Crash Narratives and Predictability *

John Cotter† Richard McGee‡ Xiaomeng Wang§

Abstract

This paper documents the superior predictive power of recently proposed Crash Narratives for US stock market volatility. Compared to other volatility predictors, Crash Narratives are found to perform better during periods of high volatility and NBER recessions. They also produce accurate out-of-sample forecasts of the conditional variance. By employing the conditional variance forecasts based on Crash Narratives, we reexamine the strong return predictability of the variance risk premium, and develop a notably profitable volatility-managed portfolio strategy. This strategy outperforms the traditional volatility-managed portfolio and buy-and-hold strategies, yielding larger alphas and Sharpe ratios, with cumulative returns at least 25% greater than those of alternative strategies.

Keywords: Stock Market Crash, Narratives, Natural Language Processing, Volatility Forecasting, Volatility-Managed Portfolio

JEL Classification: G11, G12, G17, E44

*We thank William N. Goetzmann, Jerry Parwada, and Sophie Moinas for helpful comments. We gratefully acknowledge the support of Science Foundation Ireland under grant numbers 16/SPP/3347, 13/RC/2106_P2, 17/SPP/5447 and 18/CRT/6224.

†john.cotter@ucd.ie. Smurfit Graduate Business School, University College Dublin, Dublin, Ireland

‡richard.mcgee@ucd.ie. Smurfit Graduate Business School, University College Dublin, Dublin, Ireland

§Corresponding author. xiaomeng.wang@ucdconnect.ie. Smurfit Graduate Business School, University College Dublin, Dublin, Ireland

1 Introduction

Understanding how market crash concerns influence market volatility is crucial for investors when developing risk management and portfolio strategies. Traditionally, market crash concerns have been quantified mainly through option-based tail-risk measures (Andersen et al., 2021; Bollerslev et al., 2011, 2015). However, a significant gap remains between these traditional measures of extreme crash concerns and the real concerns of average investors, as traditional economic and stock market data often fail to capture or retain rapidly changing sentiments (Manela and Moreira, 2017). In contrast, concerns about stock market crashes are persistently reported in financial newspapers, providing an alternative data source for measuring these concerns.

Goetzmann et al. (2022) have introduced a new approach to narrow this gap: a high-order measure of Crash Narratives extracted from Wall Street Journal (WSJ) news articles. In this paper, we conduct an empirical analysis to evaluate the relative importance of Crash Narratives in predicting US stock market volatility. Specifically, we compare Crash Narratives with three types of predictors, including tail risk, text-based volatility, and uncertainty measures. To further explore their practical implications, we examine whether this new measure improves the out-of-sample forecast of conditional variance, the performance of volatility-managed portfolios, and the predictability of the variance risk premium on market excess returns.

We begin with the assumption that the time variation in Crash Narratives extracted from the business press is a good proxy for the evolution of investors' concerns about market crashes. The timeliness and rich narrative content of business news data are significant advantages over standard financial data sets. These advantages arise from sophisticated human processing of complex contexts. In comparison with traditional measurements of crash concerns, narratives are non-parametric and not limited by strict model assumptions, when utilized to identify investors' beliefs. Natural Language Processing methods can uncover information from this rich and unique text data set. Crash Narratives are measured by semantic similarities between news articles and articles published in a window around the 1987 US market crash, using the

Doc2Vec model. A key advantage of this high-dimensional semantic space model is its ability to extract information from unstructured data.

In existing asset pricing papers, market crash concerns, typically measured by tail-risk measures, have been shown to predict market volatility (Wachter, 2013). Unlike traditional tail-risk measures, which are empirically estimated through complex procedures using option data, Crash Narratives measure the recall of the pivotal historical market crash (1987 Crash) in the financial press. These indices rise in response to reporting around disturbances (e.g. stock market crashes, wars, significant international political events, and financial crisis), as stories are fundamentally concerned with state changes intermediated by crash events (Bybee et al., 2024; Goetzmann et al., 2022). Disaster models illustrate that state changes lead to variations in market volatility (Gabaix, 2012; Gourio, 2012; Wachter, 2013). Motivated by these papers, we study whether fluctuations in Crash Narratives encode information about future market volatility.

We find that Crash Narratives effectively predict future market volatility[1]. More importantly, our paper is the first to shed light on the outperformance of Crash Narratives relative to alternative text-based volatility measures, traditional tail risk measures, and uncertainty measures. We demonstrate that the superior predictive power of Crash Narratives is primarily due to their stronger performance during periods of high volatility and the NBER recession.

Our work has three main contributions: first, we offer a more accurate estimate of the Conditional Variance: we include Crash Narratives in the Conditional Variance forecast model from Corsi (2009): a Heterogeneous Autoregressive model. We find that integrating Crash Narratives markedly improves the accuracy of Conditional Variance forecasts. The out-of-sample Mean Absolute Error of the forecasting model with Crash Narratives decreased by approximately 10% compared to the non-estimated model.

Second, we construct portfolios that scale monthly returns by the inverse of their con-

[1]We also replicate Crash Narratives constructed by the Financial Times news articles and a combination of the Wall Street Journal and Financial Times articles. We also examine their predictability and obtain similar results, which are available on request.

ditional variance. These volatility-managed SP500 portfolios decrease risk exposure during periods of high conditional variance and vice versa. We find that the performance of these portfolios can be improved using the conditional variance forecasting models with Crash Narratives. Among a group of volatility-managed portfolios, Crash Narratives-based trading strategies produce the largest alpha (10.1%) and the highest annualized Sharpe ratio (around 1). We begin with an investment of $1 in 2006. By December 2023, the volatility-managed S&P 500 portfolios, on the basis of the conditional variance estimated from Crash Narratives, accumulated to approximately $15—an improvement of at least 25% compared to all other strategies.

Third, we estimate the variance risk premium using the conditional variance forecast derived from Crash Narratives. Our new proxy demonstrates superior forecasting power for short-term returns, with Adjusted R2 increasing by about 3%. The strong predictability of the variance risk premium lends credence to the asset pricing framework of Bollerslev et al. (2009) and Campbell and Cochrane (1999); and Drechsler and Yaron (2011), which attributes variations in equity risk premiums to counter-cyclical shifts in risk aversion.

This paper relates and contributes to three strands of literature. The first strand involves empirical and theoretical research on measuring market crash concerns. Recent approaches range from option-based tail-risk measures (Bollerslev et al., 2015) to text-based methods that use news data (Baker et al., 2021; Goetzmann et al., 2022), with evidence suggesting that traditional tail-risk measures may not fully capture the evolving and heterogeneous nature of crash fears (Baker et al., 2021; Manela and Moreira, 2017). Second, a large and fruitful literature focuses on improving the accuracy of volatility forecasts by adding new predictors, such as macroeconomic variables (Chiu et al., 2018; Conrad and Loch, 2015a, 2015b; Engle et al., 2013; Paye, 2012), sentiment indices (Lee et al., 2002; Seo and Kim, 2015), textual measures (Behrendt and Schmidt, 2018), uncertainty measures (Asgharian et al., 2023; Goodell et al., 2020), and the CBOE volatility index (Fernandes et al., 2014). For example, the Macroeconomic Uncertainty Index captures uncertainty about future macroeconomic conditions (Jurado et al., 2015; Ludvigson et al., 2021). Our contribution is the first to use a news-based measure of market crash concerns—Crash Narratives—to predict realized volatility.

Broadly, our paper contributes to the rapidly growing literature in finance that uses text as data. Several textual analysis approaches are employed to quantify information from newspapers. A popular approach is to create a topic-specific compound full-text search statement and then to count the resulting number of articles normalized by a measure of normal word count. Two leading examples of this approach are the news-based Economic Policy Uncertainty Index (EPU) by Baker et al. (2016) and the Equity Market Uncertainty (EMV) index proposed by Baker et al. (2021). Advances in Natural Language Processing have enabled the extraction of narrative topics from news articles (Bybee et al., 2023, 2024; Dierckx et al., 2021; Dim et al., 2023; Hirshleifer et al., 2023), and Goetzmann et al. (2022) quantify Crash Narratives using semantic analysis. In this paper, we compare text-based measures and demonstrate that Crash Narratives provide superior volatility predictions.

The paper proceeds as follows. Section 2 reviews the background on measuring market crash concerns and textual analysis. Section 3 describes the news data to construct Crash Narratives, the stock market data to calculate realized volatility, and alternative volatility predictors. Section 4 tests the hypothesis that Crash Narratives are important predictors of market realized volatility and presents the models for monthly Conditional Variance forecasts. Section 5 reports the in-sample and out-of-sample forecasting results. Section 6 constructs volatility-managed portfolios based on Crash Narratives. Section 7 assesses the Variance Risk Premium's predictability on market excess return. Section 8 concludes.

2 Background

Estimating the fear of disasters is challenging but crucial because these rare events can significantly influence investors' beliefs about the stock market and, therefore, shape their investment strategies. It has been documented that market participants, including institutional investors, pay a "fear premium" to insure against disaster shocks in asset pricing (Barro, 2006; Bollen and Whaley, 2004; Bollerslev and Todorov, 2011; Driessen and Maenhout, 2007;

Gourio, 2012; Han, 2008; Rietz, 1988; Wachter, 2013; Welch, 2016).

Recent years have witnessed a significant evolution in research on measuring market crash concerns, with various methodologies reflecting the complex and diverse nature of market sentiment. A significant segment has focused on tail-risk measures (Andersen et al., 2015; Bollerslev and Todorov, 2011; Bollerslev et al., 2015; Drechsler, 2013; Gao et al., 2019; Kelly and Jiang, 2014; Wachter, 2013). For example, Left Tail Volatility, developed by Bollerslev et al. (2015), always serves as an important indicator for market crash concerns. Specifically, it measures the expected (risk-neutral) return volatility due to a ten-standard deviation or larger downward move in the S&P 500 Index within a one-week period. However, these tail risk measures cannot comprehensively and timely capture crash concerns of most market participants, stemming from the fact that market crash concerns are determined by events whose heterogeneous nature varies depending on the period considered and are characterized by a non-uniform and unknown distribution among economic agents (Baker et al., 2021; Manela and Moreira, 2017).

More recently, the role of media narratives in shaping market concerns has gained prominence. Shiller (2014) proposes that the market fluctuates with different mindsets and that the aggregate price changes in the stock market reflect different perceptions. Shiller (2017) coined the term "narrative economics" to describe the study of how stories, explanations, and justifications of events we tell ourselves and others shape individuals' behaviors and their decision-making and drive economic fluctuations. He references the narrative concerning the stock market collapse on 28 October 1929 as a crash narrative, marking it as the inaugural narrative of the Great Depression. Baker et al. (2021) illustrates that newspaper articles mirror these mindsets and their shifts over time, and they drive market volatility.

Several textual analysis approaches are used to quantify information from news articles. A popular approach is to create a topic-specific compound full-text search statement and then to count the resulting number of articles normalized by a measure of normal word count. This creates a univariate time series that can be used in least squares regression analysis. A key advantage of this method is that articles are highly likely to be pertinent to the topic of interest. However, this approach depends heavily on the judgement of the econometrician, since the selection of key words for specific topics is determined by them. Two leading examples of this approach are the news-based Economic Policy Uncertainty Index (EPU) by Baker et al. (2016) and the Equity Market Uncertainty (EMV) index proposed by Baker et al. (2021). The second approach uses machine learning models to estimate the relationship between news coverage and market volatility. This method addresses the challenge of effectively managing the large dimensionality of the feature space, allowing the data to inform the analysis with minimal human intervention. Manela and Moreira (2017) utilized this approach to develop a monthly news-based implied volatility (NVIX) measure using abstracts and headlines from front-page articles of the Wall Street Journal. However, a potential drawback is its reliance on a relatively small text corpus.

The development of Natural Language Processing has stimulated the use of analytical methods to detect degrees of "narrativity". The Natural Language Processing approach not only lets the data speak for itself (not rely on the econometricians' key words selection and judgment), but also depends on larger text corpus (the full text and all news articles related to a stock market). Several papers have thus far applied topic models to extract narratives from news (Bybee et al., 2023, 2024; Dim et al., 2023; Hirshleifer et al., 2023, 2025). For example, Bybee et al. (2024) find that news narratives align with various economic activities and can predict aggregate stock market returns. Dierckx et al. (2021) demonstrate that a novel topic model, which combines Doc2Vec and Gaussian mixture models, outperforms traditional topic models, such as Latent Dirichlet Allocation, in VIX prediction. Goetzmann et al. (2022) quantify the crash narrative based on newspaper articles published after global market crashes using a high-dimensional semantic space model (Doc2Vec) and then use the crash narrative to measure market crash concerns.

3 Data

In this section, we present the data to calculate realized volatility and conditional variance (Section 3.1). We introduce the news data to construct Crash Narratives and alternative volatility predictors for comparative analysis (Section 3.2). The sample period spans from 31 May 1996 to 31 December 2023. Finally, Sections 3.3 and 3.4 show summary statistics and time-series properties.

3.1 Realized Volatility and Conditional Variance

For the stock market data, we download the daily SP500 index from 31 May 1996 to 31 December 2023. The monthly realized volatility is calculated as the square root of the sum of the squared daily SP500 returns within a month.

3.2 Crash Narratives and Other Predictors

We follow the methodology outlined in Goetzmann et al. (2022) to construct Crash Narratives[2]. Our construction relies on data from a leading financial newspaper, the Wall Street Journal (WSJ). We search the digital archives of the WSJ to obtain articles that contain the "stock market" or are categorically tagged under the subject "stock market". We use the ProQuest database to conduct a search within the Wall Street Journal (WSJ) spanning 31 May 1996, through 31 December 2023 [3]. This search yields a total of 154989 articles. Given the limitation of the search term regarding a word count restriction, we further devise a filter to select articles encompassing no less than 200 words. We download the full text of these articles, including their headlines, leading paragraphs, and topic tags. Articles from 1 October 1987 to 31 October 1987 are downloaded from the Factiva database, as WSJ news articles prior to 31 May 1996 are not accessible in our ProQuest database [4]

Based on Goetzmann et al. (2022), we construct Crash Narratives through a four-step procedure. First, we trained the Doc2Vec model using all WSJ articles from 1987 to 2023 (see A for details of the Doc2Vec model.)[5]. In the second step, we compute the average

[2]Since Crash Narratives are not available for direct download, we replicate the approach detailed in Goetzmann et al. (2022).

[3]The search term for the WSJ is: (((stock NEAR/5 market) OR SU(stock) AND WC > 200 OR SU(securities)) AND la.exact("English")) AND (publication.exact("Wall Street Journal") AND bdl(1000001)).

[4]These articles are acquired from the Factiva Database using the search term: stock NEAR5 market and WC > 200 or ns=M11 and WC > 200, where "ns=M11" signifies that the articles belong to the subject of the equity market. This search term in the Factiva database yields search results that are nearly identical to those obtained from the ProQuest database.

[5]The training process encompasses 100 iterations, based on a feature vector size of 250, a maximum distance

cosine similarity of each article to those published in the WSJ from October 20 to October 23, 1987. Higher average cosine similarities indicate greater textual similarity. However, due to language evolution over time, articles tend to exhibit higher similarities to recent periods regardless of narrative content. To address this, we account for the effects of language evolution by recalculating similarities in step three. This adjustment is based on the average cosine similarity of each article to articles published from October 5 through October 9, 1987, capturing a pre-crash baseline. The adjusted similarity is calculated as the difference between the natural logarithm of one plus the average cosine similarities. Finally, we average these adjusted similarity scores for articles from the same day. This is called *Crash Narratives*.

For comparison, we apply several alternative text-based volatility measures, traditional tail-risk measures, and uncertainty measures as volatility predictors, as shown in Table 1.

[Insert Table 1 here]

Both Equity Market Volatility and Economic Policy Uncertainty are downloaded from Economic Policy Uncertainty's website. Left-Tail Volatility is acquired from the Chicago Board Options Exchange. Macroeconomic Uncertainty Index is available from Ludvigson's website. [6]

3.3 Summary Statistics

Table 2 presents the summary statistics of Crash Narratives and alternative predictors. Daily Crash Narratives are calculated as the average of adjusted cosine similarities between WSJ stock market news articles published on a given day and those from October 20–23, 1987.

[Insert Table 2 here]

of 5, a minimum word frequency of 10 across all articles, a sub-sampling threshold of 10^{-5}, and a negative sample of 5.

[6]Manela and Moreira (2017) developed a monthly news-based implied volatility (NVIX) measure using abstracts and headlines from front-page articles of the Wall Street Journal. NVIX can be downloaded from Moreira's website, while the monthly time series are available only till March 2016. We examine the predictability of NVIX on realized volatility for the overlapping time frame and compare the results with Crash Narratives. The results support our findings and are available upon request.

Crash Narratives$_t$ = 0 indicate that the market sentiment expressed by the WSJ on a given day shows no particular concern about a market crash, as the average cosine similarities between the stock market articles on day t and those during the 1987 crash are identical to the average cosine similarities between stock market articles on day t and those from the benchmark period (two weeks prior to the 1987 crash). The cosine similarities quantify not only word choice but also language and semantic structure similarities. Articles with the same language and semantic structure (that is, narrative) suggest that they describe the same events and tell the same story. $0 <$ Crash Narratives$_t < 1$ indicate that the day's sentiment is modestly more aligned with the crash period than with the benchmark. The marginally negative mean (e.g., -0.48 for Crash Narratives) means that, on average, cosine similarities between stock market articles and those from the 1987 Crash are slightly lower compared to the cosine similarities between stock market articles and those from the benchmark period.

A highly negative crash narrative value (Crash Narratives$_t \leq -1$) suggests that most stock market news articles on that day are irrelevant to market crashes, and instead tell stories considerably analogous to those from the benchmark period, thus a lower possibility of market crash.

Conversely, a highly positive crash narrative value (Crash Narratives$_t \geq 1$) means that the media perceive an increased likelihood of a market crash, since the average cosine similarities between the stock market articles on day t and those during the 1987 crash are much higher compared to the average cosine similarities between the stock market articles on day t and those from the benchmark period, suggesting that substantially stock market articles published on day t are more relevant to the 1987 crash than with stories during the benchmark period. This is because articles with high adjusted similarity scores are similar not only in word choice but also in semantic context to those published during crash periods. For example, the highest WSJ Crash Narrative score was 3.24 occurred on July 31, 2011, when the financial markets were dominated by concerns about the US debt ceiling crisis and the potential for a government default. The five largest Crash Narratives were observed during the COVID-19 pandemic (March 17, 2020, the S&P 500 plunged nearly 12%, one of its worst single-day losses since 1987's Black Monday on March 16, 2020), the Subprime Crisis (October 14, 2007, when the subprime mortgage market in the US was collapsing, and October 12, 2008, the Dow Jones Industrial Average experienced one of its largest single-week declines in history on October 10, 2008), and the European Sovereign Debt Crisis (April 2, 2010, when Greece announced plans to seek European Union and International Monetary Fund assistance to manage its sovereign debt crisis.

3.4 Time-series Properties

Figure 1 shows that stock market crashes, wars, significant international political events, and financial crises play an important role in shaping Crash Narratives. The spikes of Crash Narratives suggest that journalists have broadly employed previous narratives used after the 1987 crash. Despite variations in context, they aim to alert readers' towards market crashes by recalling not only the previous crash events, but also the fears during the crash periods. Thus, Crash Narratives are reasonable proxies for investor's market crash concerns.

[Insert Figure 1 here]

We also plot alternative volatility predictors alongside US realized volatility in Figure 1. Most predictors, excluding Macroeconomic Uncertainty Index, are highly volatile, with peaks during the 2007–2008 financial crisis and the recent COVID-19 pandemic, similar to the pattern observed in realized volatility. Notably, the most pronounced divergence between Left-Tail Volatility and realized volatility occurs in the periods leading up to financial crises, such as the Asian Financial Crisis and the Subprime Crisis. During these times, realized volatility, Crash Narratives, and Equity Market Volatility begin to rise, whereas Left-Tail Volatility remains low. Additionally, compared to Equity Market Volatility and Crash Narratives, the two uncertainty measures (Economic Policy Uncertainty and Macroeconomic Uncertainty Index) show less similarity to the RV pattern. Economic Policy Uncertainty is more sensitive to political events than stock market events. In contrast, Macroeconomic Uncertainty Index is smoother compared to realized volatility, as Jurado et al. (2015) explains that much of the variability in the stock market is not driven by changes in uncertainty across the broader economy.

The heat map in Figure 2 shows the contemporaneous correlation matrix of monthly realized volatility, VIX, Crash Narratives, and other predictors. As expected, Crash Narratives are significantly and positively correlated with realized volatility. In particular, the correlation between realized volatility and Crash Narratives (0.75) is higher than those between realized volatility and Left Tail Volatility, Economic Policy Uncertainty, and Macroeconomic Uncertainty Index (0.64, 0.41, and 0.55, respectively). Furthermore, the correlation between VIX and Crash Narratives (0.77) is higher than between realized volatility and Crash Narratives, suggesting the forward-looking nature of Crash Narratives, similar to VIX.

[Insert Figure 2 here]

4 Methodology

In this section, we begin by testing the hypothesis that time variation in market crash concerns is a key predictor of future equity market volatility. In Section 4.1, we perform a simple predictive analysis, and a sub-sample analysis, and use quantile regressions to consider volatility distributions. Next, applying Principal Component Analysis, we assess whether Crash Narratives provide additional information in realized volatility prediction compared to the other four alternative predictors. Finally, in Section 4.2, we present the models for monthly Conditional Variance forecasts and discuss the estimation methods and evaluation metrics.

4.1 Simple Predictive Analysis, Sub-Sample and Quantile Analysis, Principal Component Analysis

We examine the monthly predictability of Crash Narratives on realized volatility. The monthly time series are obtained by averaging daily time series within a specific month. The full sample period is June 1996 to December 2023.

We then compare the predictability of Crash Narratives with alternative predictors in Table 1 and perform sub-sample comparisons. First, previous research has demonstrated that the text-based measures - Economic Policy Uncertainty and Equity Market Volatility - possess significant predictive power for market volatility (Fang et al., 2020; Li et al., 2023; Ma et al., 2022, 2023; Su et al., 2017). Second, traditional measures of market fear are widely recognized for their ability to predict future realized volatility. Bollerslev et al. (2015) demonstrates that Left Tail Volatility, which is derived from the option surface and associated with compensation for left jump tail risk, represents separate state variables that drive future market volatility. Third, following Knightian's definition and the early studies on uncertainty by Bernanke (1983) and Dixit and Pindyck (1994), academics and practitioners have sought to objectively quantify economic uncertainty, aiming to reflect the prevailing uncertainty in the process of decision making by economic agents. Investors in financial markets, who are concerned about uncertainty, respond by evoking a gradual slowdown or sometimes a sharp decline in asset price returns. This response simultaneously leads to a spike in volatility (Pástor and Veronesi, 2013), as several theoretical and empirical literature has shown that stock market volatility is driven by uncertainty (Engle et al., 2013; Fisher et al., 2022; Jurado et al., 2015; Paye, 2012). We apply two uncertainty measures: Economic Policy Uncertainty and Macroeconomic Uncertainty Index.

We estimate Crash Narratives using the Doc2Vec model, which captures the semantic meanings of news articles. Our approach not only lets the data speak for itself, in contrast to Economic Policy Uncertainty and Equity Market Volatility (they rely on the econometricians' key words selection and judgment), but also relies on larger text corpus, the full text and all news articles related to a stock market in the sample period, rather than comparing abstracts and headlines of front-page articles used by the NVIX. These methodological advantages lead us to hypothesize that Crash Narratives can capture distinct pieces of information. The monthly realized volatility (RV) predictive regressions in the full sample and four subsamples are shown as below:

$$RV_{t+1} = \beta_0 + \beta_1 X_t + \epsilon_{t+1}, \tag{1}$$

where $X \in$ [Crash Narratives, Left-Tail Volatility, Equity Market Volatility, Economic Policy

Uncertainty, Macroeconomic Uncertainty Index].

Stock market crashes are typically linked to significant and unforeseen spikes in market volatility. To explore the likelihood of variations in the relationship between Crash Narratives and market volatility across its distribution, we use a quantile regression approach. Therefore, we are able to explore whether (or not) the predictability of Crash Narratives is most pronounced for volatility at the highest quantile. This method was introduced by Koenker and Bassett (1978) and is widely used to unravel the dependence structure between the financial and economic variables. The monthly realized volatility (RV) quantile regressions are shown below:

$$Q_{RV_{t+1}}(\tau|X_t) = \phi(\tau) + \lambda(\tau)X_t, \tag{2}$$

where $X \in$ [Crash Narratives, Left-Tail Volatility, Equity Market Volatility, Economic Policy Uncertainty, Macroeconomic Uncertainty Index] and $\tau \in$ [0.1, 0.25, 0.5 0.75, 0.9], representing 10th, 25th, 50th, 75th and 90th quantile.

