News Implied Volatility and Disaster Concerns

Asaf Manela  Alan Moreira*

First draft: September 2012
This draft: October 2015

Abstract
We construct a text-based measure of uncertainty starting in 1890 using front-page articles of the Wall Street Journal. News implied volatility (NVIX) peaks during stock market crashes, times of policy-related uncertainty, world wars and financial crises. In US post-war data, periods when NVIX is high are followed by periods of above average stock returns, even after controlling for contemporaneous and forward-looking measures of stock market volatility. News coverage related to wars and government policy explains most of the time variation in risk premia our measure identifies. Over the longer 1890–2009 sample that includes the Great Depression and two World Wars, high NVIX predicts high future returns in normal times, and rises just before transitions into economic disasters. The evidence is consistent with recent theories emphasizing time variation in rare disaster risk as a source of aggregate asset prices fluctuations.

JEL Classification: G12, C82, E44

Keywords: Text-based analysis, implied volatility, rare disasters, equity premium, return predictability, machine learning

*Washington University in St. Louis, amanela@wustl.edu; and Yale University, alan.moreira@yale.edu. We thank Fernando Alvarez, Jacob Boudoukh, Diego García (discussant), Armando Gomes, Gerard Hoberg, Bryan Kelly (discussant), Ralitsa Petkova (discussant), Jacob Sagi (discussant), Jesse Shapiro, Chester Spatt (discussant), Paul Tetlock, seminar participants at Ohio State and Wash U, and conference participants at the AFA meetings, NBER BE meetings, Rothschild Caesarea, SFS Cavalcade, and Texas Finance Festival for helpful comments.
1 Introduction

Looking back, people’s concerns about the future more often than not seem misguided and overly pessimistic. Only when these concerns are borne out in some tangible data, do economists tip their hat to the wisdom of the crowds. This gap between measurement and the concerns of the average investor is particularly severe when rare events are concerned. In this case, concerns might change frequently, but real economic data often makes these concerns puzzling and unwarranted. This paper aims to quantify this “spirit of the times”, which after the dust settles is forgotten and only hard data remains to describe the period. Specifically, our goal is to measure people’s perception of uncertainty about the future, and to use this measurement to investigate what types of uncertainty drive aggregate stock market risk premia.

We start from the idea that time-variation in the topics covered by the business press is a good proxy for the evolution of investors’ concerns regarding these topics.\(^1\) We estimate a news-based measure of uncertainty based on the co-movement between the front-page coverage of the Wall Street Journal and options-implied volatility (VIX). We call this measure News Implied Volatility, or NVIX for short. NVIX has two useful features that allow us to further our understanding of the relationship between uncertainty and expected returns: (i) it has a long time-series, extending back to the last decade of the nineteen century, covering periods of large economic turmoil, wars, government policy changes, and crises of various sorts; (ii) its variation is interpretable and provides insight into the origins of risk variation. The first feature enables us to study how compensation for risks reflected in newspaper coverage has fluctuated over time, and the second feature allow us to identify which kinds of risk were important to investors.

We rely on machine learning techniques to uncover information from this rich and unique text dataset. Specifically, we estimate the relationship between option prices and the frequency of words using Support Vector Regression. The key advantage of this method over Ordinary Least Squares is its ability to deal with a large feature space. We find that NVIX predicts VIX well out-of-sample, with an R-squared of 0.34 and root mean squared error of 7.52 percentage points. When we replicate our methodology with realized volatility instead of VIX, we find that it works well even as we go decades back in time, suggesting word-choice is fairly stable over this period.\(^2\)

\(^1\)This approach is consistent with the Gentzkow and Shapiro (2006) empirically supported model of news firms.
\(^2\)We analyze word-choice stability and measurement error in Section 2.3. One could potentially improve on
Asset pricing theory predicts that fluctuation in options implied volatility is a strong predictor of stock market returns as it measures fluctuation in expