Conditional Predictability in Market Returns: How Economic Policy Uncertainty Shapes the Forecasting Power of the Implied Volatility Spread

John Cotter[a], Thomas Conlon[a], Yike Wang[a]

[a]*Michael Smurfit Graduate Business School, University College Dublin, Dublin, Ireland*

Abstract

Given the increasing prominence of economic uncertainty in global financial markets, this paper evaluates the impact of the Economic Policy Uncertainty Index (EPU) on the predictive capacity of the Implied Volatility Spread (IVS) for market premiums. Our findings reveal that IVS's forecasting power varies significantly across EPU states, demonstrating weaker performance under low EPU conditions and stronger predictability during high EPU periods, a pattern observed at both daily and monthly frequencies. Incorporating EPU as a dynamic proxy enhances IVS-based forecasts, improving utility gains by 2.5% and highlighting the practical value of adapting predictive models to uncertainty.

Keywords:
Implied Volatility Spread, Economic Policy Uncertainty, Conditional Forecasting, Market Excess Return

1. Introduction

One approach to predicting market returns is to utilize forward-looking information from options (Manaster and Rendleman Jr, 1982; Jackwerth and Rubinstein, 1996; Easley et al., 1998; Bakshi et al., 2003). Among the various predictors derived from options, the Implied Volatility Spread (IVS) stands out as a significant predictor for market premium[1] (Han and Li, 2021; Cremers and Weinbaum, 2010; Easley et al., 1998; An et al., 2014; Atilgan et al., 2015). The IVS, defined as the difference between the implied volatilities of call and put options with identical strike prices and maturities, reflects imbalances in supply and demand dynamics, often driven by informed traders. For instance, investors with superior information about future price movements may disproportionately trade one side of the market (calls or puts), resulting in persistent volatility differences (Easley et al., 1998). Importantly, IVS is not merely a reflection of trading pressures but also encapsulates forward-looking information, capturing market participants' expectations about future price movements and volatility, thereby demonstrating robust predictive capabilities for stock returns across both short-term and long-

[1]Following Goyal et al. (2024), we use the term "market premium" to refer to the market return minus the risk-free rate.

term horizons (Atilgan et al., 2015; Cremers and Weinbaum, 2010). Recent research by Han and Li (2021) highlights the aggregation of IVS across individual stocks as an effective tool, demonstrating its superior forecasting ability compared to other derivatives-based predictors.

While prior studies, such as Han and Li (2021), have established IVS as a robust predictor, they fail to address how its forecasting power varies across different economic states. This gap highlights a lack of understanding regarding the dynamic nature of IVS's predictive ability under changing economic conditions. With economic uncertainty gaining increasing prominence in global financial markets[2], this question has become particularly relevant. Uncertainty has drawn significant attention from both investors and policymakers in recent years (Ludvigson et al., 2021; Brogaard and Detzel, 2015; Bali et al., 2017; Wang et al., 2022), further emphasizing the importance of investigating the interaction between economic uncertainty and market predictors like IVS.

Such an inquiry is further motivated by evidence that market return predictability is highly sensitive to fluctuations in economic conditions (Fernández et al., 2023), underscoring the need to evaluate whether IVS maintains its forecasting power in periods of heightened uncertainty. For instance, numerous studies have demonstrated that predictive models and variables tend to perform better during periods of recession compared to expansions (Henkel et al., 2011; Neely et al., 2014; Rapach et al., 2010). Moreover, Zhu and Timmermann (2022) highlight that even strong models can lose predictive power under specific economic conditions. This variability highlights the importance of evaluating the robustness of forecasting tools, such as IVS, within the dynamic uncertain environments. In particular, EPU, as a proxy for uncertainty, offers a unique lens through which to investigate the stability of IVS's predictibilities. EPU, as a forward-looking measure of economic uncertainty, provides a more dynamic and actionable proxy for economic conditions.

We select EPU as a proxy for uncertainty for several reasons. First, uncertainty-based measures like EPU are more practical for real-time decision-making than static classifications of economic states, such as recession-based definitions, which are typically determined ex-post and lack applicability in real-time contexts. For example, recent studies already adopt similar uncertainty indices for practical analysis. Fernández et al. (2023) utilize a financial uncertainty index derived from VXO to research the predictability of technical predictors in different uncertainty conditions, while Birru and Young (2022) construct uncertainty measures based on stock market volatility and trading volume to research the sentiment's predictability at different periods. These studies further illustrate the application of uncertainty proxies for environmental conditions in forecasting market returns.

Second, the EPU index, based on textual analysis of news data, is related to future macroeconomic conditions and asset prices (Baker et al., 2016; Wang et al., 2022; Fang et al., 2019). Text-based uncertainty measures, such as EPU, can encapsulate information like sentiment or policy signals that

[2]For example, the Economic Policy Uncertainty Index (EPU) surpassed levels observed during the 2008 Global Financial Crisis in the wake of the COVID-19 pandemic in 2020 (Baker et al., 2020), while the VIX surged above 70, its highest level since 2008.

may not yet be fully priced into the market. This sets EPU apart from traditional measures like market volatility, which rely on historical price data and capture only short-term market fluctuations (Ludvigson et al., 2021). Furthermore, EPU's extensive influence across various facets of the economy extends beyond merely capturing political uncertainties, establishing it as a key indicator of overall financial environment uncertainty.

Finally, and most importantly, EPU and IVS are both fundamentally linked to forward-looking macroeconomic information, albeit through distinct mechanisms. EPU has been shown to predict future macroeconomic conditions (Baker et al., 2016), while IVS's forecasting ability stems from its incorporation of macro-economic information(Han and Li, 2021). This overlap raises the hypothesis that EPU may be related to the information embedded in IVS and, consequently, its ability to forecast market movements. However, this relationship has yet to be empirically validated, leaving an important gap in the literature. Moreover, while both IVS and EPU significantly relate to market returns, their low correlation suggests they capture distinct dimensions of information, further supporting the potential for EPU to act as a moderating factor in IVS predictability.

Utilizing IVS data from OptionMetrics and EPU data spanning from January 1996 to June 2019, we examine the effect of EPU on the predictive power of IVS across both daily and monthly frequencies. By conducting in-sample and out-of-sample analyses, this study aims to uncover the influence of EPU on IVS's forecasting capability over both short- and long-term horizons.

From the interaction regression results and marginal effect analysis, we confirm the presence of an interaction effect at both daily and monthly frequencies. Building on this, the grouped regression analysis demonstrates that the predictive power of IVS varies across different EPU levels. Specifically, IVS shows diminished predictive capability under low EPU conditions and enhanced accuracy during high EPU periods, a pattern consistently observed across both daily and monthly forecasting horizons. Further investigation into various thresholds for defining EPU levels validate the robustness of these findings, underscoring the observed predictive discrepancies.

We further extend the analysis to out-of-sample (OOS) predictability, revealing that IVS's performance remains consistently poor during periods of low EPU. At the daily level, in particular, IVS demonstrates no OOS predictability compared to historical mean forecast methods. These findings highlight that the predictive information carried by IVS is both unstable and conditionally dependent on the state of EPU. Building on these insights, we present a refined model as an example to demonstrate the economic implications of our findings. In this model, benchmark forecasts are applied under low EPU conditions, while IVS-based forecasts are used otherwise. This approach ultimately achieves a 2.5% improvement in utility gain.

The contributions of this paper are threefold. First, it introduces the concept of conditional predictability in the forward-looking information embedded within the IVS, highlighting an asymmetry in its forecasting power across different levels of EPU. Specifically, IVS exhibits stronger predictive capabilities under high EPU conditions and weaker performance under low EPU conditions. This finding is particularly significant in the current context of heightened economic uncertainty and contributes to the literature by providing new insights into the predictive information carried by IVS. Secondly, this research also echoes long-standing concerns about the instability of models in predicting market returns(Zhu and Timmermann, 2022; Rossi, 2013; Fernández et al., 2023), which found that model performance typically fluctuates over time and correlates with environment conditions. Thirdly, our research provides an illustration to demonstrate the economic implications of our findings. We incorporate benchmark forecasts under low EPU conditions while using the original IVS forecasts for other periods. This approach aligns with the underlying rationale of Zhu and Timmermann (2022) and Giacomini and White (2006), which emphasize the importance of dynamically selecting models based on related indicators. By incorporating this strategy, our study contributes to the literature by demonstrating how addressing conditional predictability can result in significant improvements in economic outcomes.

Lastly, this study deepens the understanding of EPU's role in financial research by establishing its importance as an uncertainty indicator for examining conditional predictability. While previous studies have highlighted metrics such as financial uncertainty (Fernández et al., 2023) and stock market uncertainty (Birru and Young, 2022), our research underscores that EPU also serves as a valuable tool for assessing the conditional predictability of the options market. By demonstrating EPU's influence on the forward-looking information captured by IVS, this study broadens the application of EPU beyond traditional research, emphasizing its ability to link macroeconomic uncertainty with market-level forecasts.

The remainder of the paper is structured as follows. Section 2 reviews the relevant literature, highlighting the background and identifying the research gap. Section 3 outlines the data, while Section 4 details the methodology and presents empirical results on the relationship between EPU and the predictive power of IVS. Section 5 illustrates the economic implications of our findings, and Section 6 addresses robustness checks to validate the results. Finally, Section 7 concludes with key insights, implications, and directions for future research.

2. Literature Review

2.1. Predictive Information of Implied Volatility Spread

Option market data is often considered to lead the stock market due to its incorporation of forward-looking information embedded in option prices (Manaster and Rendleman Jr, 1982; Jackwerth and Rubinstein, 1996). Consequently, using indicators derived from the options market to forecast stock returns or volatility has become a prominent research area (Easley et al., 1998; Bandi and Perron, 2006; Cremers and Weinbaum, 2010; Bakshi et al., 2003; Cremers and Weinbaum, 2010), with many studies reporting encouraging results. Among these indicators, the IVS has emerged as a robust predictor of market excess returns (Han and Li, 2021).

The IVS measures the difference in implied volatility between call and put options, originating from deviations in the theoretical put-call parity relationship in options markets (Stoll, 1969). In perfect market conditions, put-call parity ensures a specific price relationship between call and put options:

$$C - P = S - PV(K),$$

where $PV(.)$ represents the present value of the strike price. According to the Black-Scholes model, implied volatility (IV) is the value that maintains equilibrium in this relationship. For European options, the put-call parity condition dictates that the implied volatilities of call and put options should be identical, even when option prices deviate from the Black-Scholes formula:

$$IV^{Call} = IV^{Put}.$$

However, empirical evidence shows that deviations from put-call parity frequently occur in real-world options markets (Cremers and Weinbaum, 2010). These deviations manifest as differences in implied volatilities between calls and puts, known as the implied volatility spread:

$$IVS = IV^{Call} - IV^{Put}.$$

Such deviations often signal relative mispricing: a higher implied volatility for calls suggests they are priced relatively higher than puts, and vice versa. The IVS encapsulates the informational content of these pricing differences, often attributed to informed trading activities. Easley et al. (1998) provide a foundational framework, suggesting that informed traders, by buying calls or selling puts, drive up the implied volatility of calls relative to puts, reflecting positive information about future stock prices. Cremers and Weinbaum (2010) extend this by identifying IVS as an anomaly linked to put-call parity deviations, emphasizing its predictive power for market premiums. They attribute this capability to informed trading and market liquidity (reflected in trading volume), which facilitates price discovery of option and enhances the IVS's forecasting strength.

Further studies, such as Atilgan et al. (2015), confirm that IVS captures short-term price pressures from informed traders, making it a reliable predictor of short-term market returns. More recently, Han and Li (2021) developed an aggregated IVS across stock options to measure collective deviations from put-call parity. Their findings demonstrate IVS's predictive power over both short- and long-term horizons. This research underscores that IVS's forecasting strength arises from both informed trading activities and macroeconomic information embedded in option markets. Collectively, this body of literature positions IVS as a key anomaly in financial research, offering valuable insights into equity return forecasts.

2.2. Market Predictability and Different Environment State

Predicting stock returns has long been a central focus of financial research.[3] While the proportion of returns that can be predicted is relatively small, it holds significant importance for both theoretical and practical applications. Over the past three decades, numerous advancements, including the introduction of new predictive variables, refinements in factor asset pricing theories (Fama and French, 1993, 2015), and the adoption of advanced predictive techniques (Kelly and Gu, 2020; Leippold et al., 2022; Avramov et al., 2023), have demonstrated that stock returns are partially predictable.

Even a limited degree of return predictability can yield meaningful benefits for investors. For instance, Campbell (2000) argue that even small levels of predictability can significantly enhance the utility of risk-averse investors. Similarly, Rapach and Zhou (2013) highlight that modest predictability in returns—though insufficient to disrupt market efficiency—can lead to substantial utility gains. This underscores the practical relevance of identifying and analyzing predictive components in returns.

Various methodologies have been developed to evaluate and compare predictive models. These include frameworks for in-sample and out-of-sample testing, such as those proposed by Campbell and Thompson (2008) and **?**. These approaches provide robust tools for assessing the accuracy and reliability of forecasting models and play a crucial role in validating predictive variables.

However, the evaluation methods reveal that stock return predictability is not static but varies significantly across different economic conditions. Research shows that return predictability strengthens during recessions. For example, Rapach et al. (2010) demonstrate that combination forecasts using macroeconomic variables outperform historical averages during downturns. Similarly, Henkel et al. (2011) find that economic variables offer superior predictive performance during recessions, whereas historical averages are more reliable in stable economic conditions. These findings align with the conclusions of Dangl and Halling (2012), who observe that predictability is predominantly concentrated in recessionary periods, and Neely et al. (2014), who show that technical indicators are more effective during downturns than in periods of economic expansion. To explain these phenomena, Cujean and Hasler (2017) propose an equilibrium model linking return predictability to the macroeconomic environment.

Despite the strong focus on recessions, this variable is often defined ex-post, making it difficult to apply in real-time market forecasting. To address this limitation, recent research has shifted toward using uncertainty as a more dynamic and actionable measure of economic conditions. For instance, Fernández et al. (2023) demonstrate that predictive models perform better during periods of low financial uncertainty, while their effectiveness diminishes under high-uncertainty conditions. Similarly, Birru and Young (2022) find that heightened uncertainty amplifies the predictive power of sentiment indexes, with a one-standard-deviation increase in uncertainty enhancing sentiment's

[3]The predictability of such returns does not contradict market efficiency. As assumed in recent empirical methods, such as factor models, market efficiency is not determined by whether returns are predictable, but by whether risk-adjusted returns are zero.

forecasting ability by two to four times. This evidence underscores the importance of integrating dynamic measures of uncertainty into predictive models, as these variables offer a more practical and forward-looking approach to understanding market predictability.

In summary, while stock return predictability is well-documented, its effectiveness is heavily influenced by the surrounding economic environment. Investigating the stability of emerging predictive variables, such as the IVS, across different economic regimes is crucial for their practical application and for advancing our understanding of return predictability.

2.3. EPU and the Predictive information of IVS

Uncertainty is a pervasive feature of global economic development (Cascaldi-Garcia et al., 2023) and has increasingly garnered the attention of investors and policymakers (Ludvigson et al., 2021; Brogaard and Detzel, 2015; Bali et al., 2017; Wang et al., 2022). Among various metrics for uncertainty, the Economic Policy Uncertainty Index (EPU), developed by Baker et al. (2016), has been widely recognized as a key indicator for assessing uncertainty in economic policy. Unlike traditional measures of market volatility, such as the historical conditional variance, EPU captures the broader macroeconomic and policy-driven aspects of uncertainty based on daily newspaper content rather than historical price data, offering a forward-looking perspective (Chiang, 2019; Pástor and Veronesi, 2013).

Baker et al. (2016) reports that the Economic Policy Uncertainty Index (EPU) is closely associated with distinct investment behaviors and stock price volatility, observable from both microeconomic and macroeconomic perspectives. This correlation is particularly pronounced in politically sensitive industries such as healthcare, defense, infrastructure, and finance. Furthermore, Gulen and Ion (2016) found that EPU leads companies to delay significant investment decisions that affect financial outcomes and future cash flows, exerting downward pressure on stock prices. EPU has also been demonstrated to significantly influence other areas, such as crude oil, Bitcoin, and volatility(Wang et al., 2022; Fang et al., 2019; Yfanti and Karanasos, 2022). Thus, EPU serves as a forward-looking uncertainty proxy, closely linked to comprehensive aspects of economy.

The relationship between EPU and the predictive power of forward-looking financial indicators, such as the IVS, warrants closer examination EPU can serve as a crucial metric for analyzing market predictability under varying economic environments. Unlike recession indicators, typically identified retrospectively and difficult to apply in real-time decision-making, EPU offers a forward-looking perspective, reflecting uncertainty as it unfolds. This makes EPU particularly valuable for assessing dynamic market conditions. Furthermore, in contrast to volatility measures derived from historical price data, EPU is constructed directly from textual analysis of daily news. This unique feature allows it to capture a broader range of forward-looking signals, such as expectations, sentiment, and potential disruptions, which may not yet be priced into financial assets. Its extensive application in financial research highlights its robustness as a measure of policy-driven uncertainty and its utility for market forecasting.

Another key reason for linking EPU with the predictive information of IVS is their shared underlying factors. IVS has been shown to possess significant forecasting ability, partially due to its ability to encapsulate macroeconomic information Han and Li (2021) . As EPU is also linked to future economic conditions (Baker et al., 2016), it is reasonable to hypothesize that EPU may directly affect the informational content embedded in IVS. This suggests a potential interaction where EPU serves as a contextual filter, modulating IVS's ability to forecast market returns.

Moreover, by investigating this interplay, we can bridge the gap in understanding how policy uncertainty influences the predictive dynamics of forward-looking financial indicators like IVS. While research on the relationship between EPU and option market information is relatively scarce, Yfanti and Karanasos (2022) is one of the few studies that incorporates the EPU index as a variable to enhance the predictive power of implied volatility for forecasting volatility. However, their analysis does not explore this interaction in depth.

3. Data

3.1. IVS

We calculate the implied volatility spread using data on call and put option implied volatilities (IV) sourced from the database of OptionMetrics. The IVS for stock i at time t is defined as:

$$IVS_{i,t} = IV_{i,t}^{Call} - IV_{i,t}^{Put}.$$

To obtain the aggregate IVS for the entire market at time t, we compute the average across all N stocks:

$$IVS_t = \frac{1}{N}\sum_{i=1}^{N} IVS_{i,t}.$$

This formulation captures deviations from put-call parity at the individual stock level and aggregates them to reflect market-wide dynamics. Consistent with Han and Li (2021), we focus on at-the-money-forward (ATMF) standardized options with an expiration of approximately 30 days. Furthermore, following Atilgan et al. (2015), Cremers and Weinbaum (2010), and Han and Li (2021), we truncate the IVS data at the 0.1% level to mitigate the influence of extreme outliers that may distort statistical results.

We compute the IVS on both daily and monthly frequencies to cater to different research perspectives:

1. **Daily IVS**: For each trading day, we calculate the IVS using IV data from actively traded call and put options for each stock. These individual stock-level IVS values are then averaged cross-sectionally to obtain the daily aggregate IVS.

2. **Monthly IVS**: To construct the monthly aggregate IVS, we follow the methodology outlined in Han and Li (2021), computing the average of the daily IVS_t values from the last five trading days of each month. This approach minimizes short-term noise and provides a more stable and representative measure of IVS. The use of end-of-month data reflects market participants' expectations for future developments, particularly in anticipation of significant events such as earnings announcements or macroeconomic updates.

   To ensure data quality and reliability, we impose the following criteria for stock inclusion in the monthly IVS computation:

   - The stock must have at least 12 trading days of at-the-money (ATM) call and put option data observed within the month.(Han and Li, 2021).

This criterion excludes illiquid or inactive options, which may not fully reflect prevailing market conditions (Han and Li, 2021). By focusing on actively traded options, we ensure that the computed IVS more accurately captures market participants' expectations and sentiments.

3.2. EPU

The Economic Policy Uncertainty (EPU) index[4], developed by Baker et al. (2016), measures a country's economic risk stemming from government policy changes, elections, and debates over economic policies. It has drawn widespread attention from both policymakers and scholars. The index is constructed as a weighted average of three components: (i) the frequency of articles referencing policy uncertainty in ten major U.S. newspapers, (ii) uncertainty regarding tax code changes, and (iii) discrepancies in monetary and fiscal policy forecasts. These components were selected because uncertainties in critical economic policies, such as regulation, taxation, monetary policy, and government budgeting, can significantly affect business operations (Baker et al., 2016).

The EPU index provides a distinct perspective on uncertainty by focusing exclusively on policy-related news. This distinguishes it from traditional measures such as the VIX or realized volatility, which primarily capture market fluctuations that are already reflected in asset prices. Unlike recession indicators, which are identified ex post, the EPU index can be updated daily or monthly, offering a timely and practical tool for empirical analysis.

Another reason for incorporating EPU in assessing market conditions and evaluating the predictive ability of IVS is its close connection to macroeconomic information. The predictive power of IVS originates from macroeconomic signals embedded within its structure, while EPU reflects policy-driven uncertainty that often precedes macroeconomic developments and provides valuable insights into future economic conditions. This temporal precedence of EPU suggests that it may indirectly influence the informational content of IVS by shaping market participants' perceptions and expectations regarding macroeconomic conditions. Furthermore, the extensive application of EPU in financial

[4]Data are available at the Economic Policy Uncertainty Index website: https://www.policyuncertainty.com.

research highlights its robustness as a measure of policy-driven uncertainty. While policy-oriented, its effects extend to macroeconomic activity and various asset returns, making it a comprehensive indicator of overall economic conditions.

3.3. Stock Market Return

For market returns, we use the value-weighted returns of all CRSP firms incorporated in the U.S. and listed on the NYSE, AMEX, or NASDAQ, with a CRSP share code of 10 or 11 at the beginning of the month. The one-month Treasury bill rate serves as the proxy for the risk-free rate. Both datasets are sourced from the Kenneth R. French website.[5]

3.4. Descriptive Statistical Analysis

The dataset spans 8,613 firms and 5,912 trading days from January 4, 1996, to June 28, 2019. On average, 3,394 firms are included per trading day, with data available at both daily and monthly frequencies.

Figure 1 illustrates the time-series trends of the aggregate IVS and EPU indices, with NBER-defined recessions shaded in gray. While the IVS remains relatively stable over time, the EPU exhibits significant spikes during economic downturns, such as the Dot-com bubble (2000–2002) and the Global Financial Crisis (2008–2009). These spikes indicate the heightened policy uncertainty that often accompanies economic distress. Notably, the two indices are not perfectly synchronized, capturing distinct dimensions of market information. This differentiation highlights the complementary nature of the two indicators in explaining market dynamics.

The distribution of EPU, however, is right-skewed[6], which poses challenges for regression analysis, such as potential biases in parameter estimation (Liu and Lin, 2014). To address this, a natural logarithmic transformation (*ln(EPU)*) is applied, which normalizes the distribution and ensures more reliable statistical inference. Consistent with standard practices, high and low levels of EPU are defined as one standard deviation above and below the mean, respectively, to capture periods of extreme policy uncertainty.

Table 1 summarizes key variables at daily and monthly frequencies. Firstly, the IVS exhibits similar characteristics on both daily and monthly scales: nearly 75% of the IVS values are negative, with a small variance. This reflects the structural bias of the options market, where puts often exhibit higher implied volatility than calls due to risk-hedging demand. The standard deviation of IVS is low, indicating minimal variability over time. Secondly, *ln(EPU)* values exhibit greater variability compared to IVS, highlighting its sensitivity to changes in market conditions. Thirdly, the average monthly excess market return is 0.64%, closely aligned with the daily return of 0.03%. The standard

[5] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[6] See Appendix for the detailed distribution of EPU.

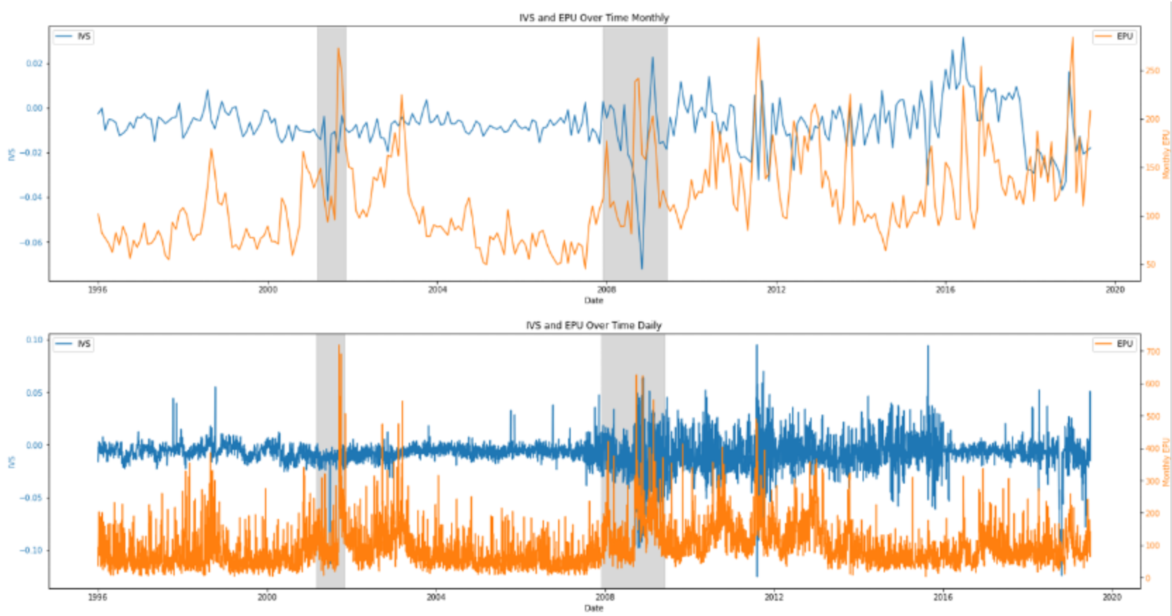

Figure 1: Time-series of Aggregated IVS and EPU During Economic Cycles.

This graph describes the variations over time in the Implied Volatility Spread (IVS) represented by the blue line, and the Economic Policy Uncertainty (EPU) indicated by the orange line. Gray-shaded areas represent NBER-defined recessions. The data spans from January 4, 1996, to June 28, 2019.

Table 1: Descriptive Statistics of Key Variables

	Monthly			Daily		
	IVS	MKT_RF(%)	ln(EPU)	IVS	MKT_RF(%)	ln(EPU)
N	282	282	282	5912	5912	5912
Mean	-0.01	0.64	4.69	-0.01	0.03	4.37
Std	0.01	4.44	0.37	0.01	1.18	0.63
Min	-0.07	-17.23	3.80	-0.13	-8.95	1.56
25%	-0.01	-1.96	4.43	-0.01	-0.48	3.97
50%	-0.01	1.19	4.68	-0.01	0.07	4.39
75%	0.00	3.45	4.96	0.00	0.60	4.80
Max	0.03	11.35	5.65	0.09	11.35	6.58

*This table summarizes the descriptive statistics of key variables at daily and monthly frequencies. IVS represents the implied volatility spread, calculated as the difference between call and put implied volatilities and aggregated across stocks. MKT_RF denotes the value-weighted market excess return over the risk-free rate, and *ln(EPU)* is the logarithmic transformation of the Economic Policy Uncertainty index. The statistics include the sample size (N), mean, standard deviation (*Std*), and key percentiles, capturing the central tendency and variability of each variable from January 4, 1996, to June 28, 2019.

deviation of market returns is notably higher on a monthly scale (4.44%) compared to the daily scale (1.18%), reflecting the compounding effect of daily return volatility over time.

To facilitate cross-variable comparisons and ensure consistency in the subsequent regression analysis, all input variables are standardized. This step is particularly important given the substantial differences in scales and variances across the variables. For instance, while IVS is constrained within a narrow range, *ln(EPU)* and market returns exhibit broader distributions, potentially skewing regression outcomes if left unstandardized.

4. Methodology and Empirical Results

4.1. In Sample Analysis

4.1.1. Interaction Regression

First, the baseline forecast model examining the relationship between market excess returns and the IVS is as follows:

$$R_{t+1,h} = a + b * IVS_t + \epsilon_{t+1} \tag{1}$$

ϵ_{t+1} is the error term, capturing the unexplained variation in excess returns. Specifically, for the daily dimension, h is set at 1, 2, 3, and 4 days. For the monthly dimension, h is set as 1, 3, 6, and 12 months. $R_{t+1,h}$ represent excess return over the horizon h following period t, in line with Han and Li (2021), elaborated as follows:

$$R_{t+1,h} = \frac{1}{h} \sum_{s=t+1}^{t+h} r_s \tag{2}$$

where r_s represents the excess market return at time s, calculated as the market return minus the risk-free rate. This approach aligns with recent studies, such as Han and Li (2021), which adopt similar methodologies for forecasting market returns. To explore how EPU moderates the relationship between IVS and market returns, the forecast model is extended to include an interaction term, as follows:

$$R_{t+1,h} = a + b_1 \cdot IVS_t + b_2 \cdot EPU_t + b_3 \cdot (IVS_t \cdot EPU_t) + b_4 \cdot Control\_Variable_t + \epsilon_{t+1} \tag{3}$$

where EPU_t is the level of Economic Policy Uncertainty at time t, and a, b_1, b_2 and b_4 are coefficients to be estimated, where b_3 captures the interaction effect between IVS_t and EPU_t.

This specification allows us to assess not only the predictive power of IVS and EPU but also the extent to which EPU moderates the relationship between IVS and future returns. This approach is particularly relevant for investigating whether different uncertainty conditions amplifies or mitigate the forward-looking information embedded in IVS. We also account for potential confounding influences by including controls for sentiment indicators as well as stock market volatility (VIX), and a recession dummy variable (equal to 1 during recession periods as defined by NBER). The interaction analysis is conducted on the full sample, allowing the inclusion of ex post identified indicators such as recession periods. By introducing these controls in a stepwise manner, we isolate the independent impact of EPU on the relationship between IVS and market premiums.

Table 2: The Moderating Role of Economic Policy Uncertainty in the Predictive Power of Implied Volatility Spread Across Forecast Horizons

	Daily Horizon				Monthly Horizon			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IVS	2.200***	1.468***	1.393***	1.449***	1.089***	0.680**	0.579*	0.685**
	(0.631)	(0.510)	(0.516)	(0.510)	(0.295)	(0.324)	(0.316)	(0.322)
IVS×ln(EPU)		1.318*	1.302*	1.370**		0.552*	0.561*	0.526*
		(0.700)	(0.702)	(0.696)		(0.322)	(0.310)	(0.318)
ln(EPU)		0.617	0.741**	0.343		0.367	0.446*	0.373
		(0.400)	(0.369)	(0.346)		(0.261)	(0.252)	(0.260)
Recession_dummy			-2.124				-1.704	
			(1.989)				(1.221)	
VIX				0.802				-0.098
				(0.794)				(0.287)
Observations	5910	5910	5910	5906	280	280	280	280
Adjusted R^2(%)	0.9	1.2	1.2	1.2	5.3	6.7	7.5	6.4

* This table presents the interaction regression results assessing whether Economic Policy Uncertainty (EPU) moderates the predictive relationship between the Implied Volatility Spread (IVS) and market excess returns. IVS serves as the primary predictor, with its interaction with ln(EPU) capturing the moderating effect, controlling for the recession dummy (NBER-defined) and VIX. The analysis spans two horizons: daily (Columns 1–4), and monthly (Columns 5–8), using a sample period from January 4, 1996, to June 28, 2009. Coefficients are reported with standard errors in parentheses. Statistical significance is denoted by $*, **, ***$ at the 10%, 5%, and 1% levels, respectively. The dependent variable, market excess returns, is adjusted for different horizons using equations in 2, scaled to monthly returns based on 22 trading days, and includes a one-day lag to control auto-correlation following Han and Li (2021). Additionally, the multi-variable regression above pass the multicollinearity diagnostics tests(all 1/VIF smaller than 1.9), confirming that no multicollinearity exists among the explanatory variables.

Table 2 reports the interaction regression results across daily and monthly forecast horizons, examining the moderating role of *ln(EPU)* on the predictive power of the IVS for market excess returns. The interaction term IVS × ln(EPU) is included to capture the moderating effect of EPU on the predictive power of the IVS. Firstly, the coefficients on IVS are positive and statistically significant across all regressions, confirming its predictive ability for market excess returns. As IVS is standardized, the reported values indicate the change in market excess returns for a 1 percentage point increase in IVS. For instance, at daily or monthly horizon, a 1 percentage point increase in IVS leads to a 2.2 or 1.089 percentage increase in market excess returns. Even after controlling for EPU and other factors, IVS's predictive ability remains significant, though slightly reduced. Notably, the adjusted R^2 increases consistently from the daily to the monthly horizons, rising from 0.9% in the daily models to 5.3% in the monthly models. This suggests that the models explain a greater proportion of variation in returns as the forecast horizon extends, potentially reflecting the compounding effects of predictive signals over time.

Secondly, the significance of the interaction term highlights the moderating effect of EPU on IVS's predictive ability. The coefficients for the interaction term in regressions (2), (3), and (4) suggest that a one-unit increase in EPU enhances the effect of IVS on market excess returns by 1.301 to 1.370 percentage at the daily level. At the monthly level, the interaction term coefficients in regressions (6), (7), and (8) indicate an increase of 0.526 to 0.552 percentage per one-percentage rise in EPU. These findings demonstrate that EPU moderates the relationship between IVS and market returns, with higher EPU amplifying the predictive power of IVS. In other words, when EPU is elevated, market returns exhibit greater sensitivity to IVS, enhancing its predictive signal. The results remain robust after controlling for recessionary periods and the VIX. Furthermore, the coefficients for *recessiondummy* and *VIX* are not significant in any specification, indicating that once IVS and EPU are included, recessionary periods (as defined by NBER) and VIX do not contribute substantial additional explanatory power to the model. Besides, diagnostic tests confirm the absence of multicollinearity, with all Variance Inflation Factor (VIF) values below 1.9, ensuring the robustness of the regression specifications.

These results highlight the crucial role of EPU in enhancing IVS's predictive power, providing new insights into return predictability under varying economic conditions.

4.1.2. Marginal Effect

To deepen our understanding of how EPU influences the predictive power of the IVS, we analyze the marginal effects of IVS under three specific EPU levels. This analysis builds on interaction regression models (2) and (6), allowing us to visualize how changes in IVS predict market excess returns across different uncertainty conditions. While most econometricians can infer the interaction effects from regression coefficients, marginal effect analysis provides an intuitive and graphical representation of the dynamic relationship between IVS and EPU, offering unique insights into this interaction's practical implications. This approach has been employed in studies such as Williams (2012) and Li and Tang (2010) to interpret interaction terms effectively. Specifically, we examine three standardized levels of ln(EPU)—low ($\leq$ mean $-$ std), medium (= mean), and high ($\geq$ mean + std)—and their influence on IVS's predictive ability. These levels allow us to uncover patterns of interaction that are often masked in broader analyses, thereby enhancing our understanding of how EPU moderates the predictive power of IVS.

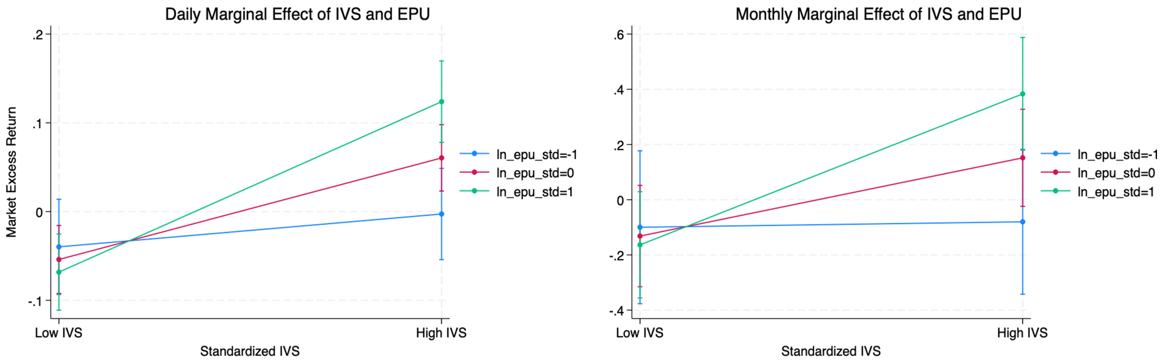

Figure 2: Marginal Effect Analysis of IVS and EPU at Daily (Left) and Monthly (Right) Horizons. This figure shows the marginal effects of the Implied Volatility Spread (IVS) on market excess returns under different levels of Economic Policy Uncertainty (EPU). Here, ln_EPU_std = -1 (blue line),ln_EPU_std = 0 (red line) and ln_EPU_std = 1 (green line) represent low , medium and high uncertainty EPU levels, respectively. The steeper slope for high EPU underscores a stronger positive impact of IVS on market returns, while a flatter slope at low EPU indicates less predictive accuracy. The 95% confidence intervals, narrower during high uncertainty and wider during low uncertainty, reflect the stability of predictions across daily (left panel) and monthly (right panel) dimensions. This highlights how extremes in EPU significantly influence the effectiveness of IVS in forecasting market returns.

Figure 2 presents the marginal effects of IVS on market returns at daily and monthly horizons across varying EPU levels. In both panels, the green line ($\ln(\text{EPU})_{\text{std}} = 1$) reflects high uncertainty, the blue line ($\ln(\text{EPU})_{\text{std}} = -1$) represents low uncertainty, and the red line ($\ln(\text{EPU})_{\text{std}} = 0$) corresponds to medium uncertainty.

Firstly, the observed non-parallel lines in both the monthly and daily graphs further confirm the existence of an interaction effect. The steeper slope for red line indicates that higher uncertainty amplifies the positive impact of IVS on market returns. In contrast, the flatter slope for green line suggests that under low EPU, IVS has a weaker influence on market excess returns. Secondly, confidence intervals (CI), represented by vertical lines at each estimated point, further underscore this distinction. At low EPU levels, the broader confidence intervals indicate greater variability and less consistent predictive performance. As seen in both panels, the green line highlights a gradual transition, underscoring the progressive amplification of IVS's predictive power as EPU increases. **Notably, at low EPU levels, the predictive relationship between IVS and market returns appears nearly flat. This result suggests that during periods of low uncertainty, IVS may struggle to provide reliable forecasts of market returns, highlighting its limited effectiveness in stable environments. These findings highlight that the strength and reliability of IVS's predictive relationship with market returns are highly contingent on the prevailing level of EPU.**

This section offers an initial perspective on how different levels of EPU influence the predictability of IVS, serving as a visual complement to the interaction regression models. In the subsequent in-sample analysis, we will statistically validate these differences to provide a more robust confirmation of the observed patterns.

4.1.3. Grouped Regression for Different Levels of EPU

Building on these findings, the grouped regression analysis takes a more systematic statistical approach. By classifying ln(EPU) into distinct groups (low, medium, and high[7]), it examines the predictive power of IVS within each group, focusing on the significance of the coefficient b and changes in adjusted R^2. This allows for a more detailed evaluation of how different EPU levels influence IVS's predictability, thereby extending the insights derived from the marginal effects analysis. It also provides a discrete perspective for analyzing the relationship between IVS and EPU, complementing the global and linear approach of interaction regression analysis. For each group, we estimate the regression model specified in Equation 1. This method enables us to evaluate how IVS's predictive power, reflected in the coefficient b, statistical significance and adjusted R^2, systematically varies across different EPU levels. Both IVS and EPU were standardized using z-scores within their respective groups to ensure comparability of coefficients across different EPU levels. This standardization normalizes the variables to have a mean of zero and a standard deviation of one within each group, eliminating potential biases arising from baseline differences in IVS levels. As a result, the observed variations in predictive power can be attributed to differences in EPU levels, rather than underlying disparities in IVS across groups.

Table 3: Grouped Regression Analysis of Daily Ln(EPU): Partitioned by Deviation from the Mean

Forecast Horizons (Daily)	Low	Medium	High
1 Day	IVS: 0.402 R^2: -0.116	IVS: 1.486*** R^2: 0.492	IVS: 5.421** R^2: 2.195
2 Day	IVS: 0.348 R^2: -0.129	IVS: 1.326*** R^2: 0.704	IVS: 3.785*** R^2: 1.956
3 Day	IVS: 0.078 R^2: -0.210	IVS: 1.234*** R^2: 0.921	IVS: 2.922*** R^2: 1.871
4 Day	IVS: 0.661* R^2: 0.318	IVS: 0.979*** R^2: 0.773	IVS: 2.772*** R^2: 3.002

*The table presents the coefficients of IVS predictors for market excess returns and the adjusted R^2 percentages obtained from a grouped regression analysis, categorized by varying levels of Ln(EPU). Specifically, values less than one standard deviation below the mean are classified as the Low level group, values greater than one standard deviation above the mean are deemed the High level group, and values that fall in between are categorized as the Medium group. The distribution of the three groups is as follows: medium 69.33%, high 16.24%, and low 15.43%. The analysis assesses the IVS's ability to forecast market premiums across different horizons (1 to 4 days). To control for autocorrelation, the one-day lag of market excess return is included in each regression, following Han and Li (2021). The dataset covers the period from January 4, 1991, to June 28, 2019.

[7] we classify ln(EPU) into three groups: low (ln(EPU) $\leq$ mean $-$ 1 standard deviation), high (ln(EPU) $\geq$ mean $+$ 1 standard deviation), and medium (values in between).

Table 4: Grouped Regression Analysis of Monthly Ln(EPU): Partitioned by Deviation from the Mean

Forecast Horizons (Monthly)	Low	Medium	High
1 Month	IVS: 0.820* R^2: 2.784	IVS: 0.569** R^2: 1.091	IVS: 2.952*** R^2: 24.959
3 Month	IVS: 0.648** R^2: 10.326	IVS: 0.781*** R^2: 11.593	IVS: 1.958*** R^2: 40.721
6 Month	IVS: 0.148 R^2: -0.195	IVS: 0.606*** R^2: 12.571	IVS: 0.958*** R^2: 27.020
12 Month	IVS: 0.181 R^2: -1.477	IVS: 0.267*** R^2: 5.710	IVS: 0.308*** R^2: -0.035

*The table presents the coefficients of IVS predictors for market excess returns and the adjusted R^2 percentages obtained from a grouped regression analysis, categorized by varying levels of Ln(EPU). Specifically, values less than one standard deviation below the mean are classified as the Low level group, values greater than one standard deviation above the mean are deemed the High level group, and values that fall in between are categorized as the Medium group. The distribution of the three groups is as follows: medium 66.33%, high 18.24%, and low 16.63%. The analysis assesses the IVS's ability to forecast market premiums across different horizons (1 to 6 months). To control for autocorrelation, the one-day lag of market excess return is included in each regression, following Han and Li (2021). The dataset covers the period from January 4, 1991, to June 28, 2019.

As table 3, 4 shown, the main finding is the IVS predictability has substantial difference across the different EPU levels. Specifically, daily IVS exhibits limited predictive power when EPU is at a low level for forecast horizons of 1 to 3 days. For the 4-day horizon, although the IVS coefficient is significant at the 10% level, its significance is lower compared to other EPU groups. Besides, its coefficient is only 0.661, considerably lower than the 2.772 coefficient observed in the high EPU group, which is significant at the 1% level. Furthermore, the adjusted R^2 for the low EPU group is just 0.318%, indicating a relatively low explanatory power, compare to high EPU level. Taken together, while there is some predictive ability at the 4-day horizon under low EPU conditions, it is much weaker than in other groups. Conversely, during periods of high EPU, the IVS coefficients are larger, and the adjusted R^2 values are significantly higher, indicating a stronger predictive power. This suggests that IVS becomes more effective at forecasting market premiums under elevated uncertainty, whereas its predictability declines in more stable, low-uncertainty environments. For monthly horizons, at low EPU levels, IVS retains some forecasting ability, though with lower significance and a relatively low adjusted R^2. At the 6- and 12-month horizons under low EPU levels, IVS exhibits a lack of predictive power, as evidenced by the insignificant IVS coefficients and negative adjusted R squared. In contrast, at high EPU levels, the IVS coefficient is highly significant, and the forecasted adjusted R^2 reaches a high level for horizon 1-6 months. This indicates that the predictive power of IVS is notably stronger when EPU is high.

Overall, under low levels of EPU, the predictive power of IVS for market excess returns is limited or negligible. In contrast, at high levels of EPU, IVS demonstrates significantly stronger predictive capability, as reflected by highly significant regression coefficients and the highest adjusted R^2 value.

Moreover, there is no standardized approach in the literature for defining different

levels of uncertainty conditions. For instance, Birru and Young (2022) investigate the impact of stock market uncertainty on sentiment index predictability, defining low and high uncertainty at the 10% threshold.[8] Similarly, Fernández et al. (2023) examine the variation in the predictability of technical variables across different levels of financial uncertainty, directly categorizing low and high uncertainty as below or above the mean. However, these studies do not further explore whether alternative definitions of high and low uncertainty thresholds might influence their results. To ensure robustness, this study adopts multiple thresholds for defining EPU levels, ranging from the lowest and highest 10% to 50% of ln(EPU), in increments of 5%. The threshold analysis reinforces our main finding by demonstrating that IVS exhibits limited or no predictive power under low EPU conditions, while its predictive capability significantly strengthens under high EPU conditions, thereby enhancing the robustness of the results across varying EPU levels (see Appendix for details of the threshold analysis).

4.2. Out-of-sample Analysis

4.2.1. OOS Predictability in Different EPU levels

Based on in-sample analysis, we infer that the predictive capability is weaker during periods of low EPU and stronger when EPU is high. To test the robustness of this finding in an out-of-sample (OOS) context, we examine whether the OOS predictive power of IVS varies across different levels of EPU. Specifically, we assess the OOS predictive performance during high and low EPU periods using the OOS R^2(Campbell and Thompson, 2008). A simple interpretation of R^2_{oos} is one minus the ratio of the mean squared prediction errors (MSPE) between the competitive model and the benchmark model. Usually, the historical mean average return often serves as the benchmark model in prior analyses(Han and Li, 2021; Dong et al., 2022; Fernández et al., 2023). A positive R^2_{oos} suggests that the predictive regression boasts a lower average MSPE in comparison to benchmark forecast return. The definition of R^2_{oos} is:

$$R^2_{OOS,c} = 1 - \frac{\sum_{t=1}^{T} I_t^c (r_t - \hat{r}_t)^2}{\sum_{t=1}^{T} I_t^c (r_t - \bar{r}_t)^2}, \quad \text{for } c = \text{low EPU period, other EPU Period, all period}$$

where I_t^c is an indicator variable equal to 1 if the observation at time t falls within regime c, and 0 otherwise. $\hat{r}_t$ denotes the forecasted market excess return at time t based on data up to $t-1$. The historical mean model serves as a benchmark, providing the estimate $\bar{r}_t$ for the same preceding period $t-1$. The statistical significance of R^2_{oos} is assessed using the Clark and West (CW) test (Clark and West, 2007), including Newey and West (NW) test (Newey and West 1987) as adjustments for autocorrelation and heteroskedasticity.

[8] At the 10% threshold, the lowest 10% and highest 10% of ln(EPU) values are classified as "low" and "high," respectively, with the remaining values categorized as "medium."

Table 5: Monthly Out-of-Sample Predictability of Recursive Forecast Model Across Total, Low EPU, and Other Periods

	Total Period		Low EPU Period		Other Period	
Horizon (Monthly)	R^2_{OOS}	CW	R^2_{OOS}	CW	R^2_{OOS}	CW
1	2.895	2.284	2.404	0.974	2.911	2.267
3	9.832	2.082	15.719	2.174	9.688	2.053
6	4.766	2.480	-42.916	-0.627	5.454	2.544
12	-1.124	2.152	-30.723	-1.298	0.095	2.377

*This table presents the monthly out-of-sample (OOS) predictive performance of the recursive forecast model for the entire OOS period, as well as for periods when EPU is low (ln(EPU) $\leq$ mean1 standard deviation) and for all other periods. The OOS R^2 represents the assessment of out-of-sample predictability, following the method outlined by Campbell and Thompson (2008). The CW statistic refers to the Clark-West test statistic (Clark West, 2007). A positive OOS R^2 and a CW statistic greater than 1.645 (the 10% significance level) indicate significant out-of-sample predictability for the anomaly. The sample period covers January 4, 1996, to June 28, 2019, with the training period ending on December 31, 2001; the remaining period is used for OOS evaluation.

We employ a recursive window method to generate predictive values across horizons (1–4 days, 1–12 months). The initial training window spans data from January 1, 1996, to December 31, 1999 (4 years), and expands by one observation at each step to incorporate the most recent data while retaining all prior observations. This approach ensures that the model utilizes all available historical information at each point in time for out-of-sample predictions.

From Tables 5 and 6, we observe that OOS predictive power exists across 1–4 day and 1–6 month horizons for the entire evaluation period, as indicated by positive and significant. For the 12-month horizon, there appears to be no OOS predictive power over the entire period. However, after excluding the low EPU period, a positive and significant positive and significant R^2_{OOS}.

Table 6: Daily Out-of-Sample Predictability of Recursive Forecast Model Across Total, Low EPU, and Other Periods

	Total Period		Low EPU Period		Other Period	
Horizon (Daily)	R^2_{OOS}	CW	R^2_{OOS}	CW	R^2_{OOS}	CW
1	0.671	2.720	-1.285	-0.207	0.786	2.758
2	0.898	3.200	-0.434	0.375	0.978	3.197
3	1.197	3.728	-0.741	0.367	1.308	3.738
4	1.466	3.753	1.618	2.025	1.458	3.633

*This table presents the daily out-of-sample (OOS) predictive performance of the recursive forecast model for the entire OOS period, as well as for periods when EPU is low (ln(EPU) ¡ mean1 standard deviation) and for all other periods. See notes of table 5 for details.

At the daily frequency, IVS demonstrates no OOS predictability during the 1–3 day horizon and the low EPU period, whereas other periods exhibit higher and significant OOS predictive power. Similarly, at the monthly frequency, although the low EPU period still demonstrates OOS predictability, its predictive power remains relatively weak compared to other periods. Additionally, the 4-day and 3-month horizon tests deviate from the general pattern of diminished OOS predictive capability during low EPU periods. This deviation may stem from the limited sample size characteristic of low EPU periods, which could introduce elements of chance and noise interference. To address this issue, we follow the approach outlined by Clark and West (2007), incorporating the squared differences between predictions from competitive and benchmark models as a penalty to the Mean Squared Error (MSE). This adjustment mitigates noise and offers a more accurate and statistically reliable basis for assessment compared to directly comparing R^2_{OOS} values.

4.2.2. Statistical Hypothesis Test for the Decrease in IVS OOS Predictive Power at Low EPU Levels

Building on the observed variations in OOS predictability across different EPU levels, this section conducts a more rigid analysis by introducing a statistical hypothesis test. Specifically, we compare the OOS predictive power of IVS between low EPU periods and other periods using an indicator variable I_t, which takes the value of zero during low EPU periods and one otherwise. This approach aims to determine whether EPU levels significantly influence the OOS predictability of IVS. Our hypothesis is following:

$$H_0 : R^2_{oos}|_{I_t=1} \leq R^2_{oos}|_{I_t=0} \text{ against } H_1 : R^2_{oos}|_{I_t=1} > R^2_{oos}|_{I_t=0} \tag{4}$$

The null hypothesis represents that predictability increases or not change when EPU is low, while the opposite hypothesis means the decrease of the predictability when EPU is low. The test for whether R^2_{oos} is significant is based on the statistical test CW test introduced by Clark and West (2007):

Let $\hat{e_1} = (r_t - \bar{r}_t)$, $\hat{e_2} = (r_t - \hat{r}_t)$ be the one step forward forecast errors of the benchmark model and our own competitive model, respectively. To eliminate the impact of noisy and over-fitting problems, we following Clark and West (2007) to use adjusted mean squared prediction error (adjusted MSE):

$$f = \hat{e_1}^2 - \hat{e_2}^2 + (\bar{r} - \hat{r})^2 \tag{5}$$

.

To evaluate whether OOS predictability varies across different levels of EPU, we analyze the relationship between the adjusted MSE and periods categorized by EPU levels:

$$f_{t|t-1} = \mu + \gamma I_t + \epsilon_t \tag{6}$$

where I_t denotes a dummy variable that captures different EPU levels. In this way, the hypothesis will be transformed to the test on

$$H_0 : \gamma \leq 0 \text{ against } H_1 : \gamma > 0 \tag{7}$$

If γ is found to be positive and statistically significant, it suggests that the model's out-of-sample (OOS) predictive power diminishes during low EPU periods. A positive γ implies that during low EPU periods (when $I_t = 0$), the model's predictive power is reduced, as $I_t = 0$ corresponds to a significant decrease in the predictor's effectiveness. This results in higher forecast errors for the model during these periods, which leads to an increase in the adjusted MSE and a consequent decrease in the OOS R^2.

Table 7: Statistical Test for Comparison of OOS Predictability Across Different ln(EPU) Levels

	Horizon (daily)				Horizon (Monthly)			
	1	2	3	4	1	3	6	12
γ	21.91	12.14	7.89	5.41	3.39	3.66	1.51	0.34
cw	2.45	2.35	2.79	2.28	2.35	1.91	2.72	2.28

*This table presents the hypothesis testing results on whether out-of-sample (OOS) predictability decreases during periods of low EPU, based on model 6. The test for the significance of gamma (γ) is conducted using the Clark-West (CW) test, as introduced by Clark and West (2007). With a critical value of 1.645 and a positive gamma (γ) , all tests reject the null hypothesis, allowing us to conclude that the predictive ability of the IVS decreases during low EPU periods.

As shown in Table 7, from both monthly and daily perspectives, all γ values are positive, and the corresponding CW statistics exceed 1.645, confirming their significance. Thus, we reject the null hypothesis and conclude that the out-of-sample predictive power of IVS is indeed weaker during periods of low EPU compared to other periods. This result holds consistently across various time horizons. This pattern remains consistent for the 4-day and 3-month horizons, suggesting that the observed increase in R^2_{OOS} for these horizons under low EPU in the previous chapter may be attributed to noise or random variation.