We further examine whether Crash Narratives provide additional information in realized volatility prediction compared to the other four alternative predictors. Since Crash Narratives and the other predictors are highly correlated, as shown in Figure 2, including them all directly in a regression could lead to multicollinearity issues, potentially distorting coefficient estimates and reducing the reliability of the results. To address this, we perform Principal Component Analysis on the four predictors to extract uncorrelated principal components (PCs) that retain most of the information. We then run a predictive regression that includes both Crash Narratives and the four principal components:

$$RV_{t+1} = b_0 + \sum_{i=1}^{4} b_i PCi_t + b_5 CN_t + \epsilon_{t+1}, \tag{3}$$

where CN represents Crash Narratives from the Wall Street Journal, and the PCs represent the orthogonalized components of Left-Tail Volatility, Equity Market Volatility, Economic Policy Uncertainty, and Macroeconomic Uncertainty Index. This approach mitigates multicollinearity and isolates the unique predictive power of Crash Narratives.

4.2 Conditional Variance Forecasting Analysis

4.2.1 Forecasting Model Estimation

It is now generally accepted that direct time series models based on realized volatility have more accurate predictability compared to popular GARCH and stochastic volatility models (Andersen et al., 2003; Chen et al., 2012). A good example of these models is the Heterogeneous Autoregressive model proposed by Corsi (2009). This model considers the volatility components realized over different interval sizes. It is able to reproduce the same volatility persistence observed in the empirical data as well as many of the other main stylized facts of financial data, while remaining parsimonious and easy to estimate. [7]

This model can be economically explained by the Heterogeneous Market Hypothesis by Müller et al. (1999), which recognizes the presence of heterogeneity among traders. Typically, a financial market consists of participants with a wide range of trading frequencies. The main idea of the hypothesis is that participants with varying trading frequencies recognize, respond to and generate different kinds of volatility components. To simplify, Corsi (2009) identify three primary volatility components: short-term participants who trade daily or more frequently, mid-term investors who usually adjust their portfolios weekly, and long-term market players with holding periods of at least a month. Thus, the daily Heterogeneous Autoregressive model has three heterogeneous volatility components, each of which is generated by the actions of different types of market participants.

The simplicity of the forecasting model of Corsi (2009) allows it to be easily extended in various directions. Other statistically and economically significant variables could simply be added as additional regressors. For example, Bekaert and Hoerova (2014) add VIX as a pre-

[7] The autocorrelations of squared returns exhibit strong persistence that spans long periods (months). Return distributions across different time horizons display fat tails, meaning that return probability density functions exhibit leptokurtosis with shapes dependent on the time scale, and they converge to a normal distribution very slowly as time scales lengthen. Financial data also show signs of scaling and multiscaling. Conventional GARCH and stochastic volatility models fail to capture all of these characteristics, while the Heterogeneous Autoregressive model can.

dictive variable in the HAR model, motivated by the finding that option prices as reflected in implied volatility should have information about future stock market volatility (Christensen and Prabhala, 1998). They evaluated a wide variety of state-of-the-art volatility forecasting models and found that one of the winning models is the HAR model supplemented with the squared VIX.

Following the findings in Section 4.1 and the benefits of the Heterogeneous Autoregressive model, we extend the monthly HAR model by including exogenous predictive variables, such as Crash Narratives, Equity Market Volatility, Left Tail Volatility, Economic Policy Uncertainty, Macroeconomic Uncertainty Index, which we call the HAR-X model.

First, we estimate the monthly HAR-X model for each predictor variable one at a time, with lags of four horizons of the variable (one-, three-, six-, and twelve-month) to determine how many horizons are included in the subsequent analysis. Since the parameters of the 3-, 6- and 12-month horizons for all variables are insignificant, we only include the one-month lag in the subsequent analysis.

Model 1 is the monthly HAR model. Model 2 adds Crash Narratives to Model 1, respectively. Model 3 adds Left Tail Volatility to Model 1. Model 4 incorporates the news-based Equity Market Volatility into Model 1. Models 5 and 6 include two additional uncertainty measures: Economic Policy Uncertainty and Macroeconomic Uncertainty Index, respectively, to Model 1. Our general forecasting model can be represented as follows:

$$RV_{t+1}^2 = c_0 + c_1 RV_t^2 + c_2 X_t + \varepsilon_{t+1}, \tag{4}$$

where

$$X \in \begin{Bmatrix} \text{Crash Narratives,} \\ \text{Left Tail Volatility,} \\ \text{Equity Market Volatility,} \\ \text{Economic Policy Uncertainty,} \\ \text{Macroeconomic Uncertainty Index} \end{Bmatrix}.$$

The monthly realized variance, RV_t^2, is calculated as the sum of the squared daily SP500 returns within a month, with returns presented as percentages.

Estimation noise is widely recognized to negatively affect out-of-sample forecast accuracy. Consequently, simpler models, such as the martingale model, might perform better than more complex models. Therefore, we also consider a non-estimated model: the lagged realized variance (Model 7, the model used in Bollerslev et al. (2009)). Subsequently, we evaluate multiple forecasting models to select the model with the best forecasting performance.

4.2.2 Model Evaluation Procedure

The model evaluation procedure involves assessing the in-sample and out-of-sample forecasting performance of the seven models. We estimate all models by using Ordinary Least Squares regressions with heteroskedasticity-consistent standard errors, following Newey and West (1987) [8] using monthly data from June 1996 to December 2023. We use standardized predictor variables (divided by standard deviation) to make the estimated coefficients comparable.

We report a number of metrics for a comparison of the in-sample predictions of the models, including: adjusted R-squared, the Bayesian Information Criterion (BIC), the F-test and Wald test. The F-test and Wald test are used to analyze the contribution of exogenous predictors, compared to only including RV_{t-1}^2.

[8] We use 3 lags for both the monthly in-sample and out-of-sample analysis. The results are qualitatively robust to the choice of lags.

Although in-sample analysis provides more efficient parameter estimates and thus more precise forecasts by using all available data, Welch and Goyal (2008) argue that out-of-sample tests are more appropriate to avoid the in-sample overfitting issue. The out-of-sample forecasts for the various conditional variance models are based on a rolling estimate window of 120 months. More specifically, we use data from June 1996 to May 2006 as the initial estimation period so that the prediction evaluation period spans June 2006 to December 2023. The initial in-sample estimation period balances between the desire to have enough observations to accurately estimate the initial parameters and the desire for a relatively long out-of-sample period for evaluation. We also compare the out-of-sample predictions of the models with the non-estimated model. To quantify out-of-sample forecast accuracy, we first apply two different criteria: Root Mean Square Error (RMSE), Mean Absolute Error (MAE). Comparative analysis of forecast error measures across different models is performed using the Diebold and Mariano (1995) test. We apply this test to examine whether the RMSE and MAE of the estimated models vary significantly from those of the non-estimated model (Model 7).

Welch and Goyal (2008) show that the historical average is a very stringent out-of-sample benchmark, and forecasting models generally fail to outperform the historical average. Hence, we also evaluate the out-of-sample predictive performance based on the widely used R^2_{OOS} (Campbell and Thompson, 2008) and Mean-Squared Prediction Error (MSPE) statistic (Clark and West, 2007). The R^2_{OOS} measures the proportional reduction in the MSPE for the predictive regression relative to the historical average benchmark:

$$R^2_{OOS} = 1 - \frac{\sum_{t=1}^{T} RV_t^2 - \widehat{RV}_t^2}{\sum_{t=1}^{T} RV_t^2 - \overline{RV}_t^2}, \tag{5}$$

where $\widehat{RV_t}$ is the fitted value from predictive regressions of variance estimated recursively with information available at time t-1. $\overline{RV_t}$ denotes the historical average benchmark estimated through period t-1. R^2_{OOS} lies in the range $(\infty, 1]$. $R^2_{OOS} > 0$ means that the predictive models outperform the historical average in terms of MSPE. As a second diagnostic, we use the MSPE statistic of Clark and West (2007) (the CW test). Under the null hypothesis, the historical average MSPE is less than or equal to that of the predictive regression. The test is one-sided (upper-tail), with the alternative that the historical average MSPE exceeds the predictive regression MSPE. Equivalently, we test H0: $R^2_{OOS} \leq 0$ against H1: $R^2_{OOS} > 0$.

5 Forecasting Results

In this section, we report that Crash Narratives can predict future market volatility and outperform alternative text-based volatility measures, traditional tail risk measures, as well as uncertainty measures. In Section 5.1, our subsample analysis reveals that the superior predictive power of Crash Narratives is primarily due to their stronger performance during periods of high volatility and NBER recession. Moreover, quantile regressions indicate that the predictive power of Crash Narratives increases with higher quantiles, suggesting that Crash Narratives may amplify predictability in more volatile periods. In Section 5.2, we find that the model with Crash Narratives produces a more accurate out-of-sample forecast of the Conditional Variance compared to models without Crash Narratives.

5.1 Simple Predictive Analysis, Sub-Sample and Quantile Analysis, Principal Component Analysis

The predictive results are presented in Table 3. The first column shows that Crash Narratives are statistically significant in forecasting realized volatility. In particular, Crash Narratives exhibit the highest coefficient of 0.624 among all predictors, and a 1% increase in Crash Narratives is associated with a subsequent 0.624% increase in realized volatility for the following month. Whilst all of the predictors are significant, the regression of realized volatility on the lagged Crash Narratives yields the highest adjusted R-squared value of 33.8%.

[Insert Table 3 here]

In addition, Crash Narratives exhibit higher adjusted R-squared compared to Equity Market Volatility in forecasting realized volatility, supporting our hypothesis that Crash Narratives provide more comprehensive information on future market volatility. [9] Left Tail Volatility

[9]Both Crash Narratives and Equity Market Volatility show a higher adjusted R-squared compared to NVIX, consistent with explanations of Baker et al. (2021), who attribute Equity Market Volatility's superior performance to the advantage of using a much larger text corpus, which enhances volatility prediction.

has the second lowest adjusted R-squared of 19.1%, approximately two-thirds of that of Crash Narratives, indicating that our text-based measure of market crash concerns has stronger predictive power compared to the traditional option-based tail risk measure. The two uncertainty measures also exhibit much lower adjusted R-squared values, particularly the Economic Policy Uncertainty, which is only 5.5%, demonstrating that the policy uncertainty accounts for a very small fraction of future volatility variance. Overall, these findings suggest that Crash Narratives capture the largest portion of variance in future market volatility.

The second and third columns of Table 3 demonstrate that the stronger predictive power of Crash Narratives, compared to alternative predictors, is primarily attributable to their better performance during periods of high volatility. The adjusted R-squared values of Crash Narratives are lower in the subsample that excludes the NBER recession periods (fourth column) compared to the full sample and higher in the NBER recession subsample (fifth column). In contrast, the adjusted R-squared of Left Tail Volatility in the recession are lower compared to the full sample. These results suggest that Crash Narratives capture more information about future market volatility during recession periods, while Left Tail Volatility falls short.

Table 4 shows that the predictability of Crash Narratives in realized volatility varies across different market volatility conditions. For example, the 90th quantile represents the upper tail of the volatility distribution, where the volatility is at an extremely high level relative to the overall sample; while 10th quantile represents the lower tail of the volatility distribution, where the volatility is at a low level relative to the overall sample. Significant coefficients for Crash Narratives are observed across all quantiles, with pseudo-R2 increasing as the quantiles rise. The pseudo-R2 value in the 90th quantile is double that of the 10th quantile. This shows that Crash Narratives are considerably more effective in predicting market volatility during periods of high volatility compared to periods of lower volatility. Crash Narratives have higher pseudo-R2 compared to other predictors in all quantiles. Compared to the Left-Tail Volatility, the traditional option-based measure of market crash concerns, the news-based Crash Narratives can capture more information about future realized volatility according to the higher pseudo-R2 at all quantiles. It is also noteworthy that the pseudo-R2s of the two uncertainty measures are much lower than the others, indicating that political and macroeconomic uncertainties play a less important role than financial market predictors.

