QUEEN MARY UNIVERSITY OF LONDON

Essays on Option-Implied Measures of Risk

Mina Mirshahi

Thesis submitted in partial fulfillment

of the degree of Doctor of Philosophy (PhD) in Finance

in the

School of Economics and Finance

Queen Mary University of London

United Kingdom

August 2023

To my dad

Declaration of Authorship

I, Mina Mirshahi, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Mina Mirshahi

August, 2023

Details of Collaboration

Author contributions and additional collaborators are listed below for each chapter:

Chapter 1: This Chapter is a joint work with Prof George Skiadopoulos. All tasks related to data gathering, preparation as well as statistical analyses are solely my own work. The write-up and presentation of results have been carried out collectively with my co-authors.

Chapter 2: This Chapter is a joint work with , Dr Elise Gourie and Mo Wang. All tasks related to data gathering, preparation as well as statistical analyses are solely my own work. The write-up and presentation of results have been carried out collectively with my co-authors.

Chapter 3: This Chapter is my own work. All tasks related to data gathering, preparation as well as statistical analyses are solely my own work. The write-up and presentation of results have been carried out by myself.

Abstract

This thesis comprises three studies on extracting the information embedded in option prices.

In Chapter One, we propose a new predictor to forecast U.S. real economic activity (REA) by utilising the information embedded in equity option prices. We construct our equity option-based predictor by applying standard and recent data reduction methods, to the cross-section of computed option-implied expected returns of the underlying stocks. Our predictor forecasts REA both in- and out-of-sample setting even after controlling for common REA predictors and considering their persistence. We find a robust negative relationship between the option-implied predictor and REA. We show that individual stocks contain some additional predictive power that is not being captured neither by the index option-implied expected return, nor by standard factors.

In Chapter Two, we analyze the impact of Federal Open Market Committee (FOMC) meetings on S&P 500 option markets. We document that volatility and tail risks as well as their premiums increase in the week before meetings, whereas option liquidity decreases. We show that these findings are stronger in weaker economic conditions.

In Chapter Three, we review variables available for extracting information from equity option prices. We consider option-implied volatilities, implied risk-neutral skewness and implied expected returns. We discuss how these measures have been used in the literature and document the usefulness of their information in forecasting future prices, asset allocation strategies, corporate events, and the state of the economy. In addition, we comprehensively analyse the cross-sectional predictive power of several option-implied variables for forecasting future equity returns. We show that option-implied measures

demonstrate statistically and economically significant predictability for future stock returns.

Acknowledgments

First and foremost, I would like to thank my supervisors, Prof. George Skiadopoulos and Dr Elise Gourier for their continued support, precious guidance and passionate encouragement. I would not have been able to complete my research without their enormous kindness and devotion to the supervision and our collaborative research projects.

I would like to express my sincere appreciation and gratitude to my examiners, Dr. Ilaria Piatti and Professor Pedro Saffi, for their invaluable support and contributions throughout the examination process of this thesis. Their expertise, guidance, and thoughtful feedback have been instrumental in shaping the final outcome of this research.

During my doctoral studies at Queen Mary, I had the opportunity to discuss my research with several faculty members and visitors of the School of Economics and Finance. Their feedback guided me to substantially improve my thesis. In this context, I would like to express my gratitude to Prof. Jason Sturgess, Dr Konstantinos Zachariadis, Dr Alessandra Bonfiglioli, Dr Andrea Tesei, Prof Francis Breedon, Dr Renato Faccini, and Dr Tatsuro Senga. I would also like to thank the administrative staff for their assistance.

A special recognition goes to Dr Eirini Konstantinidi, my co-author Mo Wag and my friends at QMUL, especially, Haiqa Ali, Ali Ebadi and Fulvia Marotta. I will always feel privileged for sharing parts of this journey with you. You made it really happen.

Parts of this dissertation were presented at different stages at the following conferences where I received valuable comments: the 7th Young Finance Scholars (2021) and European Financial Management Association (2019 Annual Meeting). I would like to thank the organizers and the participants at these seminars and conferences. Specifically, I would like to thank to Prof Mark Shackleton, Dr Andreas Neuhierl, Prof Lubos Pastor and Dr Juliana Salomao.

I also wish to acknowledge the support from my employer, Alastair Stuart at St James Place, and Nima Zarrabi at Invesco, who enabled me to concentrate on my PhD study.

I would like to thank my family specially my lovely sister, Parisa, and my grandmother, Maryam, my amazing friends, Elham, Mona and Bahareh, for their enormous support and encouragement. I am hugely thankful to my husband, Nima, who understands and supports my struggle with research and life.

I dedicate this work to my lovely heavenly father, Mohammad Mirshahi. He was my rock, support and courage. I am forever grateful to him for inspiring me in every step of my life. He was my whole life and I hope to be able to make him proud someday.

Contents

List of Tables

List of Figures

Chapter 1

Predicting Real Economic Activity with Individual Option-Implied Expected Returns

1.1 Introduction

The provision of reliable forecasts of real economic activity (REA) has long been one of the principal challenges for economists and of primary importance to corporates, investors and policymakers. There has been considerable research on forecasting economic activity using measures based on the prices of stocks, bonds and commodities. These variables, however, failed to predict the economic downturn from 2007 to 2009 and led econometricians and macroeconomists to question the adequacy of these predictors. Therefore, they highlight the necessity of developing a precise REA predictor using financial markets' information (Ng and Wright (2013)).

In this paper, we propose a new option-based predictor which uses for the first time information from *individual* equity option prices, that is, option-implied underlying stocks' expected returns. The stock's expected return is a natural candidate as a predictor of REA. Campbell and Cochrane (1999) show that in a standard log-linear representative agent model, the stock's expected returns are counter-cyclical. Surprisingly, the previous literature has not examined its predictive ability. This may be due to the econometric challenges in estimating stocks' expected returns using historical data.[1] Our purpose is to estimate an updated, forward-looking expected return, and recent research bypasses these obstacles by providing methods which use information extracted from option prices. Options are forward-looking and are expected to possess informational content about the future state of the economy, thus, being natural candidates for forecasting REA. We investigate the predictive ability of the expected returns of the individual optionable stocks extracted from the respective equity market option prices to forecast the growth of the U.S. REA. Bakshi et al. (2011), and Faccini et al. (2019) have recently proposed using the information from *index* option prices to predict REA and found that index option prices should be incorporated in the list of variables that predict economic growth. Individual equity and index options are different and serve different purposes and trading patterns.[2] To the best of our knowledge, our paper is the first that investigates whether individual equity option prices may also convey useful information to predict REA.

We compute the individual stocks' expected returns via Martin and Wagner (2019) formula. This formula expresses the expected return on a stock in terms of the risk-neutral variance of the market portfolio, the risk-neutral variance of the individual constituent stocks and the value-weighted average of the constituent stocks' risk-neutral variances.[3] All three quantities of risk-neutral variance are calculated directly from market equity option prices and therefore help us to exploit the information embedded in option price quotes to forecast future REA. In addition, this approach is distinct from more conventional methods, such as the risk premium estimates using historical excess returns, the CAPM and the Fama-French three-factor model, due its particular

1. Typically, the expected returns are estimated using historical time series of assets return (Damodaran (2019)). However, this is not optimal since the historical approach is backwards-looking and only relies on the belief that the future will be similar to the past, which is not always the case. As outlined by Duan and Zhang (2014), investors are subject to higher uncertainty in the volatile phase- i.e., during a crisis. Hence, the forward-looking risk premium will increase. However, historical average returns cannot adequately reflect market conditions when such an increase in market volatility is temporary. During the crisis period, the realized stock returns are negative; therefore, the historical approach produces a negative risk premium.

2. Bollen and Whaley (2004) document that trading in S&P 500 index options contributes to the hedging demand of investors- i.e., buying index puts as portfolio insurance against market declines. While Lakonishok et al. (2007) show that trading in stock options involves naked positions.

3. We also compute the individual stocks' expected returns via Chabi-Yo et al. (2022)s' method, which takes into account the entire shape of risk-neutral density using higher moments of the risk-neutral distribution. The predictive power of the option-based expected returns predictor remains regardless of the method used to compute the option-implied expected returns.

features.[4] Firstly, it does not require accounting data. Second, it is based on observed market prices and can be implemented in real-time. Third, it is model-free and does not need any historical information to estimate its parameters. We focus our analysis on firms that have been constituents of the S&P 500 index from July 1998 to May 2019. We compute the individuals' option-implied expected returns for different horizons (h =1, 2, 3, 6 and 12 months).

Once we compute the stocks' option-implied expected returns, we construct our option-based predictor by relying on data reduction methods. We employ Principal Component Analysis (PCA) and the Instrumented Principal Component Analysis (IPCA), separately, to assess whether results may differ depending on the method used to reduce dimensionality. PCA extracts factors that capture the co-movement of expected returns. The first PCA factor explains almost 70% of the variance of h-month(s) expected returns (h = 1,2,3,6,12). PCA assumes a static correlation coefficient between expected returns and the latent factor- i.e., static loadings. In addition, PCA method does not incorporate other data beyond returns. For these reasons, we also utilise the Instrumented Principal Component Analysis (IPCA) as proposed by Kelly and Pruitt (2019). The IPCA, first, brings the wealth of characteristics into account by considering them as an instrumental variables which help predict loadings. Then, armed with loadings, it estimates factors which inherently depend on observable firms' characteristics.[5] Mapping the characteristics on loadings provide a link between expected returns and firms' characteristics. Therefore, IPCA enables a factor model to accommodate the conditional information of stock characteristics into the expected returns. The correlation between the 1-month first IPCA factor and the 1-month first PCA factor is 96%, yet not perfectly correlated.

Next, we investigate whether PCA/IPCA factor predicts future U.S. REA both in and out of the sample. We test the predictability of these factors over different forecasting horizons while controlling for well-known predictors of REA. To evaluate the robustness

4. see Lewellen (2015) and Duan and Zhang (2014).

5. The estimation objective in the IPCA method is similar to PCA- i.e., both estimate factors and loadings by concentrating on the common variation of the system of variables. The objective function in both models is to minimise the sum of squared model errors. However, the weight applied to each squared residual is static in PCA whereas it is time-varying and depends on the characteristics of each firm in IPCA.

of our results, we employ eight different proxies of REA, namely, industrial production (IP), non-farm payroll (NFP), retail sales (RS), housing starts (HS), capacity utilisation (CU), unemployment rate (UR), Chicago Fed National Activity Index (CFNAI) and Aruoba–Diebold–Scotti (ADS).

Our results from the in-sample analysis suggest that the first principal component of expected returns obtained from either PCA or IPCA method is a statistically significant predictor of REA, especially for the short-term horizons and contains information that other financial predictors have not already incorporated. We reject the null hypothesis of no predictability after addressing econometric concerns. Specifically, we conduct statistical inference tests of our regression results using the instrumental variable test following Kostakis et al. (2015) to deal with predictors with unknown degree of persistence. We find a negative relationship between the first PCA (IPCA) factor and the future REA. Some previous studies propose that consumption smoothing is the reason for the relation between expected returns and REA (e.g., Balvers et al. (1990); Chen (1991) and Loflund and Nummelin (1997)). Consumption-smoothing models predict that expected real returns are counter-cyclical.

For each REA index, we estimate the out-of-sample (OOS) forecasts recursively using an expanding window. At each point in time, we construct the PCA/IPCA predictor recursively to avoid any look-ahead bias. We compare our forecasts obtained from a full model (which contains our options-implied PCA or IPCA factor along with a set of standard predictors of REA) with forecasts obtained from a nested model (which does not contain the option-implied factor in the set of its predictors) by computing the out-of-sample R^2 measures following Campbell and Thompson (2008). Our results suggest that the inclusion of PCA or IPCA factors in predictive models improves OOS forecasts. In addition, we show that the IPCA factor is preferred to the PCA factor to estimate OOS forecasts in longer horizons.[6]

In our robustness analysis, first, we compare the predictive ability of the PCA factor obtained from the variance-based expected returns of Martin and Wagner (2019) with

6. Firms characteristics that are employed in the IPCA method are mainly calculated based on 12 months. This may be why the predictive model with the IPCA factor estimates better OOS forecasts than the PCA factor in longer horizons.

risk-neutral-based expected returns proposed by Chabi-Yo et al. (2022).[7] We found that the predictive power of the option-based expected returns predictor remains regardless of the method used to compute the option-implied expected returns. Second, we investigate whether the option-implied PCA/IPCA factors computed from the stock option-implied expected returns outperform (i) the previously established pricing factors, such as those from Fama and French (2015) and (ii) the index option-implied expected return. To address this, we compare the OOS forecasts of the full model with PCA/IPCA factor and show that this model outperforms the nested model, which contains the market risk premium (SVIX) measure of Martin (2017) or the five systematic risk factors of Fama and French (2015). Our findings imply that even though the derived PC factors may be interpreted as "level" factors that proxy the market, individual stocks contain some additional predictive power that is not captured by the index expected return nor by standard factors that describe the cross-section of expected returns. Equity and index options are dissimilar and serve different purposes and trading patterns.[8]

The rest of the paper is organised as follows. Section 1.2 reviews the related literature. Section 1.3 describes our data. Section 1.4 describes the procedure we employ to extract the expected return factors. Section 1.5 explains our prediction models to forecast REA. Section 1.6 summarises our results. Section 1.7 reports our additional analysis results and Section 1.8 concludes.

1.2 Related Literature

This paper relates to three strands of literature. First, it relates to papers that employ financial variables to estimate real economic activity. Several studies have found that the spread between the short and long-term maturity U.S. Treasury bills (term-spread) has a strong predictive relation to aggregate economic activity. (e.g., Harvey (1989), Chen (1991), Stock and Watson (1989), Friedman and Kuttner (1991), Estrella and Hardouvelis (1991)). Some papers document the predictive power of credit spreads to forecast REA. Bernanke (1983) finds that during the interwar period, the Baa-corporate

7. Expected returns computed from Chabi-Yo et al. (2022)s' method considers the entire risk-neutral distribution of returns and accounts for higher RN moments rather than only the second moment.

8. Lemmon and Ni (2014) document that trading activity in equity options is linked to individual investors' beliefs and past market returns, whereas a hedging demand motivates index options trades.

and Treasury bond spread predicts industrial production growth. During the postwar period, however, Friedman and Kuttner (1991) find that the paper-bill spread (defined as the difference between the interest rates on commercial paper and Treasury bills) has predictive power for real growth. Gilchrist and Zakrajsek (2012) show that the credit spreads obtained from the spread between corporate bonds and treasuries bonds with same maturities are a powerful predictor of output growth. Ng and Wright (2013) find that term spreads were reasonable predictors of economic activity during 1970s and 1980s, whereas credit spreads show better predictive ability in recent years. Another group of papers, such as Liew and Vassalou (2000) suggest that asset pricing factors contain significant information about future REA growth. Likewise, Stock and Watson (2003) find that asset prices of some countries significantly predict those countires' output growth. Bakshi et al. (2012) document that the Baltic dry index (BDI) growth rate forecasts the global economic activity growth. Moreover, Hong and Yogo (2012) propose that commodity futures open interest is a predictive factor for the real economy. On the other hand, Allen et al. (2012) used the Chicago Fed National Activity Index (CFNAI) to forecast economic downturns. Kelly and Jiang (2014) utilise monthly firm-level price crashes and found that their measure negatively predicts real economic activity. While this study exploits the information in the cross-section of stock prices, we extract the information embedded in the cross-section of equity options.

Another strand of literature deals with the estimation of expected returns. The average of historical realised returns are often used for estimating expected return. Damodaran (2019) documents that this method does not reflect transient rises in the market's volatility and produces a negative risk premium, particularly during a crisis when the realised stock return is negative. On the other hand, several studies, such as Welch (2000) and Fernández (2009) interviewed academics, traders or financial managers to obtain their opinions on expected returns. However, this approach is time-consuming, requires a very long prediction horizon, and faces sample selection bias. Other studies extract the information embedded in option prices to estimate expected returns. For example, Santa-clara and Yan (2010) derive the risk premia as a function of volatility and jump intensity implied from options while Duan and Zhang (2014) derive it as a function of physical moments and option-implied risk aversions. Buss and Vilkov (2012) employ information extracted from equity and index option prices to estimate correlations and build option-implied betas. Using the implied betas, they derive linear factor models and hence expected returns. Martin (2017) derives a lower bound on the market premium in terms of simple risk-neutral variance (SVIX) that can be computed from index option prices. Ross (2015) and Schneider and Trojani (2019) estimate market return from state prices i.e., the product of risk aversion (the pricing kernel) and the risk-neutral probability distribution using the recovery theorem. These studies focus on the market return. Extracting the information embedded in individual options, Martin and Wagner (2019) and Kadan and Tang (2019) propose a forward-looking lower bound on the expected rates of return on the individual assets. Their proposed lower bounds reflect idiosyncratic and systematic risk and rely on risk-neutral variances. Chabi-Yo et al. (2022) consider higher moments of the stocks' risk-neutral distribution to extract option-implied expected returns. González-Urteaga et al. (2021) combine the results of Martin (2017) and the literature on the stochastic discount factor and extract the expected returns of stocks. We choose the method proposed by Martin and Wagner (2019) to extract individual stocks' expected returns using option prices. This estimation is computed from the observed current index and stock option prices and does not require any historical data. More importantly, this method is parsimonious in terms of the required inputs and can be computed precisely from the cross-section of market and individual option prices that are easily available. In addition, expected returns obtained from this method correspond to the market risk and indicate a significantly positive market risk premium. They are also align with several stock characteristics, including value, size, profitability, and momentum.[9]

This paper is also related to studies that have investigated the predictive power of option-implied variables for REA. Surprisingly, there is very little literature on whether the information inferred from option prices can predict REA. Bakshi et al. (2011), Faccini

9. In the robustness analysis, we compare the predictive ability of expected returns extracted from Martin and Wagner (2019) with those obtained from Chabi-Yo et al. (2022). Our results suggest that Chabi-Yo et al. (2022) 's method, which takes into account the higher moments of risk-neutral distribution, does not improve the predictive power of implied-expected returns.

et al. (2019) and Buss et al. (2019) are the only papers which explore this.[10] Bakshi et al. (2011) show the forward variance extracted from S&P 500 index options predicts economic activity. Faccini et al. (2019) find that investor's implied relative risk aversion extracted from index option prices predicts REA. Buss et al. (2019) document that implied correlation acts as a leading procyclical state variable and negatively forecasts unemployment. We add to this literature by analysing the information embedded in options at the stock level, namely individual stocks expected returns.

1.3 Data

1.3.1 Option Data

We compute the individual stocks' expected returns via Martin and Wagner (2019) formula. This formula calculates the expected return on a stock based on the risk-neutral variance of the market, the risk-neutral variance of the individual stock, and the value-weighted average of stocks' risk-neutral variance. These measures are computed directly using observed prices of options on both the index and the individual stocks. We collect the required variables from OptionMetrics Ivy DB(OM) through the Wharton Research Data Services (WRDS) and Compustat. Our sample spans January 1996 to May 2019.

We first obtain information on the S&P500 index and its constituents from Compustat. Using the lists of index constituents, we searched the OptionMetrics database for the entire universe of firms that formed the S&P 500 index during our sample period. Where available, we obtain the time series of implied volatility surface data for these individual

10. Other studies analyse the forecasting power of option-implied information for stock returns. For example, An et al. (2014) find that increases in implied volatilities of ATM put and call options have different indications, forecasting low and high future stock returns, respectively. Conrad et al. (2013) examine the relationship between RN moments and realised returns and find a negative, albeit not statistically significant. Xing et al. (2010) find that stocks with the steepest implied volatility smirks show worst earnings shocks in subsequent months. Cremers and Weinbaum (2010) show that the deviation from put-call parity arising from expensive puts possesses information about future prices of underlying stocks and predicts abnormally negative returns. Rehman and Vilkov (2012) and Stilger et al. (2016) show that the RN skewness is positively related to future stock return. Gkionis et al. (2021) document that high RNS, arising from expensiveness OTM calls relative to OTM puts, contains positive information regarding the underlying stock. Other studies focus on options trading activities. For example, Pan and Poteshman (2006) show a negative relationship between the put-call ratio and future stock returns. Hu (2013) find that the stock exposure imbalance generated by option trading is positively related to next-day stock returns. Johnson and So (2012) find that a low volume trading in options markets compared to the stock market predicts low stock performance. Kostakis et al. (2011) show that the information option-implied distributions can predict market timing.

firms from the Optionsmetrics volatility surface (VS) file. The VS file contains implied volatilities for standardised equity options for standard maturities- i.e., 30, 60, 91, 182, and 365 calendar days, on a delta grid, accompanied by implied strike prices. We remove options that their deltas or implied volatilities are missing. Moreover, we collect daily index and stock prices as well as number of outstanding shares of each firm from OptionMetrics's underlying price file to compute their market capitalisations. We collect forward prices for each firm from OM standardised option price file. In addition, Zero yield curve data and projected dividend yield are also obtained from the OM Zero Curve File and Index Dividend Yield file.

Using equity prices and volatility surface data, we compute the risk-neutral variance (SVIX) of all individual firms that have been constituents of S&P 500 index from 1996 to 2019 for horizons of one, two, three, six, and twelve months. From these time series, on each day, we focus only on the SVIX of firms that have been included in SPX and calculate the value-weighted average of the risk-neutral variance of the market. As summarised in Panel A of Table 1.1, we calculate almost 2.5 million individual stocks's SVIX for each horizon; We cover 1,092 firms over our sample period from January 1996 to May 2019. For each horizon, we collect data for 482 firms on average per day, suggesting that we capture slightly more than 96% of the firms included in the S&P 500 index.

[Table 1.1 about here.]

1.3.2 Real Economic Activity Data

We collect monthly data for eight different measures which proxy for U.S. REA from July 1998 to May 2019. First, we use Industrial Production (IP), a measure which quantifies the output of the industrial sector of the economy. The industrial sector comprises manufacturing, mining, and utilities. Second, we focus on Nonfarm payrolls (NFP) which represents the number of employees in the nonfarm sectors in the U.S. economy. Third, we consider real retail sales (RS), defined as the consumer demand for finished goods. RS is measured by the purchases of durable and non-durable goods over a specific period. Fourth, we employ housing starts (HS), defined as the total of new private-owned houses.

Fifth, we use capacity utilisation (CU), which measures the extent to which resources in corporations and factories are employed to produce goods in manufacturing, mining, and gas and electric utilities in the United States. Sixth, we consider the unemployment rate (UR), which represents the number of unemployed people as a ratio of the labour force. The latter is the sum of unemployed and those in paid or self-employment.

Our seventh proxy for the U.S. REA is the Chicago Fed National Activity Index (CFNAI), which is a monthly index created to measure overall economic activity and related inflationary pressure. Finally, our last proxy is the Aruoba–Diebold–Scotti (ADS) index (Aruoba et al. (2009)), which tracks business conditions. ADS is composed of six economic indicators, including monthly payroll employment, weekly initial jobless claims, personal income fewer transfer payments, manufacturing and trade sales, industrial production, and quarterly real GDP.

We acquire all these proxies from the Federal Reserve Economic Data (FRED) database held by the Federal Reserve Bank of St. Louis(FRED), except for ADS, which we collect from the Philadelphia Fed webpage. We denote the monthly values of these proxies by $Y_{t,i}$ ($i = 1$ for IPI, 2 for NFP, 3 for RS, 4 for HS, 5 for CU, 6 for UR, 7 for CFNAI and 8 for ADS).

Next, we calculate the h-month overlapping log growth rates of NFP, IP, HS, RS, CU and UR -i.e., $REA_{t+h,i} = lnY_{t+h,i} - lnY_{t,i}$, for $h = 1, 2, 3, 6, 12$ months and $i = 1$ to 6. By construction, the values of CFNAI and ADS denote growth or recession and therefore, we do not need to calculate the growth rates for these two proxies. Therefore, we define $REA_{t+h,i} = Y_{t+h,i}$ for $i = 7$, 8.

1.3.3 Control Variables

We obtain data on various variables that have been studied by previous literature and have shown a predictive ability to predict REA. We use these variables as control variables in our predictive regression models. First, we consider lagged REA in our regression model-i.e., $REA_{i,t}$ is the growth of the ith REA proxy over the period $t-1$ to t-i.e., the length of the period is h month(s). Second, we collect monthly treasury bond and bill rates from the FRED website to measure the term spread (TS), as the difference between the 10-year Treasury bond rate and the three-month Treasury bill rate. Third, we employ credit spread (CR), the spread between the Moody's AAA and BAA corporate bonds yields.

Fourth, we use OTM options data to calculate the value of the forward variance $FV_{t,t+1}$ between t and $t+1$ at time t- i.e., the forward variance with a h-month horizon. We follow Bakshi et al. (2011) and calculate $FV_{t,t+1}$ as follows:

$$H_{(t,t+1)} = e^{-r(h/12)} + \int_{K<S_t} \omega(K)Put_t(K)dK + \int_{K>S_t} \omega(K)Call_t(K)dK \tag{1.1}$$

where $C_t(K)$ and $P_t(K)$ are the prices of a call and put options on SPX index, respectively, at time t, with time to expiry (h) month(s) and strike price K. Moreover, $e^{r(h/12)}$ is the price at time t of a riskless discount bond with unit face value and time-to-maturity of h month(s).

$$\omega(K) = -\frac{\frac{8}{\sqrt{14}}cos(arctan(1/\sqrt{7}) + \frac{\sqrt{7}}{2}ln(\frac{K}{S_t})}{\sqrt{S_t}K^{3/2}} \tag{1.2}$$

We define the forward variance at time t between t and $t+1$, as:

$$FV_{t,t+1} = -lnH(t,t+1) \tag{1.3}$$

Our fifth control variable is investor's implied relative risk aversion (IRRA). We extract IRRA using Kang et al. (2010) formula:

$$\frac{\sigma^2_{p,t}(\tau) - \sigma^2_{q,t}(\tau)}{\sigma^2_{q,t}(\tau)} \approx \gamma\sigma_{q,t}(\tau)\theta_{q,t}(\tau) + \frac{\gamma^2}{2}\sigma^2_{q,t}(\tau)(k_{q,t}(\tau) - 3) \tag{1.4}$$

where γ is the coefficient of the relative risk aversion of the representative agent. The formula estimates the spread between the risk-neutral variance $(\sigma^2_{q,t}(\tau))$ and physical variance $(\sigma^2_{q,t}(\tau))$ as a function of the investor's relative risk aversion (γ) by considering a power utility function. In this formula $\theta_{q,t}(\tau)$ and $k_{q,t}(\tau)$ are the RN skewness and kurtosis, respectively. We calculate the S&P 500 RN moments with a τ =1-month horizon following Bakshi et al. (2003) formulae (see Appendix 1.A). Following Faccini et al. (2019), we apply the generalised method of moments employing a 30-months rolling window. Provided that our options data set is from January 1996 to May 2019, we obtain the IRRA time series from July 1998 to May 2019.

The left axis of Figure 1.A.1 depicts the movements of forward variance, and the right axis displays the changes of the U.S. implied risk aversion (IRRA) over July 1998 to May 2019. Two remarks are in order in case of IRRA. First, IRRA values span from 3 to 6.5. This is similar to range of IRRA computed by the previous literature. For example, Faccini et al. (2019) report IRRA values ranging from 2.27 to 9.55. Second, we can see that the U.S. IRRA and forward variance rises substantially in 2008 during the financial crisis and begin declining after that. These patterns could result from the quantitative easing monetary policy implemented by the Fed from 2008 to 2014.

[Figure 1.A.1 about here.]

Finally, we consider five factors of Fama-French (5FF) as our sixth set of control variables.[11] These factors include the size of firms (SMB), book-to-market values (HML), expected return on the market (mrkt), profitability factor (RMW) and investment factor (CMA). These factors capture the different dimensions of systematic risk and are expected to explain the cross-section of stock's average returns with a linear relationship. We mainly use these factors to assess whether the sources of systematic risks in individual stocks' returns are adequate to predict the REA or whether the forward-looking option-implied risk factors contain additional predictive information.

1.4 Constructing the Option-Implied Predictor

In this section, we discuss the procedure to construct our option-implied predictor. We follow Martin and Wagner (2019) formula, which derive the expected return on a stock in terms of the stock's excess risk-neutral variance and the risk-neutral variance of the market. We then apply principal component analysis (PCA) and instrumented principal component analysis (IPCA) to produce our proposed option-implied predictor.

11. We obtain 5FF from https://mba.tuck.dartmouth.edu

1.4.1 Formula for the Individual's Expected Return

The inputs to the Martin and Wagner (2019) formula are the following three measures of risk-neutral variance:

$$\begin{aligned} SVIX^2_{m,t} &= var^*_t(\frac{R_{m,t+1}}{R_{f,t+1}}) \\ SVIX^2_{i,t} &= var^*_{i,t}(\frac{R_{i,t+1}}{R_{f,t+1}}) \\ \overline{SVIX}^2_t &= \Sigma_i \omega_{i,t} SVIX^2_{i,t} \end{aligned} \tag{1.5}$$

where $R_{m,t+1}$ is the market gross return from time t to $t+1$, $R_{i,t+1}$ is the gross return on stock i from time t to $t+1$ and $R_{f,t+1}$ is the gross riskless rate from time t to $t+1$. The length of the period from time t to time $t+1$ depends on the forecasting horizon of interest-i.e., h month(s). $SVIX^2_{m,t}$ denotes the measure of market risk-neutral variance, $SVIX^2_{i,t}$ is the RN variance at the stock level and $\overline{SVIX}^2_t$, measures the average risk-neutral stock volatility.

Martin (2017) claims that the $SVIX^2_{m,t}$ index can forecast the equity premium. The prices of index options determine this measure:

$$SVIX^2_{m,t} = \frac{2}{R_{f,t+1}S^2_{m,t}}[\int_0^{F_{m,t}} Put_{m,t}(K)dK + \int_{F_{m,t}}^{\infty} Call_{m,t},(K)dK] \tag{1.6}$$

where $S_{m,t}$ denotes the price of the market portfolio. $F_{m,t} = R_{f,t+1}(S_t - \bar{D}_{m,t})$ is the forward price of stock as of time t for delivery at time $t+1$ and $\bar{D}_{m,t}$ is the the present value of dividends paid between times t and $t+1$. $Put_{m,t}(K)$ and $Call_{m,t}(K)$ are prices of European index call and put options with strike prices K.

The corresponding SVIX measure at the individual stock level is obtained using individual stock option prices:

$$SVIX^2_{i,t} = \frac{2}{R_{f,t+1}S^2_{i,t}}[\int_0^{F_{i,t}} Put_{i,t}(K)dK + \int_{F_{i,t}}^{\infty} Call_{i,t}(K)dK] \tag{1.7}$$

where the subscripts i denote that the reference asset is stock i.

Finally, using $SVIX^2_{i,t}$ for all firms in the index available at time t and $\omega_{i,t}$ be the market-capitalization weight of stock i in the index, we compute the risk neutral average stock variance index as $\overline{SVIX}^2_t = \Sigma_i \omega_{i,t} SVIX^2_{i,t}$

We compute the three measures of risk-neutral variance for horizons (i.e., option maturities) of 1, 3, 6, 12, and 24 months. We then remove a few extreme outliers in our data that do not meet the monotonicity property of $SVIX_{i,t}$ across horizons described above[12].

Then, for each optionable stock in our sample, we construct the expected returns in excess of the market following Martin and Wagner (2019).

$$\frac{\mathbb{E}_t\, R_{i,t+1} - R_{m,t+1}}{R_{f,t+1}} = \frac{1}{2}(SVIX^2_{i,t} - \overline{SVIX}^2_t) \tag{1.8}$$

Exploiting the results of Martin (2017), specifically, $E_t R_{m,t+1} - R_{f,t+1} = R_{f,t+1} SVIX^2_{m,t}$, we have:

$$\frac{\mathbb{E}_t\, R_{i,t+1} - R_{f,t+1}}{R_{f,t+1}} = SVIX^2_{m,t} + \frac{1}{2}(SVIX^2_{i,t} - \overline{SVIX}^2_t) \tag{1.9}$$

Finally, we define the expected return of stock i as:

$$\mathbb{E}_t\, R_{i,t+1} = R_{f,t+1} \times [SVIX^2_{m,t} + \frac{1}{2}(SVIX^2_{i,t} - \overline{SVIX}^2_t)] + R_{f,t+1} \tag{1.10}$$

1.4.2 Principal Component Analysis

We apply Principal Component Analysis (PCA) to the cross-section of options implied expected returns obtained in the previous section to construct our predictor. This method transforms our extensive set of firms into a smaller one that still contains most

12. If the underlying asset does not pay dividends, call prices are rising with time-to-maturity. Assuming this is not cancelled by the opposing effect of increased interest rates $R_{f,t+1}$ over longer horizons, $SVIX_{i,t}$ should be expected to be monotonic in horizon length. In the daily data, we come up with 2,837,127 firm-day observations after withdrawing 21,438 observations based on nonmonotonicity. In the case of end-of-the-month data, we end up with 135,343 observations after removing 1,012 observations based on nonmonotonocity

of the information in the large set. However, before applying the PCA, we must deal with the missing data in our sample.

As reported in Table 1.1, our sample consist of 1,092 firms. Panel A of Table 1.2 shows that "Services" is the most represented industry in our sample, which forms 16% of our sample. The data set is unbalanced: almost 45% of it is missing for three reasons. First, the volatility surface data was unavailable on some days for some firms; therefore, calculating the SVIX for those firms was impossible. Second, according to Martin and Wagner (2019), the expected returns on individual stocks are determined relative to the market index; each day, the expected return on the individual stock is reported only if the firm is incorporated in the market index (in our case SPX). Third, the SVIX measure for some firms was not monotonic with respect to the horizon on some days.

Following Martin and Wagner (2019), we only consider stocks that formed the S&P 500 index at each point in time during the sample span (July 1998- May 2019)[13] and have no missing values during our sample period. This results in a balanced panel of expected returns for 125 firms with 251 monthly observations. Panel B of Table 1.2 suggests that "Services" is the most represented industry in our reduced-balanced sample. Moreover, we compare the characteristics of the firms included and excluded from the final sample. The left panel of Table 1.3 demonstrates the descriptive statistics of 125 firms for which we have a full data history and has been constituents of SPX for the entire sample. The right panel of Table 1.3 reports the characteristic of SPX constituents firms at each time t(- i.e, not only those that we have a full IVS history but also those that that would not be constituent for the entire sample). As shown in the table, the characteristics of both group are very close to each other and suggest that our sample is a good representative of SPX constituents.

[Table 1.2 about here.]

[Table 1.3 about here.]

For each horizon (h =1, 2, 3, 6 or 12), we compute the corresponding expected returns. Then, we apply PCA to 125 firms with 251 observations of expected returns for each

13. Data covers the same period that IRRA is extracted for- i.e., July 1998 to May 2019

horizon. This yields 125 PCA components for each horizon. Table 1.4 reports the amount of variance explained by the first five PCA factors for different horizons and suggest that the first PC factor explains almost 65-70% of the variance of h-month(s) expected returns ($h = 1, 2, 3, 6, 12$).

[Table 1.4 about here.]

1.4.3 Instrumented Principal Component Analysis

PCA extracts factors that capture the co-movement of variables but can only accommodate static loadings. Therefore, it cannot incorporate any information that is beyond returns and is contained in the characteristics[14] of the firms. Kelly and Pruitt (2019) proposed a novel technique for data reduction called instrumented principal components analysis or IPCA. This approach introduces observable firms' characteristics as instrumental variables and brings the informational content of characteristic to the factor models through latent time-varying loadings.

The general IPCA model specification is a system comprising N assets with L firms' characteristics over T periods. We assume that the expected return on asset i from time t to $t+1$ ($r_{i,t+1} = \mathbb{E}_t R_{i,t+1}$) computed from Equation (1.10) maps to a factor model proposed by Kelly and Pruitt (2019) as follows:

$$r_{i,t+1} = \alpha_{i,t} + \beta_{i,t} f_{t+1} + \epsilon_{i,t+1} \tag{1.11}$$

where f_{t+1} is a $(K \times 1)$ vector of K latent factors and inherently depends on lagged observed asset characteristics incorporated in the $(L \times 1)$ instrument vector $z_{i,t}$. The specification of $\beta_{i,t}$ is the main feature of the IPCA analysis and has two roles. First, it uses observable characteristics which serve as instrumental variables in the estimation of the latent factor loadings. Second, the time-varying instruments facilities the estimation of dynamic factor loadings. $\beta_{i,t}$ is estimated following the equation below:

$$\alpha_{i,t} = z'_{i,t}\Gamma_\alpha + \nu_{\alpha,i,t} \text{ and } \beta_{i,t} = z'_{i,t}\Gamma_\beta + \nu_{\beta,i,t} \tag{1.12}$$

14. Characteristics are factors that provide independent information about average returns.

The matrix Γ_β is a $(L \times K)$ vector and maps a large number of characteristics to a few risk factors. To find the Γ_β and the factor, we use the vector form of Equations (1.13) and (1.12)[15]:

$$r_{t+1} = Z_t \tilde{\Gamma} \tilde{f}_{t+1} + \epsilon_{t+1} \tag{1.13}$$

where r_{t+1} is an $(N \times 1)$ vector of individual firm returns, Z_t is the $(N \times L)$ matrix that stores the lagged characteristics of each firm, and ϵ^*_{t+1} is the residuals of individual firms. $\tilde{\Gamma} = [\Gamma_\alpha, \Gamma_\beta]$ and $\tilde{f}_{t+1} = [1, f_{t+1}]$. The estimation goal is to minimise the sum of squared model errors:

$$min_{\tilde{\Gamma},\tilde{f}} \sum_{t=1}^{T-1} (r_{t+1} - Z_t \tilde{\Gamma} \tilde{f}_{t+1})'(r_{t+1} - Z_t \tilde{\Gamma} \tilde{f}_{t+1}) \tag{1.14}$$

The values of IPCA latent factor and Γ that minimize the above equation will satisfy the first order conditions:

$$f_{t+1} = (\Gamma'_\beta Z'_t Z_t \Gamma_\beta)^{-1} \Gamma'_\beta Z'_t (r_{t+1} - Z_t \Gamma_\alpha), \tag{1.15}$$

and

$$vec(\tilde{\Gamma}') = \left(\sum_{t=1}^{T-1} [Z_t \otimes \tilde{f}'_{t+1}]'[Z_t \otimes \tilde{f}'_{t+1}] \right)^{-1} \left(\sum_{t=1}^{T-1} [Z_t \otimes \tilde{f}'_{t+1}]' r_{t+1} \right) \tag{1.16}$$

where $\tilde{\Gamma}$ is $(L \times (K+1))$ vector and is the coefficients of mapping the L instruments to the K factor loadings. We consider $K = 1$ in our analysis. These equations suggest that the IPCA factor is achieved by dynamic cross-section regression of r_{t+1} on the latent loading matrix β_t. Similarly, Γ_β, is the regression coefficient obtained from regressing returns on factors that interact with firm characteristics. There is no closed-form solution for the system of first-order conditions above, and therefore, must be solved numerically. To this end, following Kelly and Pruitt (2019) we choose an initial guess for $\tilde{\Gamma}$ equal to the second moment matrix, $\sum_t x_{t+1} x'_{t+1}$, where $x_{t+1} = Z'_t r_{t+1}$ is the time t observed returns on a set of L managed portfolios. The l^{th} component of x_t is a weighted average of stock returns with weights specified by the value of l^{th} characteristic of each firm at time t.

15. We consider the unrestricted model of Kelly and Pruitt (2019) in which $\Gamma_\alpha \neq 0$.

Provided with starting guess for $\tilde{\Gamma}$, we estimate the least squares regression based on the first-order condition (1.15) for all t. Next, having the consequent solutions for f_{t+1}'s, we estimate the least squares regression with respect to first-order condition (1.16). We repeat between estimations of (1.15) and (1.16) until convergence. The convergence is described as the point at which the maximum absolute change in any element of $\tilde{\Gamma}$ for $\tilde{f}_{t,T}$ (for all t) is smaller than 10^{-6}.

We empirically instrument the estimation of the first IPCA factor with observable characteristics data of the firms in our sample. Specifically, we focus on 125 firms in our sample that have constantly been included in the SPX index. Our sample spans July 1996 - May 2019. For each firm, we calculate 33 characteristics. These characteristics are defined exactly following Freyberger et al. (2020) and include market beta (beta), assets-to-market (a2me), total assets (assets), sales-to- assets (ato), book-to-market (bm), cash-to-short-term- investment (c), ratio of change in property, capital turnover (cto), capital intensity (d2a), earnings-to-price (e2p), fixed costs-to-sales (fc2y), cash flow-to-book (freecf), idiosyncratic volatility with respect to the FF3 model (idiovol), investment (invest), leverage (lev), market capitalization (mktcap), turnover (turn), net operating assets (noa), plants and equipment to the change in total assets (dpi2a), operating accruals (oa), operating leverage (ol), price-to-cost margin (pcm), profit margin (pm), gross profitability (prof), Tobin's Q (q), price relative to its 52-week high (w52h), return on net operating assets (rna), return on assets (roa), return on equity (roe), momentum (mom), intermediate momentum (intmom), short-term reversal (strev), long-term reversal (ltrev), sales-to-price (s2p), the ratio of sales and general administrative costs to sales (sga2s), bid-ask spread (bidask), and unexplained volume (suv). Section 1.C in the appendix describes these characteristics and their construction.

Figure 1.A.2 shows the evolution of the first factor obtained from the PCA model, the first factor obtained from the IPCA model, as well as the SVIX measure. These measures are highly correlated with SVIX, yet the correlation is not perfect; this implies that the informational content of the factors extracted from equity and index options may differ. Specifically, the correlation between SVIX and PCA factor is 87%. The correlation between SVIX and IPCA factor is 86%. The correlation between the PCA and the IPCA factor is 96%.

[Figure 1.A.2 about here.]

1.5 Predicting REA

Stock and Watson (2003) argue that stock prices reflect the expected discounted value of future earnings, making stock returns potentially useful variables for predicting output growth. We address this argument in this section by investigating whether the option-implied factors extracted from the stock level expected returns can predict subsequent growth in real economic activity

1.5.1 Empirical Setup

To identify whether stocks' $h-$month(s) expected returns predict REA growth over h forecasting horizons ($h = 1, 2, 3, 6, 12$ months), we set up our empirical investigation based on the following predictive regression:

$$REA_{i,t+h} = \beta_{i,0} + \beta_{i,1}REA_{i,t} + \beta_{i,2}F_{t,h} + \beta'_{i,3}X_t + \epsilon_{i,t+h} \tag{1.17}$$

where $REA_{i,t+h}$ is the growth rate of the ith REA proxy ($i = 1$ for IPI, 2 for NFP, 3 for RS, 4 for HS, 5 for CU, 6 for UR, 7 for CFNAI and 8 for ADS) over the period t to $t+h$. $F_{t,h}$ is the first factor obtained from PCA or IPCA methods using h-month(s) expected returns -i.e., we match the horizon of the expected returns with the forecasting horizon. X_t is a (4×1) vector of additional predictors at time t and includes credit spread, term spread, forward variance and IRRA. $\beta_{1,i}$, $\beta_{2,i}$ and $\beta'_{3,i}$ are (1×1), (1×1) and (1×4) vectors of regression coefficient parameters respectively and the ith REA proxy is the dependent variable.

In our further analysis, we also analyse the predictive power of our proposed predictor in the presence of common risk factors, which explain the cross-section of expected returns, namely the Fama and French (2015) five-factor model. This will tell us if the additional workload of extracting option-based equity returns adds some value compared to simply using asset pricing factors to predict REA. To test this, we add Fama and French (2015) five factors to the set of our control variables mentioned above and evaluate the predictive ability of our predictor in the presence of these factors. We fully discuss this model in Section 1.7.2.

Among different economic concerns that our predictive model may have, we first checked for the correlation matrix of the explanatory variables employed in our regression models. The covariance matrices of our control variables are reported in Table 1.5.

As indicated in Table 1.5, the forward variance is highly correlated with the first PCA (bigger than 0.50) for all horizons. Moreover, there is a high correlation between credit spread and the first PCA factor extracted from the expected return of 1, 2 and 3-month horizons. Moreover, as reported in Panel B of Table 1.5 the first IPCA factors obtained from h-month horizon(s) expected returns have a high correlation with forward variance and credit spread.

To alleviate multicollinearity concerns that arise from these correlations, we orthogonalise variables that have correlations bigger than 0.5 with the h-month PCA (IPCA) factor, namely forward variance and credit spread. More specifically, we regress these variables on a constant and the h-month first factor and obtain the residuals of these regressions. These residuals represent the part of the control variables orthogonal (uncorrelated) with the factor. Then, we run our predictive regressions using the first factor and residual terms as predictors.

[Table 1.5 about here.]

Secondly, to consider the possible econometric concerns arising from small samples or time-series properties of the regressors, we report the p-values for the regression coefficients in two ways: (i) Two-sided p-values of Newey and West (1994) which considers heteroscedasticity and autocorrelation of error terms and (ii) p-values of the IVX-Wald test of Kostakis et al. (2015), which considers time-series properties of predictors. The IVX-Wald test does not presume an initial assumption for the degree of persistence and allows different persistence levels for predictors. We describe the details of IVX estimator and test in Appendix 1.D.1 and 1.D.2.

1.6 Results

1.6.1 In-Sample Evidence

Table 1.6 shows the results of estimating Equation (1.17) for forecasting horizons $h =$ 1; 2; 3; 6; and 12 months employing the first PCA factor in our predictive regression model. To recall, our predictive model considers lagged REA, credit and term spread, forward and IRRA in the set of its control variables. This table reports the standardised ordinary-least-squares (OLS) coefficient estimates of the h-month(s) first PCA factor, the Newey-West and IVX-Wald test p-values of the regressor and the adjusted R^2 of the regression. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

We can obtain three key findings from Table 1.6. First, our results suggest that the first PCA is a significant predictor of most REA proxies at the short-term horizon (i.e., 1 and 2 months horizon). Specifically, the first PCA factor predicts all but one REA proxies (i.e., RS) at a one-month horizon. It also predicts six out of eight REA proxies, including IP, NFP, CU, UR CFNAI, and ADS, at a 2-months horizon. At three months horizon, the PCA factor also predicts UR and ADS. Second, extending the forecasting horizon to the longer term, our results suggest that the PCA factor can predict CU at 12 months horizon. Third, the sign of the first PCA factor coefficient is negative in all cases except for the unemployment rate. UR is a counter-cyclical measure of economic growth, and a positive regression coefficient of the first PCA factor also explains this property. These results suggest that a rise in expected return forecasts a decrease in REA.

Tables 1.A.2 to 1.A.3 in the appendix, report the regression coefficients of the control variables in our predictive model. These results show that among different control variables, forward variance and IRRA have significant predictive power to predict REA even at long-term horizons. For example, FV predicts all REA proxies except HOUS up to 3 months horizon. It also predicts 5 out of 8 proxies, including IP, NFP, CU, UR and ADS, up to 12 months horizon. On the other hand, IRRA predicts all REA proxies for up to 3 months. These results imply that index option-implied variables, especially

FV, strongly predict REA, even for a long horizon. However, even after controlling for these strong predictive variables in our model, our stock level predictor shows significant predictive power at short-term horizons. This implies that the stock-level option-based expected returns capture info which is not being captured by index-options implied variables, and this extra information is useful for predicting REA. For longer-horizon, however, one may choose forward variance to forecast REA.

The negative relationship between expected returns and REA is in an agreement with the prediction of consumption-based asset pricing models, such as Fama and French (1989) and Campbell and Cochrane (1999), which propose that the risk premia and, hence, the expected returns are high during recessions -i.e., investors become more risk-averse and require higher premiums to hold stocks during recessions.[16]

[Table 1.6 about here.]

We expand our analysis by including the second PCA factor in our predictive regression model. However, we find no significant predictive ability of this factor in forecasting REA. Additionally, we explore whether 1-month expected return factors can predict REA over longer time horizons (h = 2, 3, 6, 12). Unfortunately, these extensions do not yield improved results.

In order to compute a PCA factor from the cross section of individual expected returns, it was necessary to obtain a balanced sample. Alternatively, we derived a single factor by calculating a weighted average of the expected returns of the market index components at each time point. This approach effectively creates a proxy for the market expected return using individual options instead of relying on index options.

Table 1.7 shows the results of estimating Equation (1.17) for forecasting horizons $h = 1; 2; 3; 6;$ and 12 months employing the weighted factor in our predictive regression model. Our findings indicate that similar to the standard PCA factor, the weighted factor exhibits robust predictive power for REA within short-term horizons. For instance, even in the presence of other control variables such as IP, NFP, HOUS, UR, CFNAI, and ADS, the weighted factor continues to forecast six REA proxies at a one-month

16. Our results remain unchanged when we use risk premia of stocks instead of expected returns. This suggests that the effect of the risk-free rate is negligible.

horizon. Furthermore, at longer horizons, the factor successfully predicts UR for up to 12 months.

[Table 1.7 about here.]

Next, we incorporate the information embedded in the firms' characteristics, and proceed by evaluating the predictive capability of the IPCA factor in forecasting the REA. Table 1.8 reports the regression coefficients of estimating Equation (1.17) for forecasting horizons $h = 1$; 2; 3; 6; and 12 months using instrumented PCA factor as the main predictor of interest. Our results suggest that similar to the standard PCA factor, the first IPCA factor is a strong predictor of REA at short-term horizons. For example, at a one-month horizon, the first 1-month IPCA factor continues to predict 7 REA proxies even in the presence of other control variables. These variables include IP, NFP, HOUS, CU, UR, CFNAI and ADS. At a longer horizon, IPCA forecasts UR and ADS up to a 3-months horizon. We could not find predictive power of option-implied expected return for horizons greater than six months.

Comparing the R^2 of PCA and IPCA models suggests that both types of factors exhibit nearly the same predictive ability. These results imply that incorporating the informational content of stock characteristics into factor loading estimations does not improve the predictive ability of expected return factors extracted from options. This observation may be attributed to the high correlation between PCA and IPCA factors (92%), suggesting that options are sophisticated derivatives and may already encompass firms' characteristics.

[Table 1.8 about here.]

1.6.2 Out-of-Sample Evidence

This section assesses the forecasting ability of the PCA (IPCA) factors in a real time out-of-sample (OOS) setting from October 2007 to May 2019. This period is interesting because it covers the global financial crisis over 2007 to 2009 and includes the explosion of the U.S. housing bubble. It also includes the substantial monetary policies led by the Fed. To this end, we construct PCA/IPCA factors from $h-$month(s) expected returns for each REA proxy. Employing an expanding window, we recursively estimate Equation

(1.17)[17] Our first training sample window covers July 1998 to September 2007. Next, at each month (t), we construct $h = 1, 2, 3, 6, 12-$months ahead forecasts. Doing so, we bypass any look-ahead bias in our forecasts by utilising only the information available up to time t. Specifically, for every month $t \geq$ September 2007, we employ entire data through t to estimate the regression model (1.17) which includes our proposed factor[18] in the set of its predictor. We denote the resulting parameter estimates by $b_{1,i}$, $b_{2,i}$ and $b_{3,i}$. Then, based on (1.17), we define our full-forecasting model which contains the first PCA/IPCA factor in the set of predictors and is in the form of:

$$\mathbb{E}_t(REA^{full}_{i,t+h}) = b_{0,i} + b_{i,1}REA_{i,t} + b_{i,2}F_{t,h} + b'_{i,3}X_t \tag{1.18}$$

$\mathbb{E}_t(REA^{full}_{i,t+h})$ denotes the $h-$ months ahead out-of-sample forecasts from the full-forecasting model. $b_{1,i}$, $b_{2,i}$ and $b'_{3,i}$ are the (1×1), (1×1) and (1×4) vectors of estimated regression coefficients, respectively. The ith REA proxy ($i = 1$ for IPI, 2 for NFP, 3 for RS, 4 for HS, 5 for CU, 6 for UR, 7 for CFNAI and 8 for ADS) is the dependent variable. $F_{t,h}$ is the PCA or IPCA factor obtained from h-months expected returns. Similar to the in-sample analysis, X_t is the vector of control variables and consists of credit spread, term spread, forward variance and IRRA. Clearly, $\mathbb{E}_t(REA^{full}_{i,t+h})$ forecasts does not rely on the information beyond time t.

We also estimate regression (1.17) in a restricted form -i.e., with models which do not contain our proposed factor within the set of predictors. Specifically, we consider two alternative model specifications for benchmark (restricted) forecasts:

1. $benchmark_1$: This is an $AR(1)$ model which considers only lagged REA as the predictor of REA.

$$E_t(REA^{benchmark_1}_{i,t+h}) = b_{0,i} + b_{1,i}REA_{i,t} \tag{1.19}$$

2. $benchmark_2$: This model is in the form of:

$$E_t(REA^{benchmark_2}_{i,t+h}) = b_{0,i} + b_{1,i}REA_{i,t} + b_{i,2}X_t \tag{1.20}$$

17. We also perform the analysis using a rolling window, and our results remain unchanged.

18. obtained recursively from h-month(s) expected returns from July 1998 to $t - h$

where X_t includes the credit and term spread, forward variance, and IRRA.

Next, we compute the out-of-sample R^2 measure following Campbell and Thompson (2008) to examine the OOS forecasting performance of our proposed estimator (individual stocks expected returns PCA (IPCA) factors). The out-of-sample R^2 indicates whether the variance explained by forecasts of the full model, which includes the PCA (IPCA) factor in the set of its predictors, is more or less than the variance explained by alternative benchmark models ($benchmark_m$, $m = 1, 2$) which do not contain PCA factor in the set of its predictors. We define the out-of-sample $R^2_{i,h}$ acquired from predicting the ith REA proxy as:

$$R^2_{i,h} = 1 - \frac{var[E_t(REA^{Full}_{i,t+h}) - REA_{i,t+h}]}{var[E_t(REA^{benchmark_m}_{i,t+h}) - REA_{i,t+h}]} \tag{1.21}$$

A positive out-of-sample R^2 indicates that the full model performs better than the benchmark model; therefore, our proposed factor has superior OOS predictive ability.

Tables 1.9 and 1.10 report the out-of-sample R^2 obtained from Equation (1.21) where the full model includes PCA and IPCA factor, respectively.[19] Panels A and B of these tables report results when we consider the $AR(1)$ model in Equation (1.19) and the nested model in Equation (1.20), respectively.

Table 1.9 shows that the out-of-sample R^2 is mostly positive when we consider the PCA factor in our model -i.e., the full regression model with the PCA factor outperforms the restricted models. Therefore, incorporating the PCA factor in REA prediction models is statistically significant[20] in an out-of-sample setting. More precisely, for the case of the AR(1) model, the out-of-sample R^2 is positive in all proxies except for two cases. Exceptions occur at the 3 and 6 months horizons for CFNAI. As reported in Panel B, the full model forecasts are also better than the nested model ($benchmark_2$) for all proxies at short-term horizons (i.e., $h \leq 3$). The evidence on this is relatively weaker for longer horizons. At 6-months horizon, the full model performs better than the nested model

19. Our results remain robust if we calculate the OOS-R^2 measure using the sum of squared errors instead of the variance.

