989) and Coibion and Gorodnichenko (2015), extending the latter methodology to include the variable of uncertainty, as measured by the real uncertainty index.

The results of the Stein (1989) method confirmed that mispricing is present in the S&P 500 options market for the period from 1996 to 2022. Specifically, the elasticity of implied volatility between distant and nearby options (β_i) was found to be higher than the theoretical upper bound (β_{th}), indicating overreaction in long-maturity options' implied volatilities to changes in short-maturity options' implied volatilities.

Further analysis using the Coibion and Gorodnichenko (2015) framework revealed significant information rigidity among investors. The regression results show that forecast errors are predictable by forecast revisions, which means that investors do not fully incorporate new information into their forecasts. This underreaction suggests a systematic bias where upward revisions predict higher future realizations, implying that forecast adjustments are too conservative.

Incorporating the *uncertainty* variable into the model significantly improved its explanatory power. The interaction effect between forecast revisions and uncertainty showed that increased uncertainty increases the underestimation of forecasts. This finding highlights the critical role of uncertainty in the pricing and forecasting processes in the option market.

The practical implications of these findings were tested through a trading strategy based on forecast revisions. The strategy, which involved taking positions in straddles based on recent forecast revisions, yielded substantial average monthly returns of 8.32%. However, the high volatility related to the strategy suggest that while the returns are high, they come with significant risk. We found weak evidence that the trading strategy is generating results that are abnormal. We tested this by implementing the CAPM model and see if there is positive alpha. This resulted in a positive alpha, which was only significant at the 10% level. Investors should be cautious when implementing this trading strategy because the evidence it generates positive alpha is weak and the volatility is high.

In conclusion, this thesis contributes to the literature by proving the overreaction in long-maturity options' implied volatilities to changes in short-maturity options' implied volatilities. This thesis also shows that investors do not fully incorporate new information into their beliefs, leading to underreaction in short-term options. It also highlights the importance of considering uncertainty in forecasting models to enhance their accuracy. Future research could further explore additional factors influencing forecast errors and extend the analysis to different markets and time periods to generalize these findings.

References

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal Of Financial Economics*, *49*(3), 307–343. https://doi.org/10.1016/s0304-405x(98)00027-0

Black, F., & Scholes, M. S. (1973). The Pricing of Options and Corporate Liabilities. *Journal Of Political Economy*, *81*(3), 637–654. https://doi.org/10.1086/260062

Bordalo, P., Gennaioli, N., Ma, Y., & Shleifer, A. (2020). Overreaction in macroeconomic expectations. *The American Economic Review*, *110*(9), 2748–2782. https://doi.org/10.1257/aer.20181219

Caldara, D., Fuentes-Albero, C., Gilchrist, S., & Zakrajšek, E. (2016). The macroeconomic impact of financial and uncertainty shocks. *European Economic Review*, *88*, 185–207. https://doi.org/10.1016/j.euroecorev.2016.02.020

Coibion, O., & Gorodnichenko, Y. (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *The American Economic Review*, *105*(8), 2644–2678. https://doi.org/10.1257/aer.20110306

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal Of Financial Economics*, *7*(3), 229–263. https://doi.org/10.1016/0304-405x(79)90015-1

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987b). Expected stock returns and volatility. Journal Of Financial Economics, 19(1), 3–29. https://doi.org/10.1016/0304-405x(87)90026-2

Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring uncertainty. *American Economic Review*, *105*(3), 1177–1216. https://doi.org/10.1257/aer.20131193

Lettau, M., & Ludvigson, S. C. (2001). Consumption, Aggregate Wealth, and Expected Stock Returns. *The Journal Of Finance*, *56*(3), 815–849. https://doi.org/10.1111/0022-1082.00347

Lochstoer, L. A., & Muir, T. (2022). Volatility expectations and returns. *The Journal of Finance*, *77*(2), 1055–1096. https://doi.org/10.1111/jofi.13120

Ludvigson, S. C., Ma, S., & Ng, S. (2021). Uncertainty and business cycles: exogenous impulse or endogenous response? *American Economic Journal. Macroeconomics/American Economic Journal, Macroeconomics*, *13*(4), 369–410. https://doi.org/10.1257/mac.20190171

Maćkowiak, B., & Wiederholt, M. (2009). Optimal Sticky Prices under Rational Inattention. *The American Economic Review*, *99*(3), 769–803. https://doi.org/10.1257/aer.99.3.769

Mankiw, N. G., & Reis, R. (2002). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *The Quarterly Journal of Economics*, *117*(4), 1295–1328. https://doi.org/10.1162/003355302320935034

Sharpe, W. F. (1994). The sharpe ratio. *Journal of Portfolio Management*, *21*(1), 49–58. https://doi.org/10.3905/jpm.1994.409501

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690. https://doi.org/10.1016/s0304-3932(03)00029-1

Stein, J. C. (1989). Overreactions in the Options Market. *The Journal Of Finance*, *44*(4), 1011. https://doi.org/10.2307/2328621