[Insert Table 4 here]

These results suggest that market crash concerns transmit significantly from Crash Narratives of media to the stock market, especially during high-volatility states or recession periods. One would expect that during the post-financial crisis period, participants in the stock market would pay considerable attention to sentiment expressed by financial news, the ups and downs of which are immediately reflected in Crash Narratives. Consequently, fluctuations in market crash concerns conveyed by the major financial media can lead to significant changes in the expected volatility of the stock market during periods of turmoil.

Table 5 demonstrates that Crash Narratives remain significant predictors of future realized volatility, regardless of how many principal components are included in the regression [10]. This reinforces the finding that Crash Narratives provide unique information beyond the other four predictors. Notably, while PC1 accounts for the largest share of variance, it is insignificant when included alongside Crash Narratives. Furthermore, the difference in Adjusted R2 between the univariate regression in Table 3 and column (1) of Table 5 is only 0.4%, suggesting limited predictive value in PC1. However, the inclusion of Crash Narratives increases Adjusted R2 by 1.7%, according to the predictive performance in columns (2) and (4).

[Insert Table 5 here]

5.2 Conditional Variance Forecasting Analysis

Table 6 presents the performance statistics for seven models. The first four columns display the in-sample estimation results for the seven estimated models. The results show that including Crash Narratives significantly increases the predictive power of the HAR model, as evidenced by the significant F-test and Wald test statistics for Model 2. By comparing the

[10] PC1 accounts for the largest share of variance of the other four alternative predictors, approximately 64.1%. PC2, PC3 and PC4 explain 16.8%, 12.2% and 6.9% of the variance, respectively.

adjusted R-squared and BIC values from the in-sample estimations, it is clear that adding Crash Narratives to the HAR model is more beneficial than incorporating Left Tail Volatility, news-based Equity Market Volatility and Economic Policy Uncertainty. For example, the adjusted R-squared increases from 42.9% in the HAR model to 43.8% in the model that includes Crash Narratives.

[Insert Table 6 here]

The final three columns in Table 6 presents the out-of-sample results to predict monthly conditional variance. Models incorporating Crash Narratives (Model 2) outperform the other six models, as evidenced by the lowest out-of-sample RMSE and MAE values. The results of the DM-test show that models with Crash Narratives produce significantly lower MAE compared to the non-estimated model (the lagged realized variance). The last column shows that Model 2 has the largest positive and significant R^2_{OOS} according to the CW test. This improvement in forecast performance indicates that integrating Crash Narratives into the model markedly improves the accuracy of conditional variance predictions. It is noteworthy that while the Macroeconomic Uncertainty Index shows good in-sample results, its out-of-sample performance is the worst, indicating that the Macroeconomic Uncertainty Index fails to provide predictive information for the conditional variance beyond the sample. In conclusion, both in-sample and out-of-sample analysis demonstrate that Crash Narratives add value in predicting the US stock market conditional variance at the monthly frequency.

6 Volatility-Managed Portfolio

Understanding how market crash concerns affect market volatility is important for investors when planning their portfolio strategies. The results in Section 5 show that models with Crash Narratives produce a more accurate forecast of the Conditional Variance compared to models without Crash Narratives. Thus, in this section, we examine whether these new measures can improve volatility-managed portfolios' performance. Specifically, we construct portfolios that scale monthly returns by the inverse of their conditional variance. These volatility-managed portfolios decrease risk exposure when the conditional variance is high and vice versa. Crash Narratives-based trading strategies earn the highest alphas among a group of volatility-managed portfolios. Section 6.1 describes the construction of Volatility-Managed Portfolios, and Section 6.2 evaluates their performance.

6.1 Construction of Volatility-Managed Portfolios

To evaluate the economic significance of the forecasting model based on Crash Narratives, we convert our volatility forecast into an out-of-sample volatility timing strategy. This strategy typically involves taking conservative positions on the underlying factors during periods of high volatility and adopting more aggressively leveraged positions when volatility is low.

Research on volatility-managed portfolios has surged in recent years. Several papers document that this simple trading strategy produces large alphas and Sharpe ratios across a wide range of asset pricing factors, and investors can benefit from volatility timing (Barroso and Santa-Clara, 2015; Božović, 2024; Daniel and Moskowitz, 2016; Moreira and Muir, 2017; Wang and Yan, 2021). The high Sharpe ratios observed in volatility timing strategies can be attributed to the fact that changes in volatility are not offset by corresponding changes in expected returns. In other words, given that variance is highly predictable at short horizons and that variance predictions have a weak correlation to future returns over these horizons, a standard mean-variance investor should time volatility. Specifically, they should take more risks when the mean-variance trade-off is attractive (volatility is low), and take less risk when the mean-variance trade-off is unattractive (volatility is high).

Moreira and Muir (2017) show that the performance of volatility-managed portfolios can be further improved through the use of more sophisticated models of variance forecasting. In Section 5, we demonstrate that integrating Crash Narratives into the Heterogeneous Autoregressive model markedly improves the accuracy of Conditional Variance predictions. In this section, we compare our Crash Narratives-based timing strategy with volatility-managed portfolios based on other forecasting models in Section 5.2 and a buy-and-hold strategy.

We follow the method of Moreira and Muir (2017) and construct our volatility-managed

SP500 portfolio by scaling excess returns by the inverse of its conditional variance. On a monthly basis, our strategy increases or decreases risk exposure to the portfolio according to the variation in our measure of conditional variance. The managed portfolio is:

$$SP500_t^{\sigma} = \frac{c}{CV_t} SP500_t, \tag{6}$$

where $SP500_t$ is the buy-and-hold SP500 portfolio's excess return, $CV_t = E_t[RV_{t+1}^2]$ is a SP500's conditional variance, and the constant c controls the average exposure of the strategy. We set c such that the unconditional standard deviation of the managed portfolio is equal to that of the buy-and-hold portfolio. This means that c can be obtained from:

$$\sigma^2\left(\frac{SP500_t}{CV_t} c\right) = \sigma^2(SP500_t) \Leftrightarrow c = \frac{\sigma(SP500_t)}{\sigma(SP500/CV_t)} \tag{7}$$

We use the conditional variance based on the seven different forecasting models in Section 5.2: the Heterogeneous Autoregressive model (HAR, Model 1), the HAR model supplemented with Crash Narratives (Model 2), the HAR model with Left Tail Volatility (Model 3), the HAR model with the news-based Equity Market Volatility (Model 4), the HAR model with Economic Policy Uncertainty (Model 5), the HAR model with Macroeconomic Uncertainty Index (Model 6), and the non-estimated model that uses the realized variance in the preceding month as the conditional variance (Model 7), which is the conditional variance used in (Moreira and Muir, 2017).

We evaluate the performance of volatility-managed portfolios using three approaches. First, we plot the cumulative nominal returns of the volatility-managed SP500 portfolios, and compare them to the buy-and-hold strategy from 2006 to 2024. Second, we run a monthly time-series regression of the volatility-managed SP500 portfolio on the original portfolio:

$$SP500_t^{\sigma} = \alpha + \beta SP500_t + \varepsilon_t. \tag{8}$$

A positive intercept (α) implies that volatility timing increases Sharpe ratios relative to the original portfolio. Third, we calculate the annualized Sharpe ratios.

6.2 Performance of Volatility-Managed Portfolios

The top panel of Figure 3 shows the cumulative nominal returns of the volatility-managed S&P 500 portfolios compared to the buy-and-hold strategy from 2006 to 2023. We begin with an investment of $1 in 2006 and plot the cumulative returns of each strategy on a logarithmic scale. The volatility-managed S&P 500 portfolios exhibit relatively steady gains. By the end of the sample period, the volatility-managed S&P 500 portfolios, on the basis of the conditional variance estimated from Crash Narratives, accumulates to approximately $15, versus nearly $12 for the portfolio based on the conditional variance from Left-Tail Volatility, around $9 for the portfolio based on realized variance from the preceding month, and about $3 for the buy-and-hold strategy. The Crash Narratives–based strategy yields cumulative returns that are at least 25% higher than those of the other strategies.

[**Insert Figure 3 here**]

The lower panel of Figure 3 plots the drawdown of the strategies, defined as the ratio of cumulative returns to their historical maximum. This measure captures the extent of a decline from the peak cumulative return over time, offering a clear perspective on downside risks. Our strategies tend to assume greater risk during periods of low volatility, such as in 2022, leading to occasional losses. In contrast, significant market losses typically occur during high-volatility periods, including the 2008 financial crisis, the euro crisis, the China stock market turmoil, and COVID-19, which our strategies largely avoid. For example, in March 2020, during the COVID-19 market shock, the drawdown of the volatility-managed portfolio based on Crash Narratives was 0.89, higher than the buy-and-hold strategy (0.79) and marginally above the volatility-managed portfolio based on realized variance (0.87). These results highlight that by integrating market crash concerns reflected in financial news, the Crash Narratives-based strategy focuses on loss reduction during market crash periods. As a result, this strategy incurs the least losses during such periods. Our strategy operates by adjusting the timing of market risk exposure and mitigates losses during market crashes, rather than capitalizing on extreme market conditions as profitable option strategies often do.

[Insert Table 7 here]

Table 7 reports results from conducting a regression of the volatility-managed SP500 portfolio on the original SP500 portfolio. We note positive, statistically significant intercepts (α's) for all seven volatility-managed portfolios. The scaled SP500 portfolio based on the conditional variance forecast by the model with Crash Narratives (Model 2), $SP500^{\sigma_{\text{Crash Narratives}}}$, has the highest annualized alpha of 10.70%.

Similarly, the annualized Sharpe ratios of $SP500^{\sigma_{\text{Crash Narratives}}}$ are significantly higher than that of $SP500^{\sigma_{RV}}$ as evidenced by the Jobson and Korkie (1981) test and also higher than the annualized Sharpe ratios of the other volatility-managed portfolios. These results demonstrate that volatility forecasting models with Crash Narratives better predict future stock volatility compared to other predictors from the perspectives of volatility-managed portfolio.

The findings of volatility-managed portfolios allow investors to dynamically adjust their exposure to market risk, potentially enhancing returns during stable periods while mitigating losses during periods of high volatility. More importantly, these results underscore the value of applying crash narratives from financial news to investment strategies, offering a novel method to gain a competitive edge in portfolio management.

7 Variance Risk Premium

In this section, we estimate the variance risk premium, based on the conditional variance forecast by a model with Crash Narratives. We find that our new proxy – which explicitly takes into account market crash concerns – has superior forecasting power for short-term returns. Section 7.1 introduces the variance risk premium as a proxy for risk aversion and its calculation. Section 7.2 provides the predictive results of the variance risk premium on stock market returns.

7.1 Variance Risk Premium and Risk Aversion

There are two main hypotheses regarding the sources of the magnitude and variation of asset prices. The first hypothesis, explored in one body of literature, examines the role of cash flow volatility dynamics as a determinant of equity premiums in both the time series and cross section (Bansal and Yaron, 2004; Wu, 2001). A different body of literature has explored shocks to investors' preferences as drivers of equity prices (Abel, 1990, 1999; Bekaert et al., 2010; Brandt and Wang, 2003; Campbell and Cochrane, 1999; Menzly et al., 2004). In the consumption-based asset pricing model of Campbell and Cochrane (1999), risk aversion drives time-variation in the equity risk premium. Bekaert et al. (2009) develop a theoretical model and an empirical strategy that can accommodate both hypotheses, and use an optimal GMM estimation to assess the relative importance of each. Their findings indicate that both the conditional volatility of cash flow growth and the time-varying risk aversion are significant factors driving the variation in the equity risk premium and the conditional volatility of returns. They conclude that cash flow volatility plays a more important role in volatility, while risk aversion is more crucial for the equity risk premium.

One simple candidate for risk aversion is the Variance Risk Premium, defined as the difference between the squared VIX index and an estimate of the conditional variance of the stock market.

$$VRP_t = VIX_t^2 - CV_t \tag{9}$$

$$CV_t = E_t[RV_{t+1}^2] \tag{10}$$

, where the VIX is the implied option volatility of the S&P500 index for contracts with a maturity of one month, and RV_{t+1}^2 is the S&P500 realized variance measured over the next month.

Economically, the squared VIX is the conditional return variance when using a "risk-neutral" probability measure, while the conditional variance is based on the "physical" probability measure. The risk-adjusted measure moves probability to states with higher marginal utility (bad states), which implies that, in many realistic economic settings, the variance risk premium will be increasing in the economy's risk aversion. Generally, Variance Risk Premium utilizes objective financial market information and naturally "cleanses" option-implied volatility from the effect of physical volatility dynamics and uncertainty, leaving a measure associated with risk aversion. Hence, it always serves as an indicator of risk aversion.