20. We test the statistical significance of the out-of-sample R^2s, using Diebold- Mariano test in Section (1.G) in the Appendix.

for all proxies except RS, UR, CFNAI and ADS. However, at 12 months horizon, the full model only outperforms the nested model for NFP and HOUS.

From Table 1.10, we can see that the forecasts of the full model with the IPCA factor beat the forecasts of the benchmark models that do not contain the IPCA factor for most REA proxies. More specifically, Panel A of the table shows that the OOS-R^2 is mostly positive and suggests that the full model outperforms the AR(1) for almost all proxies up to 12 months horizon (exceptions include CFNAI at $h > 1$ months horizon). In addition, Panel B of this table suggests that the full model with the IPCA factor beats the nested model ($benchmark_2$) for NFP, HOUS, UR and ADS at a one-month horizon. At longer horizons up to 6 months, the forecasts of the full model improve-i.e., the model with the IPCA factor is better than the nested model for all proxies except CU and CFNAI at two months and UR at six months.

[Table 1.9 about here.]

[Table 1.10 about here.]

Our findings on OOS analysis complement our results on in-sample analysis and suggest that our proposed equity-level factor contains additional information in predicting REA. This information cannot be captured by other option-implied measures such as IRRA and FV, especially at the short-term horizon. Therefore, the inclusion of our predictor is statistically significant in forecasting REA and the existing list of REA predictors should be extended by option-implied equity factors.

Furthermore, our findings suggest that the forecasts of the standard PCA factor are qualitatively similar to the instrumented PCA factor in an in-sample setting. In addition, the PCA factor forecasts perform better than the IPCA factor in an OOS setting for short-term horizons.[21] These results imply that equity options may already subsume firm's characteristics and hence the IPCA does not necessarily improve predicting REA.

21. -i.e., the $OOSR^2$ which compares the full model with the PCA factors versus the nested model is positive for all REA proxies up to 3 month, whereas the respective $OOSR^2$ for the IPCA factor is negative for four proxies at 1-month horizon and for 2 proxies at 2-months horizon.

1.7 Further Analysis

1.7.1 Variance-based versus Density-based Expected Returns

Chabi-Yo et al. (2022) propose a flexible technique to build expected return bounds that consider the entire shape of the risk-neutral distribution rather than considering only the second moment as in Martin and Wagner (2019). Specifically, they suggest that:

$$E_t(R_{i,t+1} - R_{f,t+1}) = max_{\theta \in \Theta_{i,t}} \{E_t^*(R_{i,t+1}^{\theta+1})/E^*(R_{i,t+1}^{\theta}) - R_{f,t+1}\} \tag{1.22}$$

where $R_{i,t+1}$ is the the simple gross return on stock i and $R_{f,t+1}$ is the risk-free rate. The risk-neutral expectations, E^*, in (3.15) can be calculated straightly from observed option prices for a given θ where the set of Θ is equivalent to:

$$\Theta_{i,t} = \{\theta : \theta - \gamma_t \beta_{Mi} + \frac{\gamma_t \rho_t}{R_{f,t+1}} \beta_{M^2,i} + \frac{\theta(\theta-1)}{2R_{f,t+1}} B_{i^2,i} - \frac{\theta \gamma_t}{R_{f,t+1}} \beta_{Mi,i} \leq 0\} \tag{1.23}$$

Following Chabi-Yo et al. (2022) we use parameters $\gamma_t = 2$ and $\rho_t = 1$. $\beta_{i^2,i}$, $\beta_{M,i}$, $\beta_{Mi,i}$ and $\beta_{M^2,i}$ are defined as follows:

$$\begin{aligned}
&\beta_{M,i} = COV(R_{M,t+1}, R_{f,t+1})/VAR(R_{i,t+1}) \\
&\beta_{M^2,i} = COV_t((R_{M,t+1} - R_{f,t+1})^2, R_{i,t+1})/VAR_t(R_{i,t+1}), \\
&\beta_{i^2,i} = COV_t(R_{i,t+1} - R_{f,t+1})^2, R_{i,t+1}/VAR_t(R_{i,t+1}); \\
&\beta_{Mi,i} = COV_t((R_{M,t+1} - R_{f,t+1})(R_{i,t+1} - R_{f,t+1}), R_{i,t+1})/VAR_t(R_{i,t+1})
\end{aligned} \tag{1.24}$$

where $R_{M,t+1}$ is the market return. We compare the predictive power of the PCA factor obtained from expected returns extracted from Chabi-Yo et al. (2022) which considers the higher moments of the risk-neutral distribution, with those extracted from variance-based models of Martin and Wagner (2019) in an out-of-sample setting. Panel B of Table 1.11 reports the $OOS - R^2$ obtained from Equation (1.21), where the full model considers factors obtained from variance-based models and the benchmark model contains the factors obtained from Chabi-Yo et al. (2022) 's method. Looking closely at the feasible θ in Equation (1.23), we found that for 87% of the returns, the feasible θ is equal to one, and the derived bounds are based on the second moments. Therefore, the predictive power of the option-based expected returns predictor remains regardless of the method used to compute the option-implied expected returns. The positive (albeit small) out-of-sample R^2 suggest that it is possible that the estimation errors in the density-based measure are contributing to this discrepancy. The limited information captured by the density-based measure, based mostly on second moments, may lead to challenges in enhancing the predictive ability of the option-based factor.

[Table 1.11 about here.]

1.7.2 Option-Implied Factors versus Equity Factors

PCA or IPCA are alternative models to explain the variation in expected returns. A natural question is whether or not the option-implied PCA/IPCA factors outperform the traditional pricing factors such as those from Fama and French (2015). We answer this question by extending our analysis to both in- and out-of-sample settings.

First, we assess the predictive performance of option-implied factors in an in-sample setting. To this end, we augment the control variables in the predictive regression model (1.17) with 5 Fama-French factors (5FF) proposed by Fama and French (2015). More precisely, our new predictive model considers a constant, lagged REA, the extracted first PCA or IPCA factor from h-month ($h = 1, 2, 3, 6$ *or* 12 *-months*) expected returns, credit and term spread, forward variance and IRRA plus five factors of Fama-French including the size of firms (SMB), book-to-market values (HML), expected return on the market (mrkt), profitability factor (RMW) and investment factor (CMA).[22] A significant regression coefficient suggests that our proposed forward-looking stock-level option-implied risk factors contain additional predictive information relative to the 5FF.

Table 1.12 and 1.13 report the in-sample regression coefficients of PCA and IPCA factors, respectively. These tables suggest that our results are robust, and both PCA and IPCA factors have predictive power even after controlling for five Fama-French factors. We can see that at a one-month horizon, both PCA and IPCA factors predict 7 out of 8 REA proxies, including IP, NFP, HOUS, CU, UR, CFNAI and ADS. At 2-months horizon,

22. Fama-French factors capture the different dimension of systematic risk in stock prices and are expected to explain the cross-section of stock's average returns with a linear relationship.

the PCA factor remains predictive for six proxies while IPCA predicts four proxies. However, these factors become less predictive at a longer horizon, i.e., at three months horizon, both factors only predict UR, while at 12 months horizon, only the PCA factor predicts CU. These results highlight the forward-looking nature of options and show that the expected returns implied from options contain additional information over the systematic risk factors obtained from the stock market. This extra information is useful in predicting REA, especially at short-term horizons.

[Table 1.12 about here.]

[Table 1.13 about here.]

Second, we compare the performance of our proposed predictor (option-implied stocks' expected returns PCA/IPCA factor) with five factors of Fama-French (5FF) in an out-of-sample setting. To this end, we define alternative predictive regression models ($benchmark_3$) which do not contain the PCA/IPCA factor and instead contain 5FF and are in the form of:

$$E_t(REA_{i,t+h}^{benchmark_3}) = b_{0,i} + b_{1,i}REA_{i,t} + b_{i,2}FF5_t + b'_{i,3}X_t \tag{1.25}$$

where X_t includes the credit and term spread as well as the forward variance and IRRA.

We compare the forecasts of our full predictive model explained in Equation (1.18) with forecasts of Equation (1.25) and compute the OOSR^2 following Equation (1.21). Panel A and B of Table 1.16 report the OOS results and suggest that the predictive models with either of our proposed predictors outperform the alternative model with 5 Fama-French factors. More specifically, in the case of the full model with the PCA factor, the OOS R^2 is positive for all proxies at all horizons but two cases, which are six and twelve-month horizons for UR. On the other hand, the full model with IPCA outperforms the alternative model with 5FF factors for NFP, RS, HOUS, CFNAI and ADS at all horizons. Our findings imply that although the derived PC factors may be interpreted as "level" factors that proxy for the market, individual stocks contain some additional predictive power that is not captured by standard risk factors.

1.7.3 Individual Equity Options versus Index Options

Lemmon and Ni (2014) points out that individual equity options and index options are dissimilar and attract different types of investors. Individual equity options are actively traded by individual investors and are influenced by investor sentiment. In contrast, index options are primarily used by sophisticated traders for hedging purposes.

The differences between individual equity options and index options lead to distinct advantages for single stock options. Firstly, single stock options provide investors with specific exposure to a particular company. This enables them to focus their trading strategies on individual companies and express their views on the movement of those companies' stock prices. Such focused exposure proves advantageous for investors who possess in-depth knowledge or strong opinions regarding the future prospects of specific companies. Secondly, single stock options have the potential for higher returns compared to index options. Since they are directly tied to the performance of a specific company, accurately predicting the price movement of an individual stock can yield substantial profits. Thirdly, single stock options provide more flexibility in managing risk. Investors can create custom risk profiles by selecting individual stocks with specific risk-reward characteristics that align with their investment objectives and risk tolerance. This allows for more precise hedging or speculative strategies tailored to the desired risk exposure. Lastly, individual stocks often exhibit higher levels of price volatility compared to broad market indices. This creates opportunities for options traders to capitalize on larger price swings. Higher volatility generally leads to higher option premiums, and skilled options traders can take advantage of expected volatility in specific stocks by employing strategies such as straddles or strangles.

To highlight the additional value of single stocks versus index options, we begin by comparing the market expected return implied by index options (SVIX) with the weighted factor derived from single options. To do so, we regress the weighted factor obtained from individual options on the market expected return (SVIX) and take the residuals. Next, we assess the predictive ability of the residuals. Table 1.14 presents the in-sample estimated regression coefficients of the residuals of the weighted factor on the SVIX measure and shows predictive ability of the residual component to forecast IP, CU, CFNAI, and ADS in short-term horizon. Our findings suggest that single stock options offer certain advantages over index options, which can make them attractive to investors.

[Table 1.14 about here.]

Given the high correlation between the PCA factor and the SVIX (96%), another way to understand better the added value of using individual options is to run a regression of the PCA on the SVIX and then analyze the dynamics of the residual component. Therefore, we analyze the persistence in the predictive ability of the residual. Table 1.15 shows the results of estimating Equation (1.17) for forecasting horizons $h = 1; 2; 3; 6;$ and 12 months employing the residual component in our predictive regression model in the presence of the SVIX. Our results suggest that the residual component is still a significant predictor of most REA proxies (i.e., IP, NFP, CFNAI, and ADS) even after controlling for the SVIX measure. This finding implies that the individual options contain some predictive power in forecasting REA that cannot be captured by info obtained from index options.

[Table 1.15 about here.]

Furthermore, we compare the predictive performance of our proposed predictor (option-implied stocks' expected returns PCA/IPCA factor) with the index expected return (SVIX) in an out-of-sample setting. Doing so, we examine whether the individual expected returns factors contain some predictive power in forecasting REA that cannot be captured by info obtained from index options. To this end, we define an alternative predictive regression model ($benchmark_4$), which does not contain the PCA/IPCA factor and contains SVIX instead. More specifically, we compare the forecasts of our full predictive model explained in Equation (1.18), which includes the expected return factors (PCA/IPCA)[23], with the forecast obtained from the following model:

$$E_t(REA_{i,t+h}^{benchmark_4}) = b_{0,i} + b_{1,i}REA_{i,t} + b_{i,2}SVIX_t + b_{i,3}'X_t \tag{1.26}$$

We compute the OOSR^2 following Equation (1.21) and report the results in Table 1.16. Panel C and Panel D of Table 1.16 report the OOS results when the full models include

23. Augmented by a set of control variables (X_t)- namely, lagged REA, credit spread, term spread, forward variance and IRRA

PCA and IPCA, respectively, and the benchmark model contains the SVIX measure (Equation 1.26). Panel C of the table shows that the comparative performance of a predictive model with PCA factor is better than the predictive model with the market SVIX measure, especially at short-term horizon -i.e., at $h = 1$ it outperforms all proxies except RS. Panel D of this table suggests that the proposed model with the IPCA factor beats the benchmark model with the SVIX measure only for NFP, HOUS, UR, CFNAI and ADS at one-month horizon.

Our findings imply that the individual stocks' option-implied standard PCA factors contain some useful information in predicting REA that the index options cannot capture[24] and therefore, the list of REA predictors should be extended by individual's option-implied PCA factor.

[Table 1.16 about here.]

1.8 Conclusions

Campbell and Cochrane (1999) suggests that expected return is an appropriate state variable for forecasting output growth and is countercyclical. In addition, Fama and French (1989) suggest that the risk premia and, hence, the expected returns are high during recessions -i.e., investors become more risk-averse and require higher premiums to hold stocks during recessions.[25] Due to the econometric challenges in estimating expected returns, previous literature has not examined its predictive ability. We bypass these obstacles by employing methods which use information from option prices. Doing so, we exploit the information embedded in equity option prices and investigate the predictive ability of the expected returns of the individual stocks extracted from the equity option prices to forecast the growth of U.S. REA. As expected by theory, we find

24. Especially at short-term horizon. This might be due to the fact that individual options are mostly traded for speculative reasons and at short horizons, while index options are used for hedging reasons and longer horizon. More precisely, individual options offer the potential for high returns within a short time frame -i.e, they can experience significant price movements based on company-specific news, events, or earnings reports. Therefore, options traders may take short-term positions to capitalize on these anticipated price changes. Speculative traders aim to profit from fluctuations in the price of the underlying stock and often employ strategies such as buying call or put options or engaging in more complex option strategies to generate quick profits.

25. The rationale is that as business conditions and investment opportunities change, expected returns change because investors attempt to smooth consumption over multiple periods, which generally predict a countercyclical risk premium.

a significant predictive ability of our proposed option-implied factor for forecasting REA both in and out-of-sample.

To build our proposed predictor, first, we follow Martin and Wagner (2019) to compute option-implied expected returns on stocks that have been constituents of the S&P500 index from July 1998 to May 2019. We then employ dimension reduction techniques, namely principal component analysis and instrumented principal component analysis, to construct expected return factors. We document that our expected return factor is a new predictor of U.S. real economic activity (REA) and has significant predictive power even after controlling for a well-known set of variables that has been documented by the previous literature to predict REA- namely, term spread credit spread, forward variance and IRRA. Our results are robust after addressing some econometric concerns, such as dealing with the unknown persistence degree of predictors and heteroscedasticity and auto-correlation of error terms. In addition, our findings remain robust even after contorting for the systematic risk factors of Fama and French (2015) in the in-sample analysis and imply that individual expected returns contain some predictive power in addition to the standard systematic risk factors.

We also examine the out-of-sample performance of our proposed predictive model, which contains the expected return factor within the set of its predictor. Our findings suggest that incorporating the expected return factors in predictive regression models is also statistically significant in an out-of-sample setting. In the case of forecasts obtained by the regression models, the evidence is relatively weaker for longer horizons. We also document that the OOS forecasts obtained from a predictive model with our proposed predictor are better than the OOS forecasts obtained from benchmark models, which contain the market risk premium (SVIX) measure of Martin (2017) or the five systematic risk factors of Fama and French (2015). Our findings imply that individual stocks contain some additional predictive power that is not captured by the index expected return nor by standard factors that explain the cross-section of expected returns.

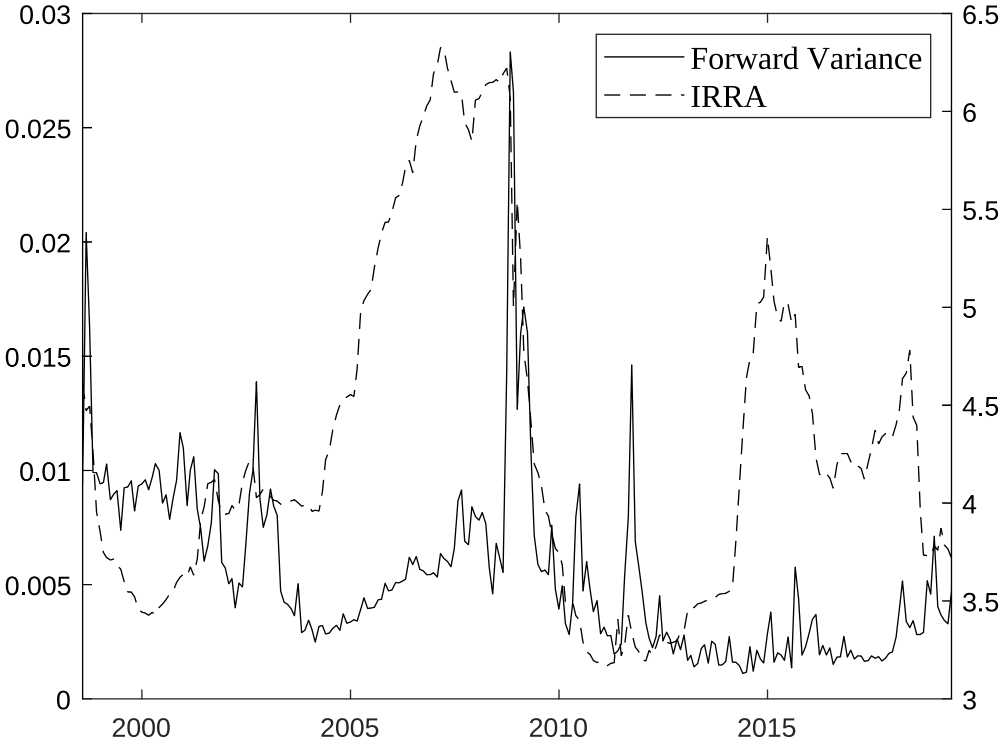


Figure 1.A.1. Evolution of the U.S. Forward Variance and Implied Risk Aversion (IRRA): The Right axis shows the evolution of IRRA while the left axis shows the evolution of forward variance over July 1998 to May 2019. We compute the IRRA time series following Faccini et al. (2019) by performing a generalised method of moments (GMM) using a rolling window estimation. We employ a rolling window with a size of 30 months to obtain the respective U.S. IRRA time series. Forward variance is constructed using index option prices as explained in section 1.3.3

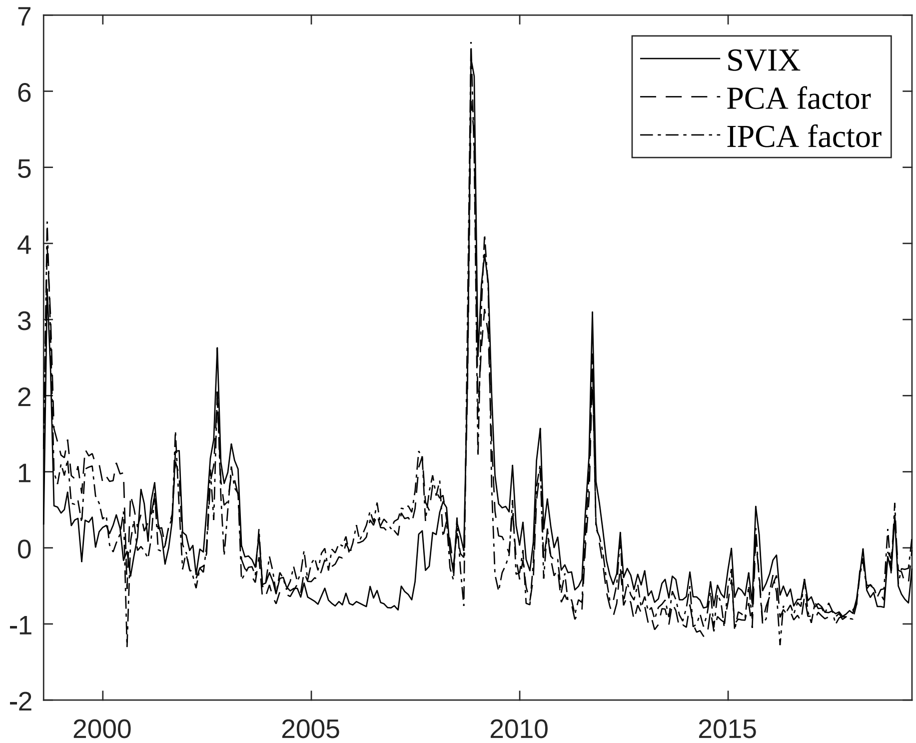


Figure 1.A.2. Evolution of the SVIX measure and PCA and IPCA factors: The figure shows the evolution of SVIX, first PCA and IPCA factor obtained from expected returns over July 1998 to May 2019. We extract the SVIX time series via Martin (2017) formula using prices of index options. We construct the PCA and IPCA factor as described in section 1.4.2 and 1.4.3 respectively.

Panel A: Daily Data for SPX firms					
horizon	30	60	90	182	365
Observations	2,837,127	2,837,127	2,837,127	2,837,127	2,837,127
Sample days	5892	5892	5892	5892	5892
Sample firms	1092	1092	1092	1092	1092
Average firms/day	482	482	482	482	482
Panel A: Monthly Data for SPX firms					
horizon	30	60	90	182	365
Observations	135,343	135,343	135,343	135,343	135,343
Sample months	281	281	281	281	281
Sample firms	1092	1092	1092	1092	1092
Average firms/month	482	482	482	482	482

Table 1.1. Sample Data: This table reports a summary of the data used in the empirical analysis. We searched the OptionMetrics database for all firms included in the S&P 500 from January 1996 to May 2019 and obtained all available volatility surface data. Using these data, we compute firms' risk-neutral variances ($SVIX^2_{i,t}$) for horizons of 1, 2, 3, 6 and 12 months. Panel A reports the total number of observations, the number of unique days and firms in our sample, and the average number of firms for which options data are available per day. Panel B reports monthly observations subsets for constituents of the S&P 500.

Panel A			
code	industry	number of firms	% of firms over industry
1	basic material	95	9%
3	Consumer Goods	97	9%
4	Financial	129	12%
5	Health care	84	8%
6	industrial Goods	69	6%
7	Services	171	16%
8	Technology	148	14%
9	Utilities	45	4%
	Unknown-from OM	220	20%

Panel B			
code	industry	number of firms	% of firms over industry
1	basic material	16	13%
3	Consumer Goods	17	13%
4	Financial	16	13%
5	Health care	13	10%
6	industrial Goods	20	16%
7	Services	25	20%
8	Technology	11	9%
9	Utilities	8	6%
	Unknown-from OM	1	1%

Table 1.2. Industrial Sector Coverage: Panel A reports the percentage of firms considering all firms in our data set (1092 firms), while Panel B reports the percentage of firms that covers each industrial sector considering firms that have been constantly in the SPX index and have no missing data (125 firms).

	Panel A: Constant constituents of SPX				Panel B: All constituents of SPX			
	mean	std	min	max	mean	std	min	max
a2me	1.66	2.92	0.06	53.01	1.94	4.47	0.00	210.03
at $*10^4$	5.72	16.76	0.04	218.76	2.64	11.20	0.00	228.12
ato$*10^2$	0.01	0.28	-10.47	3.57	0.02	0.52	-33.30	19.59
beme	0.40	0.32	-0.70	3.07	0.48	0.62	-11.69	19.70
beta	0.57	0.23	-0.11	2.26	0.61	0.29	-0.33	3.08
c	0.11	0.12	0.00	0.72	0.13	0.16	0.00	0.99
cto	1.05	0.79	0.03	6.77	1.16	0.89	0.00	18.65
$r_{12,2}$	0.14	0.30	-0.89	3.04	0.16	0.47	-0.99	12.99
$r_{12,7}$	0.08	0.21	-0.91	3.04	0.09	0.31	-0.97	11.83
$r_{1,0}$	0.01	0.09	-0.59	0.87	0.01	0.11	-0.96	2.45
$r_{36,12}$	0.32	0.53	-0.92	7.98	0.38	1.00	-0.99	77.49
d2a	0.04	0.02	0.00	0.21	0.04	0.03	0.00	0.60
dpi2a	0.05	0.13	-2.50	2.43	0.06	0.20	-2.50	10.45
e2p	0.05	0.06	-1.33	0.49	0.02	0.38	-23.51	7.36
fc2y	0.27	0.19	0.00	1.42	0.42	4.75	-0.05	422.09
$free_{cf}$	0.28	18.49	-492	764	-0.23	36.26	-3786	764
$idio_{vol}$	0.01	0.01	0.00	0.12	0.02	0.02	0.00	0.94
investment	0.09	0.23	-0.66	4.00	0.16	0.85	-1.00	55.26
lev	0.44	0.25	0.00	2.83	0.44	1.04	-73.17	40.44
lme $*10^7$	4.71	7.59	0.08	109.94	1.78	4.18	0.00	109.94
lturnover	1.49	1.16	0.13	28.07	2.02	2.18	0.00	102.96
noa	0.54	0.27	-0.30	3.74	0.61	0.64	-3.76	51.13
oa	-0.04	0.05	-0.40	0.74	-0.04	0.15	-10.86	3.14
ol	0.81	0.67	0.02	4.66	0.88	0.68	0.01	6.13
pcm	0.41	0.23	-3.37	0.98	0.35	1.67	-99.89	1.00
pm	0.16	0.15	-4.06	0.57	0.00	5.01	-457.23	0.83
prof	0.97	86.63	-2831	3193	0.06	97.43	-8993	3193
q	2.14	1.34	0.61	16.47	2.31	9.03	0.41	925.25
$rel2high_{price}$	0.84	0.16	0.06	1.00	0.81	0.19	0.02	1.00
rna	0.16	3.27	-79	55	0.30	20.06	-936	1147
roa	0.08	0.06	-0.40	0.55	0.05	0.17	-9.13	2.84
roe	0.25	18.93	-579	675	-0.23	40.06	-4251	675
s2p	0.90	0.81	0.03	8.15	1.30	2.70	0.00	90.09
sga2m	0.16	0.14	0.00	1.61	0.24	0.46	- 0.14	15.98
spread_mean $*10^{-2}$	0.32	0.59	- 0.12	6.77	0.47	0.92	- 7.29	72.84

Table 1.3. Descriptive Statistics of Firms Characteristics: Panel A provides summary statistics for the characteristics of firms employed in our analysis. The sample comprises the 125 largest US firms that have been SPX constituents consistently over the sample period July 1998 to May 2019. Panel B provides the summary statistics of all firms that have been SPX constituent, but not necessarily constantly. We cover 1,092 firms over our sample period from July 1998 to May 2019. For each horizon, we collect data for 482 firms on average per day.

	$h = 1\text{M}$	$h = 2\text{M}$	$h = 3\text{M}$	$h = 6\text{M}$	$h = 12\text{M}$
First Factor	70.20	68.81	70.01	67.16	65.09
Second Factor	7.73	8.71	9.09	10.90	12.80
Third Factor	4.91	5.34	4.98	5.34	5.28
Forth Factor	2.65	2.94	2.96	3.41	3.51
Fifth Factor	1.60	1.57	1.45	1.66	1.69

Table 1.4. Variance Explained by PCA Factors: This table reports the percentage of variance explained by the first 5 PCA factors obtained from individual stocks' options-implied expected returns, considering firms that have been constantly included in SPX for different horizons under our considerations.

Panel A - First PCA factor

	TS	CS	mrkt	SMB	HML	RWM	CMA	FV	SVIX	IRRA
1-months Factor	-0.18	0.51	-0.07	0.21	0.22	-0.10	0.02	0.97	0.87	0.20
2-months Factor	-0.22	0.51	-0.06	0.21	0.20	-0.10	0.01	0.96	0.85	0.21
3-months Factor	-0.24	0.50	-0.06	0.22	0.19	-0.10	0.03	0.95	0.83	0.20
6-months Factor	-0.29	0.46	-0.05	0.22	0.15	-0.13	0.01	0.92	0.76	0.19
12-months Factor	-0.31	0.40	-0.03	0.23	0.12	-0.15	0.00	0.89	0.70	0.17

Panel B - First IPCA factor

	TS	CS	mrkt	SMB	HML	RWM	CMA	FV	SVIX	IRRA
1-months Factor	-0.15	0.52	-0.07	0.17	0.24	-0.03	0.02	0.92	0.87	0.25
2-months Factor	-0.18	0.52	-0.06	0.17	0.22	-0.03	0.00	0.92	0.85	0.25
3-months Factor	-0.22	0.51	-0.06	0.19	0.21	-0.05	0.01	0.93	0.83	0.24
6-months Factor	-0.27	0.46	-0.04	0.20	0.15	-0.09	0.00	0.89	0.75	0.22
12-months Factor	-0.30	0.39	-0.02	0.21	0.11	-0.12	-0.02	0.86	0.68	0.19

Panel C - Control variables

	TS	CS	mrkt	SMB	HML	RWM	CMA	FV	SVIX	IRRA
TS	1.00	0.16	-0.01	-0.09	0.02	0.02	-0.04	-0.15	0.15	-0.38
CS	0.16	1.00	0.09	0.07	0.15	-0.08	-0.07	0.51	0.70	0.15
mrkt	-0.01	0.09	1.00	-0.04	-0.02	-0.34	-0.43	-0.06	-0.08	-0.02
SMB	-0.09	0.07	-0.04	1.00	-0.04	-0.29	0.03	0.22	0.16	-0.01
HML	0.02	0.15	-0.02	-0.04	1.00	-0.06	0.44	0.24	0.25	0.17
RWM	0.02	-0.08	-0.34	-0.29	-0.06	1.00	0.17	-0.11	-0.05	0.02
CMA	-0.04	-0.07	-0.43	0.03	0.44	0.17	1.00	0.04	0.03	-0.04
FV	-0.15	0.51	-0.06	0.22	0.24	-0.11	0.04	1.00	0.90	0.14
SVIX	0.15	0.70	-0.08	0.16	0.25	-0.05	0.03	0.90	1.00	-0.01
IRRA	-0.38	0.15	-0.02	-0.01	0.17	0.02	-0.04	0.14	-0.01	1.00

Table 1.5. Correlation Coefficient Between Predictors: Panel A reports the correlation coefficients between the first PCA factor extracted from the h months (i.e., h =,1, 2, 3, 6, 12) individual's expected returns and a set of control variable, namely: term spread (TS), credit spread (CR), forward variance (FV) and Fama and French 5 factors (market, SMB, HML, RMW and CMA). Panel B reports the correlation coefficient between control variables and the $h-$month first IPCA factor. Panel C shows the correlation coefficient between control variables.

		$h = 1$ month	$h = 2$ months	$h = 3$ months	$h = 6$ months	$h = 12$ months
IP	β	-0.26***	-0.22**	-0.03	0.04	0.01
		[0.00]	[0.01]	[0.58]	[0.61]	[0.48]
		(0.00)	(0.00)	(0.62)	(0.67)	(0.98)
	R2	0.25	0.41	0.48	0.41	0.32
NFP	β	-0.26***	-0.17**	-0.07	-0.09	-0.17
		[0.00]	[0.01]	[0.49]	[0.20]	[0.13]
		(0.00)	(0.00)	(0.22)	(0.19)	(0.08)
	R2	0.68	0.78	0.79	0.74	0.55
RS	β	-0.02	0.03	0.13	0.13	0.01
		[0.68]	[0.87]	[0.60]	[0.72]	[0.97]
		(0.71)	(0.71)	(0.27)	(0.35)	(0.96)
	R2	0.10	0.17	0.23	0.27	0.35
HOUS	β	-0.13*	-0.12	-0.09	-0.23	-0.18
		[0.06]	[0.15]	[0.66]	[0.41]	[0.89]
		(0.16)	(0.20)	(0.32)	(0.01)	(0.02)
	R2	0.20	0.19	0.17	0.37	0.59
CU	β	-0.30***	-0.25**	-0.10	-0.06	-0.20*
		[0.00]	[0.01]	[0.36]	[0.29]	[0.08]
		(0.00)	(0.00)	(0.09)	(0.55)	(0.10)
	R2	0.26	0.42	0.49	0.39	0.30
UR	β	0.36***	0.37***	0.24*	0.13	0.20
		[0.00]	[0.00]	[0.07]	[0.45]	[0.30]
		(0.00)	(0.00)	(0.01)	(0.20)	(0.11)
	R2	0.20	0.35	0.42	0.58	0.52
CFNAI	β	-0.35***	-0.14*	0.02	-0.01	-0.14
		[0.00]	[0.09]	[0.75]	[0.80]	[0.75]
		(0.00)	(0.01)	(0.65)	(0.93)	(0.32)
	R2	0.52	0.53	0.44	0.28	0.14
ADS	β	-0.10**	-0.18***	-0.03*	0.01	-0.15
		[0.00]	[0.00]	[0.09]	[0.60]	[0.33]
		(0.00)	(0.00)	(0.57)	(0.87)	(0.26)
	R2	0.87	0.76	0.68	0.41	0.21

Table 1.6. In-sample Prediction of REA Proxies with the First PCA Factor: This table presents the in-sample estimated regression coefficients of the first PCA factor for different U.S. real economic activity (REA) proxies. We matched the forecasting horizon with the horizon of expected returns. For each REA under our consideration, we consider predictive regression model 1.17 as explained in Section 1.5 and employ the lagged REA, term-spread, credit-spread, forward variance and IRRA in the set of our control variables. Reported entries are the standardized ordinary-least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors as well as the in-sample adjusted R^2 for the given model. The sample covers July 1998 to May 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

		h=1 month	h=2 month	h=3 month	h=6 month	h=12 month
IP	beta	-0.32***	-0.24**	-0.10	0.00	-0.06
		[0.00]	[0.03]	[0.33]	[0.40]	[0.60]
		(0.00)	(0.01)	(0.18)	(0.96)	(0.61)
	R2	0.20	0.33	0.42	0.34	0.28
NFP	beta	-0.27***	-0.18**	-0.15*	-0.16**	-0.23
		[0.00]	[0.02]	[0.06]	[0.04]	[0.11]
		(0.00)	(0.01)	(0.03)	(0.04)	(0.03)
	R2	0.67	0.76	0.76	0.70	0.53
RS	beta	-0.03	-0.03	-0.06	0.04	-0.06
		[0.67]	[0.80]	[0.57]	[0.88]	[0.87]
		(0.65)	(0.69)	(0.57)	(0.75)	(0.62)
	R2	0.09	0.10	0.15	0.25	0.35
HOUS	beta	-0.11*	-0.08	-0.07	-0.17	-0.23
		[0.06]	[0.21]	[0.53]	[0.27]	[0.80]
		(0.22)	(0.38)	(0.40)	(0.02)	(0.00)
	R2	0.20	0.19	0.17	0.37	0.61
CU	beta	-0.33***	-0.25**	-0.10	-0.02	-0.23*
		[0.00]	[0.02]	[0.28]	[0.25]	[0.08]
		(0.00)	(0.01)	(0.19)	(0.84)	(0.06)
	R2	0.22	0.36	0.45	0.34	0.28
UR	beta	0.36***	0.38***	0.36***	0.16	0.28
		[0.00]	[0.00]	[0.00]	[0.31]	[0.29]
		(0.00)	(0.00)	(0.00)	(0.11)	(0.03)
	R2	0.18	0.31	0.39	0.55	0.51
CFNAI	beta	-0.38***	-0.15	-0.05	0.03	-0.17
		[0.00]	[0.24]	[0.64]	[0.85]	[0.77]
		(0.00)	(0.02)	(0.51)	(0.78)	(0.20)
	R2	0.48	0.48	0.41	0.27	0.15
ADS	beta	-0.10***	-0.16***	-0.11**	0.02	-0.19
		[0.00]	[0.00]	[0.03]	[0.29]	[0.43]
		(0.02)	(0.01)	(0.10)	(0.85)	(0.14)
	R2	0.86	0.72	0.62	0.37	0.22

Table 1.7. In-sample Prediction of REA Proxies with the Weighted Factor This table presents the in-sample estimated regression coefficients of the weighted factor (by computing a weighted average of the expected returns of the components of the market index at each time t.) for different U.S. real economic activity (REA) proxies. We matched the forecasting horizon with the horizon of expected returns. For each REA under our consideration, we consider predictive regression model 1.17 as explained in Section 1.5 and employ the lagged REA, term-spread, credit-spread, forward variance and IRRA in the set of our control variables. Reported entries are the standardized ordinary-least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors as well as the in-sample adjusted R2 for the given model. The sample covers July 1998 to May 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

		$h = 1$ month	$h = 2$ months	$h = 3$ months	$h = 6$ months	$h = 12$ months
IP	β	-0.28***	-0.26**	-0.12	0.06	0.05
		[0.00]	[0.03]	[0.41]	[0.94]	[0.82]
		(0.01)	(0.01)	(0.04)	(0.47)	(0.65)
	R2	0.20	0.32	0.42	0.41	0.36
NFP	β	-0.24***	-0.15	-0.14	-0.06	-0.14
		[0.00]	[0.11]	[0.27]	[0.64]	[0.43]
		(0.00)	(0.01)	(0.03)	(0.36)	(0.16)
	R2	0.67	0.78	0.79	0.75	0.57
RS	β	-0.01	0.07	0.05	0.19	0.05
		[0.70]	[0.37]	[0.48]	[0.50]	[0.88]
		(0.84)	(0.26)	(0.62)	(0.14)	(0.71)
	R2	0.09	0.17	0.26	0.31	0.36
HOUS	β	-0.17**	-0.09	-0.04	-0.17	-0.13
		[0.02]	[0.21]	[0.79]	[0.51]	[0.91]
		(0.10)	(0.37)	(0.67)	(0.06)	(0.11)
	R2	0.21	0.19	0.18	0.36	0.57
CU	β	-0.31***	-0.28**	-0.13	-0.04	-0.15
		[0.00]	[0.02]	[0.34]	[0.71]	[0.38]
		(0.00)	(0.01)	(0.02)	(0.61)	(0.20)
	R2	0.22	0.35	0.44	0.39	0.32
UR	β	0.34**	0.34**	0.32**	0.12	0.17
		[0.00]	[0.01]	[0.04]	[0.65]	[0.47]
		(0.00)	(0.00)	(0.00)	(0.25)	(0.16)
	R2	0.19	0.34	0.43	0.58	0.53
CFNAI	β	-0.33***	-0.13	-0.02	0.04	-0.11
		[0.00]	[0.22]	[0.66]	[0.99]	[0.83]
		(0.00)	(0.03)	(0.74)	(0.63)	(0.43)
	R2	0.48	0.49	0.44	0.30	0.14
ADS	β	-0.08**	-0.15**	-0.11*	0.06	-0.12
		[0.01]	[0.01]	[0.09]	[0.83]	[0.82]
		(0.02)	(0.01)	(0.02)	(0.50)	(0.35)
	R2	0.86	0.73	0.65	0.44	0.22

Table 1.8. In-sample Prediction of REA Proxies with the First IPCA Factor: This table presents the in-sample estimated regression coefficients of the first instrumented PCA factor for various U.S. real economic activity (REA) proxies. We matched the forecasting horizon with the horizon of expected returns. For each REA under our consideration, we consider predictive regression model 1.17 as explained in Section 1.5 and employ the lagged REA, term-spread, credit-spread, forward variance and IRRA in the set of our control variables. Reported entries are the standardized ordinary-least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors as well as the in-sample adjusted R^2 for the given model. The sample covers July 1998 to May 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

Panel A: $OOS - R^2$: full model with PCA factor versus AR(1)								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
$h = 1$ months	0.26	0.22	0.08	0.02	0.20	0.26	0.35	0.13
$h = 2$ months	0.28	0.22	0.14	0.09	0.20	0.30	0.04	0.29
$h = 3$ months	0.17	0.24	0.14	0.10	0.09	0.27	-0.09	0.22
$h = 6$ months	0.11	0.27	0.08	0.27	0.06	0.13	-0.16	0.01
$h = 12$ months	0.19	0.24	0.31	0.41	0.09	0.23	0.10	0.14

Panel B: $OOS - R^2$: full model with PCA factor versus the nested model								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
$h = 1$ month	0.05	0.05	0.02	0.01	0.04	0.01	0.09	0.07
$h = 2$ months	0.12	0.10	0.06	0.01	0.09	0.02	0.05	0.16
$h = 3$ months	0.08	0.13	0.02	0.01	0.06	0.01	0.02	0.15
$h = 6$ months	0.04	0.10	-0.06	0.01	0.04	-0.07	-0.04	-0.02
$h = 12$ months	-0.01	0.06	-0.03	0.02	-0.01	-0.09	-0.01	-0.01

Table 1.9. Out-of-Sample Predictability of U.S. REA with PCA Factor: Panel A shows the out-of-sample R^2 obtained from the forecasts of the full model in Equation (1.18) (which contains PCA factor, lagged REA and the control variables as predictors) versus the forecasts of the AR(1) model in Equation (1.19) (which considers only lagged REA as predictors). Panel B reports the out-of-sample R^2 obtained from the forecasts of the full model versus the forecasts of the nested model ($benchmark_2$) in Equation (1.20). The nested model does not include the IPCA factor and incorporates only lagged REA and the control variables in the set of its regressors. Control variables in the full and nested model include credit spread, term spread, forward variance and IRRA.
For each U.S. REA proxy, we estimate the relevant form of Equation (1.17) for the respective full and benchmark models by considering an expanding window; the first estimation sample window covers July 1998 to September 2007. At each point in time, we estimates h = 1-,3-, 6-, 9-, and 12-months-ahead REA forecasts. The positive (negative) sign of the out-of-sample R^2 denotes that the full model that incorporates the PCA factor in the set of its predictors outperforms (underperforms) the benchmark model.

Panel A: $OOS - R^2$: full model with IPCA factor versus AR(1)								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
$h = 1$ month	0.20	0.20	0.01	0.03	0.13	0.26	0.27	0.10
$h = 2$ months	0.17	0.17	0.12	0.08	0.11	0.29	-0.02	0.21
$h = 3$ months	0.10	0.21	0.20	0.11	0.02	0.30	-0.07	0.14
$h = 6$ months	0.13	0.28	0.14	0.27	0.06	0.18	-0.09	0.07
$h = 12$ months	0.24	0.26	0.32	0.39	0.12	0.26	0.11	0.15

Panel B: $OOS - R^2$: full model with IPCA factor versus the nested model								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
$h = 1$ month	-0.04	0.02	-0.05	0.01	-0.03	0.01	-0.01	0.04
$h = 2$ months	0.01	0.05	0.04	0.01	-0.01	0.02	-0.01	0.07
$h = 3$ months	0.01	0.10	0.08	0.02	0.01	0.04	0.04	0.06
$h = 6$ months	0.06	0.12	0.01	0.01	0.05	-0.02	0.01	0.06
$h = 12$ months	0.04	0.08	-0.02	-0.01	0.02	-0.05	-0.01	0.01

Table 1.10. Out-of-Sample Predictability of U.S. REA with IPCA Factor: Panel A shows the out-of-sample R^2 obtained from the forecasts of the full model in Equation (1.18) (which contains IPCA factor, lagged REA and the control variables as predictors) versus the forecasts of the AR(1) model in Equation (1.19) (which considers only lagged REA as predictors). Panel B shows out-of-sample R^2 obtained from the forecasts of the full model versus the forecasts of the nested model ($benchmark_2$) in Equation (1.20). The nested model does not include the IPCA factor and considers only lagged REA and the control variables in the set of its regressors. Control variables in the full and nested model include credit spread, term spread, forward variance and IRRA.
For each U.S. REA proxy, we estimate the relevant form of Equation (1.17) for the respective full and benchmark models by considering an expanding window; the first estimation sample window covers July 1998 to September 2007. At each point in time, we estimates $h =$ 1-,3-, 6-, 9-, and 12-months-ahead REA forecasts. The positive (negative) sign of the out-of-sample R^2 denotes that the full model that incorporates the PCA factor in the set of its predictors outperforms (underperforms) the benchmark model.

$OOSR^2$ from a full model with variance based PCA factor versus benchmark model with RN-distribution based PCA factor								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
h=1 month	0.03	0.00	-0.01	0.01	0.03	0.00	0.01	0.10
h=2 month	0.04	0.08	0.04	0.02	0.02	0.02	0.01	0.14
h=3 month	0.04	0.12	0.10	-0.02	0.03	0.05	0.00	0.13
h=6 month	0.14	0.19	0.01	-0.06	0.11	0.13	0.02	0.14
h=12 month	0.16	0.20	0.08	0.00	0.09	0.17	0.05	0.07

Table 1.11. Variance-based versus Density-based Expected Returns: Table reports OOR^2 obtained from full model with PCA factor extracted from Martin and Wagner (2019) method versus an alternative PCA factor obtained from Chabi-Yo et al. (2022) which take into account the whole risk-neutral distribution. All predictive models are augmented by a set of control variables which consists of the lagged REA, credit spread, term spread, forward variance and IRRA.
For each U.S. REA proxy, we estimate the relevant form of Equation (1.17) for the respective full and benchmark models by considering an expanding window; the first estimation sample window covers July 1998 to September 2007. At each point in time, we estimates $h = 1$-,3-, 6-, 9-, and 12-months-ahead REA forecasts. The positive (negative) sign of the out-of-sample R^2 denotes that the full model that incorporates the PCA factor in the set of its predictors outperforms (underperforms) the benchmark model.

		$h = 1$ month	$h = 2$ months	$h = 3$ months	$h = 6$ months	$h = 12$ months
IP	β	-0.25***	-0.23**	-0.04	0.03	-0.02
		[0.00]	[0.01]	[0.62]	[0.60]	[0.64]
		(0.00)	(0.00)	(0.42)	(0.77)	(0.86)
	R2	0.24	0.41	0.48	0.41	0.33
NFP	β	-0.28***	-0.18***	-0.08	-0.10	-0.18
		[0.00]	[0.00]	[0.37]	[0.16]	[0.13]
		(0.00)	(0.00)	(0.13)	(0.09)	(0.02)
	R2	0.69	0.79	0.79	0.75	0.57
RS	β	-0.02	0.03	0.10	0.12	-0.02
		[0.64]	[0.88]	[0.71]	[0.73]	[0.74]
		(0.66)	(0.72)	(0.37)	(0.37)	(0.87)
	R2	0.09	0.17	0.23	0.27	0.37
HOUS	β	-0.13**	-0.12	-0.12	-0.24	-0.19
		[0.04]	[0.13]	[0.49]	[0.24]	[0.86]
		(0.11)	(0.17)	(0.18)	(0.00)	(0.01)
	R2	0.20	0.19	0.19	0.36	0.58
CU	β	-0.30***	-0.26***	-0.11	-0.06	-0.21*
		[0.00]	[0.00]	[0.35]	[0.28]	[0.06]
		(0.00)	(0.00)	(0.06)	(0.53)	(0.10)
	R2	0.25	0.42	0.49	0.39	0.31
UR	β	0.37***	0.38***	0.26**	0.16	0.22
		[0.00]	[0.00]	[0.02]	[0.35]	[0.29]
		(0.00)	(0.00)	(0.00)	(0.08)	(0.04)
	R2	0.19	0.36	0.44	0.59	0.54
CFNAI	β	-0.35***	-0.14**	0.00	-0.03	-0.20
		[0.00]	[0.05]	[0.65]	[0.61]	[0.51]
		(0.00)	(0.01)	(1.00)	(0.73)	(0.09)
	R2	0.51	0.54	0.45	0.27	0.18
ADS	β	-0.10***	-0.18***	-0.04	0.01	-0.18
		[0.00]	[0.00]	[0.58]	[0.78]	[0.45]
		(0.01)	(0.00)	(0.37)	(0.87)	(0.13)
	R2	0.86	0.76	0.69	0.41	0.22

Table 1.12. In-sample Prediction of REA Proxies with PCA Factor in the Presence of 5 Fama-French Factors: This table presents the in-sample estimated regression coefficients of the first PCA factor for various U.S. real economic activity (REA) proxies. We matched the forecasting horizon with the horizon of expected returns. For each REA under our consideration, we consider the predictive regression model (1.17) with the PCA factor plus a set of control variables which includes the lagged REA, credit and term spread, forward and IRRA and is augmented by 5 factors of Fama-French including the size of firms (SMB), book-to-market values (HML), expected return on the market (mrkt), profitability factor (RMW) and investment factor (CMA). Reported entries are the standardized ordinary-least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors as well as the in-sample adjusted R^2 for the given model. The sample covers July 1998 to May 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

		$h = 1$ month	$h = 2$ months	$h = 3$ months	$h = 6$ months	$h = 12$ months
IP	β	-0.27***	-0.26**	-0.12	0.04	0.03
		[0.00]	[0.02]	[0.34]	[0.92]	[0.52]
		(0.00)	(0.01)	(0.03)	(0.55)	(0.78)
	R2	0.19	0.32	0.42	0.41	0.37
NFP	β	-0.25***	-0.15	-0.14	-0.08	-0.15
		[0.00]	[0.12]	[0.20]	[0.50]	[0.50]
		(0.00)	(0.01)	(0.03)	(0.20)	(0.05)
	R2	0.68	0.78	0.79	0.76	0.58
RS	β	-0.01	0.07	0.04	0.18	0.02
		[0.71]	[0.35]	[0.60]	[0.51]	[0.56]
		(0.83)	(0.26)	(0.70)	(0.17)	(0.88)
	R2	0.08	0.17	0.26	0.31	0.38
HOUS	β	-0.17**	-0.09	-0.04	-0.19	-0.14
		[0.01]	[0.17]	[0.63]	[0.34]	[0.88]
		(0.05)	(0.31)	(0.60)	(0.02)	(0.06)
	R2	0.21	0.20	0.20	0.35	0.56
CU	β	-0.30***	-0.28**	-0.13	-0.05	-0.17
		[0.00]	[0.01]	[0.27]	[0.63]	[0.34]
		(0.00)	(0.00)	(0.02)	(0.57)	(0.18)
	R2	0.20	0.35	0.45	0.39	0.33
UR	β	0.34***	0.35**	0.33**	0.15	0.20
		[0.00]	[0.01]	[0.03]	[0.52]	[0.48]
		(0.00)	(0.00)	(0.00)	(0.10)	(0.06)
	R2	0.18	0.35	0.44	0.59	0.55
CFNAI	β	-0.33***	-0.12	-0.03	0.02	-0.16
		[0.00]	[0.17]	[0.53]	[0.84]	[0.63]
		(0.00)	(0.02)	(0.67)	(0.80)	(0.13)
	R2	0.48	0.51	0.46	0.30	0.19
ADS	β	-0.07**	-0.14**	-0.11	0.05	-0.15
		[0.02]	[0.04]	[0.11]	[0.82]	[0.95]
		(0.04)	(0.01)	(0.03)	(0.55)	(0.17)
	R2	0.86	0.73	0.66	0.44	0.23

Table 1.13. In-Sample Prediction of REA Proxies with Instrumented PCA Factor in the Presence of 5 Fama-French Factors: This table presents the in-sample estimated regression coefficients of the first IPCA factor for various U.S. real economic activity (REA) proxies. We matched the forecasting horizon with the horizon of expected returns. For each REA under our consideration, we consider the predictive regression model (1.17) with the IPCA factor plus a set of control variables which includes the lagged REA, credit and term spread, forward and IRRA and is augmented by 5 factors of Fama-French including the size of firms (SMB), book-to-market values (HML), expected return on the market (mrkt), profitability factor (RMW) and investment factor (CMA). Reported entries are the standardized ordinary-least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors as well as the in-sample adjusted R^2 for the given model. The sample covers July 1998 to May 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

		h=1 month	h=2 month	h=3 month	h=6 month	h=12 month
IP	beta	-0.14**	-0.18**	-0.15	-0.06	-0.13***
		[0.02]	[0.03]	[0.10]	[0.21]	[0.00]
		(0.05)	(0.02)	(0.08)	(0.65)	(0.61)
	R2	0.21	0.34	0.41	0.32	0.28
NFP	beta	-0.01	0.00	0.00	-0.01	0.04
		[0.86]	[0.90]	[0.91]	[0.36]	[0.12]
		(0.81)	(0.99)	(0.95)	(0.90)	(0.81)
	R2	0.67	0.76	0.76	0.69	0.54
RS	beta	-0.01	0.05	0.04	0.08	0.10
		[0.84]	[0.69]	[0.75]	[0.75]	[0.85]
		(0.85)	(0.58)	(0.66)	(0.61)	(0.56)
	R2	0.08	0.10	0.14	0.26	0.35
HOUS	beta	0.00	0.06	0.12	0.15	0.27
		[0.97]	[0.76]	[0.56]	[0.94]	[0.92]
		(0.98)	(0.46)	(0.18)	(0.07)	(0.00)
	R2	0.20	0.19	0.18	0.37	0.58
CU	beta	-0.10*	-0.14	-0.12	0.01	0.06**
		[0.09]	[0.08]	[0.16]	[0.46]	[0.03]
		(0.12)	(0.06)	(0.16)	(0.95)	(0.77)
	R2	0.22	0.36	0.43	0.32	0.28
UR	beta	0.00	-0.04	-0.08	-0.07	-0.20
		[1.00]	[0.79]	[0.65]	[0.99]	[0.93]
		(0.98)	(0.62)	(0.37)	(0.52)	(0.18)
	R2	0.17	0.31	0.39	0.55	0.53
CFNAI	beta	-0.08*	-0.03	0.01	0.07	0.10
		[0.06]	[0.36]	[0.71]	[0.84]	[0.83]
		(0.10)	(0.45)	(0.92)	(0.53)	(0.39)
	R2	0.48	0.48	0.40	0.27	0.14
ADS	beta	-0.04*	-0.08*	-0.06	0.07	0.15
		[0.10]	[0.06]	[0.16]	[0.70]	[0.58]
		(0.06)	(0.11)	(0.37)	(0.62)	(0.39)
	R2	0.86	0.72	0.61	0.36	0.21

Table 1.14. In-sample Prediction of REA Proxies with the Residual Components of the Weighted Factor on the SVIX measure: This table presents the in-sample estimated regression coefficients of the residual component (obtained from regression of the weighted factor on the SVIX) for different U.S. real economic activity (REA) proxies. We matched the forecasting horizon with the horizon of expected returns. For each REA under our consideration, we consider predictive regression model (1.17) we employ the lagged REA, term-spread, credit-spread, forward variance, SVIX, and IRRA in the set of our control variables. Reported entries are the standardized ordinary-least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors as well as the in-sample adjusted R2 for the given model. The sample covers July 1998 to May 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

		h=1 month	h=2 month	h=3 month	h=6 month	h=12 month
IP	beta	-0.43***	-0.47***	-0.30*	-0.26**	-0.32***
		[0.00]	[0.00]	[0.06]	[0.02]	[0.00]
		(0.00)	(0.00)	(0.00)	(0.11)	(0.30)
	R2	0.25	0.39	0.43	0.35	0.32
NFP	beta	-0.18**	-0.17**	-0.14*	-0.13**	-0.07***
		[0.01]	[0.04]	[0.08]	[0.01]	[0.00]
		(0.01)	(0.01)	(0.10)	(0.30)	(0.75)
	R2	0.67	0.76	0.76	0.70	0.53
RS	beta	-0.08	-0.18	-0.20	0.00	0.01
		[0.58]	[0.29]	[0.32]	[0.96]	[0.52]
		(0.33)	(0.09)	(0.14)	(0.99)	(0.97)
	R2	0.08	0.10	0.14	0.25	0.34
HOUS	beta	0.01	0.11	0.18	0.23	0.30
		[0.77]	[0.47]	[0.42]	[0.71]	[0.94]
		(0.89)	(0.27)	(0.09)	(0.07)	(0.02)
	R2	0.20	0.19	0.18	0.38	0.58
CU	beta	-0.31***	-0.32**	-0.18	-0.12*	-0.06***
		[0.00]	[0.01]	[0.26]	[0.09]	[0.00]
		(0.00)	(0.00)	(0.04)	(0.43)	(0.84)
	R2	0.24	0.39	0.45	0.34	0.30
UR	beta	0.14	0.17	0.12	0.03	-0.10
		[0.17]	[0.19]	[0.33]	[0.45]	[0.23]
		(0.10)	(0.08)	(0.26)	(0.78)	(0.53)
	R3	0.17	0.31	0.39	0.55	0.51
CFNAI	beta	-0.43***	-0.28**	-0.13	0.00	0.02
		[0.00]	[0.01]	[0.27]	[0.70]	[0.73]
		(0.00)	(0.00)	(0.15)	(0.98)	(0.91)
	R3	0.51	0.50	0.41	0.28	0.15
ADS	beta	-0.15**	-0.27***	-0.23**	-0.05	0.06*
		[0.01]	[0.00]	[0.01]	[0.14]	[0.07]
		(0.00)	(0.00)	(0.02)	(0.75)	(0.84)
	R3	0.86	0.74	0.63	0.37	0.22

Table 1.15. In-sample Prediction of REA Proxies with the Residual Components of the PCA Factor on the SVIX measure: This table presents the in-sample estimated regression coefficients of the residual component (obtained from regression of the PCA factor on the SVIX) for different U.S. real economic activity (REA) proxies. We matched the forecasting horizon with the horizon of expected returns. For each REA under our consideration, we consider predictive regression model (1.17) we employ the lagged REA, term-spread, credit-spread, forward variance, SVIX, and IRRA in the set of our control variables. Reported entries are the standardized ordinary-least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors as well as the in-sample adjusted R2 for the given model. The sample covers July 1998 to May 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

Panel A: $OOSR^2$ from model with PCA factor versus a model with 5 FF factors								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
$h = 1$ month	0.07	0.11	0.08	0.01	0.07	0.02	0.13	0.18
$h = 2$ months	0.10	0.14	0.08	0.03	0.08	-0.01	0.05	0.22
$h = 3$ months	0.06	0.17	0.05	0.01	0.05	0.01	0.00	0.13
$h = 6$ months	0.07	0.12	0.00	0.03	0.07	-0.06	0.04	0.04
$h = 12$ months	0.06	0.09	0.07	0.08	0.04	-0.06	0.03	0.11

Panel B: $OOSR^2$ from model with IPCA factor versus a model with 5 FF factors								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
$h = 1$ month	-0.01	0.09	0.01	0.01	-0.01	0.01	0.03	0.15
$h = 2$ months	-0.02	0.09	0.06	0.02	-0.02	-0.02	0.02	0.13
$h = 3$ months	-0.02	0.14	0.11	0.03	-0.02	0.05	0.02	0.04
$h = 6$ months	0.09	0.13	0.06	0.03	0.07	-0.01	0.08	0.12
$h = 12$ months	0.10	0.10	0.08	0.05	0.07	-0.02	0.03	0.12

Panel C: $OOSR^2$ from model with PCA factor versus a model with market svix								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
$h = 1$ month	0.04	0.07	-0.01	0.01	0.01	0.01	0.10	0.08
$h = 2$ months	0.12	0.10	0.04	-0.06	0.08	-0.01	0.07	0.16
$h = 3$ months	0.08	0.12	-0.03	-0.04	0.06	-0.09	0.02	0.13
$h = 6$ months	0.02	0.05	-0.16	-0.09	0.02	-0.20	-0.07	-0.07
$h = 12$ months	-0.09	-0.05	-0.11	-0.08	-0.12	-0.30	0.01	-0.02

Panel D: $OOSR^2$ from model with IPCA factor versus a model with market svix								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
$h = 1$ month	-0.05	0.04	-0.09	0.01	-0.06	0.01	0.01	0.05
$h = 2$ months	0.00	0.05	0.01	-0.06	-0.02	-0.02	0.01	0.05
$h = 3$ months	0.00	0.09	0.03	-0.02	-0.01	-0.04	0.04	0.03
$h = 6$ months	0.05	0.06	-0.09	-0.09	0.02	-0.15	-0.01	0.01
$h = 12$ months	-0.03	-0.03	-0.09	-0.11	-0.09	-0.25	0.02	0.00

Table 1.16. Individual Equity Options versus Standard Risk Factors: Panels A and B show the OOS-R^2 obtained from the full predictive model explained in Equation (1.18) which contain the PCA factor or IPCA, respectively, versus the benchmark model which contains the 5 Fama-French factors Equation (1.25). Panel C and D report OOR^2 obtained from the full model with PCA or IPCA factor, respectively, versus an alternative predictive model which contains the SVIX measure Equation (1.26). All predictive models are augmented by a set of control variables which consists of the lagged REA, credit spread, term spread, forward variance and IRRA.
For each U.S. REA proxy, we estimate the relevant form of Equation (1.17) for the respective full and benchmark models by considering an expanding window; the first estimation sample window covers July 1998 to September 2007. At each point in time, we estimates $h =$ 1-,3-, 6-, 9-, and 12-months-ahead REA forecasts. The positive (negative) sign of the out-of-sample R^2 denotes that the full model that incorporates the PCA factor in the set of its predictors outperforms (underperforms) the benchmark model.