Overall, the statistical hypothesis test reinforces the main conclusion: the predictive power of IVS weakens during low EPU periods, while it demonstrates significantly stronger predictive capability under higher EPU conditions. Notably, at the daily horizon, IVS completely loses its OOS predictability in low EPU environments, underscoring the critical role of uncertainty in enhancing its forecasting performance.

5. Economic Implications

The findings reveal that the predictive information carried by IVS is both unstable and asymmetric, depending on the state of EPU. Specifically, while IVS generally demonstrates significant out-of-sample

(OOS) predictibility—often surpassing the historical mean benchmark—its performance can fall below this benchmark during certain states at daily horizon, such as periods of low EPU. By using EPU levels at time t as a proxy for the economic state, we gain insights into the expected predictive power of IVS for subsequent periods. This approach not only aids in model selection for optimal performance but also aligns with the principles of "dynamic forecasts rotation" discussed in Zhu and Timmermann (2022) and Giacomini and White (2006). Dynamic rotation is similar to selecting the best forecast from a set of models, leveraging available moderating instruments to guide this selection process dynamically.

A straightforward application of this concept involves substituting IVS forecasts with benchmark forecasts (e.g., historical mean) during periods of low EPU, which is expressed as:

$$R_{t+1} = \begin{cases} \beta_0 + \beta_1 \cdot \text{IVS}_t + \epsilon_t & \text{if } \ln(\text{EPU})_t \geq \mu_{\ln(\text{EPU})} - \sigma_{\ln(\text{EPU})}, \\ \text{Substitute Model} & \text{if } \ln(\text{EPU})_t < \mu_{\ln(\text{EPU})} - \sigma_{\ln(\text{EPU})}, \end{cases}$$

where μ and σ represent the mean and standard deviation of the log-transformed EPU, respectively.

A key application of market return predictability is optimizing portfolio allocation between stocks and risk-free assets. We evaluate predictive models' practical implications by assessing their performance through utility gain for a mean-variance investor. Specifically, we compute the certainty equivalent return (CER) from the predictive model and compare it to a benchmark portfolio based on historical averages, as in Han and Li (2021); Dong et al. (2022). This comparison quantifies the economic value of incorporating predictive information into portfolio decisions.

We assume the utility function for a mean-variance investor is given by:

$$U = E(R_p) - \frac{\gamma}{2}\sigma_p^2,$$

where $E(R_p)$ and σ_p^2 denote the portfolio's expected return and variance, respectively, and γ represents the investor's risk aversion. The optimal weight allocated to the risky asset can be calculated using mean-variance optimization methods, as detailed in Markowitz (1952). Based on weights, we calculate portfolio performance metrics, including CER and Sharpe ratios, as shown in Table 8.

As shown in Table 8, the refined model achieves a 10.7% improvement in out-of-sample (OOS) R^2, while the Certainty Equivalent Return (CER) increases by 2.5% compared to the original model. Furthermore, the p-value of the Ledoit and Wolf test for the Sharpe ratio of both the refined and original portfolios is below 5%, indicating that both portfolios significantly outperform the benchmark. Notably, the refined model outperforms the benchmark by 1.6% more than the original model, with a smaller p-value in the Ledoit and Wolf test, highlighting its enhanced reliability and superior forecasting performance. These results highlight the adaptability of dynamic forecasting approaches to changing economic conditions and demonstrate the economic relevance of our findings through this daily horizon analysis. Figure 5 also reveals that the original IVS forecasts, which do not account for

Table 8: Performance Metrics of the Refined Model Versus Original Model

	Original Forecast	Refined Forecast	Shift
R^2_{OS}	0.67	0.74	10.7%
CW	2.72	2.76	1.4%
CER Gain	18.11	18.57	2.5%
SR	0.90	0.92	1.6%
Ledoit and Wolf Test	0.048	0.030	-

* This table describes the overall performance of the refined model, which incorporates substitute forecasts during low EPU periods. Predictions from the default yield (DFY) model, as detailed in Fama and French (1989), are used as substitutes. Evaluated metrics include Out-of-Sample R^2, Sharpe Ratio (SR), and Certainty Equivalent Return (CER) gain. All metrics are annualized, and the Shift column reflects the percentage improvement of the refined model over the original model. The Ledoit and Wolf test provides the p-value for testing whether the Sharpe ratio (SR) of the original and refined portfolios significantly differs from that of the benchmark forecast (mean historical portfolio). The one-tailed test follows the methodology outlined in Ledoit and Wolf (2008).

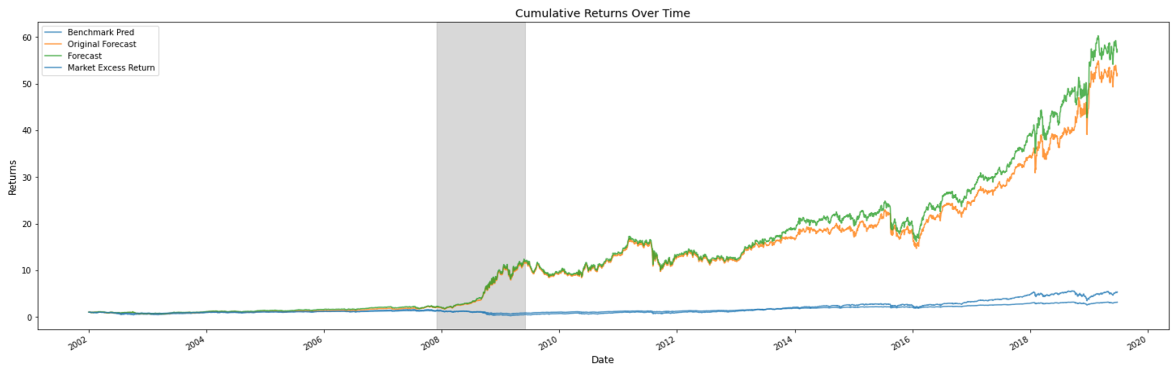

Figure 3: Cumulative Returns of Refined and Benchmark Models vs. Market Excess Returns. This figure illustrates the cumulative returns of portfolios constructed using out-of-sample predicted daily market excess returns derived from the refined model. The performance of these portfolios is compared against benchmark forecasts (mean average forecast) and actual market excess returns and the original forecasts. The original forecasts refer to predictions solely driven by the IVS model without adjustments for EPU levels. The refined model incorporates Economic Policy Uncertainty (EPU) as a moderating factor: when EPU is low (below one standard deviation below its mean), predictions are derived from the benchmark model; otherwise, predictions are based on the IVS. The sample period spans January 4, 1996, to June 28, 2019, with the initial training period ending on December 31, 2001.

EPU levels, lag behind the refined model. These findings provide further evidence of the economic significance of IVS predictability when combined with adaptive forecasting strategies.

In the future, as the literature on predictive models under varying economic conditions continues to evolve, identifying variables with strong predictive power during low EPU periods may enable the construction of even more robust models. Combining such variables with IVS in a dynamic framework could further enhance forecasting accuracy and economic outcomes. These directions present exciting opportunities for future research.

6. Robustness Test

6.1. Substituting ln(EPU) with EPU

In our main analysis, we applied a natural logarithm transformation to EPU to reduce skewness and approximate normality. In this section, we omit this transformation to evaluate whether our main findings remain robust, including the interaction regerssion, grouped regression and OOS analysis. These results can be seen in table 9,10,11.All interaction terms between EPU and IVS are significant and positive, indicating that EPU amplifies the predictive power of IVS. From the grouped regression and out-of-sample (OOS) analyses, we observe that during low EPU periods, the predictability of IVS diminishes, both for daily and monthly horizons. This additional analysis using the raw EPU data ensures that the log transformation do not introduce any unintended distortions.

7. Conclusion

Against the backdrop of increasing economic uncertainty, this study explores the impact of EPU on the predictive power of the IVS for market premiums. We validate the interaction between EPU and IVS in predicting market returns and examine how IVS's predictive power varies across distinct EPU levels. Our findings confirm that EPU amplifies IVS's forecast power, with IVS's predictive ability diminishing under low EPU and strengthening under high EPU. This pattern remains consistent across daily and monthly horizons. Moreover, varying EPU thresholds are tested to confirm the robustness of our findings.

Building on this robustness in-sample, we further validate the results out-of-sample. Comparisons between periods of low EPU and other periods reveal a significant decline in IVS's forecasting performance during low EPU periods across most horizons. At the daily level, in particular, IVS demonstrates no OOS predictability under low EPU conditions compared to historical mean forecasts. Rigorous statistical testing further confirms that IVS predictability diminishes across all horizons when EPU is low.

These findings hold significant implications. By incorporating EPU as a moderating factor, we demonstrate how enhanced models can significantly improve out-of-sample predictability and deliver

Table 9: Robustness Test: Interaction Regression

	Daily Horizon			Monthly Horizon		
	(1)	(2)	(3)	(4)	(5)	(6)
IVS	1.509***	1.427***	1.505***	0.637**	0.542*	0.639**
	(0.525)	(0.530)	(0.524)	(0.308)	(0.304)	(0.307)
IVS×EPU	1.054**	1.049**	1.071**	0.591**	0.593**	0.579**
	(0.467)	(0.469)	(0.462)	(0.265)	(0.258)	(0.271)
EPU	0.611	0.752	0.296	0.395	0.471*	0.400
	(0.495)	(0.463)	(0.446)	(0.273)	(0.264)	(0.273)
Recession_dummy		-2.162			-1.671	
		(1.964)			(1.205)	
VIX			0.812			-0.047
			(0.794)			(0.289)
Observations	5910	5910	5906	280	280	280
Adjusted R^2	0.013	0.013	0.013	0.075	0.083	0.072

* This table presents the robustness test results of the interaction regression, assessing whether Economic Policy Uncertainty (EPU) moderates the predictive relationship between the Implied Volatility Spread (IVS) and market excess returns. For a more detailed discussion of the methodology and results, please refer to Table 2.

Table 10: Robustness Test with EPU: Grouped Regression

	Daily		**Monthly**	
	Low EPU	**High EPU**	**Low EPU**	**High EPU**
Horizon 1	IVS: 2.477 R^2: 1.128	IVS: 5.845** R^2: 2.269	IVS: 0.912* R^2: 8.864	IVS: 3.351*** R^2: 29.701
Horizon 2	IVS: 1.944* R^2: 1.245	IVS: 4.153*** R^2: 2.102	IVS: 0.647 R^2: 5.981	IVS: 1.446** R^2: 32.332
Horizon 3	IVS: 1.132* R^2: 0.472	IVS: 3.339*** R^2: 2.170	IVS: 0.182 R^2: -2.290	IVS: 0.971** R^2: 27.940
Horizon 4	IVS: 1.402*** R^2: 1.664	IVS: 3.139*** R^2: 3.272	IVS: 0.157 R^2: -4.579	IVS: 0.385** R^2: 1.213

*The table presents the robustness test of EPU using the grouped regression method. For further details on the methodology and results, please refer to table 3.