A series of studies have shown that the variance risk premium can predict stock market returns, as it is ascribed to non-Gaussian elements in fundamental factors and (stochastic) risk aversion in financial models (Bekaert and Hoerova, 2014; Bollerslev et al., 2009, 2012, 2014; Carr and Wu, 2009; Goyal and Saretto, 2009). Building on our enhanced conditional variance metrics for stock market returns in Section 5.2, we examine the proposition that Variance Risk Premium contains critical information on equity risk premia in univariate regressions.

We estimate the Variance Risk Premium with various conditional variance forecast models in Section 5.2. We start with univariate regressions using the Variance Risk Premium as a predictor of equity returns, utilizing monthly average observations. The dependent variable is the S&P500 excess stock returns (the S&P500 return in excess of the three-month T-bill rate). We explore one-month forecasting horizons. The use of overlapping monthly data introduces serial correlation in the residuals and is corrected in creating standard errors.

7.2 Stock Market Return Prediction Results

Table 8 presents the regression results. First, the coefficients of the Variance Risk Premium (VRP) estimated from Models 1-5 are significantly positive, indicating that VRP has predictive power for future stock returns over a one-month horizon. Second, VRPs based on Crash Narratives and Left Tail Volatility show higher adjusted R-squared values compared to the VRP based on Economic Policy Uncertainty, indicating that the left-tail measure plays a more important role in capturing risk aversion through the VRP. Third, the adjusted R-squared of the VRP based on the Macroeconomic Uncertainty Index is insignificant, which aligns with its poor out-of-sample predictive performance discussed in Section **??**. These findings suggest that our new VRPs, generated from Crash Narratives conditional variance forecasting models, better predict the aggregated equity return compared to the non-estimated model. This shows that volatility forecasting models with Crash Narratives successfully predict future stock volatility from the perspectives of VRP and return predictability.

[Insert Table 8 here]

Overall, the superior forecasting power of the variance risk premium based on Crash Narratives suggests that incorporating Crash Narratives creates a more accurate estimate of the Variance Risk Premium. The strong predictability of variance risk premium supports the asset pricing model by Campbell and Cochrane (1999) and Bekaert et al. (2009), which attributes variation in equity risk premiums to counter-cyclical changes in risk aversion. These models appear increasingly plausible as the true economic mechanism explaining time-variation in equity risk premiums.

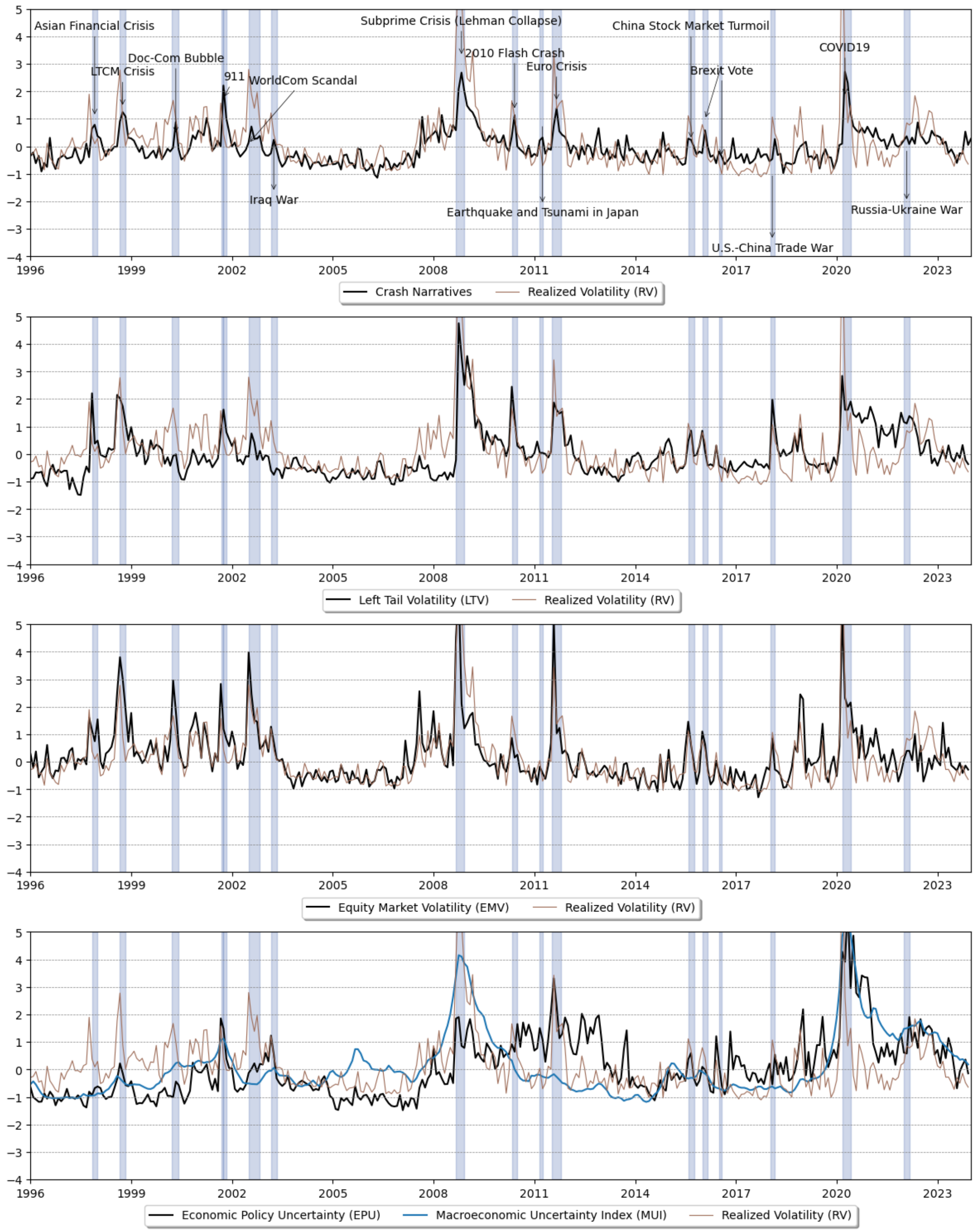
8 Conclusion

We use the text-based method proposed by Goetzmann et al. (2022), to develop a measure of market crash concerns–Crash Narratives–from the Wall Street Journal news articles. Crash Narratives exhibit a substantial increase following stock market crashes, wars, major international political events, and financial crises. This suggests that Crash Narratives are good proxies for market crash concerns.

We find that Crash Narratives predict US realized volatility. Importantly, three key pieces of evidence suggest the strong predictability of Crash Narratives in market volatility. First, the results of predictive regressions and subsample analysis demonstrate that crash narratives outperform alternative text-based volatility measures, traditional tail risk measures, as well as uncertainty measures mainly due to their stronger predictability during periods of high volatility and NBER recessions. Second, integrating Crash Narratives into the Heterogeneous Autoregressive model significantly improves both the in-sample and out-of-sample forecast accuracy of Conditional Variance. Third, Crash Narrative-based trading strategies earn the highest cumulative returns and Sharpe ratios among a group of volatility-managed portfolios.

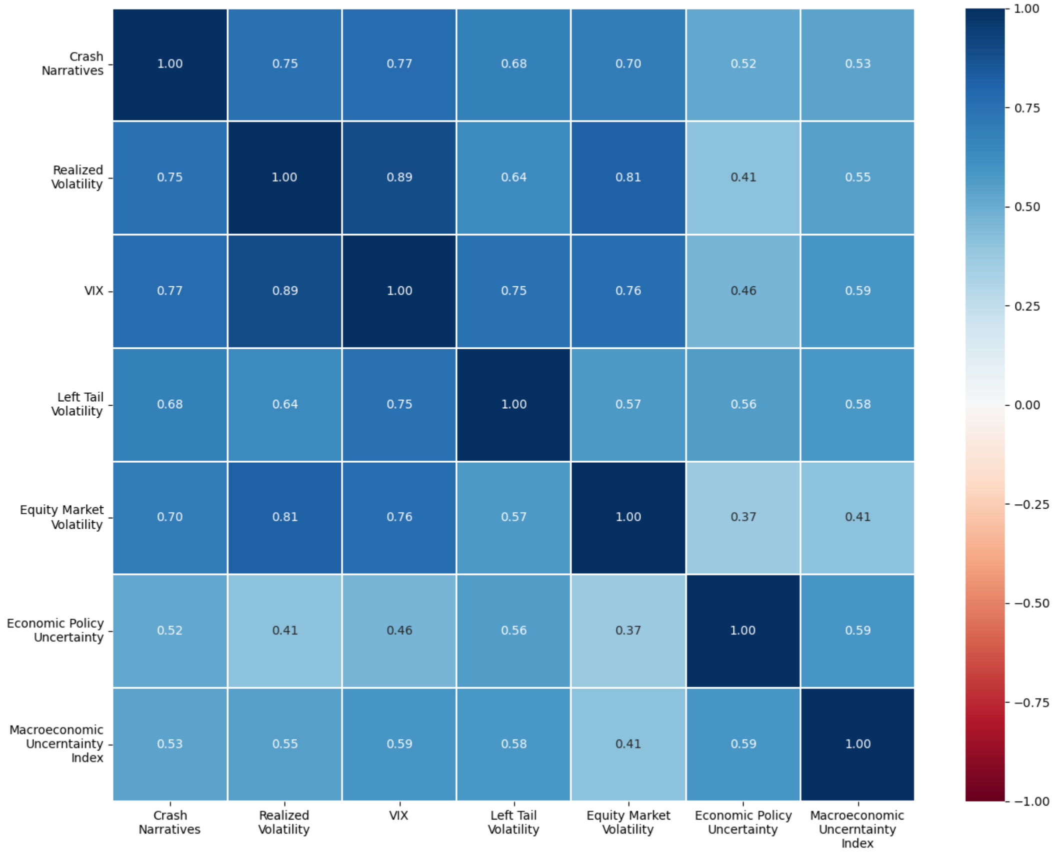
Finally, variance risk premium measures, which are based on Crash Narratives, have higher predictive power for future stock market returns, compared to previous measures. This reinforces our findings that volatility forecasting models with Crash Narratives can successfully predict future stock volatility from the perspectives of variance risk premium and return predictability. The robust predictability of the variance risk premium also corroborates the asset pricing theory proposed by Campbell and Cochrane (1999), which links fluctuations in equity risk premiums to counter-cyclical shifts in risk aversion.