Appendix - Chapter One

1.A Computation of the Risk-Neutral Moments

We compute the S&P500 risk-neutral moments following the methodology of Bakshi et al. (2003) using the volatility surface data of S&P500 index options. At time t the τ period risk-neutral skewness ($\theta_{q,t}(\tau)$) and kurtosis ($kq, t(\tau)$) of the log-return $R(t, \tau)$ distribution with horizon τ are given by

$$\theta_{q,t}(\tau) = \frac{e^{r\tau} M(3)_{t,T} - 3e^{r\tau}\mu_{t,T} M(2)_{t,T} + 2\mu_{t,T}^3}{[e^{r\tau} M(2)_{t,T} - \mu_{t,T}^2]^{3/2}} \tag{1.A.1}$$

$$k_{q,t}(\tau) = \frac{e^{r\tau} M(4)_{t,T} - 4e^{r\tau}\mu_{t,T} M(3)_{t,T} + 6e^{r\tau}\mu_{t,T}^2 M(2) - 3\mu_{t,T}^4}{[e^{r\tau} M(2)_{t,T} - \mu_{t,T}^2]^2} \tag{1.A.2}$$

where $M(n)_{t,T}(n = 2, 3, 4)$ is given by:

$$M(n)_{t,T} = \int_{S_t}^{\infty} \eta(K, S_t, n) C_t(K, T) dK + \int_0^{S_t} \eta(K, S_t, n) P_t(K, T) dK \tag{1.A.3}$$

$$\eta(K, S_t, n) = \frac{n}{K^2}[(n-1)log(\frac{K}{S_t})^{n-2} - log(\frac{K}{S_t})^{n-1}] \tag{1.A.4}$$

and

$$\mu_{t,T} = e^{r\tau} - 1 - e^{r\tau}[M(2)_{t,T}/2 + M(3)_{t,T}/6 + M(6)_{t,T}/24] \tag{1.A.5}$$

r is the riskless rate. S_t denotes the price of the asset at time t and $P_t(K, T)$ and $C_t(K, T)$ are the OTM put and call prices expiring at T.

1.B SVIX Empirical Computation

A spectrum of option prices with respect to the strike price is needed to compute the integrals in Equation (1.5). However, the required continuum of option prices is not available in empirical applications, and we have to make some choices regarding the implementation. Firstly, we follow Conrad et al. (2013) and estimate SVIX for days that a stock has more than two OTM calls and two OTM puts with the same maturity.

Moreover, following Dennis and Mayhew (2002), we employ equal numbers of OTM puts and calls for each asset on each day. Therefore, If there are only n OTMP available on day t, we need only n OTMC prices and if $N > n$ OTMC prices are available on the day t, we only use the n OTMC that are the least out-of-the-money. We also discard a few day-stock IV surface data if their OptionsMetrics's strike prices are not monotonic in deltas.

Next, we follow Jiang and Tian (2005) to impose a structure for implied volatilities. To do so, we define moneyness level as the implied strike prices provided by OM, divided by the stock price (K/S_t).

Since the SVIX formula requires only OTM-forward option prices, we follow Christoffersen et al. (2012) and keep only calls (puts) with moneyness that are greater (less) than or equal to the forward-moneyness (F/S_t). Then, we sort the puts and calls in increasing order by moneyness:

$$\frac{K_1^P}{S_t} < \frac{K_2^P}{S_t} < \frac{K_3^P}{S_t} < ... < \frac{K_{np}^P}{S_t} < \frac{F}{S_t} < \frac{K_1^C}{S_t} < \frac{K_2^C}{S_t} < \frac{K_3^C}{S_t} < ... < \frac{K_{nc}^C}{S_t},$$

where nc (np) denote the number of call (put) options.

Then, we interpolate implied volatilities using cubic piece-wise Hermite polynomial interpolation in the moneyness-IV metric to obtain a continuum of implied volatilities on a fine grid of 1,001 implied volatilities for moneyness levels between 1/3 and 3. Cubic piece-wise Hermite polynomial interpolation is adequate for interpolating between the available moneyness levels. Therefore, for values outside the available ranges, we use the implied volatility of the lowest strike price (K_1^P) for moneyness below the lowest available moneyness level. We use the implied volatility of the highest strike (K_{nc}^C) for moneyness above the highest available moneyness.

Once the continuous Implied Volatility surface is formed, we compute the BS options prices at 1,001 strike points (equally-spaced) over the moneyness from 1/3 to 3. Moneyness levels smaller than ($F_{t,T}/St$) are used to generate OTM put prices and moneyness levels greater than ($F_{t,T}/St$) are used to generate call prices. In addition, to convert IVs to option prices, we derive the projected dividend payments $D_{t,T}$ from forward price as $D_{t,T} = R_{f;t;T}S_t - F_{t;T}$. Next, we deduct the present value of $D_{t,T}$ from the stock price at time t and replace it in the modified Black-Scholes formula.

Finally, we follow Christoffersen et al. (2012) we employ trapezoidal approximation rule to numerically approximate the integrals of Equation (1.7). More specifically, on each day and for each maturity we assume that:

$$\int_0^{\infty} Q_{t,T}(K)dK = \Sigma_{i=1}^{N} \frac{Q_{t,T}(K_{i+1}) + Q_{t,T}(K_i)}{2}(K_{i+1} - Ki) \tag{1.A.6}$$

where $K_1, ..., K_N$ are the strikes of observable options, $Q_{t,T}(K_i)$ is the mid bid-ask price of an OTM option with strike K_i.[26]

1.C Characteristic Data

In this section, we define characteristic variables used in the IPCA method. We use CRSP and Compustat variable names in parentheses and report the relevant references. All these definitions[27] are directly obtained from Freyberger et al. (2020). We mainly use balance-sheet data from the fiscal year ending in year $t-1$ for returns from July of year t to May of year $t+1$ following the Fama and French (2015) timing convention.

1. "**A2ME**: Following Bhandari (1988), we define assets-to-market cap as total assets (AT) over market capitalisation as of December t-1. Market capitalisation is shares outstanding (SHROUT) times price (PRC). AT Total assets (AT) as in Gandhi and Lustig (2015).

2. **ATO**: Following Soliman (2008), ATO is Net sales over lagged net operating assets. Net operating assets are operating assets minus operating liabilities. Operating assets are total assets (AT) minus the investment, minus cash and short-term investments (CHE), and other advances (IVAO). Operating liabilities are total assets (AT) minus long-term debt (DLTT), minus common equity (CEQ), minus debt in current liabilities (DLC), minus preferred stock (PSTK), minus minority interest (MIB).

26. In Equation (1.A.6), we use put prices when their strike prices are less than the forward price and call prices when their strikes are greater than or equal to the forward price.

27. The definition of all 33 characteristic variables are directly brought from Freyberger et al. (2020). However, we compute all measures ourselves.

3. **BEME**: BEME is the ratio of the book value of equity to the market value of equity. Book equity is deferred taxes and investment tax credit (TXDITC) plus shareholder equity (SH) minus preferred stock (PS). SH is shareholders' equity (SEQ). If missing, SH is the sum of common equity (CEQ) and preferred stock (PS). If missing, SH is the difference between total assets (AT) and total liabilities (LT). Depending on availability, we use the redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for PS. The market value of equity is as of December t-1. The market value of equity is shares outstanding (SHROUT) times price (PRC). See Rosenberg et al. (1985) and James L. Davis and French (2000).

4. **Beta**: We define the CAPM beta as correlations between the expected return of stock i and the market expected return times volatilities following Frazzinia and Pedersen (2014). We calculate correlations from overlapping three-day log expected returns over a five-year period requiring at least 750 non-missing observations. We estimate volatilities using the standard deviations of daily log expected returns over a one-year horizon requiring at least 120 observations.

5. **C**: Following Palazzo (2012), cash to short-term investment is defined as cash and short-term investments (CHE) divided by total assets (AT).

6. **CTO**: Capital turnover is defined as the ratio of net sales (SALE) to lagged total assets (AT) following Haugen and Baker (1996).

7. **D2A**: Capital intensity is the ratio of depreciation and amortization (DP) to total assets (AT) as in Gorodnichenko and Weber (2016).

8. **DPI2A**: Following Evgeny Lyandres and Zhang (2008), the change in property, equipment, and plants is defined as the changes in property, plants, and equipment (PPEGT) and inventory (INVT) over lagged total assets (TA).

9. **E2P**: Following Basu (1983), we define earnings to price as the ratio of income before extraordinary items (IB) to the market capitalisation as of December t-1. Market capitalisation is the shares outstanding (SHROUT) times price (PRC).

10. **FC2Y**: Following D'Acuntoa et al. (2018), Fixed costs to sales are defined as the

ratio of selling, general, and administrative expenses (XSGS), advertising expenses (XAD), and research and development expenses (XRD) to net sales (SALE).

11. **Free CF**: Cash flow to book value of equity is the ratio of depreciation and amortisation (DP), net income (NI) less change in capital expenditure (CAPX) and working capital (WCAPCH), divided by the book value of equity defined as in the construction of BEME following Houa et al. (2012).

12. **Idio vol**: Following Ang et al. (2006), Idiosyncratic volatility is the standard deviation of the residuals from a regression of expected returns on the Fama and French (1993) three-factor model. We use one month of daily data and require at least fifteen non-missing observations.

13. **investment**: Investment is the percentage year-on-year growth rate in total assets (AT) following Cooper et al. (2008).

14. **Lev**: Following Lewellen (2015), we define Leverage as the ratio of long-term debt (DLTT) and debt in current liabilities (DLC) to the sum of long-term debt, debt in current liabilities, and stockholders' equity (SEQ).

15. **LME**: Following Fama and French (1992), Size is shares outstanding (SHROUT) times the total market capitalisation of the previous month defined as price (PRC).

16. **LTurnover**: Following Datar et al. (1998), lagged Turnover is last month's volume (VOL) divided by shares outstanding (SHROUT).

17. **NOA**: Net operating assets are the difference between operating assets minus operating liabilities scaled by lagged total assets as in Hirshleifer et al. (2004). Operating assets are total assets (AT) minus the investment and other advances (IVAO) minus cash and short-term investments (CHE). Operating liabilities are total assets (AT), minus long-term debt (DLTT), minus debt in current liabilities (DLC), minus preferred stock (PSTK), minus minority interest (MIB), minus common equity (CEQ).

18. **OA**: Following Sloan (1996), we define operating accruals as changes in non-cash working capital minus depreciation (DP) scaled by lagged total assets (TA). Non-

cash working capital is the difference between non-cash current assets and current liabilities (LCT), income taxes payable (TXP) and debt in current liabilities (DLC). Non-cash current assets are current assets (ACT) minus cash and short-term investments (CHE).

19. **OL**: Operating Leverage is defined as the sum of administrative expenses (XSGA) and cost of goods sold (COGS) over total assets as in Novy-Marx (2011).

20. **PCM**: The price-to-cost margin (PCM) is the difference between net sales (SALE) and costs of goods sold (COGS) divided by net sales (SALE) as in Bustamant and Donangelo (2017).

21. **PM**: Profit margin is defined as the ratio of operating income after depreciation (OIADP) to net sales (SALE) as in Soliman (2008).

22. **Prof**: Following Balla et al. (2015) Profitability is defined as gross profitability (GP) divided by the book value of equity.

23. **Q**: Tobin's Q is total assets (AT), the market value of equity (SHROUT times PRC), minus deferred taxes (TXDB) scaled by total assets (AT) minus cash and short-term investments (CEQ).

24. **Rel to High**: Closeness to 52-week high is the stock price (PRC) at the end of the previous calendar month divided by the previous 52-week high price following George and Hwang (2004).

25. **RNA**: The return on net operating assets is operating income after depreciation divided by lagged net operating assets following Soliman (2008). Net operating assets are the difference between operating assets minus operating liabilities. Operating assets are total assets (AT) minus the investment and other advances (IVAO) minus cash and short-term investments (CHE). Operating liabilities are total assets (AT) minus common equity (CEQ), minus minority interest (MIB), minus debt in current liabilities (DLC), minus preferred stock (PSTK), minus long-term debt (DLTT).

26. **ROA**: Return-on-assets is income before extraordinary items (IB) to lagged total

assets (AT) following Balakrishnan et al. (2010).

27. **ROE**: Return-on-equity is income before extraordinary items (IB) to lagged book-value of equity as in Haugen and Baker (1996). r_{12-2} : We define momentum as cumulative return from 12 months before the return prediction to two months before, as in Fama and French (1992).

28. r_{12-7}: The intermediate momentum is a cumulative return from 12 months before the return prediction to seven months before, as in Novy-Marx (2012).

29. r_{2-1} : Short-term return is the lagged one-month return as in JEGADEESH (1990).

30. r_{36-13} : Long-term return is the cumulative return from 36 months before the return prediction to 13 months before, as in Bondt and Thaler (1985).

31. **S2P**: Sales-to-price is the ratio of net sales (SALE) to the market capitalisation as of December following Lewellen (2015).

32. **SGA2S**: SG&A to sales is the ratio of selling, general and administrative expenses (XSGA) to net sales (SALE).

33. **Spread**: The bid-ask spread is the average daily bid-ask spread in the previous months as in JEGADEESH (1990)."

1.D IVX-Wald test

1.D.1 IVX Estimator

Assume that the following multivariate system of predictive regressions consists of regressors with arbitrary degree of persistence:

$$y_t(K) = \mu + Ax_{t-1} + \epsilon_t \tag{1.A.7}$$

where $y_t(K) = \Sigma_{i=0}^{K-1} y_{t+i}$, A is a $(m \times r)$ coefficient matrix and

$$x_{t+1} = R_n x_t + u_{t+1} \tag{1.A.8}$$

where $x_t = (x_{1t}, x_{2t}, ..., x_{rt})$ is the vector of predictors in (1.A.7), $R_n = I_r + C/n^{\alpha}$ is the autoregressive matrix with degree of persistence equal to $\alpha \geq 0$, $C = diag(c1, ..., cr)$ and n is the sample size.

The IVX method proposed by Kostakis et al. (2011) estimates Equation (1.A.7) using two-step least square equations which uses the near-stationary instruments in z_t rather than the initial predictors x_t. The intuition behind it is to build an instrumental variable with a known degree of persistence using the initial predictors of x_t with an unknown degree of persistence. To construct the IVX estimator, first, we estimate Equations (1.A.7) and (1.A.8) using ordinary least squares. Next, we build the near-stationary instruments Z_t by differencing the regressor x_t, initializing at $z_0 = 0$. More specifically, IVX construct instruments using a first-order autoregressive process with the autoregressive artificial matrix of R_{nz} and innovations Δx_t as follows:

$$z_t = R_{nz} z_{t-1} + \Delta x_t \tag{1.A.9}$$

where $R_{nz} = I_r + \frac{C_z}{n^{\beta}}$. We select $C_z = -I_r$ and $\beta = 0.95$ following Kostakis et al. (2015)

.

Once we construct our instrument variable, we apply standard instrumental variable estimation and continue with IVX estimation of A using regression system in (1.A.7)

$$\bar{A}_{IVX(K)} = \bar{Y}(K)'Z[\bar{X}(K)'Z'] = \sum_{r=1}^{n}(y_t - \bar{y}_n)\bar{z}'_{t-1}[\sum_{j=1}^{n}(x_j - \bar{x}_{n-1})\bar{z}'_{j-1}]^{-1} \tag{1.A.10}$$

where $\bar{y}_n = (1/n)\sum_{t=1}^{n} y_t$, $\bar{x}_{n-1} = (1/n)\sum_{t=1}^{n} x_{t-1}$, $\bar{Y} = (Y'_1, ..., Y'_n)$ and $\bar{X} = (X_0, ..., X'_{n-1})$ are the predictive regression matrices which are demeaned and Z_{n_k} is the instrument matrix.

1.D.2 IVX-Wald Test

The asymptotic feature of the IVX method suggests that linear restrictions on the coefficients A rendered by the system of predictive regression (1.A.7) can be examined by a standard Wald test using the IVX estimator. We test for the predictive ability of

x_t with the following null hypothesis:

$$H_0 : Hvec(A) = 0 \tag{1.A.11}$$

where $vec(A)$ is the vectorisation of A and H is a known $r \times r$ matrix, which its (i, i) element is one and its other elements are zero. This helps to test for the significance of each regressor individually.

We consider the following IVX-Wald test statistic for testing the H_0

$$W_{IVX} = (Hvec\bar{A}_{IVX)})' Q^{-1}(Hvec\bar{A}_{IVX)}) \tag{1.A.12}$$

where $vec\bar{A}_{IVX}$ is the vectorization of IVX estimator in (1.A.10) and:

$$Q_H = H[(\bar{Z}'X)^{-1} \otimes I_m]\mathbb{M}[(X'\bar{Z})^{-1} \otimes I_m]H' \tag{1.A.13}$$

$$\mathbb{M} = \bar{Z}'\bar{Z} \otimes \hat{\Sigma}_{\epsilon\epsilon} - n\bar{z}_{n-1}\bar{z}'_{n-1} \otimes \hat{\Omega}_{FM} \tag{1.A.14}$$

$$\hat{\Omega}_{FM} = \hat{\Sigma}_{\epsilon\epsilon} - \hat{\Omega}_{\epsilon u}\hat{\Omega}^{-1}_{uu}\hat{\Omega}'_{\epsilon u} \tag{1.A.15}$$

Denoting $\hat{\epsilon}_t$ the OLS residuals from regression (1.A.7) and $\hat{u}_t$ residuals from regression (1.A.8), the covariance matrices $\hat{\Sigma}_{\epsilon\epsilon}$, $\hat{\Omega}_{\epsilon u}$ and $\hat{\Omega}_{uu}$ can be estimated as follows:

$$\hat{\Sigma}_{\epsilon\epsilon} = \frac{1}{n}\sum_{t=1}^{n}\hat{\epsilon}_t\hat{\epsilon}'_t,\ \hat{\Sigma}_{\epsilon u} = \frac{1}{n}\sum_{t=1}^{n}\hat{\epsilon}_t\hat{u}'_t,\ \hat{\Sigma}_{uu} = \frac{1}{n}\sum_{t=1}^{n}\hat{u}_t\hat{u}'_t, \tag{1.A.16}$$

and

$$\hat{\Lambda}_{uu} = \frac{1}{n}\sum_{i=1}^{M_n}(1 - \frac{i}{M_n+1})\sum_{t=i+1}^{n}\hat{u}_t\hat{u}'_{t-i},\ \hat{\Omega}_{uu} = \hat{\Sigma}_{uu} + \hat{\Lambda}_{uu} + \hat{\Lambda}'_{uu} \tag{1.A.17}$$

$$\hat{\Lambda}_{u\epsilon} = \frac{1}{n}\sum_{i=1}^{M_n}(1 - \frac{i}{M_n+1})\sum_{t=i+1}^{n}\hat{u}_t\hat{\epsilon}'_{t-i},\ \hat{\Omega}_{u\epsilon} = \hat{\Sigma}_{uu} + \hat{\Lambda}'_{u\epsilon} \tag{1.A.18}$$

where M_n is a bandwidth parameter satisfying $Mn \to 1$ and $M_n/\sqrt{n} \to 0$ as $n \to \infty$. We set $M_n = n^{1/3}$ following Kostakis et al. (2015).

1.E Improving REA Predictability Using SVIX

In this section, we investigate the predictive ability of SVIX, as inferred from the index option, to predict future growth in real economic activity.

In order to examine whether SVIX forecasts REA growth over h forecasting horizons, we set up our empirical analysis based on the following predictive regression:

$$REA_{i,t+h} = \beta_{i,0} + \beta_{i,1} REA_{i,t} + \beta_{i,2} SVIX_t + \beta_{i,2} X_t + \epsilon_{i,t+h} \tag{1.A.19}$$

where $REA_{i,t+h}$ is the growth rate of a different measure of economic activity as explained in Section 1.3.2. $SVIX_{t,h}$ is the h-month SVIX measure defined in Equation (1.5) , and X_t is a vector of additional predictors at time t as explained in Section 1.3.3. More specifically, we consider all of our control variables: term spread, credit spread, forward variance and IRRA.

Our results are reported in Table 1.A.1 and show two main findings. First, we can see that SVIX predict 8 out of 9 REA proxies at 1-month horizon (all proxies except RS). At two-months horizon, SVIX predicts 6 out of 9 proxies (All proxies except RS and CFNAI). At three-months horizon, SVIX predicts IP, NFP, HOUS, UR and ADS. In addition, SVIX predicts NFP and IP at all forecasting horizon (1 to 12 month). The sign of the SVIX coefficient is significantly negative for NFP, IP, CNFAI, ADS and CU. We found a positive SVIX coefficient for the unemployment rate, which explains its countercyclical behaviour. This suggests that a rise in SVIX forecasts a decrease in REA.

[Table 1.A.1 about here.]

1.F Regression Coefficient of Control Variables

This section analyses the regression coefficient of the control variables included in our predictive regression model. Tables 1.A.2 to 1.A.3 report the regression coefficient of control variables, namely term spread, credit spread and option-based variables (IRRA and Forward Variance) for different forecasting horizons when the variable of interest is the PCA factor. These results show that FV and IRRA have significant predictive power to predict REA among different control variables. Specifically, at 1 and 2 months horizon, the regression coefficient of forward variance is significant for IP, NFP, CU, CFNAI and ADS and is positive for UR. The coefficient of IRRA is also significant for IP, HOUS, RS and UR at one and two-months horizons. In addition, the regression coefficient of FV at 6 and 12 months horizon is significantly negative for IP, NFP, CU and ADS, while it is positive for the unemployment rate. Furthermore, at six months horizon, the coefficient of IRRA is significant for estimating HS.

[Table 1.A.2 about here.]

[Table 1.A.3 about here.]

1.G Diebold-Mariano Test for Accuracy of OOS Results

In this section, we test the statistical significance of the out-of-sample R^2s, using Diebold-Mariano test. We report the results of this test in Table 1.A.4. The Diebold-Mariano test compares the forecast accuracy of two forecast methods. It defines the loss differential between the two forecasts and examine whether the two forecasts are significantly different. As shown in the Panel B of Table 1.A.4, the $OOSR^2$ is significantly positive for forecasting IP, NFP CU and ADS specially at short-term horizon, suggesting that the full model that incorporates the PCA factor outperforms the nested model which does not incorporate the factor in the set of its predictors.

In addition, we plot a time series comparison of the predicted OOS values of REA for the two models (1. Full model with the PCA and 2. Nested model without the PCAfactor) in Figures 1.A.1 and 1.A.2 over September 2007 - May 2019.

[Table 1.A.4 about here.]

[Figure 1.A.1 about here.]

[Figure 1.A.2 about here.]

1.H Compare the Predictive Power of PCA and SVIX for Future SPX Returns

In this section, we examine the predictive power of three variables for forecasting future stock market returns: PCA factor, weighted factor, and SVIX. The PCA factor is derived from individual equity options and is expected to proxy for the market's expected return. We also control for three factors from the Fama-French model: market, size (SML), and value (HML).

We conduct regression analysis using the following model to estimate the coefficients for each variable in predicting the market's expected return:

The results, shown in Table 1.A.5, indicate that neither the PCA factor, the weighted factor, nor SVIX shows a significant coefficient in forecasting market returns when controlling for the three Fama-French factors.

$$Ret_{spx,t+1} = \alpha + \beta_1 Factor_t + \beta_2 mrkt_t + \beta_3 SML_t + \beta_4 HML_t + \epsilon_t$$

[Table 1.A.5 about here.]

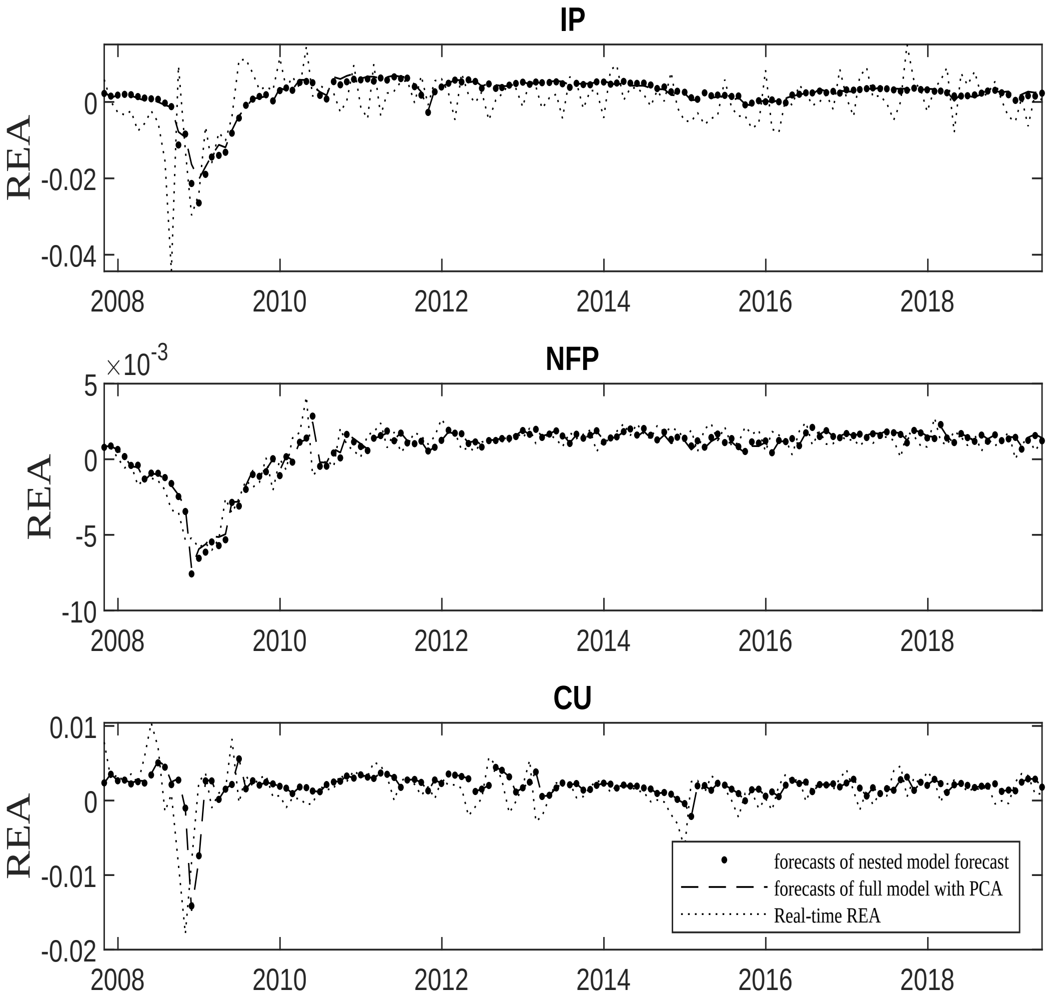


Figure 1.A.1. Evolution of OOS Forecasts of IP, NFP and CU: The figure shows the evolution of REA forecasts (IP, NFP and CU) over September 2007 - May 2019. For each REA measure, we extract the time series by performing a rolling window estimation. We use a rolling window starting on July 1998 with size 30 months.

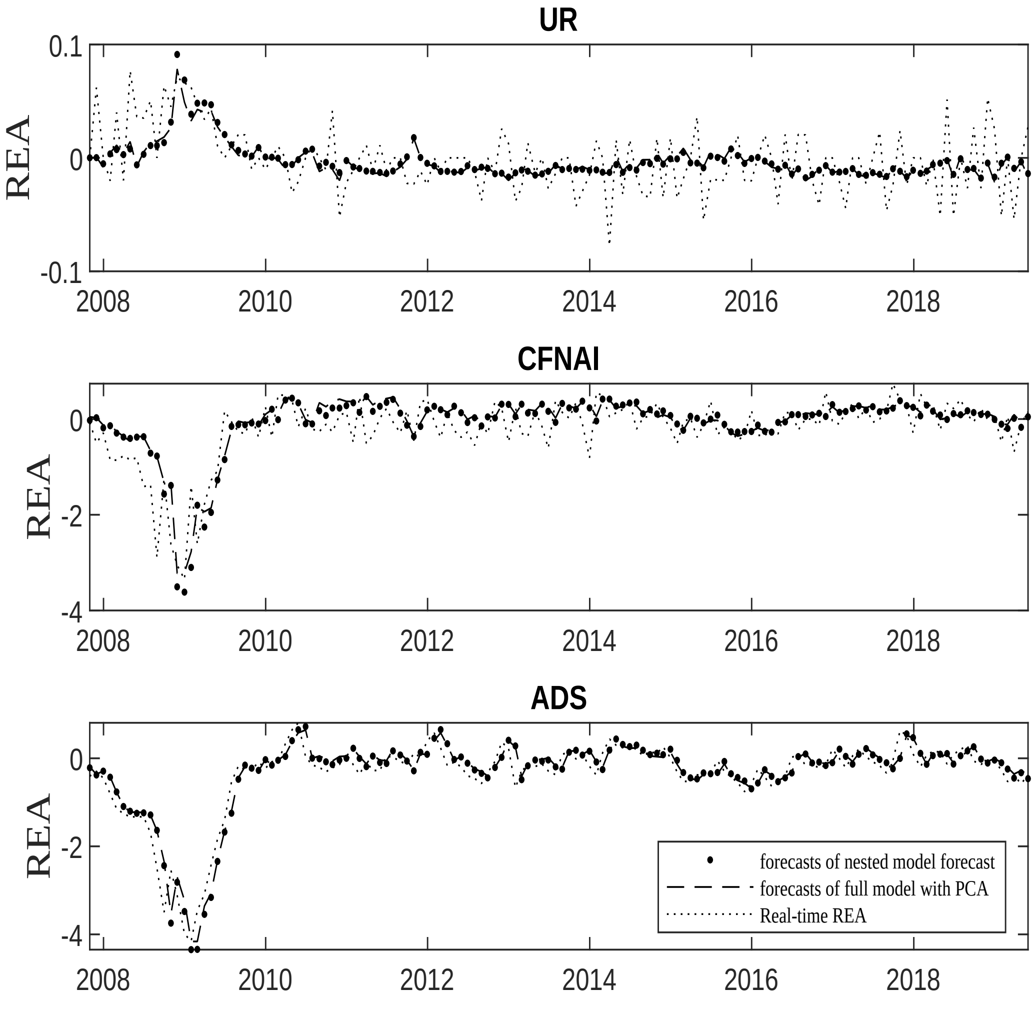


Figure 1.A.2. Evolution of OOS Forecasts of UR, CFNAI and ADS: The figure shows the evolution of REA forecasts (UR, CFNAI and ADS) over September 2007 - May 2019. For each REA measure, we extract the time series by performing a rolling window estimation. We use a rolling window starting on July 1998 with size 30 months.

		h=1 month	h=2 month	h=3 month	h=6 month	h=12 month
IP	beta	-0.38***	-0.34***	-0.17	-0.05	-0.19
		[0.00]	[0.00]	[0.16]	[0.39]	[0.02]
		(0.00)	(0.00)	(0.05)	(0.65)	(0.34)
	R2	0.21	0.32	0.40	0.31	0.27
NFP	beta	-0.29***	-0.20**	-0.18**	-0.17**	-0.20***
		[0.00]	[0.01]	[0.02]	[0.02]	[0.01]
		(0.00)	(0.01)	(0.04)	(0.11)	(0.26)
	R2	0.67	0.76	0.76	0.71	0.54
RS	beta	-0.05	-0.03	-0.07	0.07	0.03
		[0.38]	[0.75]	[0.56]	[0.91]	[0.71]
		(0.43)	(0.66)	(0.57)	(0.59)	(0.80)
	R2	0.09	0.11	0.16	0.26	0.34
HOUS	beta	-0.11**	-0.06	-0.01	-0.07	0.08
		[0.04]	[0.18]	[0.56]	[0.30]	[0.92]
		(0.28)	(0.55)	(0.87)	(0.22)	(0.37)
	R2	0.20	0.20	0.19	0.36	0.56
CU	beta	-0.38***	-0.32***	-0.14	-0.04	-0.14
		[0.00]	[0.00]	[0.17]	[0.42]	[0.00]
		(0.00)	(0.00)	(0.07)	(0.69)	(0.35)
	R2	0.22	0.34	0.43	0.32	0.28
UR	beta	0.38***	0.40***	0.36***	0.13	0.07
		[0.00]	[0.00]	[0.00]	[0.26]	[0.29]
		(0.00)	(0.00)	(0.00)	(0.24)	(0.57)
	R2	0.18	0.32	0.41	0.56	0.52
CFNAI	beta	-0.43***	-0.18	-0.05	0.06	-0.05
		[0.00]	[0.19]	[0.53]	[0.82]	[0.85]
		(0.00)	(0.01)	(0.56)	(0.64)	(0.73)
	R2	0.49	0.48	0.41	0.27	0.15
ADS	beta	-0.12***	-0.21***	-0.15**	0.05	0.01
		[0.00]	[0.00]	[0.01]	[0.52]	[0.77]
		(0.01)	(0.01)	(0.05)	(0.73)	(0.97)
	R2	0.86	0.71	0.61	0.37	0.22

Table 1.A.1. Predicting REA via SVIX: This table reports results from the estimated multiple predictor regressions as in Equation (1.A.19) for various U.S. real economic activity (REA) proxies and for an h-month forecasting horizon ($h = 1, 3, 6, 9$, and 12 months). The REA proxies considered are explained in Section 1.3.2. The multiple predictor model includes the lagged REA and the h-month SVIX measure as predictors and is augmented by a set of control variables: term spread (TS), credit spread (CR), forward variance (FV), and IRRA. We report the standardized ordinary-least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors as well as the in-sample adjusted R^2 for any given model. The sample spans July 1998 to May 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

		IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
TS	h=1	0.06	-0.11	-0.05**	-0.02	0.17**	0.07	0.02	0.01
		[0.27]	[0.01]	[0.49]	[0.89]	[0.02]	[0.18]	[0.55]	[0.69]
		(0.38)	(0.01)	(0.35)	(0.74)	(0.01)	(0.25)	(0.72)	(0.75)
	h=2	0.11	-0.02	0.04	0.00	0.22**	0.05	0.07	0.05
		[0.10]	[0.60]	[0.72]	[0.97]	[0.01]	[0.55]	[0.17]	[0.21]
		(0.10)	(0.60)	(0.40)	(0.97)	(0.00)	(0.40)	(0.12)	(0.22)
	h=3	0.14	0.03	0.10	0.01	0.22	0.02	0.10	0.10*
		[0.13]	[0.82]	[0.35]	[0.92]	[0.06]	[0.80]	[0.18]	[0.05]
		(0.07)	(0.50)	(0.15)	(0.85)	(0.01)	(0.67)	(0.11)	(0.03)
CS	h=1	-0.26***	-0.22***	-0.08	0.03	-0.18**	0.18**	-0.34***	-0.12*
		[0.00]	[0.00]	[0.51]	[0.60]	[0.03]	[0.02]	[0.00]	[0.08]
		(0.00)	(0.00)	(0.26)	(0.73)	(0.02)	(0.00)	(0.00)	(0.03)
	h=2	-0.25**	-0.17**	-0.18	0.00	-0.14	0.20*	-0.19***	-0.21**
		[0.02]	[0.01]	[0.19]	[0.91]	[0.26]	[0.07]	[0.00]	[0.02]
		(0.02)	(0.01)	(0.02)	(0.97)	(0.14)	(0.02)	(0.01)	(0.00)
	h=3	-0.16	-0.16**	-0.27	0.02	-0.05	0.20	-0.14*	-0.21*
		[0.39]	[0.04]	[0.17]	[0.93]	[0.80]	[0.18]	[0.06]	[0.04]
		(0.14)	(0.05)	(0.01)	(0.88)	(0.67)	(0.04)	(0.14)	(0.01)
FV	h=1	-0.28***	-0.16***	-0.13	-0.06	-0.28***	0.22***	-0.25***	-0.13***
		[0.00]	[0.00]	[0.12]	[0.55]	[0.00]	[0.00]	[0.00]	[0.00]
		(0.00)	(0.00)	(0.06)	(0.22)	(0.00)	(0.03)	(0.01)	(0.00)
	h=2	-0.37***	-0.20***	-0.27**	-0.06	-0.35***	0.28***	-0.26***	-0.27***
		[0.00]	[0.00]	[0.02]	[0.48]	[0.00]	[0.00]	[0.00]	[0.00]
		(0.00)	(0.00)	(0.00)	(0.42)	(0.00)	(0.00)	(0.00)	(0.00)
	h=3	-0.35***	-0.22***	-0.32**	-0.03	-0.32**	0.28**	-0.21**	-0.33***
		[0.00]	[0.00]	[0.03]	[0.65]	[0.02]	[0.02]	[0.02]	[0.00]
		(0.00)	(0.00)	(0.00)	(0.71)	(0.00)	(0.00)	(0.00)	(0.00)
IRRA	h=1	-0.25***	-0.18***	-0.28***	-0.23***	-0.17***	0.22 ***	-0.20***	-0.11***
		[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
		(0.01)	(0.00)	(0.00)	(0.00)	(0.06)	(0.00)	(0.01)	(0.00)
	h=2	-0.29***	-0.18***	-0.39***	-0.39***	-0.20**	0.26**	-0.21**	-0.21***
		[0.00]	[0.00]	[0.00]	[0.00]	[0.02]	[0.01]	[0.01]	[0.00]
		(0.02)	(0.00)	(0.00)	(0.00)	(0.10)	(0.00)	(0.01)	(0.01)
	h=3	-0.27**	-0.18***	-0.45***	-0.45**	-0.18*	0.27**	-0.20*	-0.25***
		[0.01]	[0.00]	[0.00]	[0.01]	[0.08]	[0.03]	[0.07]	[0.00]
		(0.05)	(0.01)	(0.01)	(0.00)	(0.18)	(0.01)	(0.04)	(0.02)

Table 1.A.2. Regression coefficient of control variables at short-term horizon forecast: This table presents the regression coefficients of the control variables in predictive model (1.17) for various U.S. real economic activity (REA) proxies and for long-term horizon (h = 6, 12). The main variable of interest in the model is the PCA factor, and control variables include term spread (TS), credit spread (CS), forward variance (FV) and IRRA. For each control variable at each horizon, we report the ordinary least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors, as well as the in-sample adjusted R^2 for any given model. The sample spans July 1998 to June 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

		IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
TS	h=6	0.17*	0.12	0.07	-0.03	0.27*	-0.13	0.08	0.09*
		[0.04]	[0.13]	[0.61]	[1.00]	[0.03]	[0.45]	[0.51]	[0.09]
		(0.21)	(0.13)	(0.60)	(0.70)	(0.06)	(0.16)	(0.54)	(0.48)
	h=12	0.07**	0.16***	0.12	-0.01***	0.24	-0.26	0.08	0.05
		[0.01]	[0.00]	[0.63]	[0.98]	[0.00]	[0.10]	[0.69]	[0.26]
		(0.70)	(0.39)	(0.25)	(0.95)	(0.16)	(0.12)	(0.41)	(0.66)
CS	h=6	-0.06	-0.07	-0.10	0.09	0.08	-0.03	0.09	0.03
		[0.45]	[0.13]	[0.57]	[0.98]	[0.97]	[0.72]	[0.59]	[0.37]
		(0.61)	(0.45)	(0.55)	(0.35)	(0.48)	(0.82)	(0.58)	(0.86)
	h=12	-0.11**	-0.05**	0.03	0.18	0.15	-0.18	0.17	0.25
		[0.05]	[0.07]	[0.86]	[0.96]	[0.76]	[0.72]	[0.76]	[0.68]
		(0.37)	(0.72)	(0.81)	(0.09)	(0.09)	(0.15)	(0.08)	(0.05)
FV	h=6	-0.38***	-0.28***	-0.12	0.02	-0.36***	0.26**	-0.12	-0.26**
		[0.00]	[0.00]	[0.50]	[0.88]	[0.00]	[0.04]	[0.43]	[0.01]
		(0.00)	(0.00)	(0.24)	(0.74)	(0.00)	(0.00)	(0.13)	(0.00)
	h=12	-0.29***	-0.27***	-0.05	0.13	-0.26***	0.24**	-0.02	-0.09*
		[0.00]	[0.00]	[0.69]	[0.68]	[0.00]	[0.01]	[0.67]	[0.07]
		(0.00)	(0.00)	(0.43)	(0.01)	(0.00)	(0.02)	(0.66)	(0.08)
IRRA	h=6	-0.34**	-0.23***	-0.47*	-0.67**	-0.23	0.25	-0.24	-0.33**
		[0.01]	[0.00]	[0.07]	[0.01]	[0.11]	[0.22]	[0.31]	[0.01]
		(0.11)	(0.12)	(0.03)	(0.00)	(0.27)	(0.10)	(0.21)	(0.13)
	h=12	-0.51	-0.34	-0.54	-0.73	-0.33	0.32	-0.30	-0.40
		[0.05]	[0.16]	[0.16]	[0.41]	[0.54]	[0.51]	[0.60]	[0.06]
		(0.05)	(0.16)	(0.02)	(0.00)	(0.16)	(0.15)	(0.14)	(0.09)

Table 1.A.3. Regression coefficient of control variables at long-term horizon forecast: This table presents the regression coefficients of the control variables in predictive model (1.17) for various U.S. real economic activity (REA) proxies and for short-term horizon ($h = 1, 2\ 3$) . The main variable of interest in the model is the PCA factor and control variables includes: term spread (TS), credit spread (CS), forward variance (FV) and IRRA. For each control variable at each horizon, we report the ordinary least-squares (OLS) coefficient estimates, IVX-Wald (within squared brackets) and Newey-West (within brackets) p-values of each one of the predictors as well as the in-sample adjusted R^2 for any given model. The sample spans July 1998 to June 2019. We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively.

Panel A:OOSR2: full model with PCA factor versus AR(1)								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
h=1 month	0.26*	0.22*	0.08	0.02	0.20	0.27*	0.35*	0.13*
	(1.72)	(2.38)	(0.51)	(0.21)	(1.42)	(3.29)	(1.99)	(1.97)
h=2 month	0.28	0.22**	0.14	0.09	0.20	0.30***	0.04	0.29***
	(1.44)	(3.22)	(1.07)	(0.86)	(1.11)	(3.91)	(0.17)	(2.72)
h=3 month	0.17**	0.24**	0.14	0.10	0.09	0.27***	-0.09	0.22
	(2.29)	(2.77)	(1.14)	(1.32)	(0.97)	(3.71)	(-0.48)	(2.34)
h=6 month	0.11	0.27***	0.08	0.27***	0.06	0.13**	-0.16	-0.01
	(1.02)	(2.72)	(0.87)	(3.11)	(0.57)	(2.10)	(-0.99)	(-0.03)
h=12 month	0.19	0.24***	0.31	0.41***	0.09	0.23**	0.10	0.14
	(1.47)	(2.92)	(4.16)	(3.60)	(0.92)	(5.08)	(0.99)	(1.31)

Panel B:OOSR2: full model with PCA factor versus the nested model								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
h=1 month	0.05	0.04	0.02	0.00	0.04	0.01	0.09	0.07
	(0.98)	(1.03)	(0.72)	(-0.24)	(1.34)	(0.42)	(1.01)	(0.99)
h=2 month	0.12**	0.10*	0.06	0.00	0.09**	0.02	0.05	0.16***
	(2.36)	(1.72)	(1.47)	(-0.11)	(2.42)	(1.12)	(0.91)	(2.88)
h=3 month	0.08*	0.13*	0.02	0.00	0.06*	0.00	0.02	0.15*
	(1.84)	(1.59)	(0.52)	(0.03)	(1.87)	(-0.15)	(0.39)	(1.92)
h=6 month	0.04	0.10	-0.06	0.00	0.04	-0.07	-0.04	-0.02
	(1.05)	(1.34)	(-1.06)	(0.13)	(1.33)	(-1.13)	(-1.01)	(-0.34)
h=12 month	-0.01	0.06	-0.03	0.03	-0.01	-0.09	-0.01	-0.01
	(0.58)	(1.12)	(-1.03)	(1.39)	(-0.09)	(-1.13)	(-0.58)	(-0.58)

Panel C: OOSR2: full model with PCA factor versus model with SVIX								
	IP	NFP	RS	HOUS	CU	UR	CFNAI	ADS
h=1 month	0.04	0.07	-0.01	0.00	0.02	0.01	0.10	0.08
	(0.46)	(1.30)	(1.24)	(0.59)	(0.93)	(0.57)	(1.03)	(1.27)
h=2 month	0.13**	0.10*	0.04**	-0.06	0.08**	-0.01	0.07	0.16***
	(2.04)	(1.61)	(1.64)	(-0.21)	(2.39)	(0.39)	(1.47)	(3.25)
h=3 month	0.08**	0.12	-0.03	-0.04	0.06**	-0.08	0.02	0.13**
	(2.05)	(1.57)	(0.96)	(-0.31)	(2.08)	(-0.74)	(0.89)	(2.37)
h=6 month	0.02	0.05	-0.16	-0.09	0.02	-0.20	-0.07	-0.07
	(1.19)	(1.39)	(-0.88)	(-0.46)	(1.35)	(-1.62)	(-0.58)	(-0.34)
h=12 month	-0.09	-0.05	-0.11	-0.08	-0.12	-0.30	0.01	-0.02
	(0.39)	(0.14)	(-0.05)	(-0.97)	(0.05)	(-2.28)	(1.12)	(0.48)

Table 1.A.4. Out-of-Sample Predictability of U.S. REA with PCA Factor with Diebold-Mariano (DM) test: This table compares the OOS^R2 and their corresponding DM test statistic obtained from comparing the forecasts of full predictive model explained in Equation (1.18) which contain the PCA factor, versus the AR(1) model (in panel A), versus forecasts of the nested model (in Panel B), versus forecasts of the full model with alternative predictive model which contains the SVIX measure (in Panel C) and forecasts of full model with FF5 (in Panel D).

	PCAF	Mrkt	SMB	HML
coef	0.00	-0.05	0.06	-0.04
t-stat	-0.55	-2.61	1.87	-1.10
NW p-value	0.58	0.01	0.06	0.27
IVX pvalue	0.83	0.02	0.09	0.31

	weightedF	Mrkt	SMB	HML
coef	-0.20	-0.05	0.06	-0.04
t-stat	-0.37	-2.60	1.83	-1.12
NW p-value	0.71	0.01	0.07	0.26
IVX pvalue	0.66	0.01	0.08	0.35

	SVIX	Mrkt	SMB	HML
coef	-0.10	-0.05	0.06	-0.05
t-stat	-0.16	-2.58	1.79	-1.20
NW p-value	0.87	0.01	0.07	0.23
IVX pvalue	0.78	0.01	0.08	0.34

Table 1.A.5. Predictive Power of PCA and SVIX for Future SPX Returns: This table reports the regression coefficient of forecasting market expected return with PCA factor or the SVIX measure controling for 3 factors of fama-french, namely market, SML and HML using the followingregression model.

$$Ret_{spx,t+1} = \alpha + \beta_1 Factor_t + \beta_2 mrkt_t + \beta_3 SML_t + \beta_4 HML_t + \epsilon_t$$

Chapter 2

Do Options Markets React to FOMC meetings?

2.1 Introduction

Macroeconomic events have been shown to have a substantial impact on stock markets and are, in general, accompanied by a rise in risk and return.[1] However, very few papers have investigated the impact of these events on options markets. It is not trivial that options markets should react similarly to stock markets. Under the hypothesis that options traders are better informed than stock traders,[2] they may better anticipate macroeconomic events' outcomes. However, Kelly et al. (2016) find that elections and summits lead to an increase in priced variance and tail risk, which suggests that it is not the case.

This paper uses as treatments options that are traded before FOMC meetings and expiring after the meetings. These options are compared to other options whose lives do not span FOMC meetings, to measure the impact of these meetings on volatility and tail risks, and on the demand for options.

1. Daniel et al. (2002), and Ozoguz (2009) suggest that there is a negative relationship between the level of macroeconomic uncertainty and asset valuations.

2. Options are contracts used to hedge against or speculate/take a view on uncertainty about the future prices of a wide range of financial assets and commodities. Option prices contain information about the investors' expectations about the future prices of these underlying assets. Hull (2000)

We consider FOMC meetings in the United States between 2006 and 2016. We investigate whether option-implied risk-neutral moments, risk premiums and trading activity indicators are significantly different for options whose lives span an FOMC meeting and for other options. We find that the implied volatility, its slope on the left of the ATM level as well as the variance risk premium (VRP) are larger for options whose lives span FOMC meetings. These findings are consistent with Kelly et al. (2016) and suggest that investors are willing to pay for protection against the price, downside tail risk and variance associated with these events. Higher prices may also compensate for the risk that market makers bear for trading with informed traders. These higher prices are associated with an increase in the bid-ask spread and more generally a decrease in the liquidity of options (open interest and trading volume). More specifically, the net demand to open positions on options decreases while the net demand to close positions increases. These results are statistically significant across all options' moneyness levels.

A time series analysis of risk-neutral moments and trading activity indicators around FOMC meetings reveals that the impact of these events on options markets are most pronounced in the week before the event and last until a few days after the event.