Table 11: Robustness Test with EPU: Statistical Test for Comparison of OOS Predictability Across Different EPU Levels

	Horizon (daily)				Horizon (Monthly)			
	1	2	3	4	1	3	6	12
γ	23.62	13.11	8.49	6.20	2.50	3.12	1.43	0.29
cw	2.67	2.63	3.08	2.58	1.61	1.50	2.46	1.97

* This table presents the robustness test of EPU by the statistical test comparing the out-of-sample (OOS) predictability across different EPU levels. For further details on the methodology and results, please refer to table 7.

greater economic gains. For example, substituting IVS forecasts with benchmark forecasts during periods of low EPU improves utility gain by 2.5%.

This study highlights several promising directions for future research. Identifying additional predictors with stronger performance during low EPU periods and combining them with IVS forecasts for high EPU periods could result in more robust and adaptive forecasting models. Furthermore, the use of EPU as a proxy for environmental conditions could extend to examining the robustness of other predictive models. Additionally, this research could be applied to different industry indices and asset classes, such as commodities and exchange rates, to explore EPU's broader role in financial forecasting. Finally, future studies could assess the stability of other predictive models under varying economic conditions, using EPU as a moderating factor.

References

An, B.-J., 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.

Atilgan, Y., Bali, T. G., and Demirtas, K. O. (2015). Implied volatility spreads and expected market returns. *Journal of Business & Economic Statistics*, 33(1):87–101.

Avramov, D., Cheng, S., and Metzker, L. (2023). Machine Learning vs. Economic Restrictions: Evidence from Stock Return Predictability. *Management Science*, 69(5):2587–2619.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring Economic Policy Uncertainty*. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Baker, S. R., Bloom, N., and Davis, S. J. (2020). Covid-induced economic uncertainty. *NBER Working Paper*, 26983.

Bakshi, G., Kapadia, N., and Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1):101–143.

Bali, T. G., Brown, S. J., and Tang, Y. (2017). Is economic uncertainty priced in the cross-section of stock returns? *Journal of Financial Economics*, 126(3):471–489.

Bandi, F. M. and Perron, B. (2006). Long memory and the relation between implied and realized volatility. *Journal of Financial Economics*, 82(2):341–377.

Birru, J. and Young, T. (2022). Sentiment and uncertainty. *Journal of Financial Economics*, 146(3):1148–1169.

Brogaard, J. and Detzel, A. (2015). The Asset-Pricing Implications of Government Economic Policy Uncertainty. *Management Science*, 61(1):3–18. Publisher: INFORMS.

Campbell, J. Y. (2000). Asset pricing at the millennium. *Journal of Finance*, 55(4):1515–1567.

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.

Cascaldi-Garcia, D., Sarisoy, C., Londono, J. M., Sun, B., Datta, D. D., Ferreira, T., Grishchenko, O., Jahan-Parvar, M. R., Loria, F., Ma, S., et al. (2023). What is certain about uncertainty? *Journal of Economic Literature*, 61(2):624–654.

Chiang, T. C. (2019). Economic policy uncertainty, risk and stock returns: Evidence from g7 stock markets. *Finance Research Letters*, 29:41–49.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, 138(1):291–311.

Cremers, M. and Weinbaum, D. (2010). Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis*, 45(2):335–367.

Cujean, J. and Hasler, M. (2017). Why does return predictability vary over time? *Journal of Finance*, 72(6):2717–2754.

Dangl, T. and Halling, M. (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics*, 106(1):157–181.

Dong, X., Li, Y., Rapach, D. E., and Zhou, G. (2022). Anomalies and the expected market return. *The Journal of Finance*, 77(1):639–681.

Easley, D., O'hara, M., and Srinivas, P. S. (1998). Option volume and stock prices: Evidence on where informed traders trade. *The Journal of Finance*, 53(2):431–465.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

Fang, L., Bouri, E., Gupta, R., and Roubaud, D. (2019). Does global economic uncertainty matter for the volatility and hedging effectiveness of bitcoin? *International Review of Financial Analysis*, 61:29–36.

Fernández, M. F., Henry, Ó., Pybis, S., and Stamatogiannis, M. P. (2023). Can we forecast better in periods of low uncertainty? the role of technical indicators. *Journal of Empirical Finance*, 71:1–12.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6):1545–1578.

Goyal, A., Welch, I., and Zafirov, A. (2024). A comprehensive 2022 look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, page hhae044.

Gulen, H. and Ion, M. (2016). Policy uncertainty and corporate investment. *The Review of financial studies*, 29(3):523–564.

Han, B. and Li, G. (2021). Information content of aggregate implied volatility spread. *Management Science*, 67(2):1249–1269.

Henkel, S. J., Martin, J. S., and Nardari, F. (2011). Time-varying short-horizon predictability. *Journal of Financial Economics*, 99(3):560–580.

Jackwerth, J. C. and Rubinstein, M. (1996). Recovering probability distributions from option prices. *The journal of Finance*, 51(5):1611–1631.

Jenkins, D. G. and Quintana-Ascencio, P. F. (2020). A solution to minimum sample size for regressions. *PloS one*, 15(2):e0229345.

Kelly, B. T. and Gu, S. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5):2223–2273.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

Leippold, M., Wang, Q., and Zhou, W. (2022). Machine learning in the Chinese stock market. *Journal of Financial Economics*, 145(2, Part A):64–82.

Li, J. and Tang, Y. (2010). Ceo hubris and firm risk taking in china: The moderating role of managerial discretion. *Academy of Management Journal*, 53(1):45–68.

Liu, M. and Lin, T.-I. (2014). A skew-normal mixture regression model. *Educational and Psychological Measurement*, 74(1):139–162.

Ludvigson, S. C., Ma, S., and Ng, S. (2021). Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response? *American Economic Journal: Macroeconomics*, 13(4).

Manaster, S. and Rendleman Jr, R. J. (1982). Option prices as predictors of equilibrium stock prices. *The Journal of Finance*, 37(4):1043–1057.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). Forecasting the equity risk premium: The role of technical indicators. *Management Science*, 60(7):1772–1791.

Pástor, L. and Veronesi, P. (2013). Political uncertainty and risk premia. *Journal of financial Economics*, 110(3):520–545.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies*, 23(2):821–862.

Rapach, D. E. and Zhou, G. (2013). Forecasting stock returns. *Handbook of Economic Forecasting*, 2:328–383.

Rossi, B. (2013). Exchange rate predictability. *Journal of economic literature*, 51(4):1063–1119.

Stoll, H. R. (1969). The relationship between put and call option prices. *The Journal of Finance*, 24(5):801–824.

Wang, X., Li, J., and Ren, X. (2022). Asymmetric causality of economic policy uncertainty and oil volatility index on time-varying nexus of the clean energy, carbon and green bond. *International Review of Financial Analysis*, 83:102306.

Williams, R. (2012). Using the margins command to estimate and interpret adjusted predictions and marginal effects. *The Stata Journal*, 12(2):308–331.

Yfanti, S. and Karanasos, M. (2022). Financial volatility modeling with option-implied information and important macro-factors. *Journal of the Operational Research Society*, 73(9):2129–2149.

Zhu, Y. and Timmermann, A. (2022). Conditional rotation between forecasting models. *Journal of Econometrics*, 231(2):329–347.

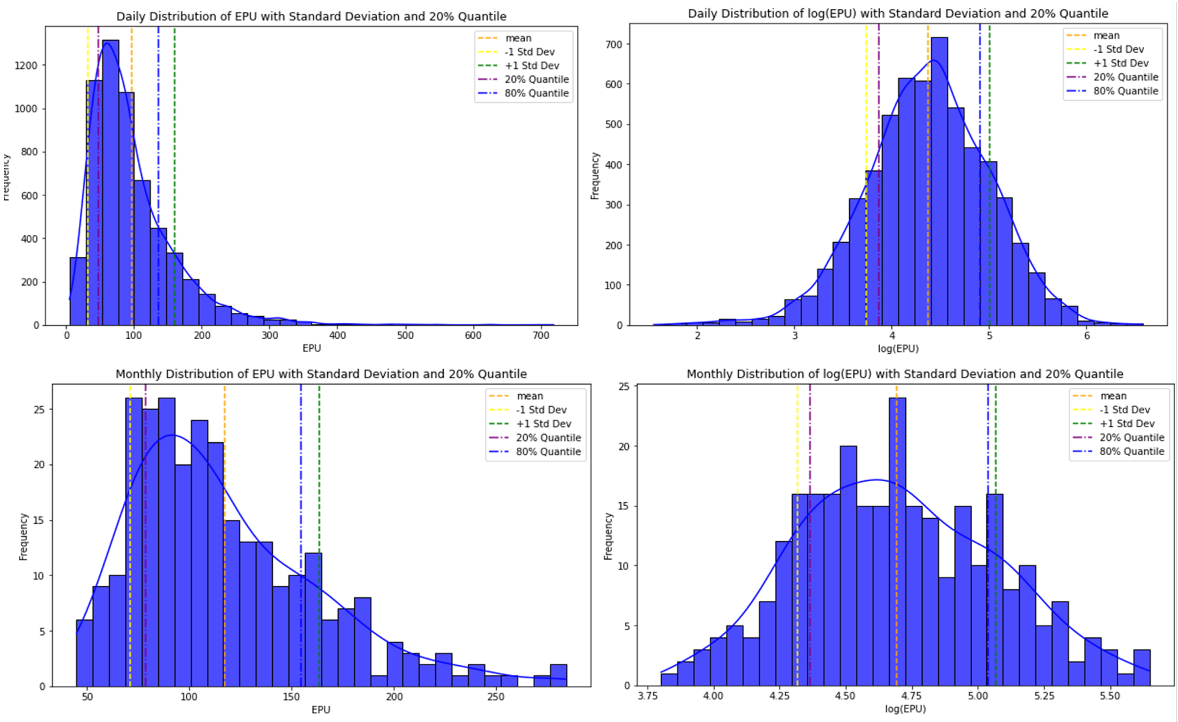

Figure 4: Distribution of EPU. This figure illustrates the distribution of EPU (left panels) and ln(EPU) (right panels) at both daily (top panels) and monthly (bottom panels) frequencies. The histograms represent the frequency of observations, while the overlaid lines show the smoothed density estimates. Key statistical markers, including the mean (yellow dashed line), one standard deviation above and below the mean (green and purple dashed lines), and the 20th and 80th percentiles (black and blue dashed lines), are highlighted to emphasize critical thresholds used in the empirical analysis. Notably, the distributions of EPU are right-skewed, whereas ln(EPU) exhibits a more symmetric shape, supporting the use of log transformation to normalize the data for robust statistical modeling. The monthly distributions reflect a smaller sample size, resulting in greater variability compared to the daily data.

8. Appendix

8.1. Data Supplementary Note

Figure 4 presents the distributions of EPU and its log transformation ln(EPU) across daily and monthly frequencies. The raw EPU data exhibit significant right skewness, especially at the daily level, indicating periods of extreme policy uncertainty. The log transformation normalizes these distributions, resulting in a more symmetric shape suitable for statistical analysis. Monthly data, due to a smaller sample size, show greater variability but follow a consistent pattern with the daily data. These characteristics justify the use of ln(EPU) in the main analysis to ensure robust and reliable results.