Figure 1: Crash Narratives, Alternative Volatility Predictors and Realized Volatility



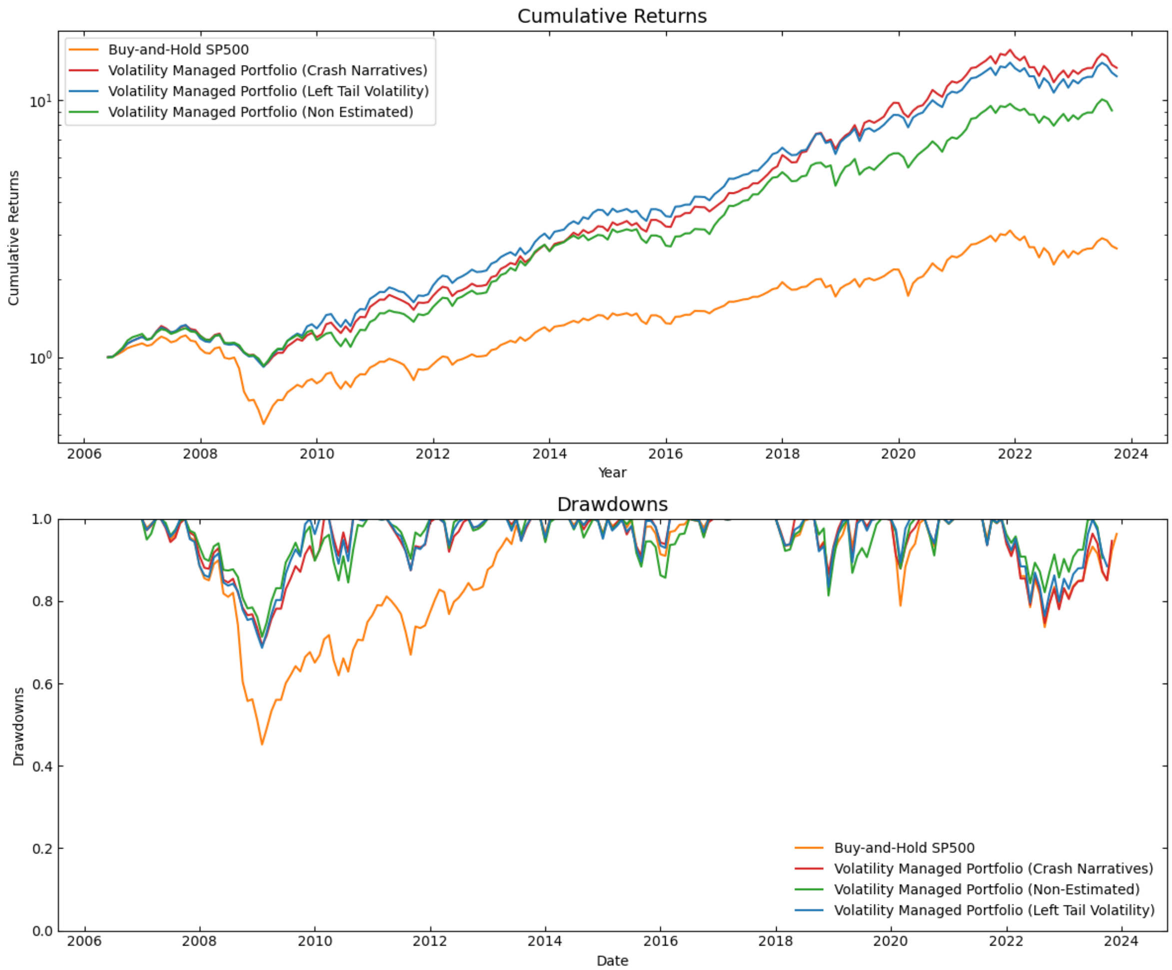
The figure displays monthly averages of Crash Narratives, monthly Equity Market Volatility (EMV), Left Tail Volatility (LTV), Economic Policy Uncertainty (EPU), Macroeconomic Uncertainty Index (MUI), and monthly Realized Volatility from June 1996 through December 2023. Crash Narratives are the cosine similarities of the Wall Street Journal news articles published on day t with those published between October 20–23, 1987 minus the cosine similarity with those published between October 5–9, 1987. All measures are standard normalized. The shaded areas highlight the periods associated with significant events that lead to the spikes of crash narratives.

Figure 2: Correlation among Crash Narratives, Alternative Volatility Predictors and Market Volatility



The heat map displays monthly contemporaneous correlations of Realized Volatility (RV), VIX, Crash Narratives, Equity Market Volatility (EMV) Left Tail Volatility (LTV), Economic Policy Uncertainty (EPU) and Macroeconomic Uncertainty Index (MUI).

Figure 3: Cumulative Returns of the Volatility-Managed SP500 Portfolios



The top panel plots the cumulative returns to a buy-and-hold strategy versus three volatility-managed strategies for the SP500 portfolio from 2006 to 2023. Volatility Managed Portfolio (Crash Narratives) is the volatility-managed S&P 500 portfolio based on the conditional variance estimated from Crash Narratives. Volatility Managed Portfolio (LTV) is the portfolio based on the conditional variance estimated from Left-Tail Volatility. Volatility Managed Portfolio (Non Estimated) is the portfolio based on realized variance from the preceding month, the method used in Moreira and Muir (2017). The y-axis is on a log scale and all strategies have the same unconditional monthly standard deviation. The log scale is used to normalize returns, allowing for proportional comparisons across strategies and making compounded growth appear linear for clearer visualization of performance trends. The lower panel shows the drawdown of each strategy, drawdown $= \frac{\text{cumulative return}}{\text{expanding maximum}}$, where the expanding maximum is the highest cumulative return achieved at any time prior to or at a specific date.

Table 1: Descriptions of Alternative Volatility Predictors

Predictor:	**Description:**	**Source:**	**Forecasting Applications:**
Text-based volatility measures:			
Equity Market Volatility	This index is constructed by using scaled frequency counts of newspaper articles that contain selected terms: "economic", "economy" or "financial"; "Stock Market", "equity", "equities", or "Standard and Poors"; "Volatility" "volatile", "uncertain", "uncertainty", "risk", or "risky".	Baker et al. (2021)	Wen et al. (2024)
Tail-risk measure:			
Left-Tail Volatility	This index measures the expected volatility due to a ten-standard deviation, or larger, down move in the SP500 Index over one week. It provides market participants with access to a robust tail risk measure estimating the expected (risk-neutral) return volatility stemming from large negative price moves over short horizons.	Bollerslev et al. (2015)	Dierkes et al. (2024)
Economic uncertainty measures			
Economic Policy Uncertainty	This index reflects the frequency of articles in 10 leading U.S. newspapers that contain the following terms: "economic" or "economy"; "uncertain" or "uncertainty"; and one or more of "Congress", "deficit", "Federal Reserve", "legislation", "regulation", or "White House".	Baker et al. (2016)	Ma et al. (2022), Asgharian et al. (2023), Li et al. (2023), Wen et al. (2024).
Macroeconomic Uncertainty Index	This index is defined as the common component in the time-varying volatilities of h-step ahead forecast errors obtained from a vector autoregressive model with a large number of macroeconomic series that include variables from three categories: real activity, price, and financial, following the method of Jurado et al. (2015).	Ludvigson et al. (2021)	Ma et al. (2022), Asgharian et al. (2023), Wen et al. (2024).

This table displays detailed descriptions and sources of alternative volatility predictors.

Table 2: Summary Statistics of Crash Narratives and Alternative Volatility Predictors

	mean	std	min	25%	50%	75%	max
Crash Narratives	-0.483	0.578	-3.252	-0.854	-0.542	0.195	3.244
Left Tail Volatility	8.60	3.67	0.385	6.39	7.69	9.96	41.6
Equity Market Volatility	19.9	7.97	8.03	14.9	17.9	22.3	69.8
Economic Policy Uncertainty	115	39.2	57.2	88.1	108	133	350
Macroeconomic Uncertainty Index	65.3	10.5	52.6	58.4	62.5	68.1	121.8

This table displays summary statistics of Crash Narratives, Equity Market Volatility (EMV) Left Tail Volatility (LTV), Economic Policy Uncertainty (EPU) and Macroeconomic Uncertainty Index (MUI). Crash Narratives (CN) are the cosine similarities of the Wall Street Journal news articles published on day t with those published between October 20–23, 1987 minus the cosine similarity with those published between October 5–9, 1987. The alternative volatility predictors are all expressed as index levels in percentages.

Table 3: Crash Narratives Predict Realized Volatility

Dependent Variable:		RV_{t+1}				
Sample:		Full	High RV	Low RV	Excl. Recession	Recession
CN	β_1	0.624***	0.545***	0.032	0.517***	0.728**
	t_{nw}	6.462	4.081	1.250	8.199	2.577
	Adj. $R^2(\%)$	33.8	26.2	0.3	20.0	31.6
LTV	β_1	0.544***	0.435***	0.026	0.426***	0.486***
	t_{nw}	5.596	4.112	0.915	5.604	3.925
	Adj. $R^2(\%)$	19.1	13.7	0.1	11.6	17.3
EMV	β_1	0.587***	0.462***	0.055	0.452***	0.684***
	t_{nw}	5.843	3.396	1.511	9.046	3.893
	Adj. $R^2(\%)$	31.7	22.3	1.9	20.0	38.0
EPU	β_1	0.227***	0.297***	0.011	0.130**	0.510*
	t_{nw}	2.700	2.607	0.703	2.106	1.819
	Adj. $R^2(\%)$	5.5	8.7	0.1	2.3	10.3
MUI	β_1	0.421***	0.418***	0.026	0.225**	0.776***
	t_{nw}	3.317	3.066	1.222	2.418	2.896
	Adj. $R^2(\%)$	21.0	22.0	0.3	5.6	40.9
No. Obs.		330	165	165	302	28

Reported are monthly Realized Volatility (RV) predictive regressions in the full sample and four subsamples: $RV_{t+1} = \beta_0 + \beta_1 X_t + \epsilon_{t+1}$, where $X \in$ [CN, LTV, EMV, EPU, MUI]. CN represents Crash Narratives from the Wall Street Journal. LTV is the Left-Tail Volatility. Other text-based volatility measures include Equity Market Volatility (EMV), and Economic Policy Uncertainty (EPU). MUI is the macroeconomic uncertainty index. The full sample period is June 1996-December 2023. High RV periods are months when RV is higher than the median. Low RV periods are months when RV is lower than the median. Recessions are defined by NBER business cycles. Each row and column represent a different univariate regression based on various predictors. t_{nw} is Newey and West corrected t-statistics with number of lags equal to the size of the volatility forecasting window. $_{*}$, $_{**}$, and $_{***}$indicate 10%, 5%, and 1% significance levels.

Table 4: Crash Narratives predict Realized Volatility (Quantile Regression)

Dependent Variable:		RV_{t+1}				
Quantile:		Q(0.1)	Q(0.25)	Q(0.5)	Q(0.75)	Q(0.9)
CN	λ	0.292***	0.338***	0.549***	0.669***	0.916***
	t	8.195	9.381	13.79	12.92	6.306
	Pseudo $R^2(\%)$	6.9	12.9	19.6	24.4	27.6
LTV	λ	0.255***	0.307***	0.497***	0.624***	0.694***
	t	8.110	7.785	9.614	6.675	3.786
	Pseudo $R^2(\%)$	5.4	7.7	11.7	13.2	16.3
EMV	λ	0.264***	0.375***	0.507***	0.809***	1.03***
	t	8.466	10.21	13.84	14.23	11.35
	Pseudo $R^2(\%)$	10.5	15.2	19.5	21.2	25.7
EPU	λ	0.033	0.054	0.141***	0.252***	0.322**
	t	0.971	1.394	3.040	4.292	2.452
	Pseudo $R^2(\%)$	0.5	0.5	1.6	4.3	5.2
MUI	λ	0.112***	0.235***	0.299***	0.448***	0.741***
	t	3.159	7.385	7.315	6.635	10.389
	Pseudo $R^2(\%)$	2.2	5.3	8.0	10.8	17.4
No. Obs.		330	330	330	330	330

Reported are monthly Realized Volatility (RV) quantile regressions: $Q_{RV_{t+1}}(\tau|X_t) = \phi(\tau) + \lambda(\tau)X_t$, where $X \in$ [CN, LTV, EMV, EPU, MUI]. CN represents Crash Narratives from the Wall Street Journal, and $\tau \in [0.1, 0.25, 0.50.75, 0.9]$. LTV is the Left-Tail Volatility. Other text-based volatility measures include Equity Market Volatility (EMV), and Economic Policy Uncertainty (EPU). MUI is the macroeconomic uncertainty index. Each row and column represents a different regression. The sample period is June 1996-December 2023. The quantile used are the 10th, 25th, 50th, 75th, and 90th percentiles, respectively. Pseudo R^2 is a measure of goodness-of-fit that compares the sum of absolute deviations of the fitted model to a null model, indicating how well the quantile regression explains the variation in the dependent variable. t is t-statistics with number of lags equal to the size of the volatility forecasting window. $_*$, $_{**}$, and $_{***}$indicate 10%, 5%, and 1% significance levels.

Table 5: Crash Narratives predict Realized Volatility (Principal Component Analysis)

Dependent Variable:	RV_{t+1}				
	(1)	(2)	(3)	(4)	(5)
CN	0.519***		0.327**	0.296**	0.290**
	[3.389]		[2.213]	[2.197]	[2.181]
$PC1$	0.081	0.325***	0.171**	0.186**	0.189**
	[0.959]	[5.828]	[2.220]	[2.484]	[2.452]
$PC2$		0.369***	0.255***	0.265***	0.268***
		[4.957]	[3.464]	[4.169]	[4.258]
$PC3$		-0.304***		-0.290***	-0.290***
		[-2.825]		[-2.833]	[-2.834]
$PC4$		0.105			0.089
		[0.717]			[0.649]
Adj. $R^2(\%)$	34.2	39.3	37.1	40.9	41.0
No. Obs.	330	330	330	330	330

Reported are monthly Realized Volatility (RV) predictive regressions in the full sample: $RV_{t+1} = b_0 + \sum_{i=1}^{4} b_i PCi_t + b_5 CN_t + \epsilon_{t+1}$. CN represents Crash Narratives from the Wall Street Journal. PCs are the four Principal Components of the Left-Tail Volatility, Equity Market Volatility, Economic Policy Uncertainty and Macroeconomic Uncertainty Index. The full sample period is June 1996-December 2023. Each column represent a regression. Newey and West corrected t-statistics are in brackets. The number of lags equal to the size of the volatility forecasting window. $_*$, $_{**}$, and $_{***}$indicate 10%, 5%, and 1% significance levels.