Our results are in agreement with Handa et al. (2006), who show that bid-ask spreads are correlated with information asymmetry in a limit order book market. In addition, our results are consistent with Erenburg and Lasser (2009), who find that the bid-ask spreads of index-linked securities increase around macroeconomic news releases in a limit order book market. Furthermore, Lucca and Moench (2015) show that the trading activities on mini S&P 500 futures decrease before FOMC meetings. They document that the FOMC decisions provide information about the future economic outlook and interest rates; therefore, pre-FOMC returns carry a premium that equity investors require for bearing systematic risk associated with these events. Although FOMC members may share some information through interviews and speeches between meetings, they avoid these discussions in the week before FOMC meetings[3] and hence, systematic risk is likely to be high these days, the window during which we observe higher prices, tail risk, and lower market liquidity.

3. a time period known as the blackout period

Further, we examine the existence of directional trading before the FOMC events. Following Bernanke and Kuttner (2005), we decompose monetary policy shocks (mps) into their expected and unexpected components and assess their relationship with tail risk and the demand for protection (proxied by RN-skewness and put-call ratio). We find that the expected component of monetary policy shocks is positively related to the average put-call ratio on the days before FOMC meetings and negatively related with the RN-skewness. Therefore, a positive shock to downside risk and to the put-call ratio predicts an increase in the Fed rate. It implies that when investors become more worried about a decline in stock market prices, they predict a contractionary monetary policy- i.e., future short-term rates to increase. These results align with Bernanke and Kuttner (2005), who find that an increase in the Fed fund decreases the 30-day index returns. They are also in line with David and Veronesi (2014) equilibrium model, suggesting that when the tail-risk index rises, it indicates that the economy is moving to a booming regime. Consequently, the forward-looking Taylor rule anticipates a tightening of monetary policy back to normal levels- i.e., higher future interest rates. We also find a negative relationship between the 30-days implied volatility of S&P500 options and mps. This finding suggests that the monetary authority reacts to a volatile stock market by lowering interest rates. Therefore, investors anticipate lower interest rates from the Fed during periods of high volatility.

Finally, employing the Composite Leading Indicator (CLI) published by the OECD as a measure of economic conditions, and a measure of political uncertainty (EPU) proposed by Baker et al. (2013), we analyse whether the effect of FOMC meetings depends on political uncertainty and economic conditions. We find that they do. Indeed, consistent with Kelly et al. (2016), we find that the implied volatility and VRP differentials on the week preceding FOMC meeting are larger when the uncertainty is higher, or the economy is weaker. Also, the open net demand is higher in the week before the event when the economy is weaker. This is consistent with the Garleanu et al. (2009) demand pressure hypothesis. The higher demand pressure increases the implied volatility of options and hence option prices.

The rest of this paper is organised as follows: Section 2.2 reviews the related literature. Section 2.3 describes our data; Section 2.4 quantifies the effect of FOMC meetings; Section 2.5 examines the timing of option trading activities and changes on pricing moments around FOMC meetings. Section 2.6 considers how option traders are informed; Section 2.7 analyses the role of economic conditions on the effect of uncertainty and Section 2.8 briefly concludes.

2.2 Related Literature

We build on several strands of literature. Some papers studied the financial effect of monetary policy events on different markets. Using a daily event window, Bernanke and Kuttner (2005) find that the effect of unexpected monetary policy actions on expected excess returns explains most significant part of the changes on stock prices. Lucca and Moench (2015) show that returns drift upwards on the day before FOMC meetings, regardless of the direction of the monetary policy surprise. Neuhier and Weber (2018) find that stock returns start shifting upwards 25 days before expansionary monetary policy announcements, and this increase continues after the policy decision for another 15 days. Cieslak et al. (2018) find that the equity premium is higher in even weeks between FOMC meeting cycles. Their study, together with Vissing-Jorgensen (2019), also implies that FOMC announcements have a larger stock market impact than other types of macroeconomic announcements. Savor and Wilson (2013) show that stock returns are considerably higher on macroeconomic announcement days such as FOMC days. We contribute to this literature by studying the effect of FOMC meetings on options markets in terms of pricing moments and find that implied volatility and tail risk extracted from options increase before the meeting.

This paper also contributes to literature that studies liquidity dynamics around announcements, i.e., Cremers et al. (2022) find that traders of options on individual stocks are more active than other market participants both immediately before and immediately after news releases. However, Erenburg and Lasser (2009) find that in a limit order book market, the bid-ask spreads of index-linked securities increase around macroeconomic news releases. Bollerslev et al. (2016) find that the trading volume spike after FOMC announcements is related to the level of disagreement among investors. In this paper, we find that the bid-ask spread of index options increase before the meeting, resulting in lower trading activities.

Some papers have suggested that options traders were more informed and skilled than stock traders and argued that the options markets facilitates the price discovery process in the equity market. Skinner (1997) shows that introducing options trading for a specific stock improves price efficiency. If options traders hold superior information, then we should observe directional trading in the options markets. Amin (1997) examines abnormal trading volume in the options markets around the earning announcements and document directional trading evidences. Pan and Poteshman (2006) find that put-call open-buy trading volume ratios predict underlying stock returns suggesting that options traders who open new positions possess significant trading skills. Along the same line, Jin et al. (2012) analyse the unbalance between put and call trading by studying implied volatility spreads between ATM puts and calls and implied volatility skews. They show that the predictive ability of these variables is larger before earnings announcements, client and product announcements, and days with large stock price movements. Consistent with these studies, our results show that traders take advantage of the predictable movements of underlying index returns and their volatility- i.e., a positive shock to the downside-risk proxied by RN-skewness and the put-call ratio predicts higher Fed rate.

At this point, we have not performed tests to identify the channel whereby decreased liquidity in options is associated with increased implied volatility and decreased risk-neutral skewness. A few papers have linked the fundamentals of options and risk-neutral moments. Bollen and Whaley (2004) show that changes in option demand lead to changes in option prices and found that the implied volatility of index options is directly related to the buying pressure for index puts. Garleanu et al. (2009) find that demand pressure in one option contract increases its price by an amount proportional to the variance of the unchangeable part of the option. Constantinides and Lian (2021) find that the net buy of puts by public customers has negative relationship with risk-neutral variance and disaster index. The intuition is that when the risk-neutral variance and disaster index increase, public customers incline to buy more puts as insurance, but market makers become more risk averse. Therefore, both the supply and demand curves shift, however, the supply shift emerges as the driving factor in decreasing the equilibrium net buy of puts. Chen et al. (2019) show that the net buy of deep OTM puts proxies for demand for crash insurance and predicts the return on the S&P 500 index.

They suggest that tightening intermediary constraints is linked to expensive option prices, increased risk premia, illiquidity, and broker-dealer deleveraging. Therefore, positive comovements between prices and trading indicators are due to demand shocks, while negative relationships are consistent with the presence of supply shocks. Garleanu et al. (2009) show that volatility demand on days prior to earnings announcements increases option prices.

2.3 Data

2.3.1 Macroeconomic Events

We study the impact of the Federal Open Market Committee (FOMC) meetings that took place during 2006 - 2016, listed on the website of the Federal Reserve. FOMC members makes policy decisions under the dual statutory mandate of price stability and maximum sustainable employment. They regularly assemble at scheduled meetings every 42 days and sometimes at unscheduled meetings, which are typically conducted via teleconference calls. In this paper, we focus our analysis exclusively on scheduled meetings and study the options markets ahead of these meetings which are known to investors in advance.

2.3.2 Data Sources

We collect daily options on S&P 500 from 2006 - 2016, from OptionMetrics. We keep options with maturity between 5 and 45 days. We remove options with missing quotes, zero open interest or moneyness above 2 (as measured by the ratio of the strike by the spot price). We create three buckets of moneyness: out-of-the-money put and call options (OTMP, OTMC) and at-the-money options (ATM). The moneyness buckets and the number of options in each category are given in Table 2.1. Before 2006, options' expiration dates were on Saturdays immediately following the third Friday of each month. From 2006, the number of available expiration dates has increased to become close to weekly. However, the most used expiration dates remain the end-of-the-month expiration days.

To compute options' implied moments with constant maturity (30 days), we use implied volatilities for standardised equity options with 30-day expiry on a grid of the deltas, as provided by OptionMetrics.

Furthermore, we use the volume data contained in the CBOE (C2 exchange) Open-Close database. This dataset contains daily records of closing short and long open interest on all S&P 500 options traded by public customers, firm proprietary traders and market makers from the beginning of 2006 until the end of 2016. Following Garleanu et al. (2009), we assume that options "end-users"- i.e., non-market makers, are either public customers or firm proprietary traders.

2.3.3 Trading Activity Variables

To analyse the trading activities of options traders, we focus on a few key liquidity indicators. First, for each option, we consider the open interest, the volume, the bid-ask spread scaled by options' mid-price, and the put-call ratio. These indicators are calculated for options in different moneyness buckets (see Table 2.1). As options trading activity has been increasing over time, the volume and open-interest have been normalised using their average value over a rolling window of 42 days.[4]

[Table 2.1 about here.]

Second, we focus on transactions initiated by end-users (i.e., non-market makers) and on the pressure these transactions may induce on financial intermediaries. We construct the net demand measure proposed by Garleanu et al. (2009), defined as the long trading volume minus the short volume. As options are in zero net supply, the end-users net demand equals the negative of the market makers' net demand. Large demand for a particular type of option creates deviation from optimal inventory for market makers and induces demand pressure, which may, in turn, be priced. We decompose net demand into open and close net demand and normalise each of these quantities by the average of open (close) buy and sell demand, over the rolling window of 42 days:

$$OND_{i,t} = \frac{OB_{i,t} - OS_{i,t}}{(\sum_{j=1}^{42} OB_{t-j}/42 + \sum_{j=1}^{42} OS_{t-j}/42)/2} \tag{2.1}$$

where $OB_{i,t}$ ($OS_{i,t}$) is the open buy (sell) positions on option i on day t and OB_{t-j}

4. There are on average 42 days between consecutive FOMC meetings.

(OS_{t-j}) is the total open (close) buy positions on day $t-j$.

In addition, we compute the refined signed put-call ratio of Cremers et al. (2022), which focuses on open-buy contracts purchased by end-users:

$$PCR_{m,t} = \frac{P_{m,t}}{C_{m,t} + P_{m,t}} \tag{2.2}$$

where $P_{m,t}$ and $C_{m,t}$ are the numbers of put and call contracts purchased by non-market makers to open new positions, on date t and at moneyness m.

Finally, following Andreou et al. (2018), we measure the dispersion in investors' beliefs regarding the future stock price by computing the daily dispersion of options' open interest across levels of moneyness M_j, labelled DISPOI. Given a range of strikes K_j where j = 1, 2, ..., N, DISPOI is defined as:

$$DISPOI_t = \sum_{j=1}^{N} \Phi_j |M_j - \sum_{j=1}^{N} {}_j M_j| \tag{2.3}$$

where Φ_j is the ratio of open interest of options strike K_j to all available strikes, we require at least three options with positive open interest to compute DISPOI on a given day. An increase in DISPOI reflects an increase in the dispersion of investor beliefs regarding the future stock price.

2.3.4 Risk-neutral Moments and Risk Premiums

We analyse the behaviour of option prices around FOMC meetings, focusing on measures of volatility and skewness of the underlying index. More specifically, we analyse the value of option protection against three aspects of risk associated with FOMC meetings: price risk, tail risk, and variance risk.

First, we examine the price risk using the at-the-money implied volatility (ATM-IV). We compute three alternatives to this measure. First, we compute the average implied volatilities of ATM options from the option-data file in Optionsmetrics[5] using options on each day over a given window of time. We denote this measure by ATM-IV$_{M1}$. Second,

5. Moneyness buckets are given in Table 2.1.

we extract the ATM-IV of the horizon 30 days from the interpolated IV surface file in the Optionsmetrics and indicate it by ATM-IV$_{M2}$. Third, the 30-days model-free IV (MFIV) is computed following Bakshi et al. (2003) using a portfolio of OTM options. These measures are defined and discussed in Appendix 2.A.1.[6]

In addition, we examine whether variance shocks are priced. We consider two approximations of the variance risk premium (VRP). The first one, VRP$_{M1}$, calculates the difference between the implied variance of an ATM option and the expected future realised variance over the life of an option, whereby the latter results from a linear model as described in Appendix 2.A.1. The second one, VRP$_{M2}$, subtracts the realised 30-day variance from the 30-day $MFIV^2$. These two estimates of the VRP have a 78% correlation.

To measure the expensiveness of out-of-the-money put and call options separately, we compute the slope of the implied volatility on the left and right hand-sides of the at-the-money level (referred to as left and right IV slopes). To construct it, we regress the implied volatilities of options each day of a given window on their deltas. We compute two alternative measures of implied volatility slope. The first measure is constructed using the option data file in Optionsmetrics. Specifically, We aggregate all options traded each day, t, with a same time-to-maturity, τ , and build $IVS_{t,\tau}$ and denote it by IVS_{M1}. The second measure is calculated using the volatility surface file using options with constant 30-days to maturity. We denote the second measure by IVS_{M2}. We complement these measures with the risk-neutral skewness (RNS) of Bakshi et al. (2003).[7] The left (resp. right) IV slope has a correlation of -0.25 (resp. 0.14) with the risk-neutral skewness. All variables are formally defined in Appendix 2.A.1.

6. Alternative proxies for volatility such as the VIX and the simple VIX (SVIX) of Martin (2017) can be used and do not change the results. Correlations between the volatility measures we use and the VIX index in our sample are between 0.95 and 1.

7. We calculate two alternatives to this measure. The first measure is constructed using S&P500 options traded each day t, with an available time-to-maturity τ. We use the option data file of Optionsmetrics to construct our first measure. This measure is denoted by RNS_{M1}. The second measure is constructed using constant 30-days time-to-maturity options using the volatility surface file in OptionsMetrics and is denoted by RNS_{M2}.

2.4 Quantify the Effect of FOMC Meetings on Options

In this section, we quantify the effect of FOMC meetings on options whose lives span the event. Unless otherwise specified, we consider all options with maturities between 5 and 45 days. To this end, we build three regression models.

The first regression model evaluates the impact of FOMC meetings on risk-neutral moments- namely, the ATM implied volatility ($ATM - IV_{M1}$)[8], the risk-neutral skewness[9] (RNS_{M1}), the implied volatility slope (IVS_{M1}), and the variance risk premium (VRP_{M1}). These variables are constructed by aggregating all options traded each day with a given maturity. We define the dummy variable $D^{FOMC}_{t,\tau}$ which takes value 1 if there is an FOMC meeting between times t and $t + \tau$, and value of 0 otherwise. The above mentioned variables ($OptionVar_{t,\tau}$) are then regressed on $D^{FOMC}_{t,\tau}$:

$$OptionVar_{t,\tau} = \beta_0 + \beta_1.D^{FOMC}_{t,\tau} + \beta_2\tau + \epsilon_{t,\tau} \tag{2.4}$$

The computation of all variables is described in Section 2.3.3 and in Appendix 2.A.2.[10]

Table 2.2 reports the coefficient estimates of regression (2.4). We find that options whose lives span FOMC meetings have higher ATM-IV, i.e., they tend to be more expensive than other options. Their VRPs are significantly larger, indicating that (at least part of) the additional volatility is priced. Kelly et al. (2016) find a similar effect induced by earnings announcements and elections. The risk-neutral skewness is also higher, but this change is not statistically significant. In particular, estimating the model with the IV slopes reveals a substantial increase in both the left and right tail risk. Both are statistically significant.

[Table 2.2 about here.]

---

8. Pricing variables denoted by subscript $M2$, which are calculated based on fixed-term maturity (30-day) options, are not suitable for this analysis. However, we find high correlations between variables obtained from the option-data file and from the volatility-surface files. For example, ATM-IV$_{M2}$ and VRP$_{M2}$ have 98% and 78% correlation with the ATM-IV$_{M1}$ and VRP$_{M1}$ statistic, respectively.

9. RNS is not comparable across maturities; hence we calculate it using the option with 25 to 35 days to maturity.

10. To calculate standard errors, both here and in our subsequent regressions, we consider Heteroskedasticity of error terms and calculate the t-stats following White (1980)'s method.

While it is usual that OTM put options become more expensive before events involving macroeconomic uncertainty (e.g., elections or earning announcements as in Kelly et al. (2016)), the increase in OTM call prices may seem more surprising. It is consistent with the increase in underlying index prices highlighted by Lucca and Moench (2015) in the days before FOMC meetings. The associated puzzle in the underlying market is that investors do not seem to take advantage of this pattern, resulting in the low liquidity of the index. We now examine whether this low liquidity can also be found in the options markets.

The second regression model evaluates the impact of FOMC meetings on trading indicators- namely volume, open interest, bid-ask spread and net-open demand or net-close demand of options and dispersion. We denote these trading indicators ($OptionVar_i$) in our regression model. Next, we define the dummy variable D_i^{FOMC}, which is option-specific: for option i, it takes value of 1 if an FOMC meeting takes place between the time the option is traded (t_i) and the time it expires ($t_i + \tau_i$), it takes value of 0 otherwise. The treatment group contains all options whose lives span an FOMC meeting, whereas the control group contains all other options. Finally, we set up our regression model as follows.[11]:

$$OptionVar_i = \beta_0 + \beta_1.D_i^{FOMC} + \beta_2\tau_i + \epsilon_i. \tag{2.5}$$

The term $\beta_2\tau_i$ controls for the maturity fixed effect of options.

Table 2.3 presents the coefficient estimates of regression (2.5) along with the corresponding T-statistics in parentheses. The T-statistics account for the Heteroskedasticity of error terms and are calculated using the method proposed by White (1980).[12] We reject the null hypothesis of a zero coefficient (no predictability) based on the p-values of the IVX-Wald test at the 1%, 5%, and 10% levels and denote it by triple, double, and single asterisks, respectively. The results suggest that the liquidity of options covering FOMC meetings is decreasing: the trading volume, open interest and open net demand are respectively 3.19%, 0.61% and 0.33% lower for these options, whereas the bid-ask

11. We also consider applying regression (2.4) for these trading variables by averaging them over a specific window. Our results remain qualitatively the same.

12. In section 2.C of the Appendix, our robustness analysis addresses time-series concerns. Clustering standard errors by time yields consistent results, and auto-correlation of error terms has no impact on our findings regarding trading activities.

spread and close net demand are respectively 1.31% and 0.33% higher. This decrease in liquidity is observed over the three buckets of moneyness. Furthermore, the dispersion measure DISPOI increases significantly for OTM calls, indicating little consensus among OTM call option traders.

[Table 2.3 about here.]

In addition to our main analysis, we also consider the time fixed effect in our regression model. The results are presented in Table 2.A.1 in the Appendix (Section 2.B). The robustness of our findings is evident as the liquidity of all options covering FOMC meetings exhibits a consistent pattern. Specifically, the trading volume and open interest for options covering the event are 1.95% and 13.2% lower, respectively, while the bid-ask spread and close net demand show respective increases of 24.3% and 0.46%. These results highlight the consistent pattern of reduced liquidity during FOMC meetings across different measures. The decrease in trading volume and open interest suggests a decline in market activity, while the higher bid-ask spread and close net demand indicate increased trading costs and potential imbalance between buyers and sellers.

Our third regression model, quantifies the effect of FOMC meetings on put-call ratio. To do so, we look more specifically at the put-call ratio, using the definition of Cremers et al. (2022) as in Equation (2.2). Therefore, the dummy $D_{t,i}^{FOMC}$, in our regression model, is set equal to one when using options whose lives span the event, and to 0 otherwise. We do not consider the time to maturity as a control variable for this model, since we aggregate options with different maturities each day.

$$OptionVar_{t,i} = \beta_0 + \beta_1 . D_{t,i}^{FOMC} + \epsilon_{t,i} \tag{2.6}$$

The low liquidity of OTM calls compared to put options is highlighted in Table 2.4, for options whose lives span the event. This result holds for both OTM and ATM options.

[Table 2.4 about here.]

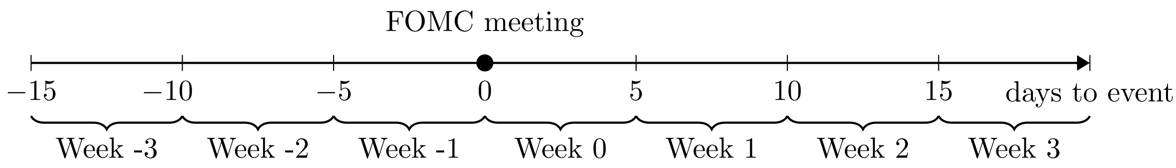
These results indicate a demand-supply imbalance of options before FOMC meetings, in line with Constantinides and Lian (2021). End-users seek to hedge the risk of FOMC meetings by buying put options, or speculate on its outcome by buying both put and call options. This demand pressure tightens constraints and induces pressure on financial intermediaries, who increase the prices of options as a response and restrict their supply. This, in turn, leads to a decreased trading of put and call options.

2.5 Option Moments and Liquidity over FOMC Cycles

In Section 2.4, we have shown that options whose lives span an FOMC event are more expensive and exhibit larger risk premiums, but lower liquidity than other options. In this Section, we analyse in more detail the timing of these phenomenons. We show that they are most prominent in the week preceding the meeting.

2.5.1 Model Design

We build on Cieslak et al. (2019) and decompose FOMC cycles into 7 weeks:



Week -1 starts four days before FOMC meeting and ends on the meeting day. Week 0 starts a day after the meeting and ends five days after, etc. Date 0 refers to the day of the scheduled FOMC meeting. For two-day meetings, date 0 is the second day, i.e., the announcement day. Weekends are omitted, and all variables under consideration are set to zero on holidays. Therefore, ten days after a meeting represent two calendar weeks after the event.

Most FOMC cycles include weeks -3, -2, -1, 0, 1, and 2, but the number of data points drops off quickly past week 2. There are 400 days in each of weeks -3 to 2, while there are only 200 days in week 3. This is mainly because FOMC meetings do not necessarily occur at regular intervals.

Each day, we calculate the volume of traded options, the open interest, the net demand, the average bid-ask spread, the average dispersion in investors' beliefs and the put-call ratio, using all options with maturity between 5 and 45 days. These measures are normalised as explained in Section 2.3.3. The return moments- i.e., ATM implied volatility, model-free IV, risk-neutral skewness and variance risk premium VRP_{M2} are computed using constant maturity (30-day) options.[13]

2.5.2 Weekly Fluctuations Over the FOMC Cycle

Figure 2.A.1 displays the fluctuations of average daily moments and risk premiums over each week of the FOMC cycle. During the week preceding FOMC meetings, consistent with our results from Section 2.4, the implied volatility is at its highest point, whereas the skewness is at its lowest. In addition, the variance risk premium is unusually high. Most of these effects disappear after the meeting. Specifically, the price and variance risk resolve after the meeting, while the tail risk remains persistent for another week.

[Figure 2.1 about here.]

Similarly, Figure 2.A.2 displays the average daily liquidity indicators over each week of the FOMC cycle, with their 5% and 95% confidence bands. The overall liquidity of the options markets is high two weeks before FOMC meetings and decreases as one gets closer to the event. Volume and open interest are the lowest one week before and the week after the meeting, and the bid-ask spread is the highest. This low liquidity is due to the low open net demand in the week preceding meetings and the high close net demand. Furthermore, the signed put-call ratio is higher during these two weeks.

[Figure 2.2 about here.]

We observe the same liquidity trends when focusing on ATM, OTM call and OTM put options, as shown in Figures 2.A.3, 2.A.4 and 2.A.5. Finally, in line with our findings in Section 2.4, the net demand to close call options is high on the week before the event. These findings confirm the ones obtained in Section 2.4.

[Figure 2.3 about here.]

[Figure 2.4 about here.]

13. To be consistent with our analysis in quantifying the effect of FOMC meetings in Section 2.4, we also construct ATM-IV, right (left) IV-slope and VRP using the option-data file in Optionmetrics, employing options with 35 to 45 days to maturity on each day.

[Figure 2.5 about here.]

[Figure 2.6 about here.]

2.5.3 Statistical Significance of Results

We test whether the averages of trading indicators or pricing moments in the weeks before the FOMC event are statistically different from those in the weeks after FOMC meeting. To this end, we regress daily variables on two dummies: one that equals 1 for days in the week preceding FOMC meetings and one that equals 1 for days in the week right after FOMC meeting. We control for the number of available maturities (NoAM) on each day.

$$OptionVar_{i,t} = \alpha + \beta_1.1\{t \in \text{week -1 }\} + \beta_2.1\{t \in \text{week 0}\} + \beta_3 \times \text{NoAM} + \epsilon_t \quad (2.7)$$

Positive (resp. negative) coefficients β_1 or β_2 suggest that the variable of interest is, on average, higher (resp. lower) during weeks -1 or 0 than during other weeks of the cycle.

As shown in Table 2.5, the implied volatility (ATM-IV) is significantly higher on days in week -1 with t-statistic of 2.80. In weeks -1 and 0, the left implied volatility slope obtained from constant maturity options is significantly higher with t-statistics of 3.61 and 2.26, respectively. Therefore, risk-neutral skewness is significantly lower by 4.75% during week -1 and 3.41% during the week after the meeting, and the VRP is higher during week -1 and drops after the meeting. To be consistent with our analysis in Section 2.4, we employ options pricing variables obtained from option-data file (variables subscripted by $M1$ as explained in Appendix) and regress daily average of these variables on week dummies.[14] Table 2.6 reports the results of the cycle analysis employing these variables. These results are consistent with our findings in Table 2.5 suggesting that the variance risk resolves after the meeting in week 0 (t-statistic of -2.23), while the tail risk (RNS) remains persistent for another week (t-statistic of -2.77).

[Table 2.5 about here.]

[Table 2.6 about here.]

14. We take the daily average of each variable using options with 25 to 35 days to maturity.

The effects of liquidity indicators are displayed in Table 2.7. The total volume on all options per day is lower on days that fall in week -1 in FOMC cycle time by 4.28%. Results of the regression on bid-ask spread and dispersion are shown in Table 2.8. In week -1, the bid-ask spread is also greater (i.e. by 0.32% on all options). We run the regression for different buckets of moneyness and observe the same results for OTM put and call options. Our results suggest that the dispersion of investors' beliefs on all, ATM and OTMP options, is high and significant in weeks -1 and 0.

[Table 2.7 about here.]

[Table 2.8 about here.]

Table 2.9 reports the significant tests on signed tradings, namely open net-demand and close net-demand. We find that in the week before the event, the net demand to close positions increases while the net demand to open positions decreases with t-stats equal to 1.88 and -1.84, respectively. Furthermore, the net demand to close the OTMC options is greater in week -1 by 8.58%. Finally, Table 2.10 reports the results of the regression on the put-call ratio and shows that the put-call ratio is higher in the week before the event by 3.6%. This increase in the put-call ratio applies for all buckets of options.

[Table 2.9 about here.]

[Table 2.10 about here.]

2.5.4 Discussion of Results

Our findings on trading activities complement our results on RN moments and suggest that around the time of the FOMC meetings, market makers are exposed to three types of risks: price, variance, and tail risk. As discussed by Lucca and Moench (2015), pre-FOMC underlying returns include a premium to compensate equity investors for bearing the uncertainty of what will be decided during the meeting. This leads investors to buy put options in the week before the meeting as a hedge against unfavorable outcomes and/or both put and call options to speculate on the outcome. The increased demand for options, combined with lower supply by constrained financial intermediaries, results in higher option prices, variance, and tail risk. Consequently, trading activity decreases, and bid-ask spreads widen.

Our results suggest that the FOMC meetings have an effect on options characteristics in the week preceding the announcement (week -1). This period is typically marked by heightened volatility and uncertainty in the options markets as traders anticipate potential policy changes or interest rate decisions by the Federal Reserve. Consequently, option premiums may increase, reflecting the higher perceived risk and uncertainty associated with the upcoming announcement. Options traders may adjust their positions or use options to hedge against potential market moves, thereby influencing options characteristics.

Some of these effects persist in the week after the meeting. One possible explanation could be that the effects of changes in Fed rates are not immediately fully incorporated in index prices. Bernanke and Kuttner (2005) found that an unanticipated 25-basis-point cut in the Federal funds rate target is associated with about a 1% increase in stock prices after 30 days.

On the other hand, Cieslak et al. (2019) show that the impact of FOMC cycles extends more broadly to the stock market, particularly in every even week of the cycle.[15] This broader impact could be attributed to the release of minutes from the previous FOMC meeting, which typically occurs three weeks after the meeting. These minutes provide insights into the discussions and debates among policymakers, offering information about the future path of monetary policy. Market participants scrutinize these minutes for clues regarding the Federal Reserve's stance and potential policy changes, leading to reactions in the stock market.

The difference in the effects of FOMC meetings on options characteristics and stock market behavior in different time frames can be attributed to the nature of policy uncertainty and its resolution. The inherent unpredictability and ambiguity surrounding the Federal Reserve's decisions, such as changes in interest rates or shifts in monetary policy, create a sense of uncertainty in financial markets. As a result, market participants

15. Our results indicate that the FOMC meetings have an impact on options characteristics in the week preceding the announcement (week -1). We also observe that these effects on the moments persist in week 3, which corresponds to week 4 in Cieslak et al. (2019)'s setting due to a slight shift in their week numbering. However, Cieslak et al. (2019) suggest that there is an effect on the stock market in every even week of the FOMC cycle. Hence, it appears that most of the effect of FOMC meetings is concentrated in weeks -1 and 3 (week 0 and 4 in Cieslak and Povala's setting). This suggests that FOMC meetings exhibit a mid-cycle effect on the stock market rather than the option market.

actively respond to this uncertainty by adjusting their strategies, resulting in heightened volatility and changes in options positions.

However, this policy uncertainty can be resolved through the actual FOMC announcements or the subsequent release of meeting minutes. Clarity is provided to the market, enabling investors to make informed decisions based on the revealed policy decisions. The resolution of policy uncertainty triggers consequential market reactions, leading to adjustments in options pricing and stock market behavior. Market participants reassess their positions and adapt their investment strategies based on the new information.

Understanding these differential effects and the nature of policy uncertainty is crucial for traders and investors to navigate the intricacies of FOMC cycles effectively and make well-informed decisions aligned with their risk appetite and investment objectives.

2.6 Directional Trading Before the Meeting

In this section, we investigate whether option traders trade on the predictable movements of FOMC policies. We decompose monetary policy shocks into an expected component and a surprise component, following Bernanke and Kuttner (2005). Let us consider a meeting taking place on day d of month m, and denote by $f^0_{m,d}$ the current-month futures rate implied federal futures funds. The one-day surprise in target rates is calculated as the difference between the rate implied by the contract expiring on the day of the meeting and the implied by the contract expiring on the day before the meeting :[16]

$$\Delta r_d^{unexpected} = \frac{D}{D-d}(f^0_{m,d} - f^0_{m,d-1}) \tag{2.8}$$

where D is the number of days in the month.

The expected component of the change in target rates is defined as the actual change minus the unexpected component:

$$\Delta r_d^{expected} = \Delta r_d^{total} - \Delta r_d^{unexpected} \tag{2.9}$$

16. These contracts are referred to as "30 Day Federal Funds Futures" and traded on the Chicago Board of Trade. They contain information on expectations of the effective Federal funds rate, averaged over the settlement. The implied futures rate is 100 minus the contract price.

We regress the total change in target rates, and its expected and unexpected individual components on trading and pricing variables:

$$mps_d = \alpha + \beta OptionVar_{d-1} + \epsilon \tag{2.10}$$

where *mps* denotes the monetary policy shock considered (total, expected or unexpected). $OptionVar_{d-1}$ denotes the average option-based variable between 15 days and 1 day before the event.[17]

Our results are reported in Table 2.11. Two interesting observations can be made. First, both components of the IV are negatively related to change in the Fed fund rate, and this relation is statistically significant. It is not surprising that expected decline in fund rates increase option prices, but the fact that the relation also holds for the unexpected change is in line with the presence of informed traders. These traders would have superior information prior to the meetings, and trade on this information, making the price of options rise. Second, there is a significant negative relationship between the RN-skewness and the total and expected component of monetary policy shocks. We also find a significant positive relationship between the put-call ratio and these shocks. These relations show that an increase in expected tail risk and in the purchase of put options relative to call options predict an increase in the Fed fund rate. This link is not surprising as Bernanke and Kuttner (2005) finds that an increase in the fed fund leads to a decrease in 30-day index returns. Interestingly, our results do not highlight any link between unexpected rate changes and tail risk.

David and Veronesi (2014)'s model can also explain our results. This model discusses that in good times, investors perceive greater downside risk in stocks than in bad times- i.e., positive fundamental news has little effect on revising investors' views about a boom regime, however, negative fundamental news may lead to a substantial downward revision of such beliefs. A significant downward revision of views in a booming regime would lead to considerably lower stock prices. Therefore, in good times, stock returns incline to be negatively skewed. The negatively skewed return distribution increases the price

17. To calculate the put-call ratios we focus our analysis on mid-term options with maturities between 25-45 days. This ensures that we have exploited the information embedded in options covering the events, and also confirms that our results from put-call ratios are comparable with RN moments obtained from constant maturity 30-days options in volatility surface file of Optionsmetrics.

of OTM put options relative to OTM call options. When the tail-risk index rises, it indicates that the economy is shifting to a booming regime with regular capacity utilisation and inflation. Consequently, the forward-looking Taylor rule anticipates tightening monetary policy back to normal levels and hence lead to higher future interest rates.

[Table 2.11 about here.]

We also found a negative relationship between implied volatilities and real and unexpected components of monetary policy shocks. This finding also suggests that the monetary authority reacts to a volatile stock market by lowering interest rates; therefore, investors expect expansionary lower Fed rates when the volatility is high. The David and Veronesi (2014) equilibrium model can explain these findings. This model suggests that when investors are uncertain about the current regime, ATMIV increases. In this situation, investors update their beliefs more rapidly, resulting in higher return volatility. During such times, expected economic growth is lower than during good times and is represented by lower expected future capacity utilisation. These views persuade the central bank to react through the forward-looking Taylor rule. Therefore, central banks decrease the target real interest rates to sustain economic stability. This explains the reason why an increase in the ATMIV is related to reduced future rates.

2.7 Role of Economic Conditions and Political Uncertainty

Pastor and Veronesi (2012) find option-implied risk measures react more strongly to political events (e.g., elections) in weak economies due to the added uncertainty. In this Section, we analyse whether our results depend on the economic conditions at the time of FOMC meetings.

We use two measures of economic conditions in the United States. First, the Composite Leading Indicator (CLI) as is a monthly index published by the Conference Board, which presents signals of turning points in business cycles and aims to predict the direction of global economic movements. We obtain monthly CLI data from the OECD website from 2006 to 2016. We standardise this measure and denote the resulting variable by *Econ*. Second, the Economic Policy Uncertainty (*EPU*) index of Baker et al. (2013) measures economic uncertainty using text analysis methods.

We regress the option-based variables (volume, open-interest, Open net-demand, ATM implied volatility, VRP and risk-neutral skewness) in the week before the event (week -1 in the cycle) on each of these indexes, $Econ$ and EPU, at the time of the event:

$$OptionVar_t = \alpha + \beta_1 Econ_t + \beta_2 EPU_t + \epsilon_t \tag{2.11}$$

As shown in Table 2.12, there is a significant negative relationship between the CLI and both the implied volatility and the variance risk premium (t-statistics of -15.12 and -5.36, respectively). A higher volatility and price of volatility preceding the meetings therefore leads to worse economic conditions on the day of the meeting. In line with this, we find a significant positive relationship between the CLI the RNS, with a t-statistic of 2.38. This positive relationship also suggests that a lower RNS leads to worse economic conditions. When focusing on policy uncertainty, we find a significant positive relationship between the EPU and the implied volatility, with a t-statistic of 6.11. Additionally, the bid-ask spread shows a positive relationship with uncertainty (EPU) and a negative relationship with economic conditions (CLI), with t-statistics of 3.06 and -6.86, respectively.

[Table 2.12 about here.]

We argue that stock overpricing is the source of this relationship. In weak economic conditions, stock overpricing may not be quickly corrected in the underlying market due to limits to arbitrage. Therefore, investors may incline to the options market to trade on their negative news or beliefs. To maximise their leverage on their negative beliefs, they sell (buy) OTM calls (puts) instead of shorting the stock. The purchase of OTM is attractive compared to directly selling the underlying stock because the options market does not expose to the potential risk that holding the stock involves. In turn, The risk-averse market makers would need to sell the underlying stock and get exposed to the downside and inventory risk. As a result, they induce an upward (downward) price pressure on OTM puts (calls) to compensate for their risk. This mechanism is consistent with the demand-based option pricing framework of Garleanu et al. (2009) and renders relatively more (less) OTM puts (calls) expensive, resulting in a lower RNS value. Subsequently, a relatively low RNS value is followed by stock underperformance if market participants perceive this options trading activity as an informative signal correct the stock overpricing.

Our additional analysis in section 2.D confirms that the inclusion of the SPX return as a control variable does not significantly alter the robustness of our findings, strengthening the evidence of the link between options trading and stock market dynamics in weak economic conditions. These insights shed light on the interplay between options characteristics and market behavior in response to economic uncertainty and provide valuable implications for investors and traders navigating uncertain market conditions.

2.8 Conclusions

Options provide valuable protection against unfavourable outcomes. Many investors also use options for speculation purposes. We analyse the value of option protection against risks associated with uncertainty in FOMC events. Furthermore, we investigate the changes in the trading behaviour of traders and the value of option protection against three types of risk corresponding to these political events: price risk, variance risk and tail risk.

We find that policy uncertainty increases option prices and therefore implied volatility of options for policy events.

Our results suggest that in the case of monetary policy news and announcements (FOMC meetings), trading activity decreases from 6 days before the event and shifts more towards put options. This happens with an increase in price, variance and tail risk of option prices, so option prices become more left-skewed and more expensive in weeks around the event. Similarly, we found that the put-call ratio increases one week before the meeting.

We also examine how option variables can predict monetary policy shocks. We find that a positive shock to downside-risk predicts an increase in the Fed rate and implies that when investors become more worried about a stock market decline, they predict that future short-term rates to increase. In addition, we found a negative relationship between implied volatilities and monetary policy shocks. These findings suggest that the monetary authority reacts to a volatile stock market by lowering interest rates.

Therefore when the volatility is high, investors expect an expansionary lower Fed rate.

Finally, we analyse the role of economic conditions in the value of option protection against risks associated with political events. We showed that our findings are more pronounced in weaker economic conditions.

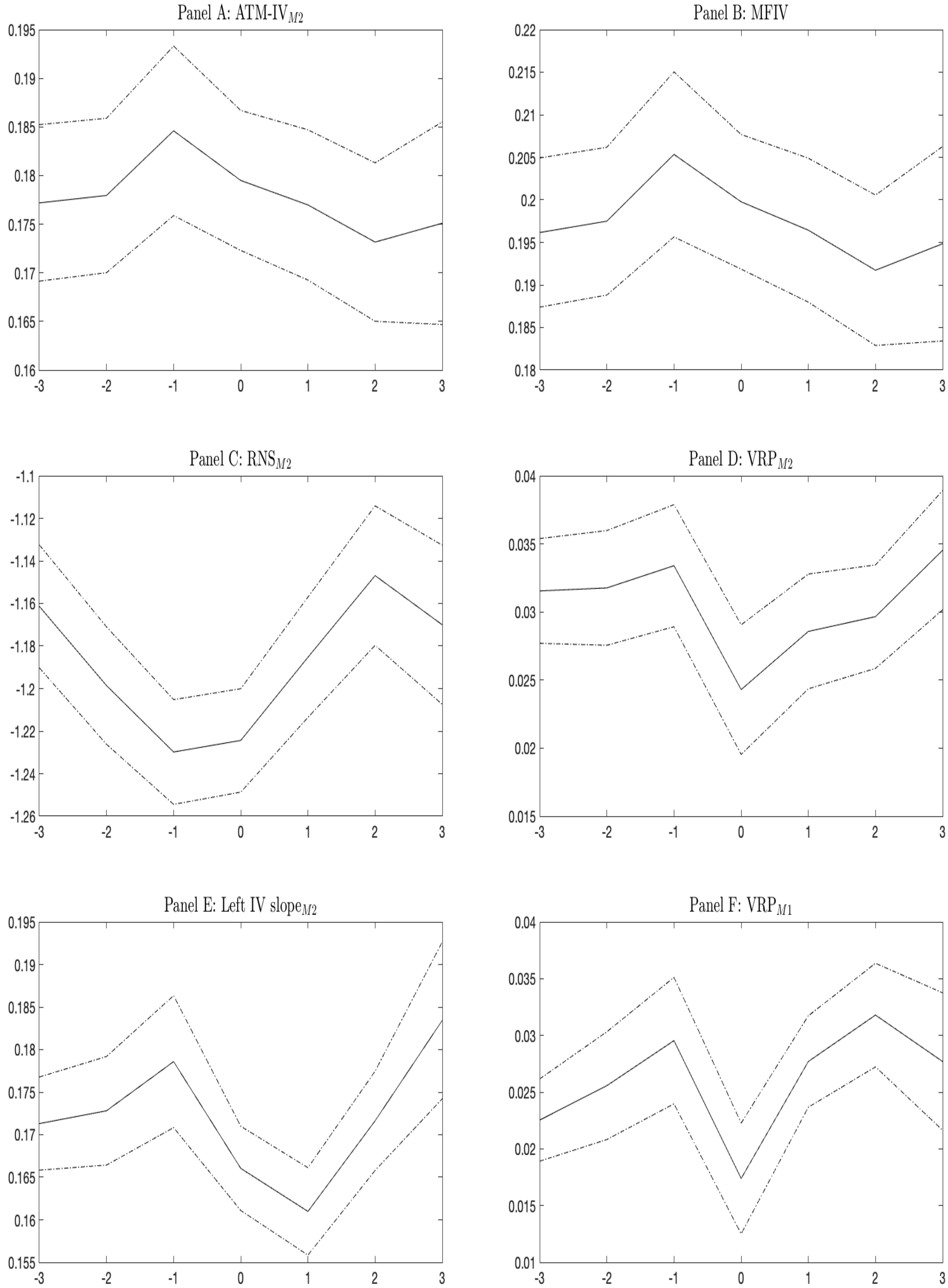


Figure 2.A.1. Option moments and premia over the FOMC cycle: This figure represents the fluctuations of pricing indicators over the FOMC cycle for all options traded between 2006 and 2016. The period covers 82 FOMC cycles. Dashed lines represent the 5% and 95% confidence intervals. Subscript $M1$ and $M2$ are explained in the Appendix 2.A.2.

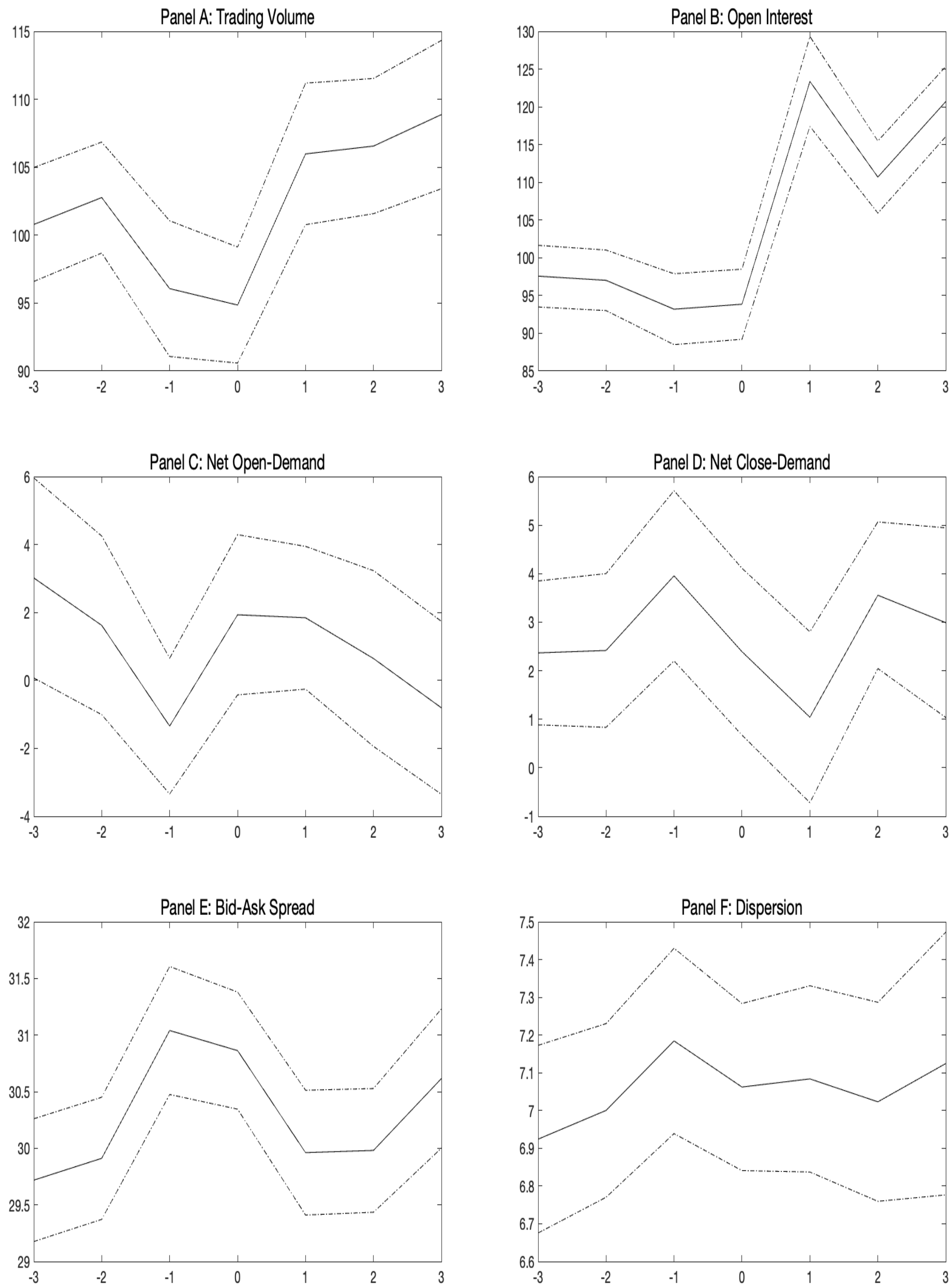


Figure 2.A.2. ALL Option liquidity over the FOMC cycle: This figure represents the fluctuations of liquidity indicators over the FOMC cycle, for all options traded between 2006 and 2016. The period covers 82 FOMC cycles. Dashed lines represent the 5% and 95% empirical confidence intervals.

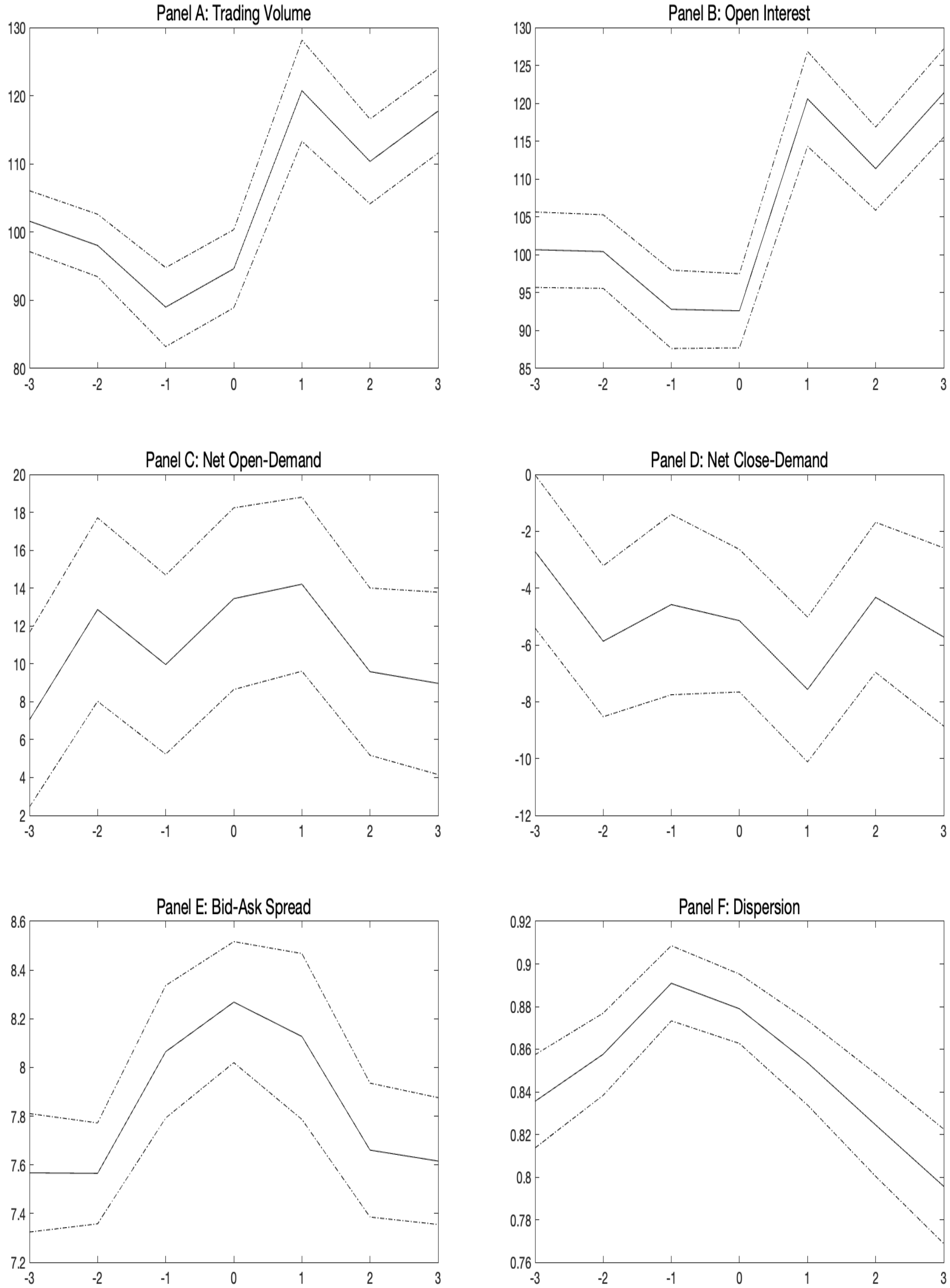


Figure 2.A.3. ATM Option liquidity over the FOMC cycle: This figure represents the fluctuations of liquidity indicators over the FOMC cycle, for ATM options traded between 2006 and 2016. The period covers 82 FOMC cycles. Dashed lines represent the 5% and 95% empirical confidence intervals.

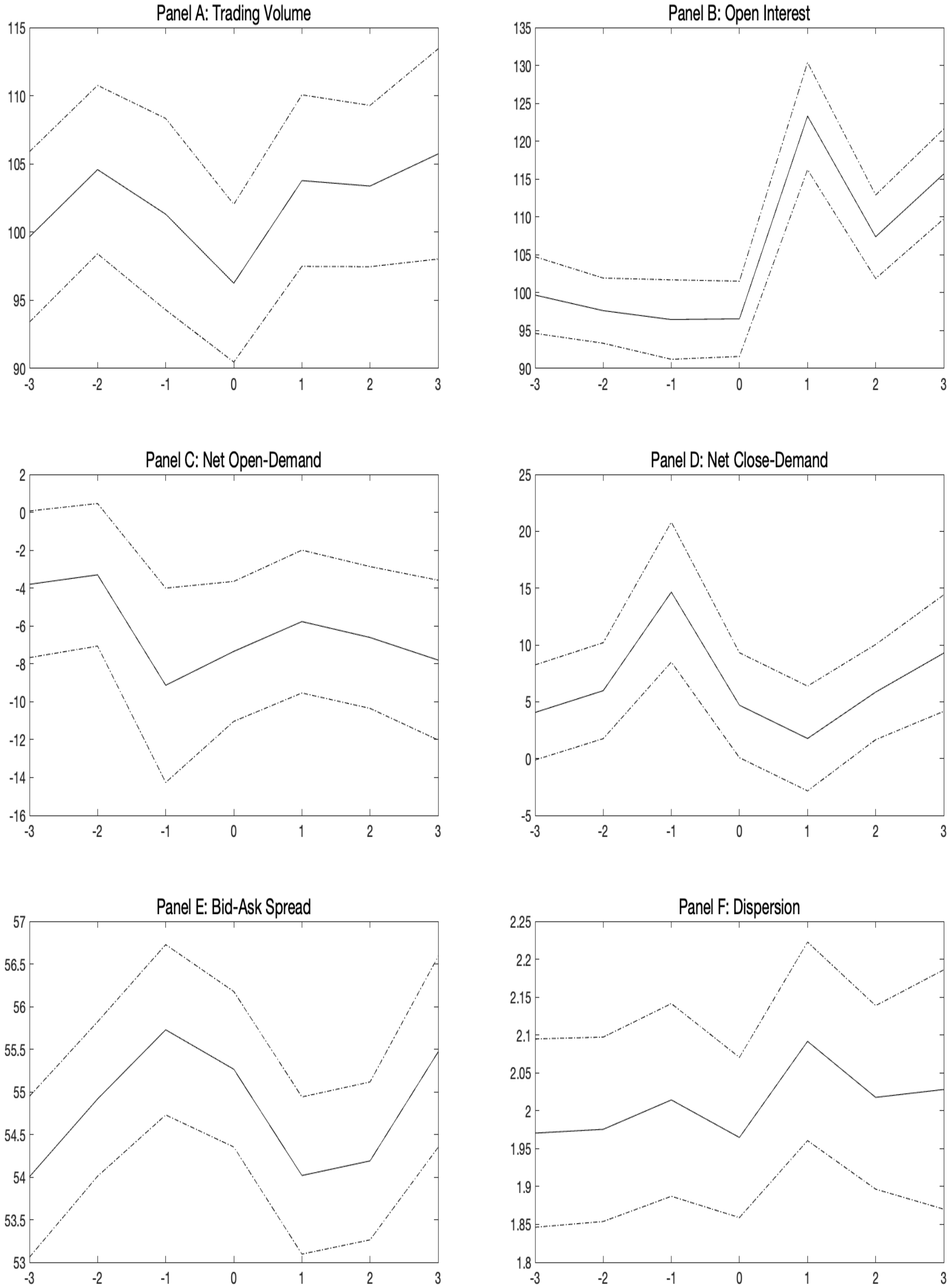


Figure 2.A.4. OTMC Option liquidity over the FOMC cycle: This figure represents the fluctuations of liquidity indicators over the FOMC cycle, for OTMC options traded between 2006 and 2016. The period covers 82 FOMC cycles. Dashed lines represent the 5% and 95% empirical confidence intervals.