8.2. Grouped Regression Analysis: EPU Threshold Sensitivity Analysis

To ensure robustness of our grouped regression, this study adopts multiple thresholds for defining EPU levels, ranging from the lowest and highest 10% to 50% of ln(EPU), in increments of 5%. In each threshold, grouped regression is conducted to analyze the predictive power of IVS. This approach enables a more comprehensive evaluation of IVS's predictive power across varying levels of economic policy uncertainty, while also testing the sensitivity of the results to different threshold definitions. The choice of the 10% as the lowest threshold is driven by the fact that the monthly forecast dataset contains up to 280 observations[9]. To ensure reliability, each group analysis requires at least 25 samples (Jenkins and Quintana-Ascencio, 2020). Therefore, to maintain consistency and robustness in our analysis, we limit our study to groupings no lower than the 10 %.

As figure 5, we observe that across all horizons and thresholds, the high EPU group consistently displays higher coefficients and adjusted R-squared values, as well as lower p-values, compared to the low EPU group. Notably, the difference in *p_diff* between high and low EPU values often surpasses the 10% significance level, indicating that IVS coefficients in the low EPU group are frequently insignificant. This finding supports our main conclusion on a daily scale, demonstrating that it holds consistently across various EPU levels. Specifically, the differences are most pronounced when the threshold is set at the 10% and 20% levels. Looking at variations across different horizons, we find that the coefficient differences are more pronounced for shorter horizons. However, p-values and adjusted R-squared values show varying patterns across thresholds, with no consistent trend across different horizons.

In the monthly dimension, we similarly observed that setting the threshold at the top 20% results in more pronounced differences across all three metrics compared to other thresholds from figure 6. Specifically, in the p-value plot, a notable drop in p-value and adjusted R-squared differences between the low and high EPU groups occurs when the threshold is set at 20%. After the 35% threshold, however, these differences diminish, and our main finding no longer holds. From the perspective of horizon, we observe that the coefficient differences are most significant at shorter horizons, particularly when the horizon is set to 1. This suggests that the impact of EPU on IVS is strongest in terms of predictive coefficients at one month horizons. However, for p-values and adjusted R-squared values, no clear trend is evident across horizons.

8.3. Bootstrap Test for Sharpe Ratio Differences

To evaluate the significance of Sharpe ratio differences, we employ the bootstrap-based method proposed by Ledoit and Wolf (2008). The null hypothesis, $H_0 : \Delta = 0$, is tested using the studentized statistic:

$$d = \frac{|\Delta|}{s(\Delta)},$$

where Δ is the observed difference in Sharpe ratios between two portfolios, and $s(\Delta)$ is its standard deviation.

[9]The sample sizes vary slightly across different monthly horizons, with 280, 277, 271, and 259 observations for the 1-, 3-, 6-, and 12-month horizons, respectively.

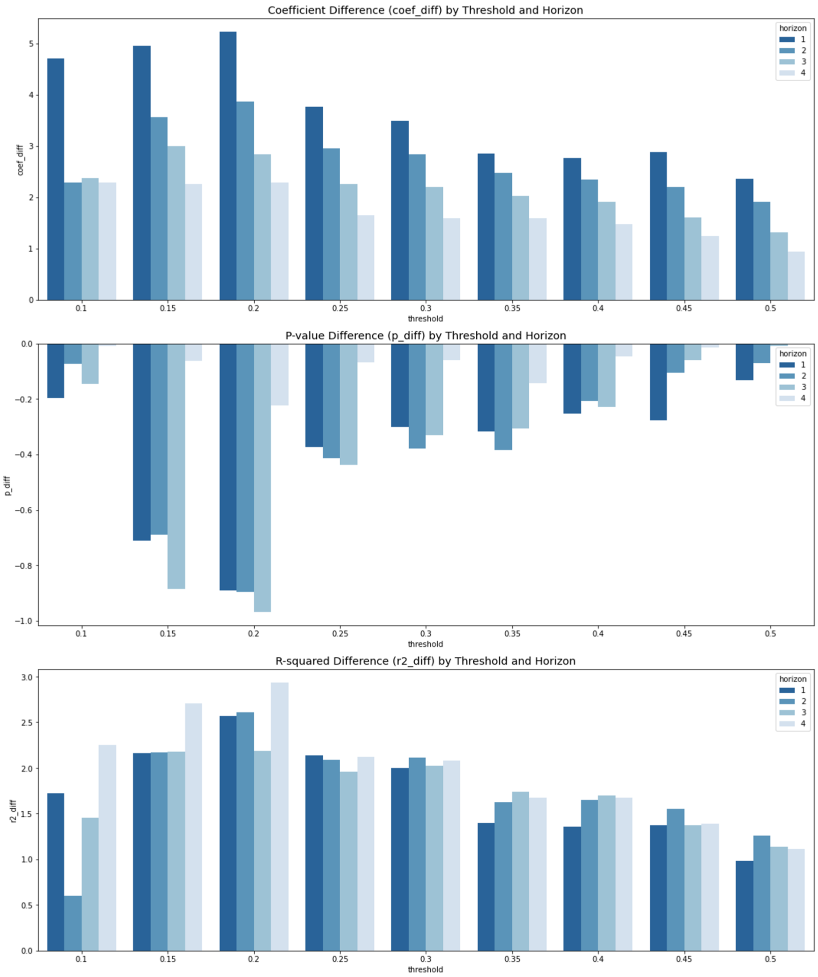

Figure 5: Statistical Differences in IVS between Low and High EPU Levels Across Thresholds and Daily Horizons. This figure illustrates the statistical differences in IVS's ability to predict market excess returns across varying thresholds, specifically by redefining high and low EPU levels based on different threshold values. Grouped regressions are performed for IVS under high and low EPU conditions to verify the robustness of our main findings against variations in EPU level definitions. The statistical differences include three metrics: coefficient difference (*coef_diff*), p-value difference (*p_diff*), and adjusted R-squared difference (*r2_diff*). This figure presents results for daily horizons from 1 to 4 days. We observe that across all horizons and thresholds, the high EPU group consistently shows increased coefficients and adjusted R-squared values, alongside lower p-values compared to the low EPU group. Notably, the range of 'p_diff' between high and low values often exceeds the 10% significance level, indicating that the IVS coefficients in the low EPU group are frequently insignificant. Each statistical measure is calculated in the same manner as shown in the tables. The data spans from January 4, 1991, to June 28, 2019.

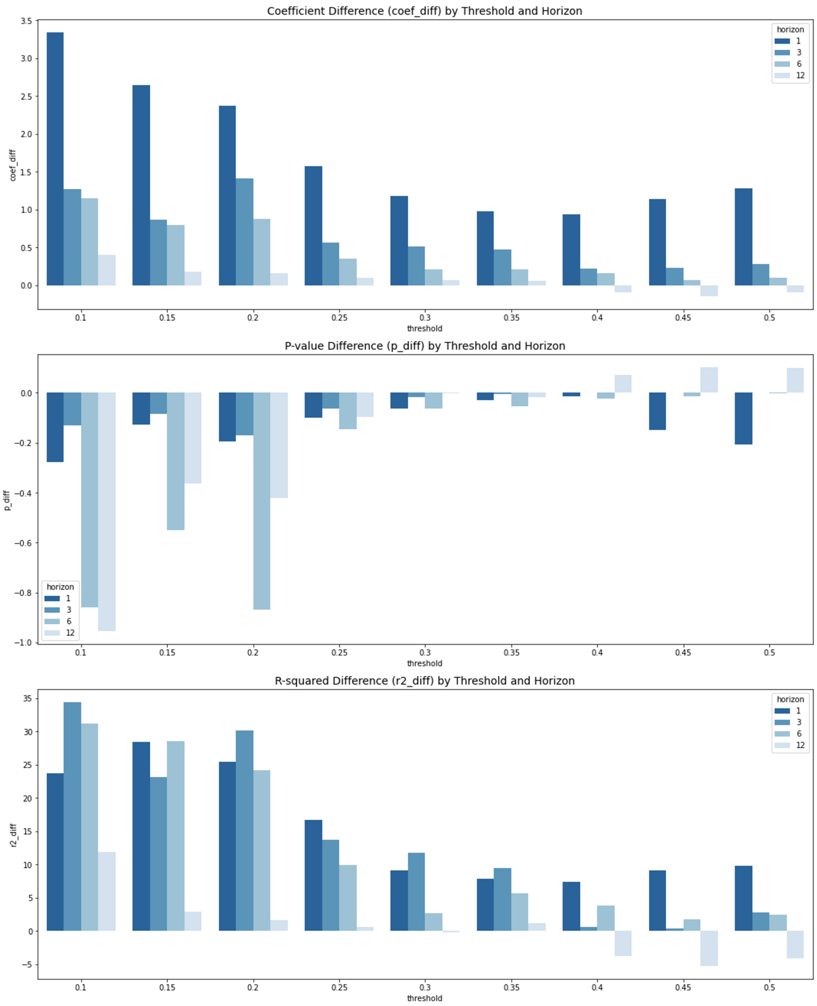

Figure 6: Statistical Differences in IVS between Low and High EPU Levels Across Thresholds and Monthly Horizons. This figure illustrates the statistical differences in IVS's ability to predict market excess returns across varying thresholds, specifically by redefining high and low EPU levels based on different threshold values. Grouped regressions are performed for IVS under high and low EPU conditions to verify the robustness of our main findings against variations in EPU level definitions. The statistical differences include three metrics: coefficient difference (*coef_diff*), p-value difference (*p_diff*), and adjusted R-squared difference (*r2_diff*). This figure presents results for daily horizons from 1 to 12 months. We observe that across all horizons and thresholds, the high EPU group consistently shows increased coefficients and adjusted R-squared values, alongside lower p-values compared to the low EPU group. Notably, the range of 'p_diff' between high and low values often exceeds the 10% significance level, indicating that the IVS coefficients in the low EPU group are frequently insignificant. Each statistical measure is calculated in the same manner as shown in the tables. The data spans from January 4, 1991, to June 28, 2019.

For each of the M bootstrap resamples, the centered studentized statistic is computed as:

$$\tilde{d}^{*,m} = \frac{|\Delta^{*,m} - \hat{\Delta}|}{s(\Delta^{*,m})}, \quad m = 1, \ldots, M,$$

where $\Delta^{*,m}$ represents the Sharpe ratio difference for the m-th bootstrap sample, and $\hat{\Delta}$ is the observed Sharpe ratio difference from the original sample.

We do the single tail test in our analysis, therefore the p-value is then calculated as:

$$\mathrm{PV} = \frac{\sum_{m=1}^{M} \mathbb{I}(\tilde{d}^{*,m} \geq d) + 1}{M + 1},$$

where $\mathbb{I}(\cdot)$ is an indicator function that equals 1 if the condition holds and 0 otherwise. This approach ensures a direct and efficient computation of the p-value, reflecting the likelihood of observing a Sharpe ratio difference as extreme as Δ under the null hypothesis.