Table 6: Conditional Variance Forecasting Model Statistics

		In-Sample				Out-of-Sample		
Forecasting Models		$Adj.R^2(\%)$	BIC	F-test	Wald-test	RMSE	MAE	R^2_{OOS}
Estimated Models								
Model 1:	HAR	42.9	1501.0			2.64	1.56**	0.418***
						(0.95)	(2.19)	(3.38)
Model 2:	HAR + CN	43.8	1501.0	5.75**	4.66**	2.63	1.54**	0.423***
						(0.98)	(2.30)	(3.41)
Model 3:	HAR + LTV	42.8	1506.4	0.43	0.19	2.64	1.55**	0.416***
						(0.84)	(2.00)	(3.33)
Model 4:	HAR + EMV	43.0	1505.2	1.57	1.94	2.65	1.56**	0.415***
						(0.90)	(2.07)	(3.36)
Model 5:	HAR + EPU	42.8	1506.8	0.05	0.08	2.66	1.59*	0.411***
						(0.87)	(1.83)	(3.30)
Model 6:	HAR + MUI	44.8	1495.0	12.0***	3.96**	2.88	1.67	0.306**
						(-0.52)	(0.69)	(1.66)
Non-estimated Model								
Model 7:	RV^2_{t-1}					2.83	1.71	0.331**
								(1.93)

This table shows the in-sample and out-of-sample statistics to evaluate performance of conditional variance forecasting models. Model 1 is the monthly Heterogeneous Autoregressive model (HAR). Model 2 adds Crash Narratives (CN) in Model 1. Model 3 adds LTV in Model 1. Left Tail Volatility (LTV) is an option-based tail risk measure. Model 4 adds EMV (Equity Market Volatility) in Model 1. Model 5 adds EPU (Economic Policy Uncertainty) in Model 1. Model 6 adds MUI (Macroeconomic Uncertainty Index) in Model 1. Model 7 is the non-estimated model. The in-sample estimations are based on 331 monthly observations. The sample period is June 1996 to December 2023. The in-sample statistics are adjusted R-squared, the Bayesian information criterion (BIC), the F-test and Wald-test statistics compared to Model 1. The out-of-sample estimations are based on a 120-month rolling window with 211 out-of-sample observations. The overall sample period is June 1996 to December 2023. The out-of-sample statistics are RMSE, MAE and R^2_{OOS}. We apply Diebold and Mariano test to examine whether the RMSE and MAE of the estimated models are significantly different from those of the non-estimated model (Model 7). Clark and West (2007) MSPE statistic that tests the null of $R^2_{OOS} \leq 0$ against the alternative $R^2_{OOS} > 0$ is applied to test for significance. The Diebold and Mariano test statistics, and Clark and West (2007) MSPE statistics are in parentheses. $_*$, $_{**}$, and $_{***}$indicate 10%, 5%, and 1% significance levels.

Table 7: Performance of Volatility-Managed SP500 Portfolios

Dependent Variable:	$SP500^{\sigma}$						
Forecasting Models:	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Conditional Variance:	CV_{HAR}	CV_{CN}	CV_{LTV}	CV_{EMV}	CV_{EPU}	CV_{MUI}	CV
β	0.93***	0.90***	0.92***	0.93***	0.92***	0.90***	0.88***
	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]
Alpha (α)	9.81***	10.09***	9.39***	9.85***	9.72***	9.54***	7.79***
	[1.47]	[1.64]	[1.49]	[1.44]	[1.48]	[1.43]	[1.82]
R^2	0.85	0.81	0.85	0.86	0.85	0.80	0.77
Annualized Sharpe Ratio	0.95	0.98	0.92	0.95	0.94	0.92	0.81
z_{JK}	1.53*	1.65**	1.24	1.60*	1.49*	1.25	

In this table, we run time-series regressions of volatility-managed SP500 portfolio on the non-managed SP500 portfolio $SP500^{\sigma}_t = \alpha + \beta SP500_t + \varepsilon_t$. The managed SP500 portfolio, $SP500^{\sigma}$, scales by the portfolio's inverse conditional variance estimated in Model 1-6: $SP500^{\sigma}_t = \frac{c}{CV_{t-1}} SP500_t$. CV_{HAR} is the Conditional Variance estimated in Model 1 in Table 6. Similarly, CV_{CN} is the Conditional Variance estimated in Model 2, where CN is Crash Narratives. CV_{LTV} is the Conditional Variance estimated in Model 3, LTV is Left Tail Volatility. CV_{EMV} is the Conditional Variance estimated in Model 4, where EMV is Equity Market Volatility. CV_{EPU} is the Conditional Variance estimated in Model 5, where EPU is Economic Policy Uncertainty. CV_{MUI} is the Conditional Variance estimated in Model 6, where MUI is Macroeconomic Uncertainty Index. CV is the Conditional Variance from the non-estimated Model 7. The data are monthly and the sample period is July 2006-December 2023. Standard errors are in parentheses. SP500 portfolio's excess returns are annualized in percent per year by multiplying monthly excess returns by 12. z_{JK} is the test stat of Jobson and Korkie (1981), to test for the difference of the Sharpe ratios of two investment strategies. z_{JK} tests whether the Sharpe ratios of Volatility Managed Portfolios based on estimated conditional variances are significantly different from that of Volatility Managed Portfolio based on the realized variance in the preceding month. $_*$, $_{**}$, and $_{***}$indicate 10%, 5%, and 1% significance levels.

Table 8: US Stock Return Univariate Predictions with Variance Risk Premium

Forecasting Models	Dependent Variable:	r^e_{t+1}						
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Model 1:	VRP_{HAR}	0.277^{**}						
		[2.388]						
Model 2:	VRP_{CN}		0.271^{**}					
			[2.101]					
Model 3:	VRP_{LTV}			0.327^{**}				
				[2.569]				
Model 4:	VRP_{EMV}				0.283^{**}			
					[2.512]			
Model 5:	VRP_{EPU}					0.269^{**}		
						[2.271]		
Model 6:	VRP_{MUI}						0.079	
							[0.913]	
Model 7:	VRP							0.127
								[1.272]
	Adj. R^2(%)	3.6	3.6	4.1	3.8	3.2	1.0	1.6
	No. Obs.	209	209	209	209	209	209	209

Reported are monthly stock return regressions based on the Variance Risk Premium (VRP): $r^e_{t+1} = a + bVRP_t + u_{t,t+1}$,. $VRP_t = VIX^2_t - CV_t$. VRP_{HAR} is based on the Conditional Variance estimated in Model 1 in Table 6. Similarly, VRP_{CN} is based on the Conditional Variance estimated in Model 2, where CN is Crash Narratives. VRP_{LTV} is based on the Conditional Variance estimated in Model 3, LTV is Left Tail Volatility. VRP_{EMV} is based on the Conditional Variance estimated in Model 4, where EMV is Equity Market Volatility. VRP_{EPU} is based on the Conditional Variance estimated in Model 5, where EPU is Economic Policy Uncertainty. VRP_{MUI} is based on the Conditional Variance estimated in Model 6, where MUI is Macroeconomic Uncertainty Index. VRP is based on the Conditional Variance from the non-estimated Model 7. Each column represents a different regression. The sample period is July 2006-December 2023. The standard errors reported in brackets are computed using Newey-West lags. $_*$, $_{**}$, and $_{***}$indicate 10%, 5%, and 1% significance levels.

A Appendix: Construction of Crash Narratives and Doc2Vec Model

We follow Goetzmann et al. (2022), who assume that both the narrative style and the choice of words adopted by business press accurately and consistently mirror the concerns of the average investor. This assumption is intuitive and aligns with a conceptual framework in which a news firm observes real-world events before deciding on the aspects to highlight in its articles, aiming to enhance its credibility and reputation. Gentzkow and Shapiro (2006) develops a theoretical model embodying this framework and provides various empirical validations supporting the model's predictions. The idea that news media content reflects reader interests is proposed by Tetlock (2007), with empirical corroboration found in Manela (2014) and Manela and Moreira (2017), underscoring the role of media in capturing and conveying the sentiment and concerns prevalent among investors.

When humans read text, their brains interpret words and extract meaning from the whole context. According to Gentzkow et al. (2019), attempts to extract meaningful data from text should likewise accommodate complex grammatical frameworks and the intricate interplay among words. As high-dimensional semantic space models have the ability to uncover information from highly unstructured data, we follow the Goetzmann et al. (2022)'s method and detect the narrativity by estimating the semantic similarities within news articles using the Doc2Vec model, introduced by Le and Mikolov (2014).

Doc2Vec contextualizes words, paragraphs, and documents through a neural network algorithm: stochastic gradient descent where the gradient is obtained via backpropagation (Rumelhart et al., 1986). It optimizes an objective function that aims to predict words based on their surrounding context and the document they belong to. This process generates vectors that capture semantic meanings of documents, allowing them to be quantitatively analyzed and compared. Essentially, the model learns to represent documents in a way that similar documents are closer in the vector space, facilitating various tasks such as predicting missing words, assessing document similarity, and classifying documents according to high-order semantics. The key advantage of this method over Bag-of-Words is its ability to use sequencing and location information within a document.

While a comprehensive exploration of Doc2Vec exceeds the limitations of this paper, we provide an intuitive overview of this method and its inherent structuring of data. Within the Doc2Vec framework, every document is mapped to a unique vector, represented by a column in matrix D; every word is mapped to a unique vector, represented by a column in a matrix W. The column is indexed by position of the word in the vocabulary. Given a sequence of training words w_1, w_2, w_3, ..., w_T, Doc2Vec maximizes the following objective:

$$\frac{1}{T}\sum_{t=k}^{T-k}\log p(w_t|w_{t-k},\ldots,w_{t+k},d) \tag{11}$$

$$\text{where } p(w_t|w_{t-k},\ldots,w_{t+k}) = \frac{e^{y_{wt}}}{\sum_i e^{y_i}}, \text{ and } y = c + U\cdot h(w_{t-k},\ldots,w_{t+k};W,D).$$

U and c are the parameters of the softmax function used to convert the raw scores into probabilities. The function h represents the concatenation or averaging of word vectors extracted from W, and it captures the context provided by the words surrounding w_t within a window of size $2k$. The vector d corresponds to the document vector from matrix D, which helps to capture the overall context and thematic structure of the document in which the word w_t appears.

By jointly optimizing the word vectors in W and the document vectors in D, Doc2Vec is able to learn distributed representations that capture both the syntactic and semantic relationships between words and documents, thereby providing a robust framework for various downstream tasks such as similarity measurement.

Narratives that appear in the news in the days following the 1987 crash can be extracted through the Doc2Vec model. This measure captures “Crash Narratives” in the media and investors’ concerns about crashes in the stock market. Using the methodology described in Goetzmann et al. (2022), we train the Doc2Vec model for the WSJ; and then generate a daily series of average similarity scores derived from the model.

The construction of Crash Narratives involves four steps. First, we trained the Doc2Vec model using all WSJ articles from 1987 to 2023[11]. In the second step, we compute the average cosine similarity of each article to those published in the WSJ from October 20 to October 23, 1987. Higher average cosine similarities indicate greater textual similarity. However, due to language evolution over time, articles tend to exhibit higher similarities to recent periods regardless of narrative content. To address this, we account for the effects of language evolution by recalculating similarities in step three. This adjustment is based on the average cosine similarity of each article to articles published from October 5 through October 9, 1987, capturing a pre-crash baseline. The adjusted similarity is calculated as the difference between the natural logarithm of one plus the average cosine similarities. Finally, we average these adjusted similarity scores for articles from the same day. This is called *Crash Narratives*.

[11]The training process encompasses 100 iterations, based on a feature vector size of 250, a maximum distance of 5, a minimum word frequency of 10 across all articles, a sub-sampling threshold of 10^{-5}, and a negative sample of 5.