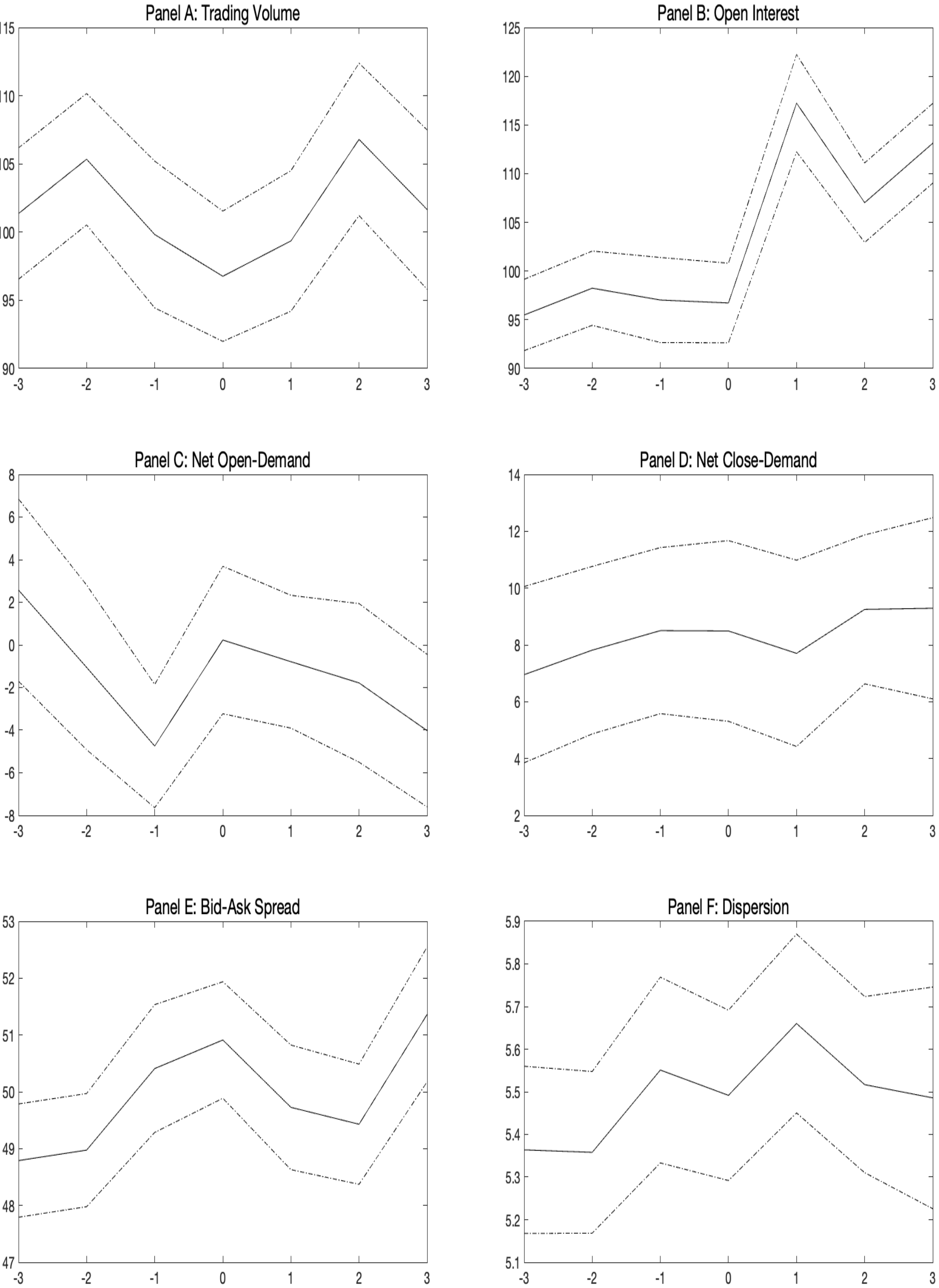


Figure 2.A.5. OTMP Option liquidity over the FOMC cycle: This figure represents the fluctuations of liquidity indicators over the FOMC cycle, for OTMP options traded between 2006 and 2016. The period covers 82 FOMC cycles. Dashed lines represent the 5% and 95% empirical confidence intervals.

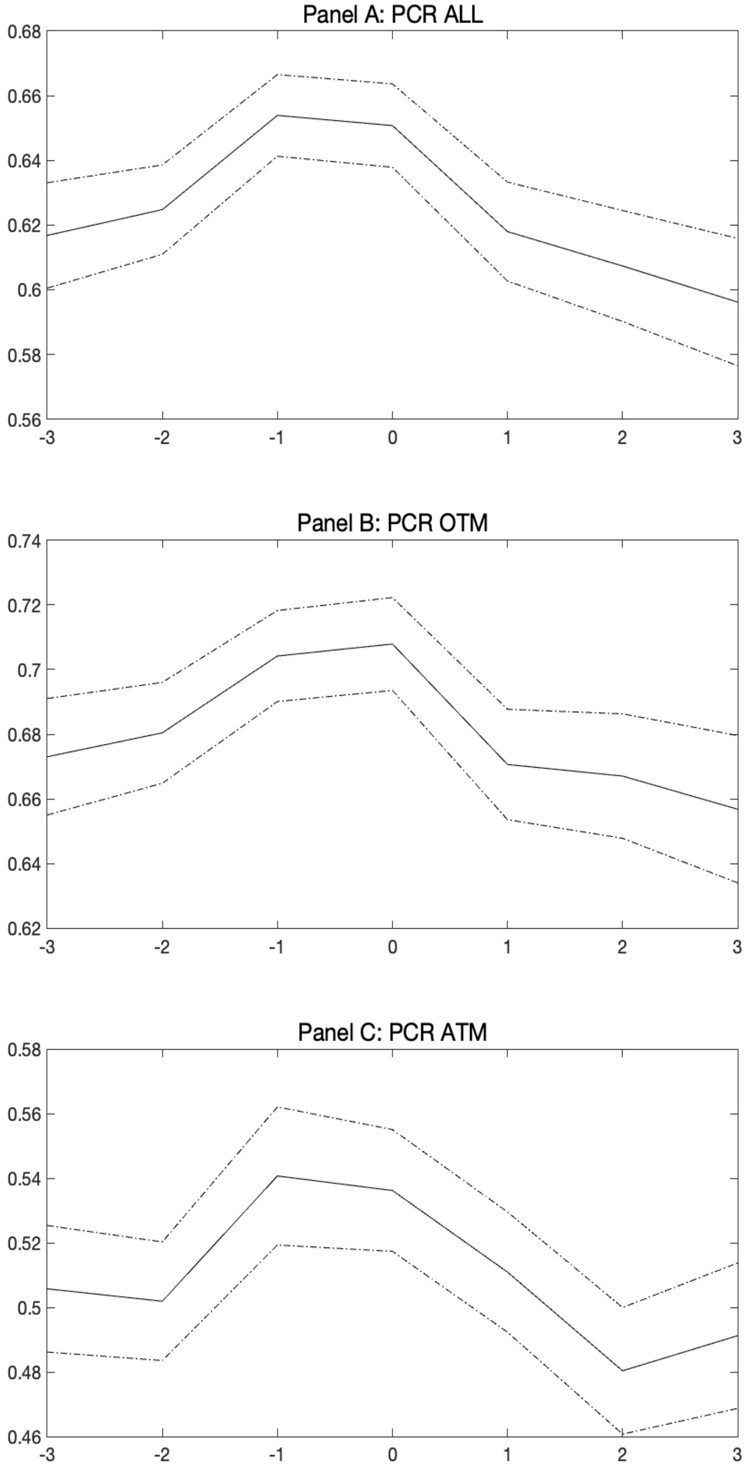


Figure 2.A.6. Put-call ratio over the FOMC cycle: This figure represents the fluctuations of put-call ratio over the FOMC cycle, for all options traded between 2006 and 2016. The period covers 82 FOMC cycles. Dashed lines represent the 5% and 95% empirical confidence intervals.

	Moneyness	Number of Obs.
ALL options	$0 <= K/S < 2$	1,083,353
ATM Options	$0.98 <= K/S < 1.02$	175,173
OTM puts	$0 <= K/S <= 0.98$	473,497
OTM calls	$1.02 <= K/S <= 2$	152,272

Table 2.1. Sample Data: This table reports the number of options in our data set for each bucket of moneyness K/S, where K denotes an option's strike and S is the underlying spot price.

	$ATM - IV_{M1}$	VRP$_{M1}$	RNS_{M1}	left IVS_{M1}	right IVS_{M1}
β_1	0.90***	0.45**	2.39	0.62***	0.66***
t-stat	(3.55)	(2.72)	(0.47)	(3.35)	(5.75)

Table 2.2. Quantifying the Effect of FOMC Meetings on Risk-Neutral Moments and Risk Premia: This table displays the estimated β_1 coefficients in regression (2.4):

$$OptionVar_{t,\tau} = \beta_0 + \beta_1.D_{t,\tau}^{FOMC} + \beta_2\tau_i + \epsilon_{t,\tau}$$

The dependent variables ($OptionVar$) are the ATM implied volatility (ATM IV) and the variance risk premium (VRP$_{M1}$), RNS, the left and right implied volatility slopes (left/right IVS),. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

		Vol	OI	BAS	OND	CND	DISP
ALL	β_1	-0.0319***	-0.0061**	0.0131***	-0.0033	0.0048	8.88
		(-13.65)	(-2.536)	(6.064)	(-1.139)	(1.618)	(1.33)
ATM	β_1	-0.0457***	-0.0016	-0.0142***	-0.0137**	0.0107*	-1.14**
		(-7.533)	(-0.2545)	(-2.737)	(-2.198)	(1.710)	(-1.96)
OTMC	β_1	-0.0366***	-0.0418***	0.0388***	0.0059	0.0032	8.00**
		(-6.009)	(-6.522)	(6.649)	(0.8830)	(0.4795)	(2.41)
OTMP	β_1	-0.0268***	-0.0063*	0.0380***	-0.0016	0.0030	-4.24
		(-7.543)	(-1.730)	(11.55)	(-0.3700)	(0.6840)	(-0.68)

Table 2.3. Quantifying the Effect of FOMC Meetings on Trading Activity Variables: This table displays the estimated β_1 coefficients. For the first five variables, the estimated β_1s come from the (2.5):

$$OptionVar_i = \beta_0 + \beta_1.D_i^{FOMC} + \beta_2\tau_i + \epsilon_i$$

The dependent variables ($OptionVar$) are the volume (VOL), open-interest (OI), bid ask spread (BAS), open net demand (OND) and close net demand (CND). D_i^{FOMC} is the dummy that if there is an FOMC meeting between quotation date and expiration date. τ_i is the days-to-maturity. For the dispersion in investors' beliefs (DISP), the estimated β_1 comes from the (2.4):

$$DISP_{t,\tau} = \beta_0 + \beta_1.D_{t,\tau}^{FOMC} + \beta_2\tau + \epsilon_{t,\tau}$$

where $D_{t,\tau}^{FOMC}$ takes the value 1 if there is an FOMC meeting between times t and $t+\tau$, τ is the days-to-maturity. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variables:	pcr.all	pcr.atm	pcr.otm
	0.0886***	0.0730***	0.0831***
	(9.580)	(11.01)	(8.163)

vcovHC co-variance matrix, t-stats in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 2.4. Quantifying the Effect of FOMC Meetings on Put-Call Ratio Measures: This table displays the estimated β_1 coefficients in the following regression:

$$OptionVar_{t,i} = \beta_0 + \beta_1.D_{t,i}^{FOMC} + \epsilon_{t,i}$$

Each date t, we divide options with different days-to-maturity into two groups. The treatment group contains the options whose remaining lives span at least one FOMC meetings. The control group contains the other options in date t. The dependent variables ($OptionVar$) are the put-call ratios of the two groups of options at each date t. $D_{t,i}^{FOMC}$ equals 1 for the treatment group and 0 for the control group. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependant Variable		M1	M2	M3
ATM.IV	week -1	0.0080^{***}		0.0086^{***}
		(2.667)		(2.803)
	week 0		0.0017	0.0033
			(0.5821)	(1.086)
MFIV	week -1	0.0093^{***}		0.0101^{***}
		(2.681)		(2.860)
	week 0		0.0028	0.0046
			(0.8171)	(1.337)
VRP	week -1	0.0029^{*}		0.0027
		(1.736)		(1.599)
	week 0		-0.0015	-0.0011
			(-0.8564)	(-0.5737)
RNS	week -1	-0.0413^{***}		-0.0475^{***}
		(-2.988)		(-3.354)
	week 0		-0.0255^{*}	-0.0341^{**}
			(-1.712)	(-2.243)
RIVS	week -1	0.0091^{***}		0.0099^{***}
		(2.736)		(2.920)
	week 0		0.0027	0.0045
			(0.8432)	(1.368)
LIVS	week -1	0.0095^{***}		0.0107^{***}
		(3.269)		(3.612)
	week 0		0.0048	0.0067^{**}
			(1.616)	(2.226)

Table 2.5. Regressions of Pricing Moments on Week Dummies: This table displays the estimated coefficients β_1 and β_2 of the following regression run on daily data:

$$OptionVar_{i,t} = \alpha + \beta_1.1\{t \in \text{week -1 }\} + \beta_2.1\{t \in \text{week 0}\} + \epsilon_t$$

where $OptionVar$ is the pricing indicator obtained from 30-days constant maturity options, namely $(ATM-IV_{M2})$, the risk-neutral skewness (RNS_{M2}), left and right implied volatility slopes $(L(R)IVS_{M2})$, and the variance risk premium (VRP_{M2}) obtained from model 2 as discussed in Appendix 2.A.1. These measures are computed using volatility surface file in Optionsmetrics using options with constant time to maturity of 30-days. In regression 1 (M1), the coefficient β_2 is set to 0, in regression 2 (M2), β_1 is set to 0, in regression 3 (R3) they are both estimated. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method. ***/**/* indicate significance at the 1/5/10% levels.

		M1	M2	M3
ATM-IV$_{M1}$	β_1	0.75		0.68
		(0.98)		(0.88)
	β_2		0.52	0.43
			(-0.73)	(-0.59)
VRP$_{M1}$	β_1	0.42		0.27
		(0.93)		(0.59)
	β_2		-0.99**	-0.95**
			(-2.35)	(-2.23)
RNS$_{M1}$	β_1	-6.44		-8.31*
		(-1.43)		(-1.84)
	β_2		-10.63**	-11.80***
			(-2.52)	(-2.77)
LIVS$_{M1}$	β_1	0.30		0.23
		(0.83)		(0.64)
	β_2		-0.45	-0.42
			(-1.32)	(-1.21)
RIVS$_{M1}$	β_1	0.76		0.67
		(1.27)		(1.10)
	β_2		-0.68	-0.59
			(-1.22)	(-1.04)

Table 2.6. Regressions of Pricing Moments on Week Dummies - Robustness Check: This table displays the estimated coefficients β_1 and β_2 of the following regression run on daily data:

$$OptionVar_{i,t} = \alpha + \beta_1.1\{t \in \text{week -1 }\} + \beta_2.1\{t \in \text{week 0}\} + \epsilon_t$$

where $OptionVar$ is the left slope of IV smile (LIVS$_{M1}$), the right slope of IV smile (RIVS$_{M1}$) or the variance risk premium (VRP$_{M1}$), obtained from model 1, as discussed in Appendix 2.A.1. These measures are calculated using option-data file in Optionsmetrics- i.e., using options quoted on each day over a given window of time. In regression 1 (R1), the coefficient β_1 is set to 0, in regression 2 (R2), β_2 is set to 0, in regression 3 (R3) they are both estimated. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method. ***/**/* indicate significance at the 1/5/10% levels.

Moneyness	Dependent Variables: / Model:	Vol M1	Vol M2	Vol M3	OI M1	OI M2	OI M3
ATM	week1	-13.53***		-14.60***	-9.376***		-11.29***
		(-4.125)		(-4.369)	(-2.902)		(-3.432)
	week0		-2.915	-5.695		-8.027**	-10.18***
		(-0.8081)	(-1.548)			(-2.505)	(-3.113)
OTMC	week1	0.3327		0.2493	-7.133**		-8.212***
		(0.0860)		(0.0632)	(-2.305)		(-2.594)
	week0		-0.4904	-0.4430		-4.172	-5.736*
			(-0.1283)	(-0.1136)		(-1.275)	(-1.713)
OTMP	week1	0.4181		0.5000	-4.627*		-5.455**
		(0.1368)		(0.1605)	(-1.793)		(-2.077)
	week0		0.3400	0.4353		-3.358	-4.397
			(0.1101)	(0.1383)		(-1.230)	(-1.581)
ALL	week1	-4.281		-4.770*	-9.923***		-11.61***
		(-1.566)		(-1.711)	(-3.563)		(-4.099)
	week0		-1.693	-2.602		-6.765**	-8.978***
			(-0.6332)	(-0.9538)		(-2.244)	(-2.922)

Table 2.7. Regressions of Unsigned Liquidity Indicators on Week Dummies: This table displays the estimated coefficients β_1 and β_2 of the following regression run on daily data:

$$OptionVar_{i,t} = \alpha + \beta_1.1\{t \in \text{week -1 }\} + \beta_2.1\{t \in \text{week } 0\} + \beta_3 * \text{NoAM} + \beta_4 * \text{DTM} + \epsilon_t$$

where $OptionVar$ is the trading volume (VOL) and open-interest (OI). NoAM is the number of available maturities on each day. DTM is the average days-to-maturity of each day. In regression 1 (M1), the coefficient β_1 is set to 0, in regression 2 (M2), β_2 is set to 0, in regression 3 (M3) they are both estimated. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method. ***/**/* indicate significance at the 1/5/10% levels.

Moneyness	Dependent Variables: Model:	Spread M1	M2	M3	dispersion M1	M2	M3
	Variables						
ATM	week -1	-0.0075		-0.0118	0.0465^{***}		0.0630^{***}
		(-0.6053)		(-0.9297)	(4.038)		(5.309)
	week 0		-0.0177	-0.0203		0.0759^{***}	0.0879^{***}
			(-1.341)	(-1.494)		(6.522)	(7.339)
OTMC	week -1	-0.0403		-0.0714	0.1046		0.1287^{*}
		(-0.5356)		(-0.9208)	(1.475)		(1.775)
	week 0		-0.1296^{*}	-0.1454^{*}		0.1037	0.1282^{*}
			(-1.663)	(-1.811)		(1.563)	(1.889)
OTMP	week -1	-0.0063		-0.0143	0.2336^{*}		0.3156^{**}
		(-0.2616)		(-0.5825)	(1.840)		(2.452)
	week 0		-0.0344	-0.0376		0.3757^{***}	0.4358^{***}
			(-1.399)	(-1.493)		(2.887)	(3.304)
ALL	week -1	0.0036		0.0032*	0.3688^{**}		0.4712^{***}
		(0.8022)		(1.6850)	(2.565)		(3.210)
	week 0		-0.0027	-0.0020		0.4541^{***}	0.5439^{***}
			(-0.6046)	(-0.4379)		(3.091)	(3.628)

vcovHC co-variance matrix, t-stats in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 2.8. Regressions of Liquidity Indicators on Week Dummies: This table displays the estimated coefficients β_1 and β_2 of the following regression run on daily data:

$$OptionVar_{i,t} = \alpha + \beta_1.1\{t \in \text{week -1 }\} + \beta_2.1\{t \in \text{week } 0\} + \beta_3 * \text{NoAM} + \beta_4 * \text{DTM} + \epsilon_t$$

where $OptionVar$ is the bid-ask spread (BAS) and dispersion measure (DISPOI). NoAM is the number of available maturities on each day. DTM is the average days-to-maturity of each day. In regression 1 (M1), the coefficient β_2 is set to 0, in regression 2 (M2), β_1 is set to 0, in regression 3 (M3) they are both estimated. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method.***/**/* indicate significance at the 1/5/10% levels.

Moneyness	Dependent Variables: Model:	DC M1	M2	M3	DO M1	M2	M3
Variables							
ATM	week -1	0.5258		0.4243	-2.496		-1.924
		(0.3009)		(0.2389)	(-0.9496)		(-0.7200)
	week 0		-0.6208	-0.5400		3.407	3.041
			(-0.3652)	(-0.3126)		(1.079)	(0.9464)
OTMC	week -1	8.559***		8.580**	-1.005		-1.452
		(2.589)		(2.568)	(-0.3606)		(-0.5135)
	week 0		-1.521	0.1125		-2.096	-2.373
			(-0.5076)	(0.0372)		(-0.8502)	(-0.9502)
OTMP	week -1	0.8521		0.8599	-2.443		-2.101
		(0.4818)		(0.4777)	(-1.412)		(-1.184)
	week 0		-0.1222	0.0416		2.218	1.818
			(-0.0582)	(0.0194)		(0.9700)	(0.7773)
Variables							
ALL	week -1	1.879*		1.859*	-2.178*		-1.988
		(1.881)		(1.834)	(-1.848)		(-1.643)
	week 0		-0.4582	-0.1040		1.386	1.007
			(-0.4094)	(-0.0915)		(0.8943)	(0.6349)

vcovHC co-variance matrix, t-stats in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 2.9. Regressions of Signed Liquidity Indicators on Week Dummies: This table displays the estimated coefficients β_1 and β_2 of the following regression run on daily data:

$$OptionVar_{i,t} = \alpha + \beta_1.1\{t \in \text{week -1 }\} + \beta_2.1\{t \in \text{week } 0\} + \beta_3 * \text{NoAM} + \beta_4 * \text{DTM} + \epsilon_t$$

where $OptionVar$ is the open net-demand (OND) and close net-demand (CND). NoAM is the number of available maturities on each day. DTM is the average days-to-maturity of each day. In regression 1 (M1), the coefficient β_2 is set to 0, in regression 2 (M2), β_1 is set to 0, in regression 3 (M3) they are both estimated. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method. ***/**/* indicate significance at the 1/5/10% levels.

	ATM.PCR			OTM.PCR			ALL.PCR		
Model:	M1	M2	M3	M1	M2	M3	M1	M2	M3
week1	0.026^{**}		0.033^{***}	0.024^{***}		0.031^{***}	0.029^{***}		0.0367^{***}
	(2.178)		(2.678)	(2.680)		(3.401)	(3.746)		(4.491)
week0		0.031^{**}	0.036^{***}		0.0354^{***}	0.041^{***}		0.032^{***}	0.039^{***}
		(2.431)	(2.853)		(3.712)	(4.184)		(3.709)	(4.335)

Table 2.10. Regressions of Put-Call Ratio on Week Dummies: This table displays the estimated coefficients β_1 and β_2 of the following regression run on daily data:

$$OptionVar_{i,t} = \alpha + \beta_1.1\{t \in \text{week -1 }\} + \beta_2.1\{t \in \text{week 0}\} + \beta_3 * \text{NoAM} + \epsilon_t$$

where $OptionVar$ is the put-call ratio (PCR) across different level of moneyness. NoAM is the number of available maturities on each day. DTM is the average days-to-maturity of each day. In regression 1 (M1), the coefficient β_2 is set to 0, in regression 2 (M2), β_1 is set to 0, in regression 3 (M3) they are both estimated. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method. ***/**/* indicate significance at the 1/5/10% levels.

Panel A: signed put-call ratios									
	PCR-ALL			PCR-OTM			PCR-ATM		
	total	unexp	exp	total	unexp	exp	total	unexp	exp
β	0.27*	-0.05	0.32**	0.30**	-0.04	0.34**	-0.02	-0.05	0.04
t-stat	(1.86)	(-1.24)	(2.07)	(2.38)	(-1.18)	(2.54)	(-0.13)	(-1.46)	(0.26)
p-value	0.07	0.22	0.04	0.02	0.24	0.01	0.90	0.15	0.79

Panel B: RN moments									
	IV			VRP			SKEW		
	total	unexp	exp	total	unexp	exp	total	unexp	exp
β	-0.39***	-0.10**	-0.29*	0.23	0.33***	-0.10	-0.13**	-0.01	-0.12**
t-stat	(-2.64)	(-2.36)	(-1.77)	(0.66)	(3.59)	(-0.27)	(-2.61)	(-0.54)	(-2.25)
p-value	0.01	0.02	0.08	0.51	0.00	0.79	0.01	0.59	0.03

Table 2.11. Do Option-Implied Moments and Liquidity Predict the Change in Fed fund Rate? This table shows the coefficient estimates of regression model:
$mps_\tau = \alpha + \beta$ Option $\text{Var}_{\tau-1} + \epsilon$
where the dependent variable is one of the components of monetary policy shocks (expected, unexpected and real) and the dependant variable is the average of put-call ratio (Panel A) or option implied moments (Panel B) on days before the meeting.
The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method. In this table, ***/**/* indicates significance at the 1/5/10% level

	VOL	OI	BAS	OND	IV	VRP	RNS
β_1	0.018	-0.008	-0.341***	-0.104**	-0.62***	-0.30***	0.14**
t-stat	(0.34)	(-1.15)	(-6.86)	(-1.91)	(-15.12)	(-5.36)	(2.38)
p-value	0.73	0.87	0.00	0.05	0.00	0.00	0.02
β_2	-0.018	-0.117**	0.1523**	0.017	0.25***	0.07	0.06
t-stat	(-0.32)	(-2.15)	(3.06)	(0.32)	(6.11)	(1.23)	(1.06)
p-value	0.74	0.03	0.00	0.74	0.00	0.22	0.29

Table 2.12. The Effect of Economic Policy Uncertainty (EPU) and Economic Condition (CLI) on Trading Factors and Pricing Moments: The table shows the slope coefficients, β_1 and β_2, of regression model:
$OptionVar_t = \alpha + \beta_1 CLI_t + \beta_2 EPU_t + \epsilon_t$
where $OptionVar$ is trading factors (columns 1-4) and implied option moments (columns 5-7) on week -1 of the FOMC cycle. We regress option-based variables on the measures of economic condition (CLI) and economic policy uncertainty index. CLI is standardized to zero mean using data from 2006 to 2016. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method.

Appendix - Chapter Two

2.A Variable Construction

2.A.1 Trading Activity Variables

2.A.1.1 Liquidity Indicators

Regarding general unsigned liquidity indicators, we focus on total volume and open interest of options. In addition, we define bid-ask spread as the difference between option trade prices scaled by their mid-price.

In addition, we focus on end-users open net demand and close net demand as measures of signed trading activities. We follow Garleanu et al. (2009) and build our signed trading indicators as follows. We obtain the data from the Chicago Board Options Exchange (CBOE). These data consist of a daily record of opening (closing) short and long open interest on all SPX for public customers and firm proprietary traders from the beginning of 2006 to the end of 2016. We compute the net demand of each of these groups of agents as the open (close) long open interest minus the short open interest. We focus our analysis on non-market-maker net demand, defined as the sum of the net demand of public customers and proprietary traders, which is equal to the negative of the market-maker net demand (since options are in zero net supply). Hence, we assume that public customers and firm proprietary traders—that is, all non-market-makers—are “end-users.

2.A.1.2 Put-Call Ratio

The signed put-call ratio of Cremers et al. (2022) is defined as:

$$PC_t^{OB} = OpenBuyPut_t/(OpenBuyPut_t + OpenBuyCall_t) \tag{2.A.1}$$

where, on date t, $OpenBuyPut_t$ and $OpenBuyCall_t$ are the numbers of put and call contracts purchased by non-market makers to open new positions. We defined the open-sell, close-buy and close-sell put-call ratios in the same manner.

2.A.2 Risk-neutral Moments and Risk Premiums

2.A.2.1 ATM implied Volatility (IV)

We compute two alternatives of ATM-IV. The first measure is constructed following Kelly et al. (2016) and averages IVs of ATM options as described in Table 2.1, over a given window of time- i.e., IVs of ATM options quoted on day t and expires on day $t+\tau$. We denote this measure by $ATM - IV_{M1}$.

In addition, we use 30-day Volatility Surface file from OptionMetrics to compute 30-days ATM-IV. We denote this measure by $ATM - IV_{M2}$. This file contains implied volatilities for standardized equity options with 30-day expiry for a grid of the delta space, which is often used as a measure of option moneyness. The implied volatility surface provides a standardized way of measuring the expensiveness of equity options with different strikes and maturities. 30-days ATM-IV measures stock price risk, and it is computed as the average implied volatility of at-the-money call and put options:

$$ATMIV = \frac{ATMIV_{call} + ATMIV_{Put}}{2} \tag{2.A.2}$$

where $ATMIV_{Call}$ is the implied volatility of the 0.5 delta call and $ATMIV_{Put}$ is the implied volatility of the 0.5 delta put.

2.A.2.2 30-day Model-Free IV (MFIV)

The 30-day model-free implied volatility (MFIV) for cycle analysis is estimated following Bakshi et al. (2003) and computed using a portfolio of OTM options. We use 30-day[18] Volatility Surface file from OptionMetrics to compute MFIV. MFIV at time t is defined as:

$$MFIV_t = \sqrt{\frac{e^{r_f(\tau)} M(2)_{t,\tau} - \mu^2_{t,\tau}}{\tau}} \tag{2.A.3}$$

18. $\tau = 30 - days$

where $M(n = 2)$ and μ are given by:

$$M(n)_{t,\tau} = \int_{S_t}^{\infty} \eta(K, S_t, n) C_t(K, t, \tau) dK + \int_{0}^{S_t} \eta(K, S_t, n) P_t(K, t, \tau) dK \tag{2.A.4}$$

$$\eta(K, S_t, n) = \frac{n}{K^2} [(n-1) log(\frac{K}{S_t})^{n-2} - log(\frac{K}{S_t})^{n-1}] \tag{2.A.5}$$

and

$$\mu_{t,\tau} = R_{f,t,\tau} - 1 - R_{f,t,\tau} [M(2)_{t,\tau}/2 + M(3)_{t,\tau}/6 + M(6)_{t,\tau}/24] \tag{2.A.6}$$

$R_{f,t}$ is the riskless rate. S_t denotes the price of the asset at time t and $P_t(K, t, \tau)$ and $C_t(K, t, \tau)$ are the OTM put and call prices. τ is the time to maturity.

2.A.2.3 Implied Volatility Slope

We compute the left and right slopes of the implied volatility smile using OTMP and OTMC options. We regress the implied volatilities of options each day of a given window on their deltas whenever at least three available options exist.

$$IV_{t,\tau} = \alpha + Slope \Delta_{t,\tau} + \epsilon \tag{2.A.7}$$

where $IV_{t,\tau}$ and $\Delta_{t,\tau}$ are the implied volatility of OTMP (OTMC) options respectively at time t with τ being the time to maturity. The coefficient *Slope* denote the left (right) implied volatility slope.

We construct two alternative measures of implied volatility slope. The first measure is constructed using the option data file in Optionsmetrics. We aggregate all options traded each day, t, with a same time-to-maturity, τ, and build $IVS_{t,\tau}$ and denote it by IVS_{M1}. The second measure is calculated using the volatility surface file employing options with constant 30-days to maturity. We denote the second measure by IVS_{M2}. We use the first measure for the difference-in-differences analysis and the second for the cycle analysis.

2.A.2.4 Risk-Neutral Skewness (RNS)

We calculate the risk-neutral skewness for our analysis using a portfolio of OTM options as described by Bakshi et al. (2003):

$$RNS_{t,\tau} = \frac{e^{r_f \tau} M(3)_{t,\tau} - 3e^{r_f \tau} \mu_{t,\tau} M(2)_{t,\tau} + 2\mu_{t,\tau}^3}{[e^{r_f \tau} M(2)_{t,\tau} - \mu_{t,\tau}^2]^{3/2}} \tag{2.A.8}$$

where $M(n)_{t,\tau}(n = 2, 3, 4)$ and μ are given by Equation (2.A.5) and (2.A.6) respectively.

We calculate two alternatives of this measure. The first measure is constructed using S&P500 options traded each day t with an available time-to-maturity τ. We use the option data file of Optionsmetrics to build our first measure. This measure is denoted by RNS_{M1}. The second measure is constructed using constant 30-days time-to-maturity options using the volatility surface file in Optionsmetrics and is denoted by RNS_{M2}.

2.A.2.5 Variance Risk Premium (VRP)

We construct VRP following Kelly et al. (2016), where the variance risk premium is defined as the difference between the implied and realized variances over a given window of time (τ). In our subsequent analysis, we denote this measure by (VRP$_{M1}$). To build this measure, the VRP between time t and $t + \tau$ is defined as the difference between implied variance of options quoted at time t and expires at time $t+\tau$ denoted by $IV_{t,t+\tau}^2$ and realized volatility of the the S&P index over the life of the option (i.e., between times t and $t + \tau$) denoted by $RV_{t,\tau}^2$.

$$VRP_{t,\tau} = IV_{t,t+\tau}^2 - R\hat{V}_{t,t+\tau}^2 \tag{2.A.9}$$

An estimate of the conditional expectation of future realized variance over the life of the option $\hat{RV}_{t,t+\tau}^2$, given the information set available at t, is calculated using the following linear model for the future realized volatility:

$$RV(t, t + \tau) = \alpha + \beta IV(t, t + \tau)^2 + \gamma RV(t - \tau, t) + \epsilon_t. \tag{2.A.10}$$

$IV(t, t + \tau)$ denotes the implied volatility of an ATM option quoted on date t which expires on date $t+\tau$, and $RV(t-\tau,t)$ is the realized variance[19] between $t-\tau$ and t. The coefficients α, β and γ are estimated using Ordinary Least Squares, using all ATM options available in our data. The expectation of future realized variance is then computed as $\hat{RV^2_{t,t+\tau}} = \hat{\alpha} + \hat{\beta} IV(t,t+\tau)^2 + \hat{\gamma} RV(t-\tau,t)$ and is replaced in Equation (2.A.9).

In addition, we construct an alternative measure of VRP, namely the 30-day variance risk premium. This measure is calculated as the difference between the 30-day MFIV and the 30-day historical volatility, where we obtain the historical volatility from the OM Historical Volatility file. In our subsequent analysis, we denote this measure by (VRP_{M2}).

2.B Two-Way Fixed-Effects

In Section 2.4, our primary objective is to compare options whose lifespans span the dates of political events with other options. For this purpose, let $OptionVar_{t,m}$ denote the option variable at time t of an option maturing at time $m > t$. We focus on $OptionVar$ for $t < \tau < m$, where τ represents the date of any such political event. To create control and treatment groups, we define control group options as those whose lifespans span the FOMC meetings, and treatment group options as all other options in our sample whose lifespans do not span the meeting.

In our setup, the treatment is assigned at the individual option level, so the group fixed effect approach suggested by Callaway and Sant'Anna (2021) is not applicable here. This is because the first assumption of their setup, the irreversibility of treatment, is not valid for our empirical setting. When the option does not cover the event at time t, we set the dummy variable equal to 0. The next day, we receive a new set of options with new prices, strikes, and maturity, and the options available on previous dates are no longer followed; they are treated as new options.

However, we do consider the time fixed effect in our analysis. Table 2.A.1 presents the coefficient estimates of regression (2.5), and we observe that our results remain robust.

19. Daily realized variances are collected from the website of the Oxford-Man Institute (http://realized.oxford-man.ox.ac.uk/).

The liquidity of all options covering FOMC meetings decreases: the trading volume and open interest decrease by 1.95% and 13.2%, respectively, for options covering the event, while the bid-ask spread and close net demand increase by 24.3% and 0.46%, respectively. These findings suggest a notable impact of FOMC meetings on option liquidity.

[Table 2.A.1 about here.]

2.C Standard Errors and Clustering

In Chapter 2 of our research, we utilised regression models following White (1980)'s method, which accounts for Heteroskedasticity in the covariance matrix of coefficient estimates within the linear regression model. Consequently, we reported White's heteroskedasticity-consistent (HC) t-statistics.

To ensure the robustness of our findings, we conducted a comprehensive analysis to address potential econometric concerns arising from the time-series properties of the regressors. First, we clustered standard errors by time and reported the t-stats of the regression coefficients for Equation (2.5) in Panel A of table 2.A.2. The results remained consistent, indicating that the time-series properties of the regressors did not significantly impact our main findings. This outcome can be attributed to our treatment group definition, which focuses on individual options and randomization at the individual option level rather than individual-time periods.

In addition, we thoroughly considered the presence of Auto-correlation in the error terms. Panel B of Table 2.A.2 displays the regression coefficients for Equation (2.5) and their t-stats, applying NNewey and West (1994)'s method, which accounts for both heteroscedasticity and autocorrelation of the error term. Importantly, our results concerning trading activities remained robust even after controlling for Auto-correlation of error terms.

Overall, by employing robust regression techniques to address potential econometric concerns such as heteroskedasticity and autocorrelation, we have reinforced the reliability and consistency of our main findings.

[Table 2.A.2 about here.]

2.D Role of Economic Uncertainty and Political uncertainty in the presence of Control Variables

In Section 2.7 we showed that option variables (in terms of RN-moments and trading activities) react more strongly in weak or uncertain conditions. In this section, we analyse the robustness of our results in the presence of control variables. To this end, following Cremers et al. (2022), we consider SP500 index return in the week proceeding the event in the set of our control variables. Therefore, our proposed regression equation is in the form of:

$$OptionVar_t = \alpha + \beta_1 CLI_t + \beta_2 EPU_t + \beta_3 Ret_{t,t-1} + \epsilon_t$$

where, on day t in the week before the event, $Option_v ar_t$ denotes the relative option variable (RN moments or trading indicators), $Ret_{t,t-1}$ is the return of SPX on day t, and CLI_t and EPU_t are the economic condition and uncertainty measure on day t respectively.

As shown in Table 2.A.3, there is a significant negative relationship between the CLI and both the implied volatility and the variance risk premium even after controlling for return on the SPX. A higher volatility and price of volatility preceding the meetings therefore leads to worse economic conditions on the day of the meeting. In line with this, we find a significant positive relationship between the CLI the RNS, with a t-statistic of 2.47. This positive relationship also suggests that a lower RNS leads to worse economic conditions. When focusing on policy uncertainty, we find a significant positive relationship between the EPU and the implied volatility (t-statistic of 1.70). Moreover, we find a positive (negative) relationship between the bid-ask spread and uncertainty (economic condition) in economic conditions, with a t-statistic of 3.06 (-6.88).

We argue that stock overpricing is the source of this relationship. In weak economic conditions, stock overpricing may not be quickly corrected in the underlying market due to limits to arbitrage. Therefore, investors may incline to the options market to trade on their negative news or beliefs. To maximise their leverage on their negative beliefs, they sell (buy) OTM calls (puts) instead of shorting the stock. The purchase of OTM is attractive compared to directly selling the underlying stock because the options market does not expose to the potential risk that holding the stock involves. In turn, The risk-averse market makers would need to sell the underlying stock and get exposed to the downside and inventory risk. As a result, they induce an upward (downward) price pressure on OTM puts (calls) to compensate for their risk. This mechanism is consistent with the demand-based option pricing framework of Garleanu et al. (2009) and renders relatively more (less) OTM puts (calls) expensive, resulting in a lower RNS value. Subsequently, a relatively low RNS value is followed by stock underperformance if market participants perceive this options trading activity as an informative signal correct the stock overpricing.

[Table 2.A.3 about here.]

Dependent Variables:	(Vol)	(OI)	(SP)	(OD)	(CD)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
indicator	-0.0195***	-0.0132	0.0243***	-0.0032	0.0046**
	(-3.104)	(-1.300)	(3.026)	(-1.278)	(2.177)
(dtm)	-0.0297***	-0.0351***	-0.0032	0.0039***	-0.0063***
	(-11.71)	(-7.297)	(-0.8023)	(3.545)	(-5.739)
Fixed-effects					
time	Yes	Yes	Yes	Yes	Yes

Table 2.A.1. Quantifying the Effect of FOMC Meetings on Trading indicators- Controlling for Time Fixed Effect: This table displays the estimated β_1 coefficients. For the first five variables, the estimated β_1s come from the (2.5):

$$OptionVar_i = \beta_0 + \beta_1.D_i^{FOMC} + \beta_2\tau_i + \epsilon_i$$

The dependent variables ($OptionVar$) are the volume (VOL), open-interest (OI), bid ask spread (BAS), open net demand (OND) and close net demand (CND). D_i^{FOMC} is the dummy that if there is an FOMC meeting between quotation date and expiration date. τ_i is the days-to-maturity. For the dispersion in investors' beliefs (DISP), the estimated β_1 comes from the (2.4):

$$DISP_{t,\tau} = \beta_0 + \beta_1.D_{t,\tau}^{FOMC} + \beta_2\tau + \epsilon_{t,\tau}$$

where $D_{t,\tau}^{FOMC}$ takes the value 1 if there is an FOMC meeting between times t and $t + \tau$, τ is the days-to-maturity. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980) method controlling for time fixed effect. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variables:	(Vol)	(OI)	(SP)	(OD)	(CD)
Model:	(1)	(2)	(3)	(4)	(5)
Panel A					
β_1	-0.0319**	-0.0061	0.0131	-0.0033	0.0048**
	(-2.572)	(-0.3009)	(1.533)	(-1.332)	(2.040)
Panel B					
β_1	-0.0319***	-0.0061**	0.0131***	-0.0033	0.0048
	(-13.65)	(-2.536)	(6.064)	(-1.139)	(1.618)

Table 2.A.2. Quantifying the Effect of FOMC Meetings on Trading Activity Variables of All options with Clustering Standard Errors and NewyWest t-stats: This table displays the estimated β_1 coefficients of Equation (2.5). The dependent variables ($OptionVar$) are the volume (VOL), open-interest (OI), bid ask spread (BAS), open net demand (OND) and close net demand (CND). The dummy takes the value equal to one if there is an FOMC meeting between quotation date and expiration date of the option. Panel A reports the NewyWest t-statistics in parantheses. Panel B, reports White's (1980) t-statistics in parantheses, clustering standard errors by time.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Vol	OI	SP	OND	IV	VRP	RNS
$beta_1$	0.02	-0.01	-0.34***	-0.11**	-0.62***	-0.31***	0.13**
t-stat	(0.34)	(-0.16)	(-6.88)	(-2.08)	(-16.89)	(-6.12)	(2.47)
p-value	0.74	0.88	0.00	0.04	0.00	0.00	0.01
$beta_2$	-0.02	-0.12**	0.15***	0.02	0.25***	0.06	0.09*
t-stat	(-0.33)	(-2.15)	(3.06)	(0.35)	(6.79)	(1.13)	(1.70)
p-value	0.74	0.03	0.00	0.72	0.00	0.26	0.09
$beta_3$	-7.62	-1.05	-2.34	-8.70**	-5.78**	-3.95	-24.62***
t-stat	(-1.97)	(-0.27)	(-0.66)	(-2.26)	(-2.23)	(-1.11)	(-6.74)
p-value	0.05	0.79	0.51	0.02	0.03	0.27	0.00

Table 2.A.3. The Effect of Economic Policy Uncertainty (EPU) and Economic Condition (CLI) on Trading Factors and Pricing Moments - Robustness: The table shows the slope coefficients, β_1 and β_2 β_3, of regression model:
$OptionVar_t = \alpha + \beta_1 CLI_t + \beta_2 EPU_t + \beta_3 Ret_{t,t-1} + \epsilon_t$
where $OptionVar$ is trading factors (columns 1-4) and implied option moments (columns 5-7) on week -1 of the FOMC cycle. We regress option-based variables on the measures of economic condition (CLI) and economic policy uncertainty index. CLI is standardized to zero mean using data from 2006 to 2016. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980) method.

Chapter 3

Review of Option-Implied Measures and their Applications

3.1 Introduction

Investors who trade according to their views and information about the future price of an asset may select the options market to exploit the information embedded in their prices. Options are financial derivatives that offer buyers the right, but not an obligation, to sell or buy an underlying asset at an agreed price and date. The price of the options written on a given asset, with different exercise prices and specific time to expiration, indicates the market assessment of the payoff over the series of strike prices.[1]

Black and Scholes (1973) proposed a pricing model to determine the fair price of European options. Reversing the (BS) formula converts option price to a real number called the implied volatility (IV). This mapping of prices to implied volatilities allows for an easier comparison of option prices across various strikes, expiries and underlying assets. BS formula remains popular among researchers and helps them realise the opportunity to obtain option-implied measures from option prices to enhance the forecast of future

1. For example, assume that put options maturing in a week with a strike $K <(80\% \times S_t)$ trade at a very high price (compared to some benchmark, e.g., last week's prices). This means some investors believe that the stock with a value of S_t today can drop 20% in one week; otherwise, people would price these options at zero value.

prices and the state of the economy.[2] For example, plotting BS implied volatilities against the strike price at a fixed maturity, known as implied volatility surface (IVS), indicates a skew or smile pattern and may indicate a skew in the risk-neutral density of future returns.

In this paper, we first review methods available to construct a smooth implied-volatility surface (IVS). These methods typically employ implied volatilities inferred from the Black-Scholes model and then interpolate in some clever ways to build a smooth implied volatility curve.[3] These methods include specifying an explicit function of IV and curve fitting techniques. We document that due to the complications involved in estimating the parameters of IVS functions (such as dealing with the unobserved variables or a limited number of options for estimations), curve fitting methods are more popular among practitioners and in empirical works. The construction of IVS is directly related to the underlying conditional return risk-neutral (RN) distribution[4] and offers a fertile ground for computing option-based variables.[5]

Second, we review a number of statistical variables inferred from option prices and provide a summary of techniques used to construct them. These variables include implied measures of volatility, implied measures of tail risk, and implied measures of risk preferences. Then, we review their application in the literature for forecasting future asset returns or the state of the economy. We document that the literature has employed option-implied volatility more extensively than other option-implied variables for two reasons. First, option-implied volatility estimation is more straightforward than option-implied higher moments. Second, when risk premia are small, option-implied

2. Option-based variables inferred from option prices may provide interesting additional information not contained in the historical data. For example, Bali et al. (2019) have investigated the information embedded in option-implied variables for expected stock returns and Chang et al. (2012) exploit its information for equity risk. Kostakis et al. (2011), DeMiguel et al. (2013), and Kempf et al. (2015) use option-implied information to improve asset allocation techniques or market timing. Amin (1997), Cao et al. (2005), Jin et al. (2012), Chan et al. (2015) examine the relative effectiveness of option-implied variables to predict corporate events- i.e., takeovers, earnings announcements, etc. Goyal and Saretto (2009), Bali and Murray (2013) examine option-based variables to predict option returns.

3. In a typical market, only a few strike prices and maturities are available; therefore, the IV curve is not always smooth.

4. For example, a smile indicates fat tails, whereas a skew shows an asymmetry in the return distribution.

5. Ulrich and Walther (2020) show that recognised option-implied metrics such as risk-neutral variance, the variance risk premium, and skewness are affected by techniques that the implied volatility surface is built.

forecasts are more informative. The literature suggests that the volatility risk premiums are smaller than the risk premiums of higher moments, particularly skewness (see Christoffersen et al. (2011)).

Next, equipped with methods and variables available to extract information from option quotes, we contribute to the literature by studying which option-implied measure forecasts future stock returns cross-sectionally. Specifically, we examine the statistical and economic predictive power of eight option-based variables to forecast future individual stock returns. These variables include model-free implied volatility (MFIV) from Bakshi et al. (2003), forward volatility (FV) from Bakshi et al. (2011), implied risk-neutral skewness (RNS) from Bakshi et al. (2003), implied options expected returns (ExRet) from Martin and Wagner (2019), implied volatility smirk from Xing et al. (2010), a measure of deviation from put-call parity (PCD) from Cremers and Weinbaum (2010), and implied-volatility slope of call (put) options, which reflects the right (left) slope of the volatility curve (left (right)-IVS) from Kelly et al. (2016).[6]

We apply formal statistical significant tests via predictive panel regression models to assess whether there are cross-sectional predictive abilities for future returns in the dynamics of option-implied measures mentioned above. Furthermore, we examine whether the forecasts of option-implied measures are economically significant using a portfolio sorting approach- i.e., for each option-based variable, we sort stocks into quintiles and compare the profitability of a portfolio that buys stocks with high option-based measure (top quintile) and sells stocks with low measure (bottom quintile).

We find that volatility measures, namely, model-free implied volatility, implied forward variance and implied expected returns, statistically and economically forecast future stock returns cross-sectionally- i.e., there is a significant positive relationship between MFIV, FV and ExRet with the cross-section of stock returns. This finding suggests that volatility risk induces an increase in future returns- i.e., traders need compensation for taking volatility risk.

Furthermore, we find a positive (negative) relationship between RNS (left-IVS) and the

6. We include smirk, PCD and left (right)-IVS in the set of our predictors to examine the forecasting ability of measures that capture specific positions on the volatility curve.

cross-section of returns. These relationships are statistically and economically significant. We argue that stock overpricing renders this positive relationship- i.e., stock overpricing may not be promptly updated in the underlying market due to arbitrage constraints. Therefore, market participants may incline to the options market to trade based on their negative information or views. To maximise the leverage on their negative beliefs, they sell (buy) out-the-money (OTM) calls (puts) instead of shorting the stock. Buying OTM puts is attractive compared to selling the underlying stock because the options market does not bear inventory risks. In turn, the risk-averse market makers have to sell the underlying stock and, therefore, should take the adverse risk of inventory. Consequently, they induce an increase (decrease) in OTM puts (calls) option prices to compensate for the risk they take.

Our paper contributes to the literature that reviews option implied measures. Jackwerth (2004) provide an overview on option-implied risk-neutral distributions. Homescu (2011) reviews methods to construct an option-implied volatility surface. Christoffersen et al. (2011) reviewed techniques available for extracting information embedded in option prices and option-implied variables that are used in forecasting. Our paper differs from these studies by reviewing additional recent measures such as implied expected returns, variance risk premiums, tail loss measures and implied risk aversions. These measures are extensions of model-free implied volatilities and risk-neutral skewness. They have been widely used in the literature to forecast the economy's future state or future returns.

In addition, our paper is related to studies which propose various options-based measures for forecasting the cross-section of stock returns. Xing et al. (2010), Cremers and Weinbaum (2010), Cremers et al. (2015) propose a variety of different option-based variables of stock mispricing. Our paper is different from these studies. First, we use newly proposed implied measures such as implied expected returns and implied-volatility slopes. In addition, we compare the predictive ability of OTM call options with OTM put options. We show that put options show more predictive power than call options and contain more informational content.

The remainder of this paper is organised as follows. Section 3.2 reviews different methods for extracting implied volatility surfaces (IVS). Section 3.3 focuses on statistical measures implied from options, namely, implied volatility measures, implied tail risk measures and implied risk preferences measures. Finally, Section 3.4 comprehensively analyses the predictive power of several option-implied variables for the cross-section of returns. Section 3.5 concludes.

3.2 Implied Volatility Surface

Black and Scholes (1973) (BS) proposed a pricing model to determine the fair price of call or a put option assuming that the asset prices are log-normally distributed. BS theoretical model relies on six parameters, including volatility, option type, price of the underlying asset, time to maturity, strike price, and risk-free rate. This model forms a landmark in modern quantitative finance. The BS formula remained popular among practitioners and has been used as a primary instrument, which maps the space of option prices to a unique number- i.e., option-implied volatility.

Specifically, BS model assumes that stock price follows a geometric Brownian motion dynamic in which the volatility is constant:

$$dS_t = rS_t dt + \sigma S_t d\omega_t$$

where r is a risk-less rate, σ is volatility, and ω_t is a Brownian motion. In this model, only volatility (σ) is unobserved and can be inversely backed out using the option's market value. The option price with a strike price of K and time-to-maturity of T can be obtained in a close form:

$$C^{BS}(T; K; S0; r; \sigma^2) = S_0 N(d) - Kexp(-rT)N(d - \sigma\sqrt{\sigma})$$

where S_0 is the spot price and $N(.)$ indicates the standard normal cumulative distribution function, and

$$d = \frac{ln(S_0/K) + T(r + \frac{1}{2}\sigma^2}{\sigma\sqrt{T}})$$

BS model implies that for any option with a observed market price of C_0^{Mkt} , there exists one volatility measure (σ) and can be inferred from market price of that option.[7]

$$C_0^{Mkt} = C_{BS}(T; K; S0; r; BSIV^2)$$

Due to the structure of the Black & Scholes formula, the implied volatility cannot be found in a closed-form formula. Many articles offered techniques for approximating the volatility with more or less accurate results. The investigations have been conducted in two directions: one group of studies attempt to find closed-form approximations for BS implied volatility- i.e., Bharadia et al. (1996) and Brenner and Subrahmanyan (1988). Another group of studies explore numerical methods such as Newton–Raphson method to approximate implied volatilities.

However, it soon became evident that the lognormality assumptions of BS are far from realistic and are invalidated by options' data.[8] Several studies such as Cox and Ross (1976), Rubinstein (1976) and Brennan (1979) proved that lognormal returns are necessary and sufficient conditions for option prices to satisfy the BS equation; However, this prediction was rejected by several papers including, MacBeth and Merville (1979), Rubinstein (1985), and Cont and da Fonseca (2002). These studies indicate that the variation of implied volatilities across term to maturity and option strike, known as " implied volatility surface" (IVS), is substantial. This plot suggests that the volatility surface is not flat and is dependant on both moneyness and time- i.e., IVS shows smile and has term structure. These studies propose that the skew or smile shape in the implied-volatility surface is correlated with the conditional non-normality distribution of the underlying return; therefore, the construction of a smooth IV surface is consistent with extracting the RNDs.

7. The option price function is strictly increasing of volatility- i.e., the vega of an option is always positive. For call options, we have:

$$C_{BS}(\sigma = 0, T, K, S) = (S_0 - Ke^{-r(T-t)})^+$$
$$C_{BS}(\sigma = \infty, T, K, S) = S_0$$

The intermediate value theorem implies that there always exists a Black Scholes implied volatility for a market price between these bounds. When considering arbitrage-free market data, we can always find implied volatility $BSIV$ such that the model price equals the market price.

8. Empirically, the implied volatility extracted from out-of-the-money put index options ($K/S_0 << 1$) are more than other options. Therefore, market prices of these options are more than the estimation price of the BS model.

An important question here is how to generate an arbitrage-free IV surface that represents option market information more accurately. Many papers have answered this problem by finding necessary or sufficient conditions for the IVS to be arbitrage free. The presence of an arbitrage opportunity will develop negative transition probabilities and negative local volatilities. This situation will generate severe mispricings such that even a slight deviation of the input parameters lead to entirely different price quotes and highly unstable greeks.

To satisfy the arbitrage restrictions, call option price should be a convex and monotonically decreasing function of strike price. Following Fengler (2009) conditions that are necessary and sufficient for the IVS to be statically free of arbitrage are:

$$
\begin{gathered}
-e^{-\int_t^T r_s ds} \leq \frac{\partial C}{\partial K} \leq 0 \\
\frac{\partial^2 C}{\partial K^2} \geq 0 \\
max\left(e^{-\int_t^T q_s ds} S_t - e^{-\int_t^T r_s ds} K, 0\right) \leq C \leq e^{-\int_t^T q_s ds} S_t \\
\frac{\partial C}{\partial T} \geq 0
\end{gathered}
$$

where S_t is the spot price, r is the risk-less rate and q is the stock's dividend rate and the valuation function of a European call at time t with strike K and maturity T is given by $C_t(K,T) = e^{-\int_t^T r_s ds} \int_0^\infty max(S_T - K, 0)\phi(t,T,S_T)dS_T$, assuming that the RN density exist and is denoted by $\phi(t,T,S_T)$.

Furthermore, in a typical market, only a few strike prices and maturities are available; therefore, the IV curve is not always smooth. Moreover, Ulrich and Walther (2020) show that the values of many famous option-implied metrics such as RN variance, skewness, and the variance risk premium depends on the methods available for smoothing the volatility surface. In the remainder of this section, we provide a survey of techniques available for constructing a smooth BS implied-volatility surfaces in more detail. Homescu (2011), categorises these approaches into two groups: model-free and model-based. We will discuss these models in more detail in the following sections.

3.2.1 Explicit Functions of IVS

Parametric or model-based techniques to extract IVS consist of studies which explicitly specify a function (e.g. polynomial functions) on Black-Scholes implied volatility and then, in some clever way, interpolate in the time dimension so as not to introduce arbitrage. These techniques have several advantages; For example, they are cheap to implement, easy to interpret, and need few computational requirements for option pricing.

These methods were first introduced by Dumas et al. (1998), who modelled the IV as a quadratic function of strike price and time to maturity (M, T). Specifically, they suggest that the following describes the daily IVS:

$$\sigma(M,T) = b_1 + b_2 M + b_3 M^2 + b_4 T + b_5 M \times T$$

Later, Gonçalves and Guidolin (2006) employ a same setting to the IV logs using moneyness instead of strike. Chalamandaris and Tsekrekos (2011) argue that these models approximate the term structure of the IVS linearly at any moneyness level; they perform relatively well for constructing the IVS of options with a time to expiration of shorter than a year; however, they might be problematic when considering longer option maturities. For this reason, they develop a parametric specification which employs Nelson and Siegel (1987) factorisation to model the IVS term structure. Specifically, they introduce level, slope and curvature factors to the IVS model.

$$\sigma_i = \Sigma_{j=1}^{7} \beta_j I_{j,i} + \epsilon_i$$

where $i = 1, ..., N$ is the number of implied volatilities in the cross-section, $I_{1,i} = 1$ is the Level, $I_{2,i} = 1_{\{\Delta_i>0\}}\Delta_i^2$ is the Right "smile", $I_{3,i} = 1_{\{\Delta_i<0\}}\Delta_i^2$ is the Left "smile", $I_{4,i} = \frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i}$ is the Short–term term structure, $I_{5,i} = \frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i}$ is the Medium–term term structure, $I_{6,i} = 1_{\{\Delta_i>0\}}\Delta_i\tau_i$ is the right "smile" attenuation, $I_{7,i} = 1\{\Delta_i < 0\}\Delta_i\tau_i$ is the Left "smile" attenuation.

More recently, François et al. (2022) established a parametric representation of IV surfaces based on a linear combination of risk factors. The implied volatility $\sigma(M, \tau)$

obtained for an option with moneyness M on a specific day and time to expiry τ is modelled by:

$$\sigma(M,\tau) = \underbrace{\beta_1}_{\text{Long-term ATM IV}} + \beta_2 \underbrace{exp\big(-\sqrt{\tau/T_{conv}}\big)}_{\text{Time to maturity slope}} + \beta_3 \underbrace{\Big(M1_{(M\geq 0)} + \frac{e^{2M}-1}{e^{2M}+1}1_{(M<0)}\Big)}_{\text{Moneyness slope}}$$
$$+ \underbrace{\beta_4\big(1-exp(-M^2)\big)log(\tau/T_{max})}_{\text{Smile attenuation}} + \beta_5 \underbrace{\big(1-exp((3M)^3)\big)log(\tau/T_{max})1_{(M<0)}\big)}_{\text{Smirk}}$$

where T_{max} is the maximum maturity described by the model and T_{conv} denotes the position of a fast convexity change in the implied volatility term structure based on maturity. These parameters are determined using observed data. The IV specification of François et al. (2022) uses factors that are twice continuously differentiable. Therefore, their model completes the shape of the density in the tails such that risk-neutral probability densities are smooth and add up to one over moneyness.

3.2.2 Curve-Fitting Methods in constructing IVS

Parametric methods described above are easy to implement if the collection of traded options includes a sufficient spectrum of strike prices and maturities. For instance, we can apply the BS model for each available option to infer the implied volatility from the option price. If option prices are available for all exercise prices and time-to-expirations, the inferred implied volatilities intrinsically form a smooth surface. However, in a typical market, only a few strike prices and maturities are available. In this situation, researchers proposed curve-fitting methods to back out the implied volatility surface from given option quotes.

Curve-fitting methods are implemented in two steps. First, an interpolation procedure is performed in the moneyness dimension to provide a smooth curve for the BS implied volatilities. This procedure produces implied volatility as a function of moneyness for each maturity (T). Next, the moneyness is set fixed and an interpolation method is performed in the maturity dimension. The two steps result in a smooth implied volatility surface.

Splines usually employed in the literature to interpolate implied volatilities include piecewise linear interpolation, regular cubic splines, Cubic B-splines and Thin splines.[9] These techniques have minimal restriction on the shape of IVS and hence follow the data better. Campa and Chang (1998) propose using a natural spline to ensure that the fitted IVS function is flexible and smooth. Bliss and Panigirtzoglou (2002) proposed the use of cubic spline to match the implied volatility with the option delta. Kahale (2004) suggests an interpolation technique based on piecewise convex polynomials employing the BS formula. Figlewski (2008) proposed fitting Generalised Extreme Value (GEV) distributions to the tails. In addition, for strikes beyond the series of available options, Jiang and Tian (2005) proposed use of extrapolation, while Andersen and Bondarenko (2009) and Andersen et al. (2011) suggest employing truncation. Benko et al. (2007) propose to use the local polynomial smoothing technique. A system of nonlinear minimisation problems is solved for a given grid of strikes and time-to-maturities to obtain local polynomial estimators. Laurini (2011) employs constrained smoothing B-splines.

Cont and da Fonseca (2002) develop another group of non-parametric IV smoothing methods. They used a nonparametric Nadaraya–Watson estimator with a Gaussian Kernel process.[10] In addition, Israelov and Kelly (2017) used principal component analysis to describe the full IV surface. Ackerer et al. (2020) used a neural network approach to construct an implied volatility surface.

3.2.3 Comparisons of Methods for Extracting IVS

Ait-Sahalia and Duarte (2003) suggest that parametric methods discussed above for IVS extraction are beneficial for many reasons. First, they allow extrapolation beyond the available data. Second, it is easy to specify a function which can satisfy the restrictions imposed by theory (i.e., increasing and concave returns, $Y = Ln(X)$)). Third, more unrestricted parametric models can be proposed and examined against other models to validate constraints imposed by the theory. Fourth, theoretical restrictions may decrease these models' estimated parameters' variance.

9. The option price function interpolation also has some additional applications. For example, using liquid option market quotations, an interpolation method is used to price nonliquid or nontraded options. It is also used in interest-rate cap markets and swaption markets.

10. The Nadaraya–Watson estimator is a weighted average of $Y1, \ldots, Yn$ utilising the set of weights $W_i(x)_{i=1}^n$ (they add to one). The set of varying weights depends on the evaluation point x.

Despite all their benefits, parametric models have disadvantages: (i) Specification errors will lead to inconsistent estimates. (ii) Any theory test is also a test for a parametric model. (iii) Empirically, the slow convergence rate of estimators is a significant issue. The reason behind it is mainly due to a small number of observations and liquidity issues which can be substantial. On the other hand, the nonparametric methods do not need large samples to fulfil the restrictions. For these reasons, nonparametric methods are often used empirically, at least as a first step in analysing the data.

3.3 Option-Implied Measures

The foundation of the Chicago Board Options Exchange motivated investors to reverse the process of option pricing and infer implied measures from observed option prices instead. The BS formula remained popular among practitioners and serves as a suitable and simple instrument which maps option prices to a single actual number- i.e., option-implied measure. Converting option prices to statistical implied measures allows comparing option prices with different strike prices and expiration dates for various underlying assets. For example, a skew or smile shape in the BS implied volatility surface is directly linked to the conditional non-normality distribution of the underlying's return as opposed to the BS assumptions. A smile indicates fat tails in the return distribution, while a skew shows an asymmetry in the return distribution. Furthermore, the variation of implied volatility smile over different maturities and times implies the variation of the return distribution across these features. These empirical facts suggest that option prices and variables implied from them may be forward-looking instruments and may incorporate market expectations over the option's maturity.

In addition to BS implied variables, practitioners have proposed several measures implied from options. In this section, we review a number of these metrics and discuss their superior information about the future state of the economy.

3.3.1 Measures of Volatility

We cannot directly observe volatility. Therefore, the literature has proposed different proxies to measure it. These models are categorised into three groups—first are model-based approaches, which formulate the volatility process as an ARCH or stochastic volatility model. Second are realised volatilities computed from summation of high frequency squared realised returns. Third are forward-looking implied volatilities (IV) inferred from options quotes.

Option-implied volatility is particularly the most well-known implied measure in the literature and is extensively used in asset pricing. Implied volatility inferred from option quotes offers reasonable estimates for future volatility. This section, reviews IV metrics and their extensions that forecast future state of economy such as stock prices and macroeconomic and corporate events.

3.3.1.1 Model-Free Implied Volatility

Breeden and Litzenberger (1978) suggest that a series of OTM options can define any payoff function that is twice continuously differentiable. This observation was later established formally in Green and Jarrow (1987) and Nachman (1988). More specifically, these studies suggest that the price of any hypothetical claim with payoff $H(S)$ is:

$$\begin{aligned}\mathbb{E}(e^{-rt}H[S]) &= (H[\bar{S}]-\bar{S}H_S[\bar{S}])e^{-rt}+H_S[\bar{S}]S(t)\\ +\int_{\bar{S}}^{\infty}H_{SS}[K]C(t,\tau,K)dK &+ \int_0^{\bar{S}}H_{SS}[K]P(t,\tau,K)dK\end{aligned} \tag{3.1}$$

where $H_S[\bar{S}]$ ($H_{SS}[K]$) denotes the first (second) order derivative of the payoff with respect to S estimated at the strike price ($\bar{S}$).

Carr and Madan (1998) document that using Equation (3.1) one can compute volatility payoffs by holding different positions in options without necessarily being exposed to prices.[11] Therefore, a continuum of OTM option prices can estimate the future realised variance between time t and T as follows:

$$\begin{aligned}\hat{\sigma}^2_{t,T} &= Var_0\{ln(F_T/F_t)\} = E_0\{ln(F_T/Ft)-E_0[ln(F_T/Ft)]\}^2\\ &= E_0[\int_t^T \sigma_u^2 du] = 2\int_0^{\infty}\frac{C^F(0,T,K)-C^F(0,t,K)}{K^2}dK\end{aligned}$$

where $C^F(0,t,K)$ and $C^F(0,T,K)$ are the prices of call options at time zero on a forward

11. To see the proof please refer to Carr and Madan (1998)

contract with forward price K at time t and T, respectively. Therefore, If we forecast realised volatility from the current time 0 to a date, say T, $C^F(0,t,K)$ will be replaced by the intrinsic value of a call option with strike K or $max(F_0 - K)^+$ (i.e., F_0 being the forward price at time $t = 0$ and K the strike).

Britten-Jones and Neuberger (2000) document that relying on the underlying's spot price rather than the forward contract, one can forecast realised volatility from the current time 0, to date T, by the instantaneous squared return:

$$\mathbb{E}_0[\int_0^T (\frac{dS_u}{S_u})^2] = 2\int_0^\infty \frac{C(0,T,K) - max(S_0 - K, 0)^+}{K^2} dK$$

Correspondingly, Demeterfi et al. (1999) derive a price of a variance swap[12] and employ it to replicate the risk-neutral variance. More specifically, they assume that the price of an underlying follows a stochastic process as below:

$$\frac{dS_t}{S_t} = \mu dt + \sigma_t d\omega_t = X_S$$

where X_S is the strike price of a variance swap. Using Ito Lema, they suggest that

$$\frac{dS_t}{S_t} - dln(S_t) = \frac{1}{2}\sigma_t^2 dt$$

They show that with this modelling assumption and using Ito lemma, a weighted sum of OTM call and put prices replicates the risk-neutral expected variance σ_t.

$$\begin{gathered} \mathbb{E}[\int_t^T \sigma_u^2 du] = 2\,\mathbb{E}_0[\int_t^T \frac{dS_t}{S_t} - dln(S_t)] = \\ 2[rT - (\frac{S_0}{F_t} - 1) - ln(\frac{F_t}{S_0}) + \int_0^{F_t} \frac{e^{rT}}{K^2} P(t,T,K) + \int_{F_t}^\infty \frac{e^{rT}}{K^2} C(t,T,K)] \\ 2e^{r(T-t)}[\int_0^{F_t} \frac{1}{K^2} P(t,T,K) dK + \int_{F_t}^\infty \frac{1}{K^2} C(t,T,K) dK] \end{gathered} \tag{3.2}$$

where P and C are prices of put and call options with strike price K at time t expiring at T. $F_t = S_t e^{r(T-t)}$ is the forward price.

12. Variance swap is a contract that at time T pays integrated variance between time 0 and T and receives a strike price, X_S. The strike is specified such that the value of the variance swap is zero when written at time 0- i.e., $e^{-rt}\,\mathbb{E}[\frac{1}{T}\int_0^T \sigma_t^2 dt - X_S] = 0$

Letting the relative return be $R(t,T) = ln(S_T) - ln(S_t)$, Bakshi et al. (2003) (BKM) define a volatility contract which pays functions of the realized variance of an underlying asset's returns- i.e., $H(S) = R^2(t,T)$. They showed that the fair value of this contract at time t denoted by $V(t,T) = \mathbb{E}[e^{-r(T-t)}R^2(t,T)]$ can be obtained using a continuum of option prices as:

$$V(t,T) = \int_{s_t}^{\infty} \frac{2(1 - ln[\frac{K}{S_t}])}{K^2} C(t,T,K)dK + \int_{0}^{s_t} \frac{2(1 + ln[\frac{S_t}{K}])}{K^2} P(t,T,K)dK$$

Therefore, BKM define the annualised model-free implied volatility as follows:

$$MFIV_{t,i} = \sqrt{\frac{e^{r_f(T-t)}M(2)_{t,T,i} - \mu^2_{t,T,i}}{T-t}} \tag{3.3}$$

where μ is obtained as follows:

$$\mu_{t,T,i} = R_{f,t,T} - 1 - R_{f,t,T}[M(2)_{t,T,i}/2 + M(3)_{t,T,i}/6 + M(6)_{t,T,i}/24] \tag{3.4}$$

and $M(n)$ is obtained as follows:

$$M(n)_{t,T} = \int_{S_t}^{\infty} \eta(K,S_t,n)C_t(K,T)dK + \int_{0}^{S_t} \eta(K,S_t,n)P_t(K,T)dK \tag{3.5}$$

$$\eta(K,S_t,n) = \frac{n}{K^2}[(n-1)log(\frac{K}{S_t})^{n-2} - log(\frac{K}{S_t})^{n-1}]$$

$R_{f,t}$ is the riskless rate. $S_{t,i}$ denotes the price of the asset i at time t and $P_{t,i}(K,T)$ and $C_{t,i}(K,T)$ are the OTM put and call prices of asset i with exercise price K expiring at T.

BKM derive the price of the volatility contract based on the current stock price (S_t) while Demeterfi et al. (1999) derive the payoff based on the forward price (F_t). All these concepts of volatility proposed by different studies will provide identical results if the log returns are uncorrelated and have a zero mean.

Carr and Lee (2008) assume that the variance of the market index return follows an adapted process σ_t^2. The variance of returns[13] over $\tau(= T - t)$ is equal to the quadratic variation $\int_t^T \sigma_u^2 du = \langle X \rangle_t^\tau$. $\langle X \rangle$ is known as the realized variance of the returns on S and denotes the quadratic variation of $X = log(S_T/S_t)$. Then, they consider a specific form of the generic exponential claim on integrated variance- i.e., an exponential variance claim which pays $e^{\lambda \langle X \rangle_t^\tau}$ for some constant λ. Bakshi et al. (2011) adopt this modelling setup and consider an exponential claim on variance with payoff $H[S]$ have the following price at time t:

$$\mathbb{E}^Q[e^{-r\tau} H[S]] = e^{-r\tau} E^{\mathbb{Q}}\{exp(-\int_t^{t+\tau} \sigma_u^2 du)|F_t\} \tag{3.6}$$

$$= e^{-r\tau}\, \mathbb{E}^{\mathbb{Q}}\{\sqrt{\frac{8}{7}}(\sqrt{\frac{S_{t+\tau}}{S_t}}) \times cos(arctan(1/\sqrt{7}) + \frac{\sqrt{7}}{2} ln(\frac{S_{t+\tau}}{S_t}))|F_t\} \tag{3.7}$$

They extract the price of such a claim using the prices of traded call and put options:

$$V_t^\tau = \mathbb{E}[e^{-r\tau} H[S]] = e^{-r\tau} + \int_{S_t}^{\infty} \omega[K] C(t,\tau,K) dK + \int_0^{S_t} \omega[K] P(t,\tau,K) dK \tag{3.8}$$

where $C(t,\tau,K)$ and $P(t,\tau,K)$ are the prices of index call and put options at time-t with strike price K and time to expiry τ. Moreover, $e^{-r\tau}$ is the price of a risk-free discount bond at time t maturing at $T = t + \tau$, with face value equals one and,

$$\omega = -\frac{\frac{8}{\sqrt{14}} cos(arctan(1/\sqrt{7}) + \frac{\sqrt{7}}{2} ln(\frac{K}{S_t})}{\sqrt{S_t} K^{3/2}} \tag{3.9}$$

is the position in options. Finally, they define the forward variance at time t between $t + \tau$ and $t + \tau'$ (with $\tau' > \tau$) based on observed prices of exponential claims :

$$FV_{t,\tau,\tau'} = ln(V_t^{\tau'}) - ln(V_t^\tau) \tag{3.10}$$

The right-hand side (RHS) of the above equations shows that the integral of OTM option prices over a series of strike prices defines the model-free implied volatility. The

13. In the absence of price jumps

integral can be easily calculated through numerical integration methods if a set of options is available for all strike prices. However, in the real options market, only a limited number of strike prices are traded, which may result in calculation errors. Therefore, the finite available strike prices is the main implementation issue of MFIV calculation. In addition, the numerical integration of these equations (dK) should be replaced by a summation (ΔK), resulting in discretisation errors. The ΔK needs to be reasonably small to reduce these errors. Several studies suggest applying curve-fitting techniques to interpolate across given strikes to fill the gaps in strikes. Truncation errors are another implementation issue in the calculation of model-free implied volatilities. Truncation errors arise when the integration is calculated over a limited number of strike prices rather than from 0 to ∞. Jiang and Tian (2005) suggest using a flat extrapolation beyond available strike prices to overcome truncation errors.

3.3.1.2 CBOE Volatility Index (VIX)

The Chicago Board Options Exchange proposed the VIX as a forward-looking volatility index of the US stock market to measure the expected short-term US market volatility over the next thirty calendar days. This measure, known as the "fear" gauge, is a typical example of option-implied measures and has a significant negative relationship with S&P 500 return dynamics.

Whaley (1993) was the first who proposes this index as a linear composition of implied volatilities of eight at-the-money (ATM) options on the S&P 100 options with maturities nearest to one month. This version of VIX relies on the BS pricing model and assumption and is computed as follows:

$$VIX = \sigma_1(\frac{N_2 - 22}{N_2 - N1}) + \sigma_2(\frac{22 - N1}{N_2 - N1})$$

σ_1 and σ_2 denote the estimated S&P 100 ATM-IV for the two nearest maturities (just less and more) than one month and are computed as follows:

$$\sigma_1 = (\frac{\sigma_{C,1}^{X_{1,L}} + \sigma_{P,1}^{X_{1,L}}}{2})(\frac{X_{1,L} - S}{X_{1,U} - X_{1,L}}) + (\frac{\sigma_{C,1}^{X_{1,U}} + \sigma_{P,1}^{X_{1,U}}}{2})(\frac{S - X_{1,L}}{X_{1,U} - X_{1,L}})$$

where $(X_{1,L}, X_{1,U})$ and $(X_{2,L}, X_{2,U})$ are strikes less and more than the underlying price S for the first and second options with the nearest available maturity, respectively. And N_1 and N_2 are the time to expiration (in days) of the first and second options, respectively.

$$\sigma_2 = (\frac{\sigma_{C,2}^{X_{2,L}} + \sigma_{P,2}^{X_{2,L}}}{2})(\frac{X_{2,L} - S}{X_{2,U} - X_{2,L}}) + (\frac{\sigma_{C,2}^{X_{2,U}} + \sigma_{P,2}^{X_{2,U}}}{2})(\frac{S - X_{2,L}}{X_{2,U} - X_{2,L}})$$

As mentioned in the previous section,Carr and Madan (1998), Britten-Jones and Neuberger (2000) and Demeterfi et al. (1999) proposed a model-free measures of volatility. CBOE and Goldman Sachs employed their methods and published the model-free version of VIX. In 2003, S&P 500 options absorbed more attraction and became more informative than the S&P 100 options. Therefore CBOE shifted to the new options market structure and used S&P 500 options quotes to calculate VIX. Since then, the model-free VIX has been widely used for quantifying market risk, hedgeing portfolios, volatility tradings, anticipating the spot volatility dynamics or predicting future volatility. More specifically, CBOE suggests that the calculation of the VIX at time $t = 0$ follows Equation (3.2) and can be obtained by:

$$VIX = 100\sqrt{\frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{rT} O(K_i) - \frac{1}{T}[\frac{F_0}{K_0} - 1]^2}$$

where K_0 is the first strike less than F_0, $K_i = (K_{i+1} - K_i/2)$, and $O(K_i)$ is the mid bid-ask price for an out-of-the-money put or call option with strike K_i.

The theoretical model in Demeterfi et al. (1999) adopts a diffusive dynamic for the underlying; therefore, the "model-free" feature is doubtful for some researchers. The literature suggests that different specifications of the underlying dynamics result in different schemes of weighted (OTM option) in the volatility formula. For example, Martin (2011) suggests that in the existence of jumps, VIX becomes a biased measure of diffusion risk, and its bias degree is linked to the jump severity.

Andersen et al. (2015) show that the model-free implied volatility formula is still useful in the existence of jumps if the realised volatility is calculated employing squared weighted returns; therefore, the model-free volatility is jump robust. They propose the corridor implied volatility index, CIV, assuming that volatility and returns may jump together- i.e., “co-jump”. They obtain the expected value of corridor variance under the RN probability measure (the corridor variance swap rate) utilising a continuum of options with strikes between B_1 and B_2:

$$CIV = \frac{2e^{rT}}{T} \int_{B_1}^{B_2} \frac{O(K_i)}{K^2} dk \tag{3.11}$$

If $B_1 = 0$ and $B_2 = \infty$, then corridor variance turn into model-free implied variance. This suggests that the corridor variance swap is less expensive than a variance swap and allows investors to bet on potential stock price patterns. Andersen et al. (2015) suggest that it is difficult to measure option prices in the tails of the distribution. However, if we define the corridor via symmetric percentiles, the truncation points can show the significance of the right and left tails. Andersen et al. (2015) propose to consider $B_1 = K_{0.03}$ and $B_2 = K_{0.97}$ in Equation (3.11).

3.3.1.3 Lower Bound on the Expected Excess return (SVIX)

Martin (2017) derives an option-implied lower bound on the market premium in terms of simple risk-neutral variance, $(SVIX_{t,T} = var^Q(\frac{R_{t,T}}{R_{F,t,T}}))$, that can be calculated from index option prices. To derive this measure, first, consider that the conditional RN variance of a gross return $(R_{t,T} = \frac{S_T}{S_t} - 1)$ is denoted by $var^Q_{t,T}$:

$$\frac{1}{R_{F,t,T}} var_t^Q(R_{t,T}) = \frac{1}{S_t^2 R_{F,t,T}} [\mathbb{E}_t^Q\, S_T^2 - (\mathbb{E}_t^Q\, S_T)^2] \tag{3.12}$$

The first term inside the square brackets is also a “squared contract”— which is, the price of a contract which pays S_T^2 at time T and can be approximated by the price of a continuum of options:

$$\frac{1}{R_{F,t,T}} \mathbb{E}_t^Q\, S_T^2 = 2 \int_0^{\infty} C(t,T,K) dK$$

The second term in the square brackets of Equation (3.12) is obtained considering that the forward price of the underlying stock is $F_{t,T} = \mathbb{E}_t^Q S_T$

In the light of the put-call parity and combining the above equations, one can express the simple risk-neutral variance (SVIX) as follows:

$$\frac{1}{R_{F,t,T}} var_{t,T}^Q(R_{t,T}) = \frac{1}{S_t^2}[2\int_0^\infty C(t,T,K)dK - \frac{F_{t,T}^2}{R_{F,t,T}}] =$$
$$\frac{2}{S_t^2}[\int_0^{F_{t,T}} P(t,T,K)dK + \int_{F_{t,T}}^\infty P(t,T,K)dK]$$

therefore:

$$SVIX_{t,T}^2 = \frac{2}{R_{F,t,T}S_t^2}[\int_0^{F_{t,T}} P(t,T,K)dK + \int_{F_{t,T}}^\infty P(t,T,K)dK] \tag{3.13}$$

The RHS of the Equation (3.13) resembles the VIX index, and suggests that the two measures are connected. As outlined, SVIX quantifies the RN volatility of the return on the market uniquely by European option prices. SVIX puts equal weights on option prices; however, VIX is weighted by $\frac{1}{K^2}$. Obviously, VIX puts more weight on OTM puts and less on OTM calls- i.e., imposes more importance on left-tail events. Martin (2017) documents that the VIX estimates entropy, while the SVIX index estimates variance. He shows that both quantities are equally reasonable, and it depends on the analyst's objective to choose them and exploit their information- i.e., Entropy is more sensitive to the left tail, while variance is more sensitive to the right tail.

Martin (2017) employs the SVIX measure and derives a lower bound on expected index returns. More specifically, he suggests that the following identity links the expected returns and risk-neutral variances:

$$var_t^Q R_{t,T} = \mathbb{E}_t^Q R_{t,T}^2 - (E^Q R_{t,T})^2 = R_{F,t,T}\,\mathbb{E}_t^P(M_T R_{t,T}^2) - R_{F,t,T}^2$$

where $\mathbb{E}^P$ is expectation under physical measure, and M_T is the stochastic discount factor. The following identity links expected returns and risk-neutral variance:

$$\mathbb{E}_t^P R_{t,T} - R_{F,t,T} = [\mathbb{E}_t^P (M_T R_{t,T}^2) - R_{F,t,T}] - [\mathbb{E}_t^P (M_T R_{t,T}^2) - \mathbb{E}_t^P R_{t,T}] = \frac{1}{R_{F,t,T}} var^Q R_{t,T} - cov(M_T R_{t,T}, R_{t,T})$$

Considering a negative correlation condition (NCC) assumption- i.e., $cov(M_T R_T, R_T) \leq 0$, Martin proposes that the RN variance of the market offers a lower bound on the equity premium:

$$\mathbb{E}_t R_{t,T} - R_{F,t,T} \geq \frac{1}{R_{F,t,T}} var^Q R_{t,T} = SVIX_{t,T}$$

While Martin (2017) argues that an NCC should hold for the market return, Kadan and Tang (2019) extend this idea and compute a lower bound for expected stock returns assuming that NCC holds for individual stocks. Similarly, Martin and Wagner (2019) build on the SVIX measure and propose a new formula that expresses the expected return on a stock in terms of three measures, namely the risk-neutral variance of the market ($SVIX_M$), the risk-neutral variance of the individual stock ($SVIX_i$), and the value-weighted average of stocks' risk-neutral variance ($S\bar{V}IX = \sum w_i SVIX_i$). Martin and Wagner (2019) propose that the expected return on stock i from time t to T is obtained as follows:

$$\mathbb{E}_t R_{i,t,T} = R_{f,t,T} \times [SVIX_{m,t,T}^2 + \frac{1}{2}(SVIX_{i,t,T}^2 - \overline{SVIX}_{t,T}^2)] + R_{f,t,T} \tag{3.14}$$

where $SVIX_{i,t,T}^2$ is the RN variance at the stock level and $R_{f,t,T}$ is the gross riskless rate from time t to $t+1$.

In addition, Chabi-Yo et al. (2022) propose a flexible technique to build individual assets' expected excess return bounds. Their approach considers the entire shape of the RN density rather than only the second moments in Martin and Wagner (2019). Specifically, they suggest that:

$$E_t(R_{i,t,T} - R_{f,t,T}) = max_{\theta \in \Theta_{i,t}} \{E_t^*(R_{i,t,T}^{\theta+1})/E^*(R_{i,t,T}^{\theta}) - R_{f,t,T}\} \tag{3.15}$$

where $R_{i,t,T}$ is the the simple gross return on stock i and $R_{f,t,T}$ is the risk-free rate. The RN expectations, E^*, in (3.15) are calculated using option quotes for a specific θ where the set of Θ is given by:

$$\Theta_{i,t} = \{\theta : \theta - \gamma_t \beta_{Mi} + \frac{\gamma_t \rho_t}{R_{f,t,T}} \beta_{M^2,i} + \frac{\theta(\theta - 1)}{2R_{f,t,T}} B_{i^2,i} - \frac{\theta \gamma_t}{R_{f,t,T}} \beta_{Mi,i} \leq 0\} \tag{3.16}$$

Chabi-Yo et al. (2022) suggest that the parameters $\gamma_t = 2$ and $\rho_t = 1$. They define $\beta_{i^2,i}$, $\beta_{M,i}$, $\beta_{Mi,i}$ and $\beta_{M^2,i}$ as follows:

$$\begin{aligned}
&\beta_{M,i} = COV(R_{M,t,T}, R_{f,t,T})/VAR(R_{i,t,T}) \\
&\beta_{M^2,i} = COV_t((R_{M,t,T} - R_{f,t,T})^2, R_{i,t,T})/VAR_t(R_{i,t,T}), \\
&\beta_{i^2,i} = COV_t(R_{i,t,T} - R_{f,t,T})^2, R_{i,t,T}/VAR_t(R_{i,t,T}); \\
&\beta_{Mi,i} = COV_t((R_{M,t,T} - R_{f,t,T})(R_{i,t,T} - R_{f,t,T}), R_{i,t,T})/VAR_t(R_{i,t,T})
\end{aligned} \tag{3.17}$$

where $R_{M,t,T}$ is the market return.

A fixed quadratic function of returns will project the lower bound of Chabi-Yo et al. (2022) to the variance-based bound of Kadan and Tang (2019) and Martin and Wagner (2019), while payoffs of fractional or higher degree polynomials consider higher moments.

3.3.1.4 Variance Risk Premium

One application of model-free implied volatilities is that it "Synthesises" variance swap rates and provides a robust technique for measuring the variance risk premium on financial assets. A variance swap has zero market value at time zero. However, at the expiration date, the payoff to the buyer of the swap is the spread between the realised variance during the contract's life and the variance swap rate:

$$[RV_{t,T} - SW_{t,T}] \times L$$

where $RV_{t,T}$ represents the realized annualized return variance between time t and T , $SW_{t,T}$ indicates the fixed variance swap rate that is estimated at time t and paid at time T. L indicates the notional dollar value that transforms the variance into a dollar payoff. SW rates are unavailable from real data sources; therefore, the previous literature calculates the SW rates by synthesising them via a trading strategy in European options and futures. Demeterfi et al. (1999) suggest that the no-arbitrage condition implies that the variance swap rate equals the risk-neutral variance:

$$SW_{t,T} = \mathbb{E}_t^Q[RV_{t,T}]$$

Carr and Wu (2009) estimate the average variance risk premium via the average of the spread between the variance swap rate and the realised variance, $RP_{t,T} = RV_{t,T} - SW_{t,T}$ where realised variance is calculated using the forward prices:

$$RV_{t,T} = \sum_{i=1}^{T-t} (\frac{F_{t+i,T} - F_{t+i-1,T}}{F_{t+i-1,T}})^2$$

Bollerslev et al. (2009*b*) also estimates the value of the "ex-ante" variance risk premium (VRP_{exa}) for the index proxy as the spread between the 30-days implied variance (MFIV) observed on day t and the realised index variance over the last 30 calendar days.

Konstantinidi and Skiadopoulos (2016) generalize the conventional approach of Bondarenko (2014) and propose a method to measure the market $VRP^T_{t,t+h}$ from a T-maturity swap held from t to $t+h$. More specifically, they propose that the $VRP^T_{t,t+h}$ for an investment horizon h can be obtained as follows.

$$VRP^T_{t,t+h} = \mathbb{E}_t[RetSW^T_{t,t+h}]$$

where $RetSW^T_{t,t+h} = \frac{PL^T_{t,t+h}}{SW^T_{t,t+h}}$ and

$$PL^T_{t,t+h} = e^{-r(T-h)} L[\lambda RV_{t,t+h} + (1-\lambda) SW_{t+h,T+t} - SW_{t,t+T}]$$

PL is the profit and loss received from owning a swap contract for an investment horizon h, maturing at T. L is the notional dollar amount of the swap, r is the risk-less rate. $\lambda = h/T$ is the ratio of the investment horizon over the time-to-expiry of contract, and $RV_{t,t+h}$ is the asset return's realised variance from t to $t+h$.

Konstantinidi and Skiadopoulos (2016) suggest that synthesising SW from option prices could result in a bias in VRP computations. First, the conventional methods of VRP computation followed the MFIV calculation and assumed a diffusion process in the absence of jump and stochastic volatility. However, if a jump exists in the underlying dynamics, the synthetic SW rate is a biased predictor of $\mathbb{E}^Q[RV]$. Therefore, synthetic SW rates underestimate real SW if negative jumps are dominant. Second, there are computational errors in synthesising the SW rates (Jiang and Tian (2007)). Third, the MFIV estimation is untrustworthy during a crisis, especially when it is employed to measure the investor's fear. (see Andersen et al. (2015)). To address these shortcomings, Bollerslev et al. (2011) estimates the volatility risk premium considering that the underlying asset follows a stochastic process. Furthermore, Ait-Sahalia and Mancini (2020) analyse the swap rates using a model-free technique to estimate the price of a jump in swap rates.[14]

3.3.1.5 Application of Implied Volatility Measures

Several studies employed model-free implied volatility measures for prediction. For example, some studies show that the implied-volatility contains valuable information in predicting future stock returns. An et al. (2014) document that the implied volatilities of ATM call and put options have conflicting inferences. They empirically show that an upward shift in the implied volatility of call (put) options is associated with higher (lower) future returns. Conrad et al. (2013) analyse the relationship between RN variance and subsequent stock returns and find an insignificant negative relationship between the two. Ang et al. (2006) use the variations in the VIX as innovation in volatility and show that stocks which are more sensitive to innovations in aggregate volatility earn lower returns on average. Diavatopoulos et al. (2008) define the implied idiosyncratic volatilities as the spread between the implied volatility of the market and the implied volatility of the firm and document a strong positive relationship between implied idiosyncratic risk and future asset returns. Other studies, such as Navon (2014) and Cao et al. (2022), study the relationship between implied volatility changes and corporate bond returns. Both studies document that changes in implied volatilities negatively

14. If the stock price process does not jump, the VS payoff can be replicated by dynamically trading futures contracts and by taking a static position in a continuum of European options with various strike prices and the same maturity. However, if the stock price has a jump component, the replication above no longer holds. This observation makes it possible to assess whether VS rates embed a jump component and quantify its magnitude.

predict future bond returns.

Another strand of literature employs option-implied volatilities to forecast market volatility. Number of studies state that implied volatility incorporates additional information compared to historical volatility in predicting future volatility. For example, Busch et al. (2011) predict the current month's total realised variance (RV) with the lagged realised variances and BS implied volatility. They document that adding the BS implied volatility to the set of predictors increases the adjusted R^2. Kourtis et al. (2016) compare the predictive ability of a GARCH model with the S&P 500 option-implied variance and suggest that implied volatility offers better monthly realised variance forecasts. Similarly, Giot and Laurent (2007), Frijns et al. (2009), Kambouroudis et al. (2016) all show that implied-volatility is an efficient estimator of realised variance. Some studies such as Xu and Taylor (1994), Fleming et al. (1995), Blair et al. (1999) suggest that BS implied-volatility is a biased but efficient estimator of volatility. Nevertheless, some other studies suggest that RN volatilities are systematically different from their physical pairs because variance risk premiums are non-zero (e.g., Bakshi and Madan (2006), Carr and Wu (2009). These studies document that implied volatility is a weak predictor of future volatility and has limited practical information for forecasting future volatility. (Day and Lewis (1992); Lamoureux and Lastrapes (1993), Canina and Figlewski (1993), Becker et al. (2007)).

To address risk premiums, measures synthesised from implied volatility, such as VRP (defined as the difference between the model-free implied variance and the lagged realised variance), show that market risks are linked with variance risk premiums. Bollerslev et al. (2009*a*) and Drechsler and Yaron (2011) document that the one-month variance risk premium has substantial forecasting ability for future market excess returns. Bollerslev et al. (2011), Schneider and Trojani (2018), Andersen et al. (2019), among others, show that the VRP is mainly due to the tail risk, and thus the tail risk is predictable by VRP.

Other studies focus on the CAPM and try to predict betas more precisely by taking into account the forward-looking information inferred from options. French et al. (1983) estimate beta using the correlation between stock's and market's historical returns as well as option-implied volatilities of the stock and the market.

Other studies such as Martin (2017) demonstrates that options incorporate information on the lower bound of the underlying's expected return. Martin and Wagner (2019) derive a measure for estimating a stock's expected return from its implied volatility surface. Du and Kapadia (2012) construct a tail risk index from the difference between VIX and implied variance measure of Bakshi et al. (2003). Bollerslev and Todorov (2011) introduce risk measures to forecast jump risk under RN probability measure and document that a considerable part of the equity risk premium compensates jump risk.

3.3.2 Measures of Tail Risk

Empirical works on extracting option-implied skewness are divided into two classes. The first class consists of methods which define the implied skewness based on the asymmetry in the implied volatility surface. The second relies on the RN distribution of the returns and the primary result of Bakshi et al. (2003), who show that any payoff can be determined by precise positions on options. We will discuss these methods in more detail in the following sections.

3.3.2.1 Asymmetry of Implied-Volatility Surface

The literature has extensively reported that option-implied volatility differs across strike prices, which is known as the "volatility smile", "volatility smirk," or "volatility skew". These terms refer to the situation where observed implied Black-Scholes volatilities are larger for options with lower strike prices than for options with higher strike prices.

Several studies describe the implied volatility skew based on the asymmetry in the implied volatility surface. One effort to estimate volatility smirks is proposed by Bates (1991), who defines volatility smirk as the spread between the implied volatility of OTMP options and IV of OTMC options. However, options exist only for a limited number of exercise prices and therefore, the skew measure of asymmetry cannot be estimated directly. Bates (1991) suggests interpolating across strike prices to overcome this issue.

Toft and Pruck (1997) define implied volatility skew as the spread between the implied volatility of a 10 per cent in-the-money and a 10 per cent out-of-the-money option, proportional to the implied volatility of an ATM option. This measure delivers information in both the level and the slope.

Doran and Krieger (2010) define the implied skew as the spread between the average IV of options whose exercise prices are more than the current price of the underlying and the average IV of options whose exercise prices are less than the current price. This measure is known as above-minus-below (AMB).

$$AMB = \frac{(IV_{ITMP} + IV_{OTMC}) - (IV_{ITMC} + IV_{OTMP})}{2}$$

where IV_{ITMP}, IV_{OTMC}, IV_{ITMC} and IV_{OTMP} are average implied volatilities of all ITMP options, all OTMC options, all ITMC options, and all OTMP options, respectively.

Doran and Krieger (2010) also propose two more quantities, as the spread between OTM and ATM implied volatilities of call or put options.

$$COMA = IV_{OTMC} - IV_{ATMC}$$
$$POMA = IV_{OTMP} - IV_{ATMP}$$

In contrast to AMB , $COMA$ ($POMA$) uses only OTM and ATM call (put) options to capture the volatility curve asymmetry. The ATM options used in this method are options with a strike-to-spot ratio nearest to 1.00. A positive $COMA$ is related to bullish anticipations and suggests that optimistic traders increase their trading activities. However, a positive $POMA$ shows that OTM puts are more expensive than ATM options and indicates a bearish market.

Corresponding to $POMA$, Kelly et al. (2016) proposed an implied-volatility slope, which quantifies the slope of the function that relates implied volatilities to delta. The implied slope is estimated by regressing the IV of an option against its delta as:

$$IV_{t,T} = \alpha + IVSlope\Delta_{t,T} + \epsilon_{t,T} \tag{3.18}$$

where $IV_{t,T}$ and $\Delta_{t,T}$ are the implied volatility and delta of an OTM option. The regression coefficient left (right)-$IVSlope$ denote the put (call) implied volatility slope. Kelly et al. (2016) use this measure to quantify the tail risk of political uncertainty.

Another alternative measure to $COMA$, is the volatility smirk proposed by following Xing et al. (2010) as the difference between the average implied volatilities of OTM puts and ATM options. The smirk measure is computed as follows:

$$Smirk_{t,T} = IV_{t,T}^{OTMP} - IV_{t,,T}^{ATM} \tag{3.19}$$

Xing et al. (2010) define ATM options as options with a moneyness level (defined as the ratio of strike price to the current stock price) between 0.95 and 1.05. Moreover, a put option is defined as OTM when the ratio of strike price to the stock price is between 0.8 and 0.95.

Cremers and Weinbaum (2010) (and Bali and Hovakimian (2009)) define the deviation from put-call parity (PCD) as the difference between call and put implied volatilities. More precisely, this measure is the average difference in implied volatilities ($IV_{t,T,K}$), between call and put options (with the same strike price and maturity):

$$PCDt, T = \frac{\sum_{j=1}^{N} IV_{t,K_j}^{Calls} - IV_{t,K_j}^{Puts}}{N} \tag{3.20}$$

where j refers to pairs of put and call options and index strike prices. This measure quantifies the deviation from put-call parity- i.e., the IV of call and put options with identical strike price and time-to-expiration should be the same. A deviation from put-call parity- i.e., a non-zero PCD does not certainly imply an arbitrage opportunity; instead, it could be due to short-sale constraints, transaction costs, or the existence of informed traders. For example, suppose that insider traders obtain information about an increase in underlying asset prices. In that case, the prices of call options will rise while the prices of the put option will fall, resulting in positive PCD.

Doran and Krieger (2010) suggests that each measure motioned above captures a specific position of the implied volatility surface, and specific pairs can create the third measure.

Specifically, they document that IVS_{XZZ} captures the left and the middle of the put and call volatility skew, respectively. $CPIV$ describes the middle of the volatility skew. AMB reflects both tails of the volatility curve while $COMA$ concentrates on the right and centre of the volatility curve of call options, and $POMA$ focus on the left and centre of the volatility curve of put options.

3.3.2.2 Moments of RN-density

It has been widely recognised that the volatility smile is due to heavy-tailed risk-neutral distribution- i.e., there is a one-to-one relationship between the RN density function and the IV surface. Bakshi et al. (2003) (BKM) identify a significant relationship between the measure of the RN skewness and the slope of the implied volatility- i.e., negatively sloped volatility curves, where the implied volatilities of OTM puts are higher than that of ITM puts, demonstrate negative RN skewness.

Relying on the results of Carr et al. (2001) in Equation (3.1), BKM show that the cubic contract (which measures the third moment of returns) can quantify return asymmetry by positioning in OTM put and call options. Risk-neutral model-free skewness is computed using observed prices of OTM European call and put options. More specifically, they show that the RN skewness between time t and T is given by:

$$RNS_{t,T} = \frac{e^{r_f(T-t)}M(3)_{t,T} - 3e^{r_f(T-t)}\mu_{t,T}M(2)_{t,T} + 2\mu_{t,T}^3}{[e^{r_f(T-t)}M(2)_{t,T} - \mu_{t,T}^2]^{3/2}} \tag{3.21}$$

where $M(n)_{t,T}(n = 2, 3, 4)$ is given by Equation (3.5).

and $\mu_{t,T} = R_{f,t,T} - 1 - R_{f,t,T}[M(2)_{t,T}/2 + M(3)_{t,T}/6 + M(4)_{t,T}/24]$, $R_{f,t}$ is the riskless rate. S_t denotes the price of the asset at time t and $P_t(K,T)$ and $C_t(K,T)$ are the OTM put and call prices.

This method has been broadly used by researchers mainly because it is model-free and straightforward to implement. However, as discussed by Jiang and Tian (2005), one of its shortcomings is the selection of extrapolation technique to the tails of the distribution where there exists no available option price. In addition, its estimates are likely to be biased since the moments are risk-neutral rather than physical.

3.3.2.3 Implied Measures of Tail Risk and Default

Aıt-Sahalia and Lo (2000) established working on option-implied risk measures. Their work implies that Value-at-Risk (VaR) under the RN measure might reveal some parts of market risk that the physical measure cannot estimate. Therefore, they deduce that the option-implied VaR measure is more of a compliment than a substitute measure in risk management. They suggest that VaR and expected shortfall (ES) of log returns at the α level at time t are obtained by constructing a kernel estimator of the density function F_t^r of these returns:

$$VaR_t^Q = inf\{x \in R : F_t^r(x) \leq \alpha\}$$
$$ES_t^Q = \mathbb{E}_t^Q[x \in R : x \leq -VaR_t^Q]$$

In the same spirit, Vilkov and Xiao (2012) document that an option-implied conditional expectation of market loss arising from tail events, or tail loss measure (TLM), forecasts future market returns and the possibility of crashes. TLM is similar to the conditional Value-at-Risk, calculated under the RN probability and is equivalent to RN skewness. Employing the results of the Extreme Value Theory (EVT)[15] event, they utilise the second theorem of EVT, to determine the distribution of the asset price x less than a threshold h by a generalised Pareto distribution. They suggest that the average excess tail of the asset, with $\epsilon < 1$, can be computed as:

$$\mathbb{E}(h - x_t | h > x) = \frac{\beta}{1 - \epsilon}$$

where β and ϵ are the scale and the tail shape parameters, respectively.[16] Therefore, the TLM, $\Omega_{h,t}$, at time t conditional on the threshold h is defined as:

$$\Omega_{h,t} = \frac{\mathbb{E}(h - x | h > x)}{x_t}$$

15. which primarily focuses on the distributional properties of the extreme or low probability.

16. Refer to the EVT results and generalised Pareto distribution properties in the original papers of Balkema and Haan (1974)

Vilkov and Xiao (2012) infer the parameters β and ϵ from the available market prices. Therefore, all expectations and TLM are determined under the RN probability by resolving the optimisation problem using all OTM puts with strikes relatively smaller than $K_{0,t}$:

$$\epsilon_t, \beta_t = arg\ min_{\epsilon,\beta \in (RR^+)} \sum_{i=0}^{n-1} \frac{P_{i,t} - P^*_{i,t}}{P^*_{i,t}}$$

where $P_{i,t}$ is the market price of a put option with strike $K_{i,t}$ on day t, and $P^*_{i,t}$ is the price of a put option calculated based on the Extreme Value Theory:

$$P^*_{i,t} = P^*_{0,t}[\frac{\epsilon_t}{\beta_t}(K_{0,t} - K_{i,t}) + 1]^{1-1/\epsilon_t}$$

The strike $K_{0,t}$ is the threshold strike.

Another measure of tail risk is proposed by Culp et al. (2018) and Culp et al. (2021) via the concept of implied spread (IS). These papers turn Merton (1973)'s idea and construct an option implied bond, a defaultable zero-coupon bond, with an European put option $P_t(K,T)$ with strike price K and maturity date T and a risk free zero coupon bond, $Z_t(T)$, with maturity T

$$B_t(K,T) = Z_t(T) + \frac{P_(T,K)}{K}$$

Culp et al. (2021) suggest that the implied spread of a put option is the credit spread of its implied bond and reflects tail risk.

$$IS_t(K,T) = y_t(K,T) + r_t(T)$$

where $y_t(K,T) = \log B_t(K,T)/T - t$ is the continuously compounded yield of the risky zero-coupon bond and $r_t(T) = \log Z_t(T)/T - t$ is that of the risk-free zero-coupon bond.

IS is countercyclical and predicts implied bond returns. However, it does not predicts put return. They propose that the countercyclical relationship is consistent with a stochastic volatility model- i.e., Higher volatility increases put option premia, whereas jump intensity decreases put option premia.

3.3.2.4 Forecasting with Option-Implied Skewness

Implied skewness has been widely used to explain the subsequent cross-section of returns. Several studies employed implied skew based on asymmetry in BS implied volatilities. For example, Cremers and Weinbaum (2010) document that the deviation from put-call parity is more possible when the probability of informed trading of Easley et al. (1998) is high. Xing et al. (2010) look at the difference between OTMP and ATMC implied volatilities. They find that the implied volatility smirk negatively predictive firm-level cross-sectional returns. Furthermore, they find that this predictability persists for at least six months. Xing et al. (2010) and Cremers and Weinbaum (2010) find a theoretically contradictory positive relation between RN skewness and future returns. Doran and Krieger (2010) compared implied-skewness measures based on the positions they captured in the volatility surface. They show that splitting the volatility skew into parts and dealing separately with each part can be effective in forecasting future equity returns. Doran et al. (2006) report evidence supporting the hypothesis that the risk-neutral skewness, defined as the spread between implied volatilities of OTM puts and calls, has strong forecasting ability in the short-term decline of the stock market. Baltussen et al. (2011) find a negative relationship between weekly returns and all option-implied skewnes.[17]

Furthermore, several studies have widely used the implied skewness proposed by BKM. For example, Chang et al. (2012) propose a market skewness factor which commands a negative risk premium. They show that stocks which are more sensitive to the innovations in implied market skewness demonstrate lower returns on average. Rehman and Vilkov (2012) and Stilger et al. (2016) show that the RN skewness is positively related to future stock return. In addition, Borochin and Zhao (2022), Gkionis et al. (2021), and Chordia et al. (2020) find that a strong negative RNS value, due to expensive OTM puts relative to OTM calls, indicates future stock underperformance. These studies argue that stock overpricing is the reason for this positive relationship. Stock overpricing may not be promptly updated in the underlying markets; therefore, traders may incline to the options market to trade based on their negative views by buying (selling) OTM puts (calls). Consistent with Garleanu et al. (2009) demand-based option pricing frame-

17. Defined as the difference between the implied volatility of OTMP and ATMC.

work, if the risk averse market makers cannot completely hedge their positions, they would increase the price of put options to compensate for the risk they take. This, in turn, will provide a more negative RNS value. This option-implied information gradually merges into stock prices and increases the predictive relations. These findings, however, contrast with the empirical evidence supported by Conrad et al. (2013) and Dennis and Mayhew (2002). Specifically, Dennis and Mayhew (2002) document a negative relation between the RN skewness and systematic risk, beta. They claim that market risk is mirrored in the RN skewness inferred from the option quotes. Conrad et al. (2013) find that firms with less negative RN skewness earn less returns. Their results are consistent with Brunnermeier et al. (2007)'s model, which discusses the trade-off between the benefits of diversification and skewness- i.e., holding skewed securities results in a negative correlation between idiosyncratic skewness and expected returns.

Several studies deal with forecasting crashes with volatility skew. Doran et al. (2007) analyse whether the implied volatility skew pattern can forecast the market crash or spike probability. Doran et al. (2006) and Doran et al. (2007) find that option-implied skew has substantial predictive power for estimating the extent of market crash risk. Cremers et al. (2015) derive an aggregate jump and volatility risk factors by constructing a delta-neutral ATM straddle sensitive to crash risk and document that their factor requires negative risk premiums. Wang and Yen (2018) derive option-implied tail loss and then use these quantities to analyse whether option-implied tail risk contain information on the future returns of the underlying assets. Their empirical investigation indicates that tail measures inferred from S&P 500 options can predict crashes in S&P 500 index.

Some studies employed risk-neutral skewness to enhance the accuracy of the forecasts of future volatility. For example, Byun and Kim (2013) detect whether RN skewness and kurtosis have additional forecasting power on future volatility. They show that RN skewness has supplementary information compared to past realised volatilities and implied volatility- i.e., risk-neutral skewness and kurtosis measures contain some information content to predict jump variations in the heterogeneous auto-regressive realised volatility (HAR-RV) model. DeMiguel et al. (2013) suggest employing implied volatility, variance risk premium, correlation, and skewness in portfolio selection. They document that the skewness and the variance risk premium significantly enhances the portfolio's performance.

3.3.3 Measures of Risk Preferences

The complete market hypothesis suggests that the relationship between investors' risk preferences, asset return process (i.e., $R = ln(S_T/S_t)$ and state price densities are characterised as follows:

$$RRA(R) = -R\frac{U'(R)}{U''(R)} = -R\frac{m'(R)}{m(R)}$$

where RRA is the relative risk aversion function and U is the utility function of investors. The $m(R)$ is the stochastic discount factor (pricing kernel) that specifies the equilibrium asset prices.

Pioneering with Aıt-Sahalia and Lo (2000), number of other studies including Rosenberg and Engle (2002), Bliss and Panigirtzoglou (2004), Kang and Kim (2006) have focused on extracting risk preferences of investors using option quotes. Following Bakshi and Madan (2006), Kang et al. (2010) simplify this equation using the Taylor series expansion and exponential utility function. They provide a theoretical formula to compute volatility differences and extract implied RRA:

$$\frac{\sigma^2_{p,t}(\tau) - \sigma^2_{q,t}(\tau)}{\sigma^2_{q,t}(\tau)} \approx \gamma\sigma_{q,t}(\tau)\theta_{q,t}(\tau) + \frac{\gamma^2}{2}\sigma^2_{q,t}(\tau)(k_{q,t}(\tau) - 3) \tag{3.22}$$

where γ is the coefficient of the relative risk aversion of the representative agent. The formula suggests that the difference between the risk-neutral variance ($\sigma^2_{q,t}(\tau)$) and physical variance ($\sigma^2_{q,t}(\tau)$) is a function of the representative investor's relative risk aversion (γ) and higher order moments. In this formula $\sigma^2_{q,t}(\tau)$, $\theta_{q,t}(\tau)$ and $k_{q,t}(\tau)$ are the risk-neutral variance, skewness and kurtosis following Bakshi et al. (2003) obtained from Equations (3.3) and (3.21), respectively. Extracting the relative risk aversion from Equation (3.22),[18] Faccini et al. (2019) document that investor's implied relative risk aversion obtained from S&P 500 option quotes has forward-looking information and can

18. Estimate γ (IRRA) using the generalized method of moments (GMM) and a rolling window of 30 monthly observations.

predict economic conditions.

3.4 Predictive Power of Option-Implied Variables for Cross-Section of Returns

Market prices of options are forward-looking and may reflect some information regarding the future price of the underlying asset on the expiration day. Thus, an investor who trades based on her beliefs may trade in the options market to benefit from the information embedded in options (Black (1975), Easley et al. (1998)). For example, investors may choose the options market to trade based on their negative information or views by purchasing OTM puts.

A growing body of literature discusses the opportunity of exploiting forward-looking information from option prices to enhance the prediction of future asset returns- i.e., predicting abnormal stock returns using option-implied measures. As mentioned in the previous sections, number of studies, including An et al. (2014), Hu (2014), Cremers and Weinbaum (2010), Johnson and So (2012), Muravyev et al. (2013), Manaster and Rendleman (1982), and Xing et al. (2010) propose a variety of different option-based measures for predicting stock prices. However, researchers have no consensus regarding the relationship between implied measures and the stocks' returns. The main contribution of this section is to discuss the forecasting ability of option-implied predictors. We compare the predictability of eight option-implied measures under similar circumstances to reveal when they succeed in forecasting stocks' returns.

3.4.1 Data

We compute eight option-implied variables, namely, the risk-neutral model-free implied volatility (MFIV), Forward Variance (FV), Risk Neutral Skewness (RNS), Volatility Smirk (Smirk), implied volatility call-slope (IVS-Call), implied volatility put-slope (IVS-Put), put-call deviation (PCD), individual expected returns (ExRet). These variables are computed using index and individual stocks option prices. We collect the required data from OptionMetrics Ivy DB(OM) via the Wharton Research Data Services (WRDS) and Compustat. Our sample spans January 1996 to May 2019.

First, we obtain information on the S&P 500 index and its constituents from Compustat. Using the lists of index constituents, we searched the OptionMetrics database for all firms that formed the S&P 500 during our sampling period. Then, where available, we obtain the time series of implied volatility surface (IVS) data for these individual firms from the Optionsmetrics volatility surface file. The IVS file contains implied volatilities for standardized equity options with standardized maturities, including 30 calendar days, on a grid of the delta space accompanied by implied strike prices.[19] we remove options whose deltas or implied volatilities are missing. As mentioned in Section 3.2, curve-fitting methods are among the most popular methods to construct the implied-volatility surface. Therefore, we employ this method and interpolate and extrapolate the implied volatilities of OTM options in the moneyness dimension using cubic splines. Next, we convert implied volatilities to BS option prices to calculate all the required risk-neutral expectations.[20] In addition, we collect daily index and stock prices as well as data on the number of outstanding shares of each firm to compute their market capitalisations from OptionMetrics's underlying price file. Also, we collect forward prices for each firm from OM standardised option price file. Zero yield curve data is also obtained from the OM Zero Curve File.

Using equity prices and volatility surface data, we compute the risk-neutral option-implied variables of all individual firms that have been constituents of S&P 500 from 1996 to 2019 for horizons of one month. Panel A of Table 3.1 reports that we ended up with almost 2.5 million individual stock implied variables for one-month horizons, covering 1,092 firms over our sample period from January 1996 to May 2019. Over different horizons, we have data on 482 firms on average per day, indicating that we obtain almost 96% of the firms included in the S&P 500 index. Not all of these firms are constituents of the SPX index for the entire sample. Therefore, to avoid missing variables, on each day, we only focus on firms that have been included in SPX and

19. The theoretical procedures for computing option-based variables are established for European option prices. However, the exchange-listed individual U.S. equity options are Americans. To resolve this, the OM relies on the binomial tree and calculates IVs by considering early exercise premium. We follow the routine in the literature (e.g., Martin and Wagner (2019)) and rely on the IVs provided by the OM.

20. We employ a piecewise cubic Hermite to interpolate implied volatilities as a function of moneyness between available moneyness points. For ranges beyond left and right boundaries, we extrapolate using constant rule. Therefore, we fill in the moneyness range of [1/3,3], with a total of 1,001 points.

calculate the value-weighted average of the RN variance of the market resulting in a balanced panel of expected returns for 125 firms with 251 monthly observations from June 1998 to May 2019.

[Table 3.1 about here.]

3.4.2 Variables

Equipped with the dataset discussed above; first, we compute the individuals' total return as:

$$Ret_{t+1,i} = \log \frac{P_{t+1,i}}{P_{t,i}} \tag{3.23}$$

where $Ret_{t+1,i}$ is the one-month return on the ith firm at the end of month $t+1$. $P_{t+1,i}$ and $P_{t,i}$ denote the total price of ith firm at the end of months $t+1$ and t.

Next, following the literature, we parsimony and cautiously choose specific option-implied measures from the variables reviewed in previous sections and assess their predictive ability to forecast the cross-section of future returns. Due to space and time limitations, we do not consider all the option-implied variables studied in the previous sections.

Across different model-free implied volatilities reviewed in Section 3.3.1.1, we select MFIV proposed by BKM as our first option-implied predictor. MFIV of BKM is intrinsically linked to measures proposed by Carr and Madan (1998), Britten-Jones and Neuberger (2000) and Demeterfi et al. (1999). Therefore, we construct the 30-day model-free implied volatility ($MFIV_{t,i}$) of ith firm following Bakshi et al. (2003) with a portfolio of OTM options using Equation (3.3).

Considering an exponential claim on variance, Bakshi et al. (2011) assign different weights to (positions on) the OTM put and call options. Therefore, we compute the forward variance proposed by Bakshi et al. (2011) as our second option-implied predictor to examine the effect of this positioning on its predictive ability. We use OTM options data to compute at time t the forward variance, $FV_{t,T,i}$, between t and T— i.e., the forward variance with a $T-t = 1$-month horizon. We calculate $FV_{t,T}$ using Equations (3.7) to (3.10).

Further, we employ the RN skewness proposed by Bakshi et al. (2003) which measures the tail risk as our third predictor. As highlighted by Buss and Vilkov (2012), this measure is highly correlated with the tail loss measure (TLM) discussed in Section 3.3.2.3. Therefore, between the two, we focus our analysis on RN skewness and compute it following Bakshi et al. (2003) using a portfolio of OTM options via Equation (3.21) .

Our fourth and fifth option-implied predictors are the implied volatility slope of put and call options. Specifically, we calculate the left (right)-Implied volatility slope of the *ith* firm using OTM put (call) options (the slope of the implied volatility smile). Following Kelly et al. (2016), we define OTM put (call) options as options where the absolute values of their deltas are between 0.1 and 0.5. We regress the implied volatilities of options on their deltas and define the implied volatility slope as in Equation (3.18). These measures are inherently alternatives to $COMA$ and $POMA$ proposed by Doran and Krieger (2010).

Sixth, we calculate volatility smirk following Xing et al. (2010) as the difference between the average implied volatilities of OTM puts and ATM options. The smirk measure is computed following Equation (3.19). Following Xing et al. (2010), we define ATM options as options with a moneyness level (defined as the ratio of strike price to the current stock price) between 0.95 and 1.05. Moreover, a put option is defined as OTM when the ratio of strike price to the stock price is between 0.8 and 0.95.

Seventh, following Cremers and Weinbaum (2010), we define the deviation from put-call parity (PCD) as the difference between call and put implied volatilities. More precisely, this measure is the average difference in implied volatilities between call and put options (with the same strike price and maturity) and obtained via Equation (3.20).

Eighth, following Martin and Wagner (2019), we compute option-implied individuals' expected return (ExRet) using Equation (3.14). With respect to the expected returns obtained from Martin and Wagner (2019), looking closely at the feasible θ in Equation (3.16), we found that for 87% of the returns, the feasible θ is equal to one, and the derived expected returns bounds are based on the second moments- i.e., variance-based measures. Consequently, expected returns obtained from Martin and Wagner (2019) and Chabi-Yo et al. (2022) are highly correlated; hence, we focus our analysis on risk-neutral expected returns based on the second moment.

We left examining the predictive power of other option-implied variables (such as IRRA and variance risk premium) to forecast the cross-section of returns for future research.

3.4.3 Methodology

In this section, we compare the predictive ability of option-based variables introduced above to predict future stock returns using two different methodologies, namely panel regressions and portfolio sorts strategy.

- **Panel Regressions**

First, we use panel regressions to investigate whether option-implied measures can predict the one-month returns from a statistical perspective. We run panel regressions of monthly stock returns of firm i at time t to predict the one-month returns, controlling for time and firm fixed effects:

$$Ret_{t+1,i} = \alpha + \beta_1 X_{t,i} + \beta_2 Controls_{t,i} + \epsilon_{t+1,i} \tag{3.24}$$

where $Ret_{t+1,i}$ is the log-return[21] from time t to $t+1$ on firm i, α and βs denote the intercept and slope parameters, respectively. X_t is one of the forecasting variable(s) mentioned in Section 3.4.2 computed at the end of month t. $Controls_{t,i}$ is a 1×4 vector of control variables for firm i observed at month t. With this regression, we can assess not only the predictability of the implied variable but also control for firm characteristics. Therefore, we consider four control variables obtained from the equity market. These variables are documented by the previous literature to predict the cross-section of equity returns, including book-to-market ratio, lagged return, size and turnover.[22]

Table 3.2 provides summary statistics of option-implied variables and control variables in our sample. First, we compute the the averages over the cross-section for each month, and then we again average the statistics over the monthly time series. Table 3.3 shows

21. not annualized

22. Book to market is the book value of equity divided by the market value of equity. Size is the firm's market capitalisation. Turnover is the stock trade volume over the number of shares outstanding.

the cross sectional variations of option-implied variables and control variables in our sample.

[Table 3.2 about here.]

[Table 3.3 about here.]

- **Portfolio Sorts Trading Strategy**

In the previous section, we examine the statistical significance of the option-base predictors. In this section, we examine the economic significance of the in-sample predictability. To do so, we build long-short trading strategies with respect to each option implied measures. "Each month, we sort stocks into five portfolios based on each of the following option-implied measures: risk-neutral model-free implied volatility (MFIV), Forward Variance (FV), Risk Neutral Skewness (RNS), Volatility Smirk (Smirk), implied volatility call slope (IVSC), implied volatility put slope (IVSP), put-call deviation (PCD), and individual expected returns (ExRet). Following this methodology, Portfolio 1 contains firms with the lowest corresponding implied variables, and portfolio 5 includes firms with the highest implied variables. Then, we calculate returns for each portfolio over the next month. For each option-based measure, if we buy portfolio one and sell portfolio 5, the return on this long-short strategy indicates the economic significance of the sorting implied variable. In contrast to the panel regressions method, the portfolio sorting method does not impose a restricted linear relation between the option implied variables and the return. In addition, firm-level noise is reduced by grouping individual firms into portfolios.

3.4.4 Main Results

- **Results of Panel Regressions**

Panel A of Table 3.4 reports the regression coefficients of estimating Equation (3.24) which only includes the option-implied variable and a constant. However, we also consider four control variables in the regression model to distinguish the predictive power of option-implied variables from other firm characteristics. Panel B reports the regression coefficients of option-implied variables in a regression model (3.24) which controls for firms' characteristics variables.

The panel regression analysis indicates several interesting implications. First, its confirm that price, variance and tail risks are all priced in the cross-section of returns- i.e., the regression coefficient of implied variables are all positive except for the implied volatility slope of puts. Second, it shows that including the control variables does not reduce the predictive power of implied variables.

The volatility risk results are interpretable by economic theories. These theories posit that higher volatility stocks should earn higher expected returns. The MFIV, FV, and ExRet are statistically significant predictors in panel regressions documented by their $t-statistics$ of 4.77, 4.70 and 5.88, respectively- i.e., higher implied volatility results in higher future returns. These results suggest that stocks require compensation for taking high volatility risk- i.e., stock volatility may induce a risk premium and leads to a positive relationship between volatility and returns.

Furthermore, we found a strong positive relationship between RNS and returns with a $t-statistic$ of 5.27. Similar to Gkionis et al. (2021) and Buss and Vilkov (2012), we argue that stock overpricing is the source for this positive relationship. Stock overpricing may not be promptly updated in the underlying market due to arbitrage limitations. Therefore, traders may incline to the options market to trade based on their negative information or views. To maximise their leverage on their negative beliefs, they sell (buy) OTM calls (puts) instead of shorting the stock. Buying OTM is appealing compared to selling the underlying stock since the options market does not expose to the potential risks of holding stocks. In turn, the risk-averse market makers have to sell the underlying stock and, therefore, are exposed to negative inventory risks. Therefore, they induce a raising (lowering) price pressure on OTM puts (calls) to compensate for their risk. This mechanism is in agreement with the demand-based option pricing framework of Garleanu et al. (2009) and yields OTM puts (calls) relatively more (less) expensive and hence, lower RNS value. Subsequently, a low RNS value is accompanied by stock underperformance if investors recognise this procedure as an informative sign to update the stock overpricing. Similarly, we find that the information in volatility slopes is connected to future earnings shocks because firms with the steepest volatility slope of put options have the worst earnings surprises. This finding is also inconsistent with Xing et al. (2010).

Section 3.A in the appendix of the paper addresses economic concerns related to our regression model. Specifically, we implement fixed effects and clustered standard errors to tackle omitted variables bias and dependence in residuals. Tables 3.A.1 to A.3.8 evaluate the predictive ability of option-implied variables with various control variables in multiple regression models. The findings reveal that option-implied variables remain significant predictors, even after accounting for fixed effects and clustering errors.

Furthermore, we extract a single component (e.g., using PCA) from the option-implied variables and examine its predictive power for the cross-section of returns. The results are reported in Table A.3.9 in the Appendix and demonstrate that this single factor derived from option-implied measures exhibits significant predictive ability across all regression models, even after accounting for fixed effect errors and clustering in our regression model.

- **Results of Portfolio Sorts Trading Strategy**

Following the portfolio sorts strategy, we investigate whether implied option measures are priced risk factors of stock returns. Each month, we sort stocks into quintile portfolios by each of the following option-based measures: MFIV, FV, ExpRet, RNS, PCD, LIVS, RIVS, and smirk. Each quintile portfolio has 25 stocks on average.

Column 1, of Table 3.5 reports results for equal-weighted quintile portfolios formed with respect to MFIV- i.e, portfolio 1 (low), includes firms with the lowest MFIV and has a monthly return rate of 33 basis points. Portfolio 5 (high), contains firms with the highest MFIV and has a monthly return of 84 basis points. Portfolio 5 outperforms Portfolio 1 by 76 basis points per month with a $t-statistic$ of 1.89. This is in agreement with our assumption that higher volatility predicts higher earnings. Similarly, as shown in Table 3.5, portfolios with higher FV (column 2), risk-neutral skewness (column 3), and implied expected returns (column 8) earn significantly higher returns. However, if we long portfolio one and short portfolio 5 with respect to smirk (column 4) or implied volatility slope of put options (column 5), we earn lower returns- i.e., the excess return of the long-short strategy is 55 basis points and 41 basis points, respectively. Other implied variables, such as deviation from put-call parity and implied volatility slope of call options, do not show any significant predictive ability for future realised returns.

These results are consistent with panel regression coefficients and heuristically illustrate the economic significance of the predictive ability of option-implied variables for forecasting the cross-section of future returns.

[Table 3.4 about here.]

[Table 3.5 about here.]

- **Double Sort Portfolios**

We extend our analysis by using double-sorted portfolios to assess the relative importance of different option-implied measures. In previous sections, we discovered that a forecasting model incorporating risk-neutral skewness (RNS) among its predictors outperforms models with other option-implied variables in out-of-sample (OOS) forecasts. To ensure the robustness of our findings based on single sorts, while controlling for other implied variables, we conduct double-sorts.

Table 3.6 presents returns from various double sorts, where we alternate the sorting order between RNS and other characteristics. Panel A of Table 3.6 illustrates this process by first sorting assets into five quintiles based on their RNS and then further sorting them within each quintile based on model-free volatility (MFIV), forward, implied volatility slope of put options, and implied expected returns. For instance, $MFIV_1$ represents the lowest MFIV-ranked bond quintiles within each of the five RNS-ranked quintiles, while $MFIV_5$ represents the highest MFIV-ranked bond quintiles within each quintile.

The results in Panel A demonstrate a clear pattern, where the average portfolio returns increase monotonically from $MFIV_1$ to $MFIV_5$ quintile. Importantly, the return differences between quintiles 5 and 1 range from 0.23% to 0.62% per month and are statistically significant. These findings validate the economic significance of option-implied variables in predicting future returns, reinforcing the robustness of our conclusions based on double sorts and controlling for other implied variables.

[Table 3.6 about here.]

- **Effect of Transaction Costs**

In This section, we thoroughly investigate the impact of transaction costs on portfolios sorted by various option-implied measures. These measures include MFIV, Forward, skewness, Smirk, implied volatility slope of put and call options, put-call deviation, and implied expected return. Understanding how transaction costs influence the performance of these portfolios is crucial for assessing their practicality and real-world applicability. By considering transaction costs, we aim to gain a comprehensive understanding of the economic significance and viability of these option-implied measures in predicting returns in a more realistic trading environment. To estimate portfolio transaction costs, we adopt a method also used by Chordia et al. (2015), which considers trading illiquidity and portfolio turnover. Specifically, we use the bid-ask spread as an indicator of illiquidity, and the resulting transaction cost estimates are derived by multiplying the time-series average of this illiquidity measure with the time-series average of the portfolio turnover rate.

The outcomes are presented in Table 3.7, where we deduct these estimated transaction costs from the return spreads. Remarkably, our findings demonstrate that the estimated transaction costs are relatively small compared to the return spreads for the sorted portfolios. More importantly, even after accounting for transaction costs, the return spreads in the portfolios sorted by RNS, Smirk, IV SP, and implied expected return remain economically significant.

For example, when transitioning from quintile 1 to quintile 5, the average excess return on the skewness portfolios exhibits a nearly monotonic increase from -0.03% to 0.37% per month. This indicates a monthly average return difference of 0.41% between quintiles 5 and 1, with a t-statistic of 2.53, providing clear evidence of the robustness of our findings regarding the economic significance of option-implied measures in predicting returns, even when considering transaction costs.

[Table 3.7 about here.]

3.4.5 A Comparative Analysis of Forecasting Stock Returns with Option-Implied Variables

In this section, we conduct a "horse-race" comparison among different option-implied variables to determine their relative forecasting performance. To do so, we calculate the out-of-sample R^2 ($OOSR^2$) following the method proposed by Campbell and Cochrane (1999). This metric allows us to assess whether the variance explained by models using specific option-implied variables in their predictors is higher or lower than the variance explained by alternative models with other implied variables.

Panel A of Table 3.8 presents the $OOSR^2$, where the full model considers the implied variable in the column, and the benchmark model considers the implied variable in the row as part of its predictors. Panel B reports the relevant Diebold-Mariano (DM) test statistics for $OOSR^2$. The DM test statistic is derived from the out-of-sample forecast of a full model with a specific option-implied variable (in the column) compared to a benchmark model with another implied variable (in the row). A positive $OOSR^2$ indicates that the full model outperforms the benchmark model, indicating that the variable in the column has superior out-of-sample predictive power compared to the variable in the row.

Our findings reveal that the model incorporating risk-neutral skewness (RNS) exhibits significantly positive $OOSR^2$ compared to other models with different option-implied variables. This indicates that the forecasts made by the model containing RNS in its regressors outperform models with other implied variables, suggesting the superior predictive ability of RNS in our analysis

[Table 3.8 about here.]

3.5 Conclusions

Traders favour trading in different markets to exploit their informational leverage. Therefore, one market may be better than another market in price discovery. Among different markets, option markets are known for incorporating valuable information about the future state of the economy. In this paper, we study option markets, review several statistical measures inferred from individuals' option prices, and discuss different methods to obtain them. We also analyse their application and assess their information's usefulness in forecasting. Some empirical studies have found that the option-implied moments could be biased estimators of physical moments. However, others show that this bias may be negligible, and removing it may be challenging. Therefore, a bias in an option-implied moment does not prevent it from being a valuable estimator of future variables of interest.

We investigate whether the implied variables inferred from options contain useful information for predicting the underlying stock's future returns. We compare the economic and statistical predictive ability of several option-based variables to predict expected returns both in and out-of-sample. These variables include implied volatility, implied skewness, and measures that capture a specific position of the volatility curve (i.e., deviation from put-call parity proposed by Cremers and Weinbaum (2010), or implied-volatility slope measure, which reflects the slope of the volatility curve). Five of eight option-based variables, namely model-free implied volatility, forward variance, RN skewness, implied expected returns and implied volatility slope, capture most of the information in the cross-section of stock returns. These forecasting powers follow economic theories suggesting that higher volatility stocks should earn a higher expected return. In addition, our results are consistent with Garleanu et al. (2009) model, which shows that demand positively relates to option expensiveness.

Panel A: Daily Data for SPX firms					
horizon	30	60	90	182	365
Observations	2,837,127	2,837,127	2,837,127	2,837,127	2,837,127
Sample days	5892	5892	5892	5892	5892
Sample firms	1092	1092	1092	1092	1092
Average firms/day	482	482	482	482	482
Panel A: Monthly Data for SPX firms					
horizon	30	60	90	182	365
Observations	135,343	135,343	135,343	135,343	135,343
Sample months	281	281	281	281	281
Sample firms	1092	1092	1092	1092	1092
Average firms/month	482	482	482	482	482

Table 3.1. Sample Data: This table summarises the data used in the empirical analysis. We search the OptionMetrics database for all firms that have been included in the S&P 500 during the sample period from January 1996 to May 2019 and obtain all available volatility surface data. We use these data to compute firms' risk-neutral variances ($SVIX^2_{i,t}$) for horizons of 1, 2, 3, 6 and 12 months. Panel A summarises the number of total observations, the number of unique days and unique firms in our sample, as well as the average number of firms for which options data are available per day. We also compile data subsets at a monthly frequency for firms included in the S&P 500 (Panel B).

Variable	Mean	5%	25%	50%	75%	95%
Size*10^6	35.96	2.53	6.95	11.51	30.55	169.83
BM	0.35	0.06	0.18	0.29	0.48	0.84
Turnover	0.98	0.36	0.53	0.74	1.09	2.33
Ret	0.00	-0.18	-0.06	0.00	0.07	0.20
MFIV	0.12	0.08	0.10	0.12	0.14	0.18
FV	0.02	0.01	0.01	0.02	0.02	0.04
RNS	-0.25	-0.95	-0.48	-0.24	-0.02	0.49
Smirk	0.05	-0.02	0.02	0.05	0.08	0.14
IVS-Call	-0.10	-0.48	-0.15	-0.05	0.02	0.09
IVS-Put	0.18	-0.02	0.06	0.15	0.26	0.52
PC-Deviation	-0.02	-0.15	-0.06	-0.02	0.02	0.09
ExpRet	0.98%	0.53%	0.69%	0.87%	1.14%	1.83%

Table 3.2. Summary Statistics: Data are obtained from OptionMetrics. Our sample period is from May 1998 to July 2019. Variable SIZE is the firm's market capitalisation in $ billions. Variable BM is the book-to-market ratio. Variable TURNOVER is calculated as monthly volume divided by shares outstanding. Our option-based variables include Risk-neutral model-free implied volatility (MFIV), Forward Variance (FV), Risk Neutral Skewness (RNS), Volatility Smirk (Smirk), implied volatility call slope (IVS-Call), implied volatility put slope (IVS-Put), PC-Deviation (PCD) and the individual expected returns (ExRet). We first calculate the summary statistics over the cross-section for each month, and then we average the statistics over the monthly time series. For each month, there are, on average, 125 firms in the sample.

Statistic	N	Mean	St. Dev.	Min	Max
Size*10^6	30,747	52.12	78.90	0.81	1099.43
BM	30,753	0.457	0.430	0.360	5.757
Turnover	30,748	1.533	1.142	0.181	28.072
Returns	30,871	0.007	0.104	−0.853	7.965
MFIV	30,871	0.090	0.040	0.031	0.605
FV	30,871	0.011	0.012	0.001	0.358
RNS	30,871	−0.789	0.417	−10.650	2.793
Smirk	30,871	0.064	0.043	−1.357	0.805
IVS_Call	30,871	−0.033	0.155	−1.901	5.228
IVS_Put	30,871	0.219	0.166	−0.563	3.565
PC_Deviation	30,871	−0.007	0.082	−1.165	1.423
1+ExpectedRet	30,750	1.005	0.006	0.998	1.173

Table 3.3. Descriptive Statistics: This table reports the cross sectional statistics of firms' characteristics and option-implied variables. Data are obtained from OptionMetrics. Our sample period is from May 1998 to July 2019. Variable SIZE is the firm's market capitalisation in \$ billions. Variable BM is the book-to-market ratio. Variable TURNOVER is calculated as monthly volume divided by shares outstanding. Our option-based variables include Risk-neutral model-free implied volatility (MFIV), Forward Variance (FV), Risk Neutral Skewness (RNS), Volatility Smirk (Smirk), implied volatility call slope (IVSCall), implied volatility put slope (IVS-Put), PC-Deviation (PCD) and the individual expected returns (ExRet). We first calculate the summary statistics over the cross-section for each month, and then we average the statistics over the monthly time series. For each month, there are, on average, 125 firms in the sample.

Panel A. Panel regression without control variables

	MFIV	FV	RNS	Smirk	IVS-Put	IVS-Call	PCD	ExRet
coefficient	0.0701***	0.2369***	0.0075***	-0.0009	-0.0090**	-0.0078**	0.0136*	0.6130***
t-stat	(4.775)	(4.696)	(5.279)	(-0.0643)	(-2.521)	(-2.046)	(1.887)	(5.882)
Rsquared	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17

Panel B. Panel regression with control variables

	MFIV	FV	RNS	Smirk	IVS-Put	IVS-Call	PCD	ExRet
coefficient	0.0192	0.0914	0.0072***	-0.0052	-0.0135***	-0.0118***	0.0145**	0.3369***
t-stat	(1.112)	(1.589)	(5.024)	(-0.3744)	(-3.731)	(-3.017)	(2.017)	(2.772)
Rsquared	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17

Table 3.4. Predictability of Option-Implied Measures, Panel Regression: This table reports the regression results of monthly excess returns on a constant and the lagged predictive option implied variable(s) as specified in Equation (3.24). The option-based predictor is one of the followings: MFIV, which is a risk-neutral model-free implied volatility, Forward Variance (FV), Risk Neutral Skewness (RNS), Volatility Smirk (Smirk), implied volatility call slope (IVS-Call), implied volatility put slope (IVS-Put), PC-Deviation and the individual expected returns ExRet. . Panel A reports the regression coefficient of option-implied variables without controlling for firm-specific characteristics, while Panel B reports the regression coefficient of implied variable controlling for firm's effect. R-squared is the in-sample adjusted R-squared. *, **, and *** indicate the significance at the 10%, 5%, and 1% significance levels, respectively. d. The t-statistics, reported in parentheses, consider Heteroskedasticity of error terms and are calculated following White (1980)'s method. The sample period extends from July 1998 to May 2019. All data are sampled at a monthly frequency and relate to constituents of the SP 500 index.

Portfolio	MFIV	FV	RNS	Smirk	IVS-Put	IVS-Call	PCD	ExRet
1 (low)	0.33	0.31	0.25	1.17	0.94	0.95	0.56	0.28
2	0.43	0.45	0.57	0.36	0.78	0.36	0.65	0.45
3	0.59	0.58	0.55	0.43	0.70	0.47	0.64	0.56
4	0.76	0.77	0.78	0.55	0.59	0.60	0.69	0.77
5 (high)	0.84	0.88	0.87	0.62	0.53	0.68	0.66	0.89
high-low	0.51*	0.57**	0.62***	-0.55**	-0.41**	-0.27	0.10	0.61**
t-stat	(1.89)	(2.08)	(3.85)	(-2.62)	(-2.36)	(-1.32)	(0.74)	(2.22)
p-value	0.06	0.04	0.00	0.01	0.02	0.19	0.46	0.03

Table 3.5. Predictability of Option-Implied Measures, Portfolio Forming Approach: This table reports the average returns of portfolios formed based on option-based measures discussed in Section 3.4.2. These variables include model-free implied volatility (MFIV), forward variance (FV), risk-neutral skewness (RNS), put/call implied volatility slope (IVS-put(call)), volatility smirk, put-call deviation (PCD), and expected return (ExpRet) . Our sample period is from July 1998 to May 2019. At the end of each month, we sort the stocks into quintiles based on individual stock option-based variables and create a long-short portfolio by taking a long (short) position in stocks within the top (bottom) quantile. This long-short portfolio is held until the end of the subsequent month. The t-statistics and p-values for the abnormal return of long-short portfolios are reported over 251 months.

Stocks Sorted by RNS, then Controlling for:					
MFIV					
	(1) low	(2)	(3)	(4)	(5) high
$(MFIV_1)$ low	-0.19	0.43	0.34	0.46	0.48
(2)	0.21	0.57	0.24	0.70	0.87
(3)	0.16	0.56	0.43	0.72	0.87
(4)	0.37	0.64	0.64	0.86	0.99
$(MFIV_5)$ high	0.44	0.66	0.75	0.92	1.04
high-low	0.63	0.23	0.31	0.49	0.56
t-stat	1.95	0.69	1.01	1.64	1.92
p-value	0.05	0.49	0.31	0.10	0.06

Forward					
	(1) low	(2)	(3)	(4)	(5) high
$(Forward_1)$ low	-0.23	0.43	0.34	0.44	0.45
(2)	0.27	0.59	0.23	0.72	0.92
(3)	0.12	0.54	0.44	0.73	0.88
(4)	0.37	0.63	0.63	0.86	0.99
$(Forward_5)$ high	0.40	0.67	0.74	0.93	1.04
high-low	0.63	0.24	0.31	0.49	0.59
t-stat	1.95	0.72	1.00	1.65	2.01
p-value	0.05	0.47	0.32	0.10	0.05

Put_slope					
	(1) low	(2)	(3)	(4)	(5) high
$(IVSP_1)$ low	0.37	0.82	0.79	1.09	1.16
(2)	0.15	0.49	0.56	0.89	0.94
(3)	0.23	0.62	0.58	0.81	0.96
(4)	0.15	0.48	0.48	0.70	0.79
$(IVSP_5)$ high	0.28	0.47	0.45	0.63	0.72
high-low	-0.09	-0.34	-0.37	-0.18	-0.44
t-stat	-0.34	-1.18	-1.41	-0.74	-1.95
p-value	0.74	0.24	0.16	0.46	0.05

expected_return					
	(1) low	(2)	(3)	(4)	(5) high
$(ExpRet_1)$ low	-0.17	0.45	0.34	0.41	0.44
(2)	0.22	0.55	0.20	0.72	0.90
(3)	0.16	0.56	0.43	0.71	0.85
(4)	0.37	0.63	0.63	0.87	0.99
$(ExpRet_5)$ high	0.42	0.70	0.76	0.94	1.04
high-low	0.59	0.26	0.31	0.49	0.60
t-stat	1.81	0.77	1.00	1.65	2.02
p-value	0.07	0.44	0.32	0.10	0.04

Table 3.6. Bivariate Portfolios Sorted by RNS Controlling for other Option-Implied Variables: This table presents average returns for portfolios sorted by RNS measures controlling for different implied measures. At the end of each month, all the stocks in the sample are sorted into quintile portfolios based on ascending values of RNS. Then, within each quintile portfolio, stocks are further sorted based on the values of beme, lme, and turnover. Each portfolio is held for one month. The row labeled "High-Low" reports the average values of one-month-ahead returns.

	MFIV	Forward	RNS	Smirk	IVS_P	IVS_C	PCD	Exp_Ret
(1) low	0.18	0.16	-0.03	0.85	0.65	0.64	0.20	0.14
-2	0.17	0.18	0.22	0.00	0.41	0.01	0.28	0.19
-3	0.28	0.27	0.06	-0.08	0.19	-0.03	0.11	0.24
-4	0.51	0.51	0.38	0.15	0.23	0.19	0.24	0.51
(5) high	0.56	0.61	0.37	0.13	0.10	0.20	0.10	0.61
high-low	0.38	0.44	0.41	-0.72	-0.55	-0.44	-0.10	0.48
t-stat	1.40	1.62	2.53	-3.38	-3.08	-1.13	-0.66	1.75
p-value	0.16	0.11	0.01	0.00	0.00	0.23	0.51	0.08

Table 3.7. Univariate Portfolios Sorted by Option-implied Measures, considering for Transaction Costs: This table reports the average returns of portfolios formed based on option-based measures discussed in Section 3.4.2. These variables include model-free implied volatility (MFIV), forward variance (FV), risk-neutral skewness (RNS), put/call implied volatility slope (IVS-put(call)), volatility smirk, put-call deviation (PCD), and expected return (ExpRet). Our sample period is from July 1998 to May 2019. At the end of each month, we deduct the transaction costs from return and then sort the stocks into quintiles based on individual stock option-based variables and create a long-short portfolio by taking a long (short) position in stocks within the top (bottom) quantile. This long-short portfolio is held until the end of the subsequent month. The t-statistics and p-values for the abnormal return of long-short portfolios are reported over 251 months

Panel A: $OOSR^2$	MFIV	Forward	RNS	Smirk	Call Slope	Put Slope	PCD	EXP_ret	PCA factor
RMSE	14.24	14.10	13.97	14.09	14.05	14.00	14.04	14.11	14.02
MFIV	0.00%	1.93%	3.80%	2.20%	2.76%	3.43%	2.88%	1.91%	3.11%
Forward	-1.97%	0.00%	1.90%	0.28%	0.84%	1.52%	0.96%	-0.03%	1.20%
RNS	-3.95%	-1.94%	0.00%	-1.66%	-1.08%	-0.39%	-0.96%	-1.97%	-0.72%
Smirk	-2.25%	-0.28%	1.63%	0.00%	0.56%	1.25%	0.69%	-0.31%	0.93%
Call Slope	-2.83%	-0.85%	1.07%	-0.57%	0.00%	0.69%	0.12%	-0.87%	0.36%
Put Slope	-3.55%	-1.55%	0.39%	-1.26%	-0.69%	0.00%	-0.57%	-1.57%	-0.33%
PCD	-2.96%	-0.97%	0.95%	-0.69%	-0.12%	0.57%	0.00%	-1.00%	0.24%
EXP_ret	-1.94%	0.03%	1.93%	0.30%	0.87%	1.55%	0.99%	0.00%	1.23%
PCA factor	-3.21%	-1.21%	0.71%	-0.93%	-0.36%	0.33%	-0.24%	-1.24%	0.00%

Panel B: DM stats	MFIV	Forward	RNS	Smirk	Call Slope	Put Slope	PCD	EXP_ret	PCA factor
RMSE	14.24	14.10	13.97	14.09	14.05	14.00	14.04	14.11	14.02
MFIV	-	13.33	13.95	12.78	15.24	15.00	15.72	13.67	15.41
Forward	-13.33	-	10.66	2.17	6.79	9.68	7.81	-0.87	7.72
RNS	-13.95	-10.66	-	-10.69	-8.42	-3.89	-7.65	-10.41	-5.99
Smirk	-12.78	-2.17	10.69	-	14.05	13.49	14.77	-2.26	11.20
Call Slope	-15.24	-6.79	8.42	-14.05	-	9.63	4.31	-6.63	5.50
Put Slope	-15.00	-9.68	3.89	-13.49	-9.63	-	-9.16	-9.39	-4.11
PCD	-15.72	-7.81	7.65	-14.77	-4.31	9.16	-	-7.60	3.45
EXP_ret	-13.67	0.87	10.41	2.26	6.63	9.39	7.60	-	7.47
PCA factor	-15.41	-7.72	5.99	-11.20	-5.50	4.11	-3.45	-7.47	-

Table 3.8. Comparison between the Predictive Ability of option Implied variables: Top panel of the table reports the Diebold-Mariano test statistics obtained from the OOS forecast of a model with specific option-implied (in the column) variable comparing to another model with another implied variable (in the row). The bottom panel reports the $OOSR^2$ of comparing the model with the implied variable in the column versus the model with the implied variable in the row.

Appendix - Chapter Three

3.A Clustering of Standard Errors

When working with regressions in empirical finance, we need to deal with omitted variables bias and dependence in residuals. Therefore, in this section we focus on the implementation of fixed effects and clustered standard errors.

One way to tackle the issue of omitted variable bias is to get rid of as much unexplained variation as possible by including fixed effects - i.e., model parameters that are fixed for specific groups or time. In essence, each group has its own mean in fixed effects regressions. In addition, apart from biased estimators, we usually have to deal with potentially complex dependencies of our residuals with each other. Such dependencies in the residuals invalidate the i.i.d. assumption of OLS and lead to biased standard errors. With biased OLS standard errors, we cannot reliably interpret the statistical significance of our estimated coefficients.

In our setting, the residuals may be correlated across years for a given firm (time-series dependence), or, alternatively, the residuals may be correlated across different firms (cross-section dependence). One of the most common approaches to dealing with such dependence is the use of clustered standard errors. The idea behind clustering is that the correlation of residuals within a cluster can be of any form.

Tables 3.A.1 to A.3.8 assess the predictive ability of option-implied variables in the presence of some control variables in multiple regression models. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time. Our results suggest that even after relaxing the assumptions behind our standard errors, the regression coefficients of the option-implied variable are still comfortably significant as the statistics are well above the usual critical values.

The predictive ability of Forward and MFIV to forecast one-month return are reported in Table 3.A.1 and 3.A.2 and suggest that the Forward and model-free volatility are a significant predictor of one-month returns even after controlling for two-way fixed effect and clustering errors by firm.

[Table 3.A.1 about here.]

[Table 3.A.2 about here.]

The predictive ability of RNS and IVS_P to forecast one-month return are reported in Table A.3.3 and A.3.5. Our results suggest that both RNS and IV slope of put options remain as a significant predictor of one-month return even after considering a two-way fixed effect on firm and time and two-way clustering on time and firm. We couldn't find significant predictive ability of Smirk and put-call deviation to forecast one-month returns. However, Our results in Table A.3.8 suggests that the implied-expected return is a significant predictor of returns in regression models 1 to 4.

[Table 3.A.3 about here.]

[Table 3.A.4 about here.]

[Table 3.A.5 about here.]

[Table 3.A.6 about here.]

[Table 3.A.7 about here.]

[Table 3.A.8 about here.]

3.B Predictive Power of Option-Implied PCA Factor for Cross-Sectional Returns

In this Section, we extract a single component (e.g., using PCA) from the option-implied variables and examine its predictive power for the cross-section of returns. The results, reported in Table A.3.9, demonstrate that this single factor derived from option-implied measures exhibits significant predictive ability across all regression models, even after accounting for fixed effect errors and clustering in our regression model.

[Table 3.A.9 about here.]

Dependent Variable:	Returns				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	-0.0006				
	(-0.4402)				
lagged_return	-0.0047	-0.0068	-0.0356***	-0.0356***	-0.0356**
	(-0.6687)	(-0.9566)	(-4.771)	(-4.130)	(-2.339)
beme	0.0068***	0.0254***	0.0172***	0.0172***	0.0172***
	(4.851)	(9.655)	(6.725)	(3.060)	(2.766)
lme2	-0.0178**	-0.0750***	-0.0736***	-0.0736***	-0.0736**
	(-2.311)	(-5.039)	(-4.967)	(-2.635)	(-2.459)
lturnover	0.0027***	0.0023***	0.0018**	0.0018	0.0018
	(4.572)	(3.153)	(2.203)	(1.570)	(1.223)
Forward	0.0914	0.1113*	0.3779***	0.3779**	0.3779
	(1.589)	(1.871)	(4.878)	(2.536)	(1.063)
Fixed-effects					
permncode		Yes	Yes	Yes	Yes
date			Yes	Yes	Yes
Fit statistics					
Observations	30,747	30,747	30,747	30,747	30,747
R^2	0.00276	0.00733	0.18320	0.18320	0.18320
Within R^2		0.00557	0.00567	0.00567	0.00567

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3.A.1. Predictive Power of Forward Variance: This table assess the predictive ability of Forward volatility in the presence of some control variables in multiple regression models to forecast one-month returns. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time.

Dependent Variable:			Returns		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	-0.0014				
	(-0.7963)				
lagged_return	-0.0054	-0.0077	-0.0363***	-0.0363***	-0.0363**
	(-0.7559)	(-1.079)	(-4.870)	(-4.240)	(-2.476)
beme	0.0068***	0.0251***	0.0167***	0.0167***	0.0167***
	(4.813)	(9.560)	(6.511)	(2.915)	(2.636)
lme2	-0.0176**	-0.0759***	-0.0714***	-0.0714**	-0.0714**
	(-2.278)	(-5.059)	(-4.821)	(-2.581)	(-2.376)
lturnover	0.0028***	0.0024***	0.0019**	0.0019	0.0019
	(4.617)	(3.422)	(2.291)	(1.632)	(1.306)
MFIV	0.0192	0.0203	0.1213***	0.1213**	0.1213
	(1.112)	(1.115)	(4.237)	(2.395)	(1.201)
Fixed-effects					
permncode		Yes	Yes	Yes	Yes
date			Yes	Yes	Yes
Fit statistics					
Observations	30,747	30,747	30,747	30,747	30,747
R^2	0.00272	0.00725	0.18305	0.18305	0.18305
Within R^2		0.00549	0.00548	0.00548	0.00548

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3.A.2. Predictive Power of Model-Free Implied Volatility: This table assess the predictive ability of MFIV in the presence of some control variables in multiple regression models to forecast one-month returns. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time.

Dependent Variable:			Returns		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	0.0055***				
	(3.237)				
lagged_return	-0.0025	-0.0040	-0.0332***	-0.0332***	-0.0332**
	(-0.3553)	(-0.5677)	(-4.440)	(-3.706)	(-2.239)
beme	0.0070***	0.0255***	0.0172***	0.0172***	0.0172***
	(4.999)	(9.761)	(6.722)	(3.321)	(2.955)
lme2	-0.0169**	-0.0732***	-0.0701***	-0.0701***	-0.0701**
	(-2.204)	(-4.952)	(-4.741)	(-2.622)	(-2.384)
lturnover	0.0030***	0.0031***	0.0034***	0.0034***	0.0034*
	(5.678)	(4.603)	(4.613)	(3.449)	(1.798)
RNS	0.0072***	0.0083***	0.0095***	0.0095***	0.0095***
	(5.024)	(5.549)	(6.170)	(4.525)	(4.370)
Fixed-effects					
permncode		Yes	Yes	Yes	Yes
date			Yes	Yes	Yes
Fit statistics					
Observations	30,747	30,747	30,747	30,747	30,747
R^2	0.00349	0.00821	0.18359	0.18359	0.18359
Within R^2		0.00645	0.00614	0.00614	0.00614

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.3.3. Predictive Power of Risk-Neutral Skewness: This table assess the predictive ability of RNS in the presence of some control variables in multiple regression models to forecast one-month returns. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time.

Dependent Variable:			Returns		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	0.0003				
	(0.1687)				
lagged_return	-0.0065	-0.0087	-0.0419***	-0.0419***	-0.0419***
	(-0.9282)	(-1.247)	(-5.633)	(-4.847)	(-2.862)
beme	0.0067***	0.0249***	0.0176***	0.0176***	0.0176***
	(4.797)	(9.512)	(6.901)	(3.222)	(2.899)
lme2	-0.0189**	-0.0792***	-0.0698***	-0.0698**	-0.0698**
	(-2.458)	(-5.365)	(-4.718)	(-2.562)	(-2.336)
lturnover	0.0031***	0.0028***	0.0036***	0.0036***	0.0036*
	(5.847)	(4.120)	(4.809)	(3.561)	(1.880)
Smirk	-0.0052	-0.0088	0.0261*	0.0261	0.0261
	(-0.3744)	(-0.6081)	(1.808)	(1.186)	(1.198)
Fixed-effects					
permncode		Yes	Yes	Yes	Yes
date			Yes	Yes	Yes
Fit statistics					
Observations	30,747	30,747	30,747	30,747	30,747
R^2	0.00268	0.00723	0.18265	0.18265	0.18265
Within R^2		0.00547	0.00500	0.00500	0.00500

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.3.4. Predictive Power of Smirk: This table assess the predictive ability of Smirk in the presence of some control variables in multiple regression models to forecast one-month returns. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time.

Dependent Variable:			Returns		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	-0.0011				
	(-0.8480)				
lagged_return	-0.0054	-0.0075	-0.0377***	-0.0377***	-0.0377**
	(-0.7745)	(-1.074)	(-5.082)	(-4.456)	(-2.606)
beme	0.0066***	0.0246***	0.0170***	0.0170***	0.0170***
	(4.730)	(9.425)	(6.647)	(3.116)	(2.795)
lme2	-0.0156**	-0.0763***	-0.0671***	-0.0671**	-0.0671**
	(-2.016)	(-5.164)	(-4.525)	(-2.507)	(-2.301)
lturnover	0.0034***	0.0032***	0.0038***	0.0038***	0.0038**
	(6.352)	(4.714)	(5.069)	(3.669)	(1.986)
IVS_P	-0.0118***	-0.0135***	-0.0166***	-0.0166***	-0.0166***
	(-3.017)	(-3.259)	(-4.092)	(-3.425)	(-2.844)
Fixed-effects					
permncode		Yes	Yes	Yes	Yes
date			Yes	Yes	Yes
Fit statistics					
Observations	30,747	30,747	30,747	30,747	30,747
R^2	0.00297	0.00756	0.18301	0.18301	0.18301
Within R^2		0.00580	0.00544	0.00544	0.00544

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.3.5. Predictive Power of Implied-volatility Slope of Put Options: This table assess the predictive ability of IVS_P in the presence of some control variables in multiple regression models to forecast one-month returns. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time.

Dependent Variable:	Returns				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	0.0025^{*}				
	(1.740)				
lagged_return	-0.0064	-0.0086	-0.0401^{***}	-0.0401^{***}	-0.0401^{***}
	(-0.9144)	(-1.228)	(-5.408)	(-4.664)	(-2.724)
beme	0.0069^{***}	0.0250^{***}	0.0177^{***}	0.0177^{***}	0.0177^{***}
	(4.889)	(9.580)	(6.926)	(3.247)	(2.920)
lme2	-0.0204^{***}	-0.0794^{***}	-0.0705^{***}	-0.0705^{**}	-0.0705^{**}
	(-2.652)	(-5.385)	(-4.761)	(-2.552)	(-2.331)
lturnover	0.0033^{***}	0.0032^{***}	0.0036^{***}	0.0036^{***}	0.0036^{*}
	(6.236)	(4.701)	(4.874)	(3.607)	(1.933)
IVS_C	-0.0135^{***}	-0.0154^{***}	-0.0028	-0.0028	-0.0028
	(-3.731)	(-4.051)	(-0.7165)	(-0.4782)	(-0.4487)
Fixed-effects					
permncode		Yes	Yes	Yes	Yes
date			Yes	Yes	Yes
Fit statistics					
Observations	30,747	30,747	30,747	30,747	30,747
R^2	0.00313	0.00775	0.18258	0.18258	0.18258
Within R^2		0.00599	0.00491	0.00491	0.00491

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.3.6. Predictive Power of Implied-volatility Slope of Call Options: This table assess the predictive ability of IVS_C in the presence of some control variables in multiple regression models to forecast one-month returns. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time.

Dependent Variable:	Returns				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	4.87×10^{-5}				
	(0.0379)				
lagged_return	-0.0070	-0.0093	-0.0405***	-0.0405***	-0.0405***
	(-0.9994)	(-1.323)	(-5.476)	(-4.792)	(-2.789)
beme	0.0067***	0.0247***	0.0176***	0.0176***	0.0176***
	(4.791)	(9.426)	(6.910)	(3.266)	(2.931)
lme2	-0.0190**	-0.0807***	-0.0703***	-0.0703**	-0.0703**
	(-2.478)	(-5.461)	(-4.751)	(-2.551)	(-2.333)
lturnover	0.0031***	0.0027***	0.0036***	0.0036***	0.0036*
	(5.848)	(4.104)	(4.836)	(3.575)	(1.878)
PCD	0.0145**	0.0140*	0.0027	0.0027	0.0027
	(2.017)	(1.927)	(0.3487)	(0.3010)	(0.2827)
Fixed-effects					
permncode		Yes	Yes	Yes	Yes
date			Yes	Yes	Yes
Fit statistics					
Observations	30,747	30,747	30,747	30,747	30,747
R^2	0.00281	0.00733	0.18257	0.18257	0.18257
Within R^2		0.00558	0.00490	0.00490	0.00490

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.3.7. Predictive Power of Put-Call Deviation: This table assess the predictive ability of PCD in the presence of some control variables in multiple regression models to forecast one-month returns. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time.

Dependent Variable:			Returns		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	-0.3372***				
	(-2.768)				
lagged_return	-0.0033	-0.0054	-0.0362***	-0.0362***	-0.0362**
	(-0.4590)	(-0.7624)	(-4.843)	(-4.155)	(-2.354)
beme	0.0068***	0.0253***	0.0170***	0.0170***	0.0170***
	(4.825)	(9.616)	(6.668)	(2.964)	(2.685)
lme2	-0.0187**	-0.0736***	-0.0696***	-0.0696**	-0.0696**
	(-2.434)	(-4.931)	(-4.700)	(-2.489)	(-2.336)
lturnover	0.0022***	0.0017**	0.0018**	0.0018	0.0018
	(3.686)	(2.409)	(2.185)	(1.569)	(1.204)
expRet	0.3369***	0.3677***	0.8725***	0.8725**	0.8725
	(2.772)	(2.922)	(5.239)	(2.503)	(1.107)
Fixed-effects					
permncode		Yes	Yes	Yes	Yes
date			Yes	Yes	Yes
Fit statistics					
Observations	30,627	30,627	30,627	30,627	30,627
R^2	0.00287	0.00737	0.18356	0.18356	0.18356
Within R^2		0.00560	0.00583	0.00583	0.00583

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.3.8. Predictive Power of Implied Expected Returns: This table assess the predictive ability of implied expected returns in the presence of some control variables in multiple regression models to forecast one-month returns. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time.

Dependent Variable:			Returns		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	-0.0007				
	(-0.5062)				
lagged_return	-0.0020	-0.0042	-0.0336***	-0.0336***	-0.0336**
	(-0.2797)	(-0.5952)	(-4.490)	(-3.758)	(-2.261)
beme	0.0069***	0.0255***	0.0172***	0.0172***	0.0172***
	(4.913)	(9.761)	(6.754)	(3.338)	(2.969)
lme2	-0.0170**	-0.0734***	-0.0703***	-0.0703***	-0.0703**
	(-2.219)	(-4.971)	(-4.748)	(-2.619)	(-2.383)
lturnover	0.0033***	0.0031***	0.0035***	0.0035***	0.0035*
	(6.281)	(4.668)	(4.658)	(3.427)	(1.803)
PCAfactor	0.0079***	0.0081***	0.0088***	0.0088***	0.0088***
	(5.522)	(5.684)	(6.024)	(4.267)	(4.062)
Fixed-effects					
permncode		Yes	Yes	Yes	Yes
date			Yes	Yes	Yes
Fit statistics					
Observations	30,747	30,747	30,747	30,747	30,747
R^2	0.00366	0.00826	0.18354	0.18354	0.18354
Within R^2		0.00650	0.00608	0.00608	0.00608

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.3.9. Predictive Power of PCA factor: This table assess the predictive ability of PCA factor obtained from various option-implied variables (namely, MFIV, Forward, RNS, IVS_Put, IVS_Call, Smirk, PCD and implied expected return) in the presence of some control variables in multiple regression models to forecast one-month returns. In these tables, model (1) does not considers the fixed effects and clustering. Model (2) is a is a one-way fixed effect on firm. Model (3) is a is a two-way fixed effect on firm and time. Model (4) is a two-way fixed effect on time and firm and also one-way clustering by firm. Model (5) is a two-way fixed effect and two-way clustering regression model on firm and time.

Conclusions

In Chapter One, we developed a prediction regression model to assess the predictive ability of information embedded in individual optionable stocks to forecast U.S. real economic activity (REA) growth. We construct our option-based predictor by relying on data reduction methods. We find a significant negative predictive relation between our proposed option-implied predictor and REA both in and out-of-sample. The predictive ability of our option-based measure remains even after controlling for well-known REA predictors in our prediction model. Our findings also imply individual stocks contain some additional predictive power that is not captured by standard risk factors or index options. We separately employ Principal Component Analysis (PCA) and the Instrumented Principal Component Analysis (IPCA) factors to assess whether results may differ by incorporating the wealth of firms' characteristic information into our predictor. Our results suggest that options are sophisticated derivatives and may already subsume firms' characteristics; hence the IPCA does not necessarily improve predicting REA.

We left three topics related to Chapter Two for future research: (i) Extending the analysis for other countries, (ii) Developing a theoretical modelling framework to explain the negative relationship between option-implied expected return and REA (iii) Employing other data reduction techniques such as the three-pass regression filter (3PRF) of Kelly and Pruitt (2015).

In Chapter Two, we measure the impact of FOMC meetings on volatility and tail risks and the demand for options. We used treatment options traded before FOMC meetings and expiring after the meetings and compared them to other options whose lives do not span FOMC meetings. We find that the implied volatility, its slope on the left of the ATM level, and the variance risk premium (VRP) are larger for options whose lives span FOMC meetings. These findings are consistent with Kelly et al. (2016) and suggest that investors are willing to pay for protection against the price, downside tail risk and variance associated with these events. Higher prices may also compensate for the risk that market makers bear for trading with informed traders. These higher prices are associated with an increase in the bid-ask spread and, more generally, a decrease in the liquidity of options. In addition, we conduct a time series analysis of risk-neutral moments and trading activity indicators around FOMC meetings. Our findings reveal that the impacts of these events on options markets are most pronounced in the week before the event and last until a few days after the event. In addition, we found that a positive shock to downside risk and the put-call ratio predicts an increase in the Fed rate. It implies that when investors become more worried about a decline in stock market prices, they predict a contractionary monetary policy- i.e., future short-term rates to increase.

At least three topics fall beyond this chapter's scope, and we leave them for future analysis. These topics include (i) Performing tests to identify the channel whereby decreased liquidity in options is associated with increased implied volatility and decreased risk-neutral skewness, (ii) Identifying a link between options trading and pre-announcement drift reported by Neuhier and Weber (2018) and (iii) Finding a trading strategy to exploit market inefficiencies before FOMC meetings- i.e., trading variance swaps one week before the event when the VRP peaks.

In Chapter Three, we study option markets, review several statistical measures inferred from individuals' option prices, and discuss different methods to obtain them. We also analyse their application and assess their information's usefulness in forecasting. We investigate whether the implied variables inferred from options contain useful information for predicting the underlying stock's future returns. We find a statistical and economic negative predictive relation between volatility measures, namely model-free implied volatility, implied forward variance, and implied expected returns with future stock returns cross-sectionally. Furthermore, we find a positive (negative) relationship between RNS (left-IVS) and the cross-section of future returns.

Due to space and time limitations, we do not consider all the option-implied variables studied in Chapter three. We leave the examination of the predictive ability of implied-risk aversion, variance risk premium, and tail loss measure for future research.

Bibliography

Abadir, K. M. and Rockinger, M. (2003), 'Density functionals, with an option-pricing application', *Econometric Theory* **12**, 778–811.

Ackerer, D., Tagasovska, N. and Vatter, T. (2020), 'Deep smoothing of the implied volatility surface', *Corporate Governance: Disclosure* **20(3)**, 45–67.

Ai, H. and Bansal, R. (2018), 'Risk preferences and the macro announcement premium', *Econometrica* **86**, 1383–1430.

Ait-Sahalia, Y. and Duarte, J. (2003), 'Nonparametric option pricing under shape restrictions', *Derivatives* **116**, 9–47.

Ait-Sahalia, Y. and Mancini, M. K. L. (2020), 'The term structure of equity and variance risk premia', *Journal of Econometrics* **219**, 204–230.

Aït-Sahalia, Y., Wang, Y. and Yared, F. (2001), 'Do option markets correctly price the probabilities of movement of the underlying asset', *Journal of Econometrics* **102**, 67–110.

Aıt-Sahalia, Y. and Lo, A. W. (2000), 'Nonparametric risk management and implied risk aversion', *Journal of Econometrics* **94**(1-2), 9–51.

Allen, L., Bali, T. G. and Tang, Y. (2012), 'Does systemic risk in the financial sector predict future economic downturns', *The Review of Financial Studies* **25**(10), 3000–3036.

Altavilla, C., Giannone, D. and Modugno, M. (2017), 'Low frequency effects of macroeconomic news on government bond yields', *Journal of Monetary Economics* **92**, 31–46.

Alòs, E., León, J. A. and Vives, J. (2007), 'On the short-time behaviour of the implied volatility for jump-diffusion models with stochastic volatility', *Finance and Stochastics* **11**, 571–589.

Amengual, D. and Xiu, D. (2018), 'Resolution of policy uncertainty and sudden declines in volatility', *Journal of Econometrics* .

Amin, K.and C. Lee, . (1997), 'Option trading, price discovery, and earnings news dissemination.', *Contemporary Accounting Research* **14**, 153–192.

An, B., Ang, A., Bali, T. G. and Cakici, N. (2014), 'The joint cross section of stocks and options', *The Journal of Finance* **69**(5), 2279–2337.

Andersen, T. G. and Bondarenko, O. (2009), Dissecting the pricing of equity index volatility. Working Paper.

Andersen, T. G., Bondarenko, O. and Gonzalez-Perez, M. T. (2011), Coherent model-free implied volatility: A corridor fix for high-frequency vix.

Andersen, T. G., Bondarenko, O. and Gonzalez-Perez, M. T. (2015), 'Exploring return dynamics via corridor implied volatility', *Review of Financial Studies* **28**(10), 2902–2945.

Andersen, T. G., Fusari, N. and Todorov, V. (2019), 'The pricing of tail risk and the equity premium: Evidence from international option markets', *Journal of Business &amp Economic Statistics* **38**(3), 662–678.

Andreou, P., Kagkadis, A., Philip, D. and Tuneshev, R. (2018), 'Differences in options investors' expectations and the cross-section of stock returns', *Journal of Banking and Finance* **94**, 315–336.

Ang, A., Hodrick, R. J., Xing, Y. and Zhang, X. (2006), 'The cross-section of volatility and expected returns', *Journal of Finance* **61**(1), 259–299.

Aruoba, S., Diebold, F. and Scotti, C. (2009), 'Real-time measurement of business conditions', *Journal of Business Economic Statistics* **27**, 417–427.

Baker, S. R., Bloom, N. and Davis, S. J. (2013), 'Measuring economic policy uncertainty', *The Quarterly Journal of Economics* .

Bakshi, G., Cao, C. and Chen, Z. (1997), 'Empirical performance of alternative option pricing models', *Journal of Finance* **52**, 2003–2049.

Bakshi, G., Carr, P. and Wu, L. (2008), 'Stochastic risk premiums, stochastic skewness in currency options, and stochastic discount factors in international economies', *Journal of Financial Economics* **87**, 132–156.

Bakshi, G., Kapadia, N. and Madan, D. (2003), 'Return characteristics, skew laws, and the differential pricing of individual equity options', *Review of Financial Studies* **16**, 101–143.

Bakshi, G. and Madan, D. B. (2006), 'A theory of volatility spreads', *Manag. Sci.* **52**, 1945–1956.

Bakshi, G., Panayotov, G. and Skoulakis, G. (2011), 'Improving the predictability of real economic activity and asset returns with forward variances inferred from option portfolios', *Journal of Financial Economics* .

Bakshi, G., Panayotov, G. and Skoulakis, G. (2012), The baltic dry index as a predictor of global stock returns,commodity returns, and global economic activity. Chicago meeting.

Balakrishnan, K., Bartov, E. and Faurel, L. (2010), 'Post loss/profit announcement drift', *Journal of Accounting and Economics* **50**(1), 20–41.

Bali, T. G. and Hovakimian, A. (2009), 'Volatility spreads and expected stock returns', *Management Science* **55**(11), 1797–1812.

Bali, T. G., Hu, J. and Murray, S. (2019), 'Option implied volatility, skewness, and kurtosis and the cross-section of expected stock returns', *Singapore Management University Lee Kong Chian School of Business Research Paper Series* .

Bali, T. G. and Murray, S. (2013), 'Does risk-neutral skewness predict the cross-section of equity option portfolio returns?', *Journal of Financial and Quantitative Analysis* **48**, 1145 – 1171.

Balla, R., Gerakosa, J., Linnainmaaa, J. T. and Nikolaev, V. V. (2015), 'Deflating profitability', *Journal of Financial Economics* **115**, 225–248.

Baltussen, G., van der Grient, B., de Groot, W., Hennink, E. and Zhou, W. (2011), 'Exploiting option information in the equity market', *Financial Analysts Journal* **68**, 56 – 72.

Balvers, R. J., Cosimano, T. F. and Mcdonald, B. (1990), 'Predicting stock returns in an efficient market', *Journal of Finance* **45**, 1109–1128.

Barndorff-Nielsen, O. E. (1998), 'Processes of normal inverse gaussian type', *Finance and Stochastics* **2**, 41–68.

Basu, S. (1983), 'The relationship between earnings' yield, market value and return for nyse common stocks: Further evidence', *Journal of Financial Economics* **12**(1), 129–156.

Bates, D. S. (1991), 'The crash of '87: Was it expected? the evidence from options market.', *Journal of Finance* **46**, 1009–1044.

Bates, D. S. (1996), 'Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options.', *The Review of Financial Studies* **9**, 69–107.

Bates, D. S. (2000), 'Post-'87 crash fears in the sp 500 futures option market', *Journal of Econometrics* **9**, 181–238.

Beber, A. and Brandt, M. W. (2006), 'The effect of macroeconomic news on beliefs and preferences: Evidence from the options market', *Journal of Monetary Economics* **53**, 1997–2039.

Becker, R., Clements, A. E. and White, S. I. (2007), 'Does implied volatility provide any information beyond that captured in model-based volatility forecasts?', *Journal of Banking and Finance* **31**(8), 2535–2549.

Beckmeyer, H., Branger, N. and Grunthaler, T. (2021), Fed tails: Fomc announcements and stock market uncertainty, Technical report. Working Paper.

Benhamou, E., Gobet, E. and Miri, M. (2010), 'Time dependent heston model', *SIAM J. Financial Math.* **1**, 289–325.

Benko, M., Fengler, M., Härdle, W. K. and Kopa, M. (2007), 'On extracting information implied in options', *Computational Statistics* **22**, 543–553.

Bernales, A. and Valenzuela, M. (2016), 'Implied correlation and market returns', *Capital Markets: Asset Pricing & Valuation eJournal* .

Bernanke, B. (1983), 'Nonmonetary effects of the financial crisis in the propagation of the great depression,', *American Economic Review* **73**(3), 257–276.

Bernanke, B. S. and Kuttner, K. N. (2005), 'What explains the stock market's reaction to federal reserve policy', *The Journal of Finance* **60**, 1221–1257.

Bhandari, L. C. (1988), 'Debt/equity ratio and expected common stock returns: Empirical evidence', *The Journal of Finance* **43**(2), 507–528.

Bharadia, M. A., Christofides, N. and Salkin, G. R. (1996), 'Computing the black scholes implied volatility', *Advances in Futures and Options Research* **8**, 15–29.

Birru, J. and Figlewski, S. (2009), 'Anatomy of a meltdown: The risk neutral density for the s&p 500 in the fall of 2008', *Wharton Research Data Services (WRDS) Research Paper Series* **15**, 151–180.

Black, F. (1975), 'Fact and fantasy in the use of options', *Financial Analysts Journal* **31**, 36–41.

Black, F. and Scholes, M. S. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637 – 654.

Blair, B., Poon, S.-H. and Taylor, S. (1999), 'Forecasting s&p 100 volatility: the incremental information content of implied volatilities and high frequency index returns', *Journal of Econometrics* **105**, 5–26.

Bliss, R. R. and Panigirtzoglou, N. (2002), 'Testing the stability of implied probability density functions', *Journal of Banking and Finance* **26**, 381–422.

Bliss, R. R. and Panigirtzoglou, N. (2004), 'Option-implied risk aversion estimates', *Journal of Finance* **59**, 407–446.

Bollen, N. P. B. and Whaley, R. E. (2004), 'Does net buying pressure affect the shape of implied volatility functions?', *Journal of Finance* **59**, 711–753.

Bollerslev, T., Gibsond, M. and Zhou, H. (2011), 'Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities', *JournalofEconometrics* **160**, 235–245.

Bollerslev, T., Li, J. and mei Xue, Y. (2016), 'Volume, volatility, and public news announcements', *The Review of Economic Studies* **85**, 2005–2041.

Bollerslev, T., Tauchen, G. and Zhou, H. (2009*a*), 'Expected stock returns and variance risk premia', *Review of Financial Studies* **22**(11), 4463–4492.

Bollerslev, T., Tauchen, G. and Zhou, H. (2009*b*), 'Expected stock returns and variance risk premiums', *Review of Financial Studies* **22**, 4463–4492.

Bollerslev, T. and Todorov, V. (2011), 'Estimation of jump tails', *Econometrica* **79**(6), 1727–1783.

Bollerslev, T. and Zhou, H. (2002), 'Estimating stochastic volatility diffusion using conditional moments of integrated volatility', *Journal of Econometrics* **109**, 33–65.

Bollerslev, T. and Zhou, H. (2006), 'Volatility puzzles: a simple framework for gauging return-volatility regressions', *Journal of Econometrics* **131**, 123–150.

Bondarenko, O. (2003), 'Estimation of risk-neutral densities using positive convolution approximation', *Journal of Econometrics* **116**, 85–112.

Bondarenko, O. (2014), 'Why are put options so expensive?', *Quarterly Jurnal of Finance* **4**(62–75).

Bondt, W. F. M. D. and Thaler, R. (1985), 'Does the stock market overreact?', *The Journal of Finance* **40**(3), 793–805.

Borochin, P. and Zhao, Y. (2022), 'Risk neutral skewness predicts price rebounds and so can improve momentum performance', *Capital Markets: Asset Pricing & Valuation eJournal* .

Breeden, D. and Litzenberger, R. (1978), 'Prices of state-contingent claims implicit in option prices.', *The Journal of Business* **51**, 621–652.

Brennan, M. (1979), 'The pricing of contingent claims in discrete time models.', *The Journal of Finance* **24**, 53–68.

Brenner, M. and Subrahmanyan, M. G. (1988), 'A simple formula to compute the implied standard deviation', *Financial Analysts Journal* **44**, 80–83.

Britten-Jones, M. and Neuberger, A. (2000), 'Option prices, implied price processes, and stochastic volatility', *Journal of Finance* **55**, 839–866.

Brogaard, J. and Detzel, A. (2015), 'The asset pricing implications of government economic policy uncertainty', *Management Science* .

Brown, K. C., Harlow, W. V. and Tinic, S. M. (1988), 'Risk aversion, uncertain information, and market efficiency', *Journal of Financial Economics* **22**(2), 355–385.

Bruce J. Sherrick, Philip Garcia, V. T. (1996), 'Recovering probabilistic information from option markets: Test of distributional assumptions', *he Journal of Future Markets* **16**, 545–560.

Brunnermeier, M. K., Gollier, C. and Parker, J. A. (2007), 'Optimal beliefs, assetprices, and the preference for skewed returns', *American Economic Review* **97**, 159–165.

Buchen, P. W. and Kelly, M. (1996), 'The maximum entropy distribution of an asset inferred from option prices', *Journal of Financial and Quantitative Analysis* **31**, 143–159.

Buraschi, A., Kosowski, R. and Trojani, F. (2013), 'When there is no place to hide: Correlation risk and the cross-section of hedge fund returns', *Review of Financial Studies* **27(2)**, 581–616.

Busch, T., Christensen, B. J. and Nielsen, M. Ø. (2011), 'The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets', *Journal of Econometrics* **160**(1), 48–57.

Buss, A., Schoenleber, L. and Vilkov, G. (2019), 'Expected correlation and future market returns', *CEPR Discussion Paper* **DP12760**.

Buss, A., Schönleber, L. and Vilkov, G. (2017), 'Option-implied correlations, factor models, and market risk', *INSEAD Working Paper No. 2017/20/FIN* .

Buss, A. and Vilkov, G. (2012), 'Measuring equity risk with option-implied correlations', *Review of Financial Studies* **25**(10), 3113–3140.

Bustamant, M. C. and Donangelo, A. (2017), 'Product market competition and industry returns', *The Review of Financial Studies* **30**(12), 4216–4266.

Byun, S. J. and Kim, J. S. (2013), 'The information content of risk-neutral skewness for volatility forecasting', *Journal of Empirical Finance* **23**, 142–161.

Callaway, B. and Sant'Anna, P. H. (2021), 'Difference-in-differences with multiple time periods', *Journal of Econometrics* **225**(2), 200–230.

Campa, J. M. and Chang, K. (1998), 'The forecasting ability of correlations implied in foreign exchange options', *Journal of International Money and Finance* **17**, 855–880.

Campa, J. M., Chang, P. H. K. and Reider, R. (1997), 'Implied exchange rate distributions: Evidence from otc option markets', *Capital Markets: Asset Pricing & Valuation* **117-160**, 117–160.

Campbell, J. Y. and Cochrane, J. H. (1999), 'A consumption-based explanation of aggregate stock market behavior', *Journal of Political Economy* .

Campbell, J. Y. and Thompson, S. B. (2008), 'Predicting excess stock returns out of sample: Can anything beat the historical average?', *The Review of Financial Studies* **21**(4), 1509–1531.

Canina, L. and Figlewski, S. (1993), 'The informational content of implied volatility', *Review of Financial Studies* **6**, 659–681.

Cao, C., Griffin, J. M. and Chen, Z. (2005), 'Informational content of option volume prior to takeovers', *The Journal of Business* **78**, 1073–1109.

Cao, J., Goyal, A., Xiao, X. and Zhan, X. (2019), 'Implied volatility changes and corporate bond returns', *Swiss Finance Institute Research Paper* pp. 19–75.

Cao, J., Goyal, A., Xiao, X. and Zhan, X. (2022), 'Implied volatility changes and corporate bond returns', *Management Science* .

Carr, P., Geman, H., Madan, D. B. and Yor, M. (2002), 'The fine structure of asset returns: an empirical investigation', *Journal of Business* **75**, 305–332.

Carr, P. and Lee, R. (2008), Robust replication of volatility derivatives.

Carr, P., Lee, R. and Lorig, M. J. (2021), 'Robust replication of volatility and hybrid derivatives on jump diffusions', *Mathematical Finance* **31**, 1394 – 1422.

Carr, P. and Madan, D. (1998), Towards a theory of volatility trading, *in* 'Option Pricing, Interest Rates and Risk Management', Cambridge University Press, pp. 458–476.

Carr, P., Madan, D. B. and Stanley, M. (2001), Option pricing, interest rates and risk management: Towards a theory of volatility trading.

Carr, P. P., Chang, E. C. and Madan, D. B. (1998), 'The variance gamma process and option pricing model', *European Financial Review* **2**, 79–195.

Carr, P. and Wu, L. (2007), 'Stochastic skew in currency options', *Journal of Financial Economics* **86**, 465–498.

Carr, P. and Wu, L. (2009), 'Variance risk premiums', *Review of Financial studies* .

Chabi-Yo, F., Dim, C. and Vilkov, G. (2022), 'Generalized bounds on the conditional expected excess return on individual stocks', *Management Science* .

Chalamandaris, G. and Tsekrekos, A. E. (2011), 'How important is the term structure in implied volatility surface modeling? evidence from foreign exchange options.', *Journal of International Money and Finance* **30**, 623–640.

Chan, K., Ge, L. G. and Lin, T.-C. (2015), 'Informational content of options trading on acquirer announcement return', *Journal of Financial and Quantitative Analysis* **50**, 1057–1082.

Chang, B.-Y., Christoffersen, P., Jacobs, K. and Vainberg, G. (2012), 'Option-implied measures of equity risk', *Review of Finance* **16**, 385–428.

Charles J. Corrado, T. S. (1996), 'Skewness and kurtosis in sp 500 index returns implied by option prices', *The Journal of Finance* **19**, 175–192.

Chen, H., Joslin, S. and Ni, S. X. (2019), 'Demand for crash insurance, intermediary constraints, and risk premia in financial markets', *The Review of Financial Studies* **32**, 228–265.

Chen, N.-F. (1991), 'Financial investment opportunities and the macroeconomy', *The Journal of Finance* **46**(2), 529–554.

Chen, Y. and Gwati, R. (2014), 'Understanding exchange rate dynamics: What does the term structure of fx options tell us?', *Working Paper* .

Chordia, T., Lin, T.-C. and Xiang, V. (2020), 'Risk-neutral skewness, informed trading, and the cross section of stock returns', *Journal of Financial and Quantitative Analysis* **56**, 1713 – 1737.

Chordia, T., Roll, R. and Subrahmanyam, A. (2015), 'Commonality in liquidity', *Journal of Financial Economics* **117**(2), 532–534.

Christoffersen, P. F., Heston, S. and Jacobs, K. (2009), 'The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well', *Manag. Sci.* **55**, 1914–1932.

Christoffersen, P. F., Jacobs, K. and Chang, B. Y. (2011), 'Forecasting with option-implied information', *Derivatives eJournal* .

Christoffersen, P. F. and Mazzotta, S. (2005), 'The accuracy of density forecasts from foreign exchange options', *Journal of Financial Econometrics* **3**, 578–605.

Christoffersen, P., Goyenko, R., Jacobs, K. and Karoui, M. (2018), 'Illiquidity premia in the equity options market', *The Review of Financial Studies* **31**(3), 811–851.

Christoffersen, P., Jacobs, K. and Chang, B. Y. (2012), 'Forecasting with option-implied information', Handbook of Economic Forecasting.

Chung, K. H. and Zhang, H. (2014), 'A simple approximation of intraday spreads using daily data', *Journal of Financial Markets* **17**(C), 94–120.

Cieslak, A., Morse, A. and Vissing-Jorgensen, A. (2018), 'Stock returns over the fomc cycle', *Journal of Finance* .

Cieslak, A., Morse, A. and Vissing-Jorgensen, A. (2019), 'Stock returns over the fomc cycle', *Journal of Financial* **74**, 2201–2248.

Cochrane, J. H. and Piazzesi, M. (2002), 'The fed and interest rates - a high-frequency identification', *American Economic Review* **92**(2).

Collin-Dufresne, P. and Goldstein, R. S. (2001), 'Stochastic correlation and the relative pricing of caps and swaptions in a generalized-affine framework', *Columbia: Finance & Economics (Topic)* .

Conrad, J., Dittmar, R. F. and Ghysels, E. (2013), 'Ex ante skewness and expected stock returns', *The Journal of Finance* **68**, 85–124.

Constantinides, G. M. and Lian, L. (2021), 'The supply and demand of s&p 500 put options', *Critical Finance Review* **10**.

Cont, R. (2006), 'Model uncertainty and its impact on the pricing of derivative instruments', *Mathematical Finance* **16**(3), 519 – 547.

Cont, R. and da Fonseca, J. S. (2002), 'Dynamics of implied volatility surfaces', *Quantitative Finance* **2**, 45 – 60.

Cont, R. and Tankov, P. (2004), 'Non-parametric calibration of jump-diffusion option pricing models', *Journal of Computational Finance* **7**, 1–49.

Cont, R. and Tankov, P. (2006), 'Retrieving levy pro cesses from option prices: Regularization of an iii-posed inverse problem', *SIAM Journal on Control and Optimization* **45**, 1–25.

Cooper, M. J., Gulen, H. and Schill, M. J. (2008), 'Asset growth and the cross-sectionof stock returns', *The Journal of Finance* **63**(4), 1609–1651.

Cooper, N. (1999), 'The yen / dollar exchange rate in 1998 : views from options markets', *Working Paper* .

Corcuera, J. M., Guillaume, F., Leoni, P. and Schoutens, W. (2009), 'Implied lévy volatility', *Quantitative Finance* **9**, 383–393.

Corrado, C. J. (n.d.), 'Option pricing based on the generalized lambda distribution', *Derivatives* **21**(3), 213–236.

Coutant, S., Jondeau, E. and Rockinger, M. (2001), 'Reading pibor futures options smile: the 1997 snap election', *Journal of Banking and Finance* **25**, 1957–1987.

Cox, J. and Ross, S. (1976), 'The valuation of options for alternative stochastic processes', *Journal of Financial Economics* **3**, 145–166.

Cremers, M., Fodor, A. and Weinbaum, D. (2022), 'Where do informed traders trade first? option trading activity, news releases, and stock return predictability', *Management Science, Forthcoming* .

Cremers, M., Halling, M. and Weinbaum, D. (2015), 'Aggregate jump and volatility risk in the cross-section of stock returns', *The Journal of Finance* **70**, 577–614.

Cremers, M. and Weinbaum, D. (2010), 'Deviations from put-call parity and stock return predictability', *The Journal of Financial and Quantitative Analysis* **45**, 335–367.

Croce, M. M., Kung, H., Nguyen, T. T. and Schmid, L. (2012), 'Fiscal policies and asset prices', *Review of Financial Studies* **25**, 2635–2672.

Culp, C., Gandhi, M., Nozawa, Y. and Veronesi, P. (2021), Option-implied spreads and option risk premia, Technical report.

Culp, C. L., Nozawa, Y. and Veronesi, P. (2018), 'Option-based credit spreads', *American Economic Review* **108**(2), 454–488.

Damghani, B. M. (2015), 'Introducing the implied volatility surface parametrization (ivp): Application to the fx market', *Wilmott* **2015**, 68–81.

Damodaran, A. (2019), 'Equity risk premiums (erp): Determinants, estimation and implications – the 2019 edition', *New York University Stern School of Business Research Paper Series* .

Daniel, K., Hirshleifer, D. and Teoh, S. H. (2002), 'Investor psychology in capital markets: evidence and policy implications', *Journal of Monetary Economics* .

Das, S. R. and Foresi, S. (1996), 'Exact solutions for bond and option prices with systematic jump risk', *Review of Derivatives Research* **1**, 7–24.

Datar, V. T., Naik, N. Y. and Radcliffe, R. (1998), 'Liquidity and stock returns: An alternative test', *Journal of Financial Markets* **1**(2), 203–2019.

Datta, D. D., Londoño, J. M. and Ross, L. (2017), 'Generating options-implied probability densities to understand oil market events', *Energy Economics* **64**, 40–457.

David, A. and Veronesi, P. (2014), 'Investors' and central bank's uncertaintyembedded in index options', *The Review of Financial Studies* **27**(6), 1661–1716.

Day, T. E. and Lewis, C. M. (1992), 'Stock market volatility and the information content of stock index options', *Journal of Econometrics* **52**, 267–287.

De Jong, F., Driessen, J. and Pelsser, A. (2004), 'On the information in the interest rate term structure and option prices', *Review of Derivatives Research* **7**, 99–127.

Demeterfi, K., Derman, E., Kamal, M. and Zou, J. (1999), 'A guide to volatility and variance swaps', *The Journal of Derivatives* **6**(4), 9–32.

DeMiguel, V., Plyakha, Y., Uppal, R. and Vilkov, G. (2013), 'Improving portfolio selection using option-implied volatility and skewness', *Journal of Financial and Quantitative Analysis* **48**, 1813–1845.

Dennis, P. and Mayhew, S. (2002), 'Risk-neutral skewness: Evidence from stock options', *Journal Of Financial And Quantitative Analysis* **37**(3).

Derman, E. and Kani, I. (1998), 'Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility', *International Journal of Theoretical and Applied Finance* **1**, 61–110.

Diavatopoulos, D., Doran, J. S. and Peterson, D. R. (2008), 'The information content in implied idiosyncratic volatility and the cross-section of stock returns: Evidence from the option markets', *Journal of Futures Markets* **28**(11), 1013–1039.

Diebold, F. X. and Li, C. (2002), 'Forecasting the term structure of government bond yields', *Journal of Econometrics* **130**, 337–364.

Donders, W., Kouwenberg, R. and Vorst, T. (2000), 'Options and earnings announcements: an empirical study of volatility, trading volume, open interest and liquidity', *European Financial Management* **6,**, 149–171.

Doran, J. S., Carson, J. M. and Peterson, D. R. (2006), 'Market crash risk and implied volatility skewness: Evidence and implications for insurer investments', *Corporate Finance: Valuation* .

Doran, J. S. and Krieger, K. (2010), 'Implications for asset returns in the implied volatility skew', *Financial Analysts Journal* **66**, 65–76.

Doran, J. S., Peterson, D. R. and Tarrant, B. C. (2007), 'Is there information in the volatility skew?', *Journal of Futures Markets* **27**(10), 921–959.

Drechsler, I., Savo, A. and Shnab, P. (2018), 'A model of monetary policy and risk premia', *Journal of Finance* **73**, 317–373.

Drechsler, I. and Yaron, A. (2011), 'What's vol got to do with it', *Review of Financial Studies* **24**, 1–45.

Driessen, J., Maenhout, P. J. and Vilkov, G. (2009), 'The price of correlation risk: Evidence from equity options', *Journal of Finance* **64**, 1377–1406.

Driessen, J., Maenhout, P. and Vilkov, G. (2012), 'Option-implied correlations and the price of correlation risk', *Working paper, INSEAD.* .

Du, J. and Kapadia, N. (2012), Tail and volatility indices from option prices. Working paper.

Duan, J. C. and Zhang, W. (2014), 'Forward-looking market risk premium', *Management Science* **60**, 521–538.

Dubinsky, A. and Johannes, M. (2005), 'Earnings announcements and equityoptions,', *Working paper* .

Dubinsky, A. and Johannes, M. (2006), 'Fundamental uncertainty, earning announcements and equity options,', *Working paper, Columbia University* .

Duca, M. L. and Stracca, L. (2015), 'Worth the hype? the effect of g20 summits on global financial markets', *Journal of International Money and Finance* **53**, 192–217.

Duffie, D., Pan, J. and Singleton, K. (2000), 'Transform analysis and asset pricing for affine jump diffusions', *Econometrica* **68**, 1343–1376.

Dumas, B., Fleming, J. and Whaley, R. E. (1998), 'Implied volatility functions: Empirical tests', *The Journal of Finance* **53**, 2059–2106.

Dupire, B. (1994), Pricing with a smile, Vol. 7, pp. 18–20.

D'Acuntoa, F., CarolinPfluegerc, R. and Weberde, M. (2018), 'Flexible prices and leverage', *Journal of Financial Economics* **129**(1), 46–68.

Easley, D., O'hara, M. and Srinivas, P. (1998), 'Option volume and stock prices: Evidence on where infromed traders trade', *The Journal of Finance* .

Ederington, L. H. and Jae H. Lee, . (1996), 'The creation and resolution of market uncertainty: The impact of information releases on implied volatility', *Journal of Financial and Quantitative Analysis* **31**, 513–539.

Erenburg, G. and Lasser, D. (2009), 'Electronic limit order book and order submission choice around macroeconomic news', *The Review of Financial Studies* **18**(4), 172–182.

Estrella, A. and Hardouvelis, G. (1991), 'The term structure as a predictor of real economic activity', *The Journal of Finance* **46**(2), 557–576.

Evgeny Lyandres, L. S. and Zhang, L. (2008), 'The new issues puzzle: Testing the investment-based explanation', *Review of Financial Studies* **21**(6), 2825–2855.

Faccini, R., Konstantinidi, E., Skiadopoulos, G. and Sarantopoulou-Chiourea, S. (2019), 'A new predictor of real economic activity: The s&p 500 option implied risk aversion', *Management Science* **65**(10).

Fama, E. F. and French, K. R. (1989), 'Business conditions and expected returns on stocks and bonds', *Journal of Business Economic Statistics* **25**, 23–49.

Fama, E. F. and French, K. R. (1992), 'The cross-section of expected stock returns', *The Journal of Finance* **47**(2), 427–465.

Fama, E. F. and French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**(1).

Fama, E. F. and French, K. R. (1996), 'Multifactor explanations of asset pricing anomalies', *The Journal of Finance* **51**(1), 55–84.

Fama, E. and French, K. (1998), 'Value versus growth: The international evidence.', *Journal of Finance* **53**(6), 1975–1999.

Fama, E. and French, K. (2015), 'A five-factor asset pricing model', *Journal of Financial Economics* **116**(1), 1–22.

Faria, G., Kosowski, R. and Wang, T. (2022), 'The correlation risk premium: International evidence', *Journal of Banking and Finance* **136**.

Fengler, M. (2009), 'Arbitrage-free smoothing of the implied volatility surface', *Quantitative Finance* **9**, 417 – 428.

Fernández, P. (2009), 'Market risk premium used in 2008 by professors: A survey with 1,400 answers', *Political Economy: International Political Economy eJournal* .

Fernández-Villaverde, J., Guerrón-Quintana, P., Kuester, K. and Rubio-Ramírez, J. (2015), 'Fiscal volatility shocks and economic activity', *American Economic Review, American Economic Association* **105(11)**, 3352–3384.

Figlewski, S. (2008), 'Estimating the implied risk neutral density', *New York University Stern School of Business Research Paper Series* .

Fleming, J., Ostdiek, B. and Whaley, R. E. (1995), 'Predicting stock market volatility: A new measure', *Journal of Futures Markets* **15**(3), 265–302.

Forde, M., Jacquier, A. and Mijatović, A. (2010), 'Asymptotic formulae for implied volatility in the heston model', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **466**, 3593 – 3620.

Forde, M., Jacquier, A. and Mijatović, A. (2010), 'Asymptotic formulae for implied volatility in the heston model', *Proceedings: Mathematical, Physical and Engineering Sciences* **466**, 3593–3620.

Fouque, J.-P., Papanicolaou, G. and Sircar, K. R. (2000), 'Derivatives in financial markets with stochastic volatility', *Cambridge University Press* .

Francois, P., Galarneau-Vincent, R., Gauthier, G. and Godin, F. (2016), 'Venturing into uncharted territory: An extensible parametric implied volatility surface model', *working paper* .

François, P., Galarneau-Vincent, R., Gauthier, G. and Godin, F. (2022), 'Venturing into uncharted territory: An extensible parametric implied volatility surface model', *CGN: Risk Management* .

Frazzinia, A. and Pedersen, L. H. (2014), 'Betting against beta', *Journal of Financial Economics* **111**, 1–25.

French, D. W., Groth, J. C. and Kolari, J. W. (1983), 'Current investor expectations and better betas', *The Journal of Portfolio Management* **10**(1), 12–17.

Freyberger, J., Neuhierl, A. and Weber, M. (2020), 'Dissecting characteristics nonparametrically', *Review of Financial Studies* **35**(5), 2326–2377.

Friedman, B. M. and Kuttner, K. N. (1991), 'Why does the paper-bill spread predict real economic activity?', *NBER Working Papers 3879, National Bureau of Economic Research, Inc* .

Frijns, B., Tallau, C. and Tourani-Rad, A. (2009), 'The information content of implied volatility: Evidence from australia', *Journal of Futures Markets* .

Gandhi, P. and Lustig, H. (2015), 'Size anomalies in u.s. bank stock returns', *The Journal of Finance* **79**, 733–768.

Gao, P. and Qi, Y. (2013), 'Political uncertainty and public nancing costs: Evidence from u.s. gubernatorial elections and municipal bond markets', *Working paper, University of Notre Dame* .

Garcia, R. and Gencay, R. (2000), 'Pricing and hedging derivative securities with neural networks and a homogeneity hint', *Journal of Econometrics* **94**, 93–115.

Garleanu, N. B., Pedersen, L. H. and Poteshman, A. M. (2009), 'Demand-based option pricing', *Review of Financial Studies* **22**, 4259–4299.

Gemmill, G. (1992), 'Political risk and market eciency: Tests based in british stock and options markets in the 1987 election', *Journal of Banking and Finance* **16**, 211–231.

Gemmill, G. and Saflekos, A. (2000), 'How useful are implied distributions?', *Options." Journal of Derivatives* **7**, 83–98.

George, T. J. and Hwang, C. (2004), 'The 52-week high and momentum investing', *The Journal of Finance* **59**(5), 2145–2176.

Giacomini, R., Gottschling, A., Haefke, C. and White, H. L. (2008), 'Mixtures of t-distributions for finance and forecasting', *Journal of Econometrics* **144**, 175–192.

Giamouridis, D. and Tamvakis, M. N. (2002), 'Asymptotic distribution expansions in option pricing', *The Journal of Derivatives* **9**, 33–44.

Gilbert, T. (2011), 'Information aggregation around macroeconomic announcements: Revisions matter', *Journal of Financial Economics* **101**(1), 114–131.

Gilchrist, S. and Zakrajsek, E. (2012), 'Credit spreads and business cycle fluctuations', *American Economic Review* **102**, 1692–1720.

Giot, P. and Laurent, S. (2007), 'The information content of implied volatility in light of the jump/continuous decomposition of realized volatility', *Journal of Futures Markets* **27**, 337–359.

Gkionis, K., Kostakis, A., Skiadopoulos, G. and Stilger, P. S. (2021), 'Positive stock information in out-of-the-money option prices', *Journal of Banking & Finance* .

Glosten, L. R. and Milgrom, P. (1985), 'Bid, ask, and transaction prices in a specialist market withheterogenously informed traders', *Journal of Financial Economics* **14**, 71–100.

González-Urteaga, A., Nieto, B. and Rubio, G. (2021), 'Extracting expected stock risk premia from option prices and the information contained in non-parametric-out-of-sample stochastic discount factors', *Quantitative Finance* **21**(5), 713–727.

Gonçalves, S. and Guidolin, M. (2006), 'Predictable dynamics in the sp 500 index options implied volatility surface', *Journal of Business* **79**, 1591–1635.

Goodell, J. W. and Vahamaa, S. (2013), 'Us presidential elections and implied volatility: The role of political uncertainty', *Journal of Banking and Finance* **37**, 1108–1117.

Gorodnichenko, Y. and Weber, M. (2016), 'Are sticky prices costly? evidence from the stock market', *American Economic Review* **106**(1).

Goyal, A. and Saretto, A. (2009), 'Cross-section of option returns and volatility', *Journal of Financial Economics* **94**, 310–326.

Green, R. C. and Jarrow, R. A. (1987), 'Spanning and completeness in markets with contingent claims', *Journal of Economic Theory* **41**, 202–210.

Hagan, P. S., Kumar, D., Lesniewski, A. and Woodward, D. E. (2002), 'Managing smile risk', *Wilmott Magazine* **September**, 84–108.

Han, B. (2007), 'Stochastic volatilities and correlations of bond yields', *Journal of Finance* **62**, 1491–1524.

Handa, P., Schwartz, R. and Tiwari, A. (2006), 'Quote setting and price formation in an order driven market', *Journal of Financial Markets* **4**(3), 461–489.

Harris, R. D. F., Li, X. and Qiao, F. (2019), 'Option-implied betas and the cross section of stock returns', *Journal of Futures Markets* **39**, 94–108.

Harvey, C. (1989), 'Forecasts of economic growth from the bond and stock markets', *Financial Analysts Journal* **45**(5), 38–45.

Harvey, D., Leybourne, S. and Newbold, P. (1997), 'Testing the equality of prediction mean squared errors', *International Journal of Forecasting* **13**, 281–291.

Haugen, R. A. and Baker, N. L. (1996), 'Commonality in the determinants of expected stock returns', *Journal of Financial Economics* **41**, 401–439.

Healy, J., Gregoriou, A. and Hudson, R. (2017), 'Test of recent advances in extracting information from option prices', *International Review of Financial Analysis* **56**, 292–302.

Heston, S. (1993), 'A closed-form solution for options with stochastic volatility with applications to bond and currency options', *Review of Financial Studies* **6**, 327–343.

Hibbs, D. (1977), 'Political parties and maroeconomic policy,', *American Political Science Review* .

Hirshleifer, D., Hou, K., Teoh, S. H. and Zhang, Y. (2004), 'Do investors overvalue firms with bloatedbalance sheets?', *Journal of Accounting and Economics* **38**, 297–331.

Hodges, S. D. and Zhao, B. (2013), 'Parametric modeling of implied smile functions: A generalized svi model', *Review of Derivatives Research* **16**, 53–7.

Homescu, C. (2011), 'Implied volatility surface: Construction methodologies and characteristics', *Financial Engineering eJournal* .

Hong, H. and Yogo, M. (2012), 'What does futures market interest tell us about the macroeconomy and asset prices?', *Journal ofFinancialEconomics* **105**.

Houa, K., van Dijkb, M. A. and Zhang, Y. (2012), 'The implied cost of capital: A new approach', *Journal of Accounting and Economics* **53**, 504–526.

Hu, J. (2013), 'Does option trading convey stock price information?', *Singapore Management University Lee Kong Chian School of Business Research Paper Series* .

Hu, J. (2014), 'Does option trading convey stock price information?', *Journal of Financial Economics* **111**(3), 625–645.

Hull, J. (2000), *options, futures and other derivatives*, 4 edn, Prentice Hall.

Israelov, R. and Kelly, B. T. (2017), 'Forecasting the distribution of option returns', *Mutual Funds* .

Itkin, A. (2019), 'Deep learning calibration of option pricing models: some pitfalls and solutions', *arXiv: Computational Finance* .

Itô, K. (1950), 'Stochastic differential equations in a differentiable manifold', *Nagoya Mathematical Journal* **1**, 35–47.

Jackwerth, J. (2004), 'Option-implied risk-neutral distributions and risk aversion', *Charlotteville: Research Foundation of AIMR* .

Jackwerth, J. C. and Rubinstein., M. (1996), 'Recovering probability distributions from option prices', *Journal of Finance* **51**, 1611–1631.

James L. Davis, E. F. F. and French, K. R. (2000), 'Characteristics, covariances, and average returns: 1929 to 1997', *Journal of Finance* **55**, 389–406.

Jarrow, R. and Rudd, A. (1982), 'Approximate option valuation for arbitrary stochastic processes', *Journal of Financial Economics* **10**, 347–369.

JEGADEESH, N. (1990), 'Evidence of predictable behavior of security returns', *The Journal of Finance* **45**(3), 881–898.

Jiang, G. J. and Tian, Y. S. (2005), 'The model-free implied volatility and its information content', *Review of Financial Studies* **18**(4), 1305–1342.

Jiang, G. J. and Tian, Y. S. (2007), 'Extracting model-free volatility from option prices: An examination of the vix index', *Risk Management* .

Jianqing, F., Furger, A. and Xiu, D. (2016), 'Incorporating global industrial classification standard into portfolio allocation: A simple factor-based large covariance matrix estimator with high frequency data', *Journal of Business Economic Statistics* **34**, 489–503.

Jin, W., oshua Livnat and Zhang, Y. (2012), 'Option prices leading equity prices: Do option traders have an information advantage?', *Journal of Accounting Research* **50**, 401–432.

Johnson, T. L. and So, E. C. (2012), 'The option to stock volume ratio and future returns', *Journal of Financial Economics* **106**(2), 262–286.

Jondeau, E. and Rockinger, M. (2000), 'Reading the smile: the message conveyed by meth-ods which infer risk neutral densities', *Journal of Internationnal Money and Finance* **19**, 885–915.

Jong, C. D. and Huisman, R. (2000), 'From skews to a skewed-t: modelling option-implied returns by a skewed student-t', *Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering (CIFEr) (Cat. No.00TH8520)* pp. 132–142.

Jurek, J. W. and Xu, Z. (2014), 'Option-implied currency risk premia', *ERN: International Finance (Topic)* .

Kadan, O. and Tang, X. (2019), 'A bound on expected stock returns', *The Review of Financial Studies* **33**(4).

Kahale, N. (2004), 'An arbitrage-free interp olation of volatilities', *Risk Magazine* **17**, 102–106.

Kambouroudis, D. S., McMillan, D. G. and Tsakou, K. (2016), 'Forecasting stock return volatility: A comparison of garch, implied volatility, and realized volatility models', *Journal of Futures Markets* **36**, 1127–1163.

Kang, B. J. and Kim, T. S. (2006), 'Option-implied risk preferences: An extension to wider classes of utility functions', *Journal of Financial Markets* **9**(2), 180–198.

Kang, B. J., Kim, T. S. and Yoon, S. (2010), 'Information content of volatility spreads', *The Journal of Futures Markets* **30**(6), 533–558.

Kang, C. M. and Geul, L. (2016), 'Stock return predictability of out-of-the-money option trading', *SSRN Electronic Journal* .

Kaya, O. (2018), 'Economic policy uncertainty in europe. detrimental to capital markets and bank lending.', *Deutsche Bank research* .

Kelly, B. and Jiang, H. (2014), 'Tail risk and asset prices', *The Review of Financial Studies* **27**(10).

Kelly, B., Pastor, L. and Veronesi, P. (2016), 'The price of political uncertainty: Theory and evidence from the option market', *The journal of finance* .

Kelly, B. and Pruitt, S. (2015), 'The three-pass regression filter: A new approach to forecasting using many predictors', *Journal of Econometrics* .

Kelly, B. and Pruitt, S. (2019), 'Characteristics are covariances', *Journal of Financial Economics* **134**, 501–524.

Kempf, A., Korn, O. and Sassning, S. (2015), 'Portfolio optimization using forward-looking information', *Review of Finance* **19**, 467–490.

Khrapov, S. (2015), 'Option pricing via risk-neutral density forecasting', *2nd International Workshop on "Financial Markets and Nonlinear Dynamics" (FMND)* .

Kitsul, Y. and Wright, J. H. (2012), 'The economics of options-implied inflation probability density functions', *Journal of Financial Economics* **110**, 696–711.

Konstantinidi, E. and Skiadopoulos, G. (2016), 'How does the market variance risk premium vary over time? evidence from s&ampp 500 variance swap investment returns', *Journal of Banking and Finance* **62**, 62–75.

Kostakis, A., Magdalinos, T. and Stamatogiannis, M. P. (2015), 'Robust econometric inference for stock return predictability', *The Review of Financial Studies* **28**(5), 1506–1553.

Kostakis, A., Panigirtzoglou, N. and Skiadopoulos, G. (2011), 'Market timing with option-implied distributions: A forward-looking approach', *Management Science* **57**(7), 1231–1249.

Kou, S. (2002), 'A jump-diffusion model for option pricing', *Manag. Sci.* **48**, 1086–1101.

Kourtis, A., Markellos, R. N. and Symeonidis, L. (2016), 'An international comparison of implied, realized, and garch volatility forecasts', *Journal of Futures Markets* **36**, 1164–1193.

Lakonishok, J., Lee, I., Pearson, N. D. and Poteshman, A. M. (2007), 'Option market activity', *Review of Financial Studies* **20**, 813–857.

Lamoureux, C. G. and Lastrapes, W. D. (1993), 'Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities', *Review of Financial Studies* **6**, 293–326.

Laurini, M. P. (2011), 'Imposing no-arbitrage conditions in implied volatilities using constrained smoothing splines', *Applied Stochastic Models in Business and Industry* **27**, 649–659.

Leiss, M. and H.N, H. (2018), 'Option-implied objective measures of market risk', *Journal of Banking and Finance* **88**, 241–249.

Lemmon, M. and Ni, S. X. (2014), 'Differences in trading and pricing between stock and index options', *Management Science* **60**.

Lewellen, J. (2015), 'The cross-section of expected stockreturns', *Critical Finance Review* **4**, 1–44.

Li, H. and Zhao, F. (2009), 'Nonparametric estimation of state-price densities implicit in interest rate cap prices', *Review of Financial Studies* **22**, 4335–4376.

Li, J. and Born, J. A. (2006), 'Presidential election uncertainty and common stock returns in the united states', *Journal of Financial Research* .

Liew, J. and Vassalou, M. (2000), 'Can book-to-market, size and momentum berisk factors that predict economic growth?', *Journal of Financial Economics* **57**.

Liu, S., Borovykh, A., Grzelak, L. and Oosterlee, C. (2019), 'A neural network-based framework for financial model calibration', *Journal of Mathematics in Industry* **9**, 1–28.

Liu, S., Oosterlee, C. and Bohte, S. (2019), 'Pricing options and computing implied volatilities using neural networks', *Risks* **7**, 16.

Loflund, A. and Nummelin, K. (1997), 'On stocks, bonds and business conditions', *Applied Financial Economics* **7**(2), 137–146.

Longstaff, F. A., Santa-Clara, P. and Schwartz, E. S. (2001), 'The relative valuation of caps and swaptions: Theory and empirical evidence', *Journal of Finance* **56**, 2067–2109.

Lopez, J. A. and Walter, C. (2000), 'Is implied correlation worth calculating?', *Journal of Derivatives* **7**, 65–82.

Lucas, R. (1978), 'Asset prices in an exchange economy', *Econometrica* **46**, 1429–1445.

Lucca, D. O. and Moench, E. (2015), 'The pre-fomc announcement drift.', *The Journal of Finance* **70**, 329–371.

MacBeth, J. D. and Merville, L. J. (1979), 'An empirical examination of the black-scholes call option pricing model', *Journal of Finance* **34**, 1173–1186.

Mahdavi-Damghani, B. and Kos, A. (2013), 'De-arbitraging with a weak smile: Application to skew risk', *Wilmott Magazine* **64**, 40–49.

Malz, A. M. (1997), 'Option-implied probability distributions and currency excess returns', *Banking and Financial Institutions* .

Manaster, S. and Rendleman, R. J. (1982), 'Option prices as predictors of equilibrium stock prices', *Journal of Finance* **37**, 1043–1057.

Martin, I. (2011), 'Simple variance swaps', *Risk Management eJournal* .

Martin, I. (2017), 'What is the expected return on the market', *The Quarterly Journal of Economics* **132**, 367–433.

Martin, I. and Wagner, C. (2019), 'What is the expected return on a stock?', *The Journal of Finance* .

Melick, W. R. and Thomas, C. P. (1997), 'Recovering an asset's implied pdf from option prices: An application to crude oil during the gulf crisis', *Journal of Financial and Quantitative Analysis* **32**, 91–115.

Merton, R. C. (1973), 'Theory of rational option pricing', *The Review of Financial Studies* **4**, 141–183–107.

Merton, R. C. (1976), 'Option pricing when underlying stock returns are discontinuous', *Journal of Financial Economics* **3**, 125–144.

Milne, D. B. M. . F. (1994), 'Contingent claims valued and hedged by pricing and investing in a basis', *The Journal of Finance* **4**, 223–245.

Monnier, J.-B. (2013), 'Technical report : Risk-neutral density recovery via spectral analysis', *arXiv: Computational Finance* .

Muravyev, D., Pearson, N. D. and Broussard, J. (2013), 'Is there price discovery in equity options', *Journal of Financial Economics* **107**, 259–283.

Nachman, D. C. (1988), 'Spanning and completeness with options', *Review of Financial Studies* **1**, 311–328.

Navon, Y. (2014), 'The joint cross section of options and bonds', *Working Paper* .

Nelson, C. R. and Siegel, A. F. (1987), 'Parsimonious modeling of yield curves', *The Journal of Business* **60**, 473–489.

Neuhier, A. and Weber, M. (2018), 'Monetary momentum', *SSRN Electronic Journal* .

Newey and West (1994), 'Automatic lag selection in covariance matrix estimation', *Review of Economic Studies* **61**(4), 631–653.

Ng, S. and Wright, J. H. (2013), 'Facts and challenges from the great recession for forecasting and macroeconomic modeling', *Journal of Economic Literature* **51**(4), 1120–1154.

Nippani, S. and Medlin, W. B. (2002), 'The 2000 presidential election and the stock market', *Journal of Economics and Finance* **26**, 162–169.

Nitteberg, M. (2011), Implied risk-neutral densities.

Novy-Marx, R. (2011), 'Operating leverage', *Review of Finance* **15**(1), 103–134.

Novy-Marx, R. (2012), 'Is momentum really momentum?', *Journal of Financial Economics* **103**(3), 429–453.

Ozoguz, A. (2009), 'Good times or bad times? investors' uncertainty and stock returns', *The Review of Financial Studies* **22**(11), 4377–4422.

Palazzo, B. (2012), 'Cash holdings, risk, and expected returns', *Journal of Financial Economics* **104**, 162–185.

Pan, G.-G., Shiu, Y.-M. and Wu, T.-C. (2020), 'Can risk-neutral skewness and kurtosis subsume the informationcontent of historical jumps?', *Journal of Financial Markets* **100614**.

Pan, J. and Poteshman, A. M. (2006), 'The information in option volume for future stock prices', *Review of Financial Studies* **19**, 871–908.

Pantzalis, C., Stangeland, D. A. and Turtle, H. J. (2000), 'Political elections and the resolution of uncertainty: The international evidence', *Journal of Banking and Finance* .

Papanicolaou, G. C. and Sircar, K. R. (1999), 'Stochastic volatility, smile & asymptotics', *Applied Mathematical Finance* **6**, 107–145.

Pastor, L. and Veronesi, P. (2012), 'Uncertainty about government policy and stock prices', *Journal of Finance* .

Rehman, Z. U. and Vilkov, G. (2012), 'Risk-neutral skewness: Return predictability and its sources', *Working Paper* .

Rigobon, R. and Sack, B. P. (2002), 'The impact of monetary policy on asset prices', *Finance and Economics Discussion Series 2002-4, Board of Governors of the Federal Reserve System (U.S.)* .

Ritchey, R. J. (1990), 'Call option valuation for discrete normal mixtures', *Journal of Financial Research* **13**, 285–296.

Roll, R., Schwartz, E. and Subrahmanyam, A. (2010), 'O/s: The relative trading activity in options and stock', *Journal of Financial Economics* **96**, 1–17.

Rompolis, L. S. (2010), 'Retrieving risk neutral densities from european option prices based on the principle of maximum entropy', *Journal of Empirical Finance* **17**, 918–937.

Rompolis, L. and Tzavalis, E. (2008), 'Recovering risk neutral densities from option prices: A new approach', *Journal of Financial and Quantitative Analysis* **43**, 1037–1054.

Rosenberg, B., Reid, K. and Lanstein, R. (1985), 'Persuasive evidence of market inefficiency.', *Journal of Portfolio Management* **11**, 9–17.

Rosenberg, J. and Engle, R. F. (2002), 'Empirical pricing kernels', *Journal of Financial Economics* **64**(3), 341–372.

Ross, S. A. (2015), 'The recovery theorem', *Journal of Finance* **70**(2), 615–648.

Rubinstein, M. (1976), 'The valuation of uncertain income streams and the pricing of options', *The Bell Journal of Economics* **7**, 407–425.

Rubinstein, M. (1985), 'Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active cboe option classes from august 23, 1976, through august 31, 1978', *Journal of Finance* **40**, 445–480.

Rubinstein., M. (1994), 'Implied binomial trees', *Journal of Finance* **49**, 771–818.

Santa-clara, P. and Yan, S. (2010), 'Crashes, volatility, and the equity premium: Lessons from s&p 500 options', *The Review of Economics and Statistics* **92**, 435–451.

Savor, P. and Wilson, M. (2013), 'How much do investors care about macroeconomic risk? evidence from scheduled economic announcements.', *Journal of Financial and Quantitative Analysis,* **48**, 343–375.

Schneider, P. and Trojani, F. (2018), 'Divergence and the price of uncertainty$*'$, $Journal of Financial Econometrics$**17**(3), $341--396$.

Schneider, P. and Trojani, F. (2019), '(almost) model-free recovery', *Journal of Finance* **74**(1), 323–370.

Scott, L. O. (1996), 'Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: applications of fourier inversion methods.', *Mathematical Finance* **7**, 413–426.

Shimko, D. (1993), 'Bounds on probability', *Risks* **6**, 33–37.

Sialm, C. (2006), 'Stochastic taxation and asset pricing in dynamic general equilibrium', *Journal of Economic Dynamics and Control* **30**, 511–540.

Skinner, D. J. (1997), 'Do options markets improve informational efficiency?', *Contemporary Accounting Research* .

Skintzi, V. D. and Refenes, A. N. (2005), 'Implied correlation index: A new measure of diversification', *Journal of Futures Markets* **25(2)**, 171–197.

Sloan, R. G. (1996), 'Do stock prices fully reflect information in accruals and cash flows aboutfuture earnings?', *The Accounting Review* **71**(3), 289–315.

Soliman, M. T. (2008), 'The use of dupont analysis by market participants', *The Accounting Review* **83**(3), 823–853.

Stilger, P. S., Kostakis, A. and Poon, S.-H. (2016), 'What does risk-neutral skewness tell us about future stock returns?', *Management Science* **63**, 1814–1834.

Stock, J. H. and Watson, M. (2012), 'Disentangling the channels of the 2007-2009 recession', *Brookings Papers on Economic Activity.* .

Stock, J. H. and Watson, M. W. (1989), 'New indexes of coincident and leading economic indicators', *New Indexes of Coincident and Leading Economic Indicators," in NBER Macroeconomics Annual* .

Stock, J. H. and Watson, M. W. (2003), 'Forecasting output and inflation: The role of asset prices.', *Journal of Economic Literature* **41**(3), 788–829.

Stutzer, M. J. (1996), 'A simple nonparametric approach to derivative security valuation', *Journal of Finance* **51**, 1633–1652.

Toft, K. B. and Pruck, B. (1997), 'Options on leveraged equity: Theory and empirical tests', *The Journal of Finance* **52**(3), 1151–1180.

Ulrich, M. (2013), 'How does the bond market perceive government interventions?', *Working paper, Columbia University* .

Ulrich, M. and Walther, S. (2020), 'Option-implied information: What's the vol surface got to do with it?', *Review of Derivatives Research* **23**, 323–355.

Vilkov, G. and Xiao, Y. (2012), 'Option-implied information and predictability of extreme returns', *SAFE Working Paper* **Series 5**.

Vissing-Jorgensen, A. (2019), Central banking with many voices: The communications arms race.

Wang, Y.-H. and Yen, K.-C. (2018), ‘The information content of option-implied tail risk on the future returns of the underlying asset’, *Futures Markets* **38**, 493–510.

Welch, I. (2000), ‘Views of financial economists on the equity premium and on professional controversies’, *Yale School of Management Research Paper Series* **73**(4), 501–537.

Whaley, R. E. (1993), ‘Derivatives on market volatility’, *The Journal of Derivatives* **1**(1), 71–84.

White, H. (1980), ‘A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity’, *Econometrica* **48**(4), 817–838.

Whitelaw, R. F. (1994), ‘Time variations and covariations in the expectation and volatility of stock market returns’, *The Journal of Finance* **49**(2).

X.Diebold, F. and Li, C. (2006), ‘Forecasting the term structure of government bond yields’, *Journal of Econometrics* **130**, 337–364.

Xing, Y., Zhang, X. and Zhao, R. (2010), ‘What does the individual option volatility smirk tell us about future equity returns?’, *Journal of Financial and Quantitative Analysis* **45**, 641–662.

Xu, X. and Taylor, S. M. (1994), ‘The term structure of volatility implied by foreign exchange options’, *Journal of Financial and Quantitative Analysis* **29**, 57–74.

Zheng, Y., Yang, Y. and Chen, B. (2021), ‘Incorporating prior financial domain knowledge into neural networks for implied volatility surface prediction’, *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery Data Mining* .