References

Abel, A. B. (1990). Asset Prices under Habit Formation and Catching up with the Joneses. *The American Economic Review*, *80*(2), 38–42.

Abel, A. B. (1999). Risk premia and term premia in general equilibrium. *Journal of Monetary Economics*.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, *71*(2), 579–625.

Andersen, T. G., Fusari, N., & Todorov, V. (2015). The risk premia embedded in index options. *Journal of Financial Economics*, *117*(3), 558–584.

Andersen, T. G., Todorov, V., & Ubukata, M. (2021). Tail risk and return predictability for the Japanese equity market. *Journal of Econometrics*, *222*(1, Part B), 344–363.

Asgharian, H., Christiansen, C., & Hou, A. J. (2023). The effect of uncertainty on stock market volatility and correlation. *Journal of Banking & Finance*, *154*, 106929.

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

Baker, S. R., Bloom, N., Davis, S. J., & Kost, K. (2021). Policy News and Stock Market Volatility.

Bansal, R., & Yaron, A. (2004). Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *The Journal of Finance*, *59*(4), 1481–1509.

Barro, R. J. (2006). Rare Disasters and Asset Markets in the Twentieth Century. *Quarterly Journal of Economics*, *121*(3), 823–866.

Barroso, P., & Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics*, *116*(1), 111–120.

Behrendt, S., & Schmidt, A. (2018). The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility. *Journal of Banking & Finance*, *96*, 355–367.

Bekaert, G., Engstrom, E., & Grenadier, S. R. (2010). Stock and bond returns with Moody Investors. *Journal of Empirical Finance*, *17*(5), 867–894.

Bekaert, G., Engstrom, E., & Xing, Y. (2009). Risk, uncertainty, and asset prices. *Journal of Financial Economics*, *91*(1), 59–82.

Bekaert, G., & Hoerova, M. (2014). The VIX, the variance premium and stock market volatility. *Journal of Econometrics*, *183*(2), 181–192.

Bernanke, B. S. (1983). Irreversibility, Uncertainty, and Cyclical Investment*. *The Quarterly Journal of Economics*, *98*(1), 85–106.

Bollen, N. P. B., & Whaley, R. E. (2004). Does Net Buying Pressure Affect the Shape of Implied Volatility Functions? *The Journal of Finance*, *59*(2), 711–753.

Bollerslev, T., Gibson, M., & Zhou, H. (2011). Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *Journal of Econometrics*, *160*(1), 235–245.

Bollerslev, T., Marrone, J., Xu, L., & Zhou, H. (2014). Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence. *Journal of Financial and Quantitative Analysis*, *49*(3), 633–661.

Bollerslev, T., Sizova, N., & Tauchen, G. (2012). Volatility in Equilibrium: Asymmetries and Dynamic Dependencies*. *Review of Finance*, *16*(1), 31–80.

Bollerslev, T., Tauchen, G., & Zhou, H. (2009). Expected Stock Returns and Variance Risk Premia. *Review of Financial Studies*, *22*(11), 4463–4492.

Bollerslev, T., & Todorov, V. (2011). Tails, Fears, and Risk Premia. *The Journal of Finance*, *66*(6), 2165–2211.

Bollerslev, T., Todorov, V., & Xu, L. (2015). Tail risk premia and return predictability. *Journal of Financial Economics*, *118*(1), 113–134.

Božović, M. (2024). VIX-managed portfolios. *International Review of Financial Analysis*, *95*, 103353.

Brandt, M. W., & Wang, K. Q. (2003). Time-varying risk aversion and unexpected inflation. *Journal of Monetary Economics*, *50*(7), 1457–1498.

Bybee, L., Kelly, B., Manela, A., & Xiu, D. (2024). Business News and Business Cycles. *The Journal of Finance*, *79*(5), 3105–3147.

Bybee, L., Kelly, B., & Su, Y. (2023). Narrative Asset Pricing: Interpretable Systematic Risk Factors from News Text. *The Review of Financial Studies*, *36*(12), 4759–4787.

Campbell, J. Y., & Cochrane, J. H. (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, *107*(2), 205–251.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies*, *21*(4), 1509–1531.

Carr, P., & Wu, L. (2009). Variance Risk Premiums. *Review of Financial Studies*, *22*(3), 1311–1341.

Chen, H., Joslin, S., & Tran, N.-K. (2012). Rare Disasters and Risk Sharing with Heterogeneous Beliefs. *The Review of Financial Studies*, *25*(7), 2189–2224.

Chiu, C.-w., Harris, R. D., Stoja, E., & Chin, M. (2018). Financial market Volatility, macroeconomic fundamentals and investor Sentiment. *Journal of Banking & Finance*, *92*, 130–145.

Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility1. *Journal of Financial Economics*, *50*(2), 125–150.

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

Conrad, C., & Loch, K. (2015a). Anticipating Long-Term Stock Market Volatility. *Journal of Applied Econometrics*, *30*(7), 1090–1114.

Conrad, C., & Loch, K. (2015b). The variance risk premium and fundamental uncertainty. *Economics Letters*, *132*, 56–60.

Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, *7*(2), 174–196.

Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, *122*(2), 221–247.

Dierckx, T., Davis, J., & Schoutens, W. (2021). Quantifying News Narratives to Predict Movements in Market Risk. *Data Science for Economics and Finance: Methodologies and Applications*, 265–285.

Dierkes, M., Hollstein, F., Prokopczuk, M., & Würsig, C. M. (2024). Measuring tail risk. *Journal of Econometrics*, *241*(2), 105769.

Dim, C., Sangiorgi, F., & Vilkov, G. (2023). Media Narratives and Price Informativeness. *SSRN Electronic Journal*.

Dixit, A. K., & Pindyck, R. S. (1994). *Investment Under Uncertainty*. Princeton University Press.

Drechsler, I. (2013). Uncertainty, Time-Varying Fear, and Asset Prices. *The Journal of Finance*, *68*(5), 1843–1889.

Drechsler, I., & Yaron, A. (2011). What's Vol Got to Do with It. *The Review of Financial Studies*, *24*(1), 1–45.

Driessen, J., & Maenhout, P. (2007). An Empirical Portfolio Perspective on Option Pricing Anomalies*. *Review of Finance*, *11*(4), 561–603.

Engle, R. F., Ghysels, E., & Sohn, B. (2013). Stock Market Volatility and Macroeconomic Fundamentals. *The Review of Economics and Statistics*, *95*(3), 776–797.

Fang, T., Lee, T.-H., & Su, Z. (2020). Predicting the long-term stock market volatility: A GARCH-MIDAS model with variable selection. *Journal of Empirical Finance*, *58*, 36–49.

Fernandes, M., Medeiros, M. C., & Scharth, M. (2014). Modeling and predicting the CBOE market volatility index. *Journal of Banking & Finance*, *40*, 1–10.

Fisher, A., Martineau, C., & Sheng, J. (2022). Macroeconomic Attention and Announcement Risk Premia. *The Review of Financial Studies*, *35*(11), 5057–5093.

Gabaix, X. (2012). Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance. *The Quarterly Journal of Economics*, *127*(2), 645–700.

Gao, G. P., Lu, X., & Song, Z. (2019). Tail Risk Concerns Everywhere. *Management Science*, *65*(7), 3111–3130.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as Data. *Journal of Economic Literature*, *57*(3), 535–574.

Gentzkow, M., & Shapiro, J. M. (2006). Media Bias and Reputation. *Journal of Political Economy*, *114*(2), 280–316.

Goetzmann, W., Kim, D., & Shiller, R. (2022). *Crash Narratives* (tech. rep.). National Bureau of Economic Research. Cambridge, MA.

Goodell, J. W., McGee, R. J., & McGroarty, F. (2020). Election uncertainty, economic policy uncertainty and financial market uncertainty: A prediction market analysis. *Journal of Banking & Finance*, *110*, 105684.

Gourio, F. (2012). Disaster Risk and Business Cycles. *American Economic Review*, *102*(6), 2734–2766.

Goyal, A., & Saretto, A. (2009). Cross-section of option returns and volatility. *Journal of Financial Economics*, *94*(2), 310–326.

Han, B. (2008). Investor Sentiment and Option Prices. *The Review of Financial Studies*, *21*(1), 387–414.

Hirshleifer, D., Mai, D., & Pukthuanthong, K. (2023, June). War Discourse and the Cross Section of Expected Stock Returns.

Hirshleifer, D., Mai, D., & Pukthuanthong, K. (2025). War Discourse and Disaster Premium: 160 Years of Evidence from the Stock Market (J. Linnainma, Ed.). *The Review of Financial Studies*, *38*(2), 457–506.

Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring Uncertainty. *The American Economic Review*, *105*(3), 1177–1216.

Kelly, B., & Jiang, H. (2014). Tail Risk and Asset Prices. *The Review of Financial Studies*, *27*(10), 2841–2871.

Koenker, R., & Bassett, G. (1978). Regression Quantiles. *Econometrica*, *46*(1), 33–50.

Le, Q., & Mikolov, T. (2014). Distributed Representations of Sentences and Documents. *Proceedings of the 31st International Conference on Machine Learning*, 1188–1196.

Lee, W. Y., Jiang, C. X., & Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment.

Li, D., Zhang, L., & Li, L. (2023). Forecasting stock volatility with economic policy uncertainty: A smooth transition GARCH-MIDAS model. *International Review of Financial Analysis*, *88*, 102708.

Ludvigson, S. C., Ma, S., & Ng, S. (2021). Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response? *American Economic Journal: Macroeconomics*, *13*(4), 369–410.

Ma, F., Guo, Y., Chevallier, J., & Huang, D. (2022). Macroeconomic attention, economic policy uncertainty, and stock volatility predictability. *International Review of Financial Analysis*, *84*, 102339.

Ma, F., Wang, J., Wahab, M. I. M., & Ma, Y. (2023). Stock market volatility predictability in a data-rich world: A new insight. *International Journal of Forecasting*, *39*(4), 1804–1819.

Manela, A. (2014). The value of diffusing information. *Journal of Financial Economics*, *111*(1), 181–199.

Manela, A., & Moreira, A. (2017). News implied volatility and disaster concerns. *Journal of Financial Economics*, *123*(1), 137–162.

Menzly, L., Santos, T., & Veronesi, P. (2004). Understanding Predictability. *Journal of Political Economy*, *112*(1), 1–47.

Moreira, A., & Muir, T. (2017). Volatility-Managed Portfolios. *The Journal of Finance*, *72*(4), 1611–1644.

Müller, U. A., Dacorogna, M. M., Dave, R., Pictet, O. V., Olsen, R. B., & Ward, J. R. (1999, September). Fractals and Intrinsic Time - a Challenge to Econometricians.

Pástor, Ľ., & Veronesi, P. (2013). Political uncertainty and risk premia. *Journal of Financial Economics*, *110*(3), 520–545.

Paye, B. S. (2012). 'Déjà vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables. *Journal of Financial Economics*, *106*(3), 527–546.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of Monetary Economics*, *22*(1), 117–131.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536.

Seo, S. W., & Kim, J. S. (2015). The information content of option-implied information for volatility forecasting with investor sentiment. *Journal of Banking & Finance*, *50*, 106–120.

Shiller, R. J. (2014). Speculative Asset Prices. *The American Economic Review*, *104*(6), 1486–1517.

Shiller, R. J. (2017). Narrative Economics. *The American Economic Review*, *107*(4), 967–1004.

Su, Z., Fang, T., & Yin, L. (2017). The role of news-based implied volatility among US financial markets. *Economics Letters*, *157*, 24–27.

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, *62*(3), 1139–1168.

Wachter, J. A. (2013). Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility? *The Journal of Finance*, *68*(3), 987–1035.

Wang, F., & Yan, X. S. (2021). Downside risk and the performance of volatility-managed portfolios. *Journal of Banking & Finance*, *131*, 106198.

Welch, I. (2016). The (Time-Varying) Importance of Disaster Risk. *Financial Analysts Journal*, *72*(5), 14–30.

Welch, I., & Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *Review of Financial Studies*, *21*(4), 1455–1508.

Wen, D., He, M., Wang, Y., & Zhang, Y. (2024). Forecasting crude oil market volatility: A comprehensive look at uncertainty variables. *International Journal of Forecasting*, *40*(3), 1022–1041.

Wu, G. (2001). The Determinants of Asymmetric Volatility. *The Review of Financial Studies*, *14*(3), 837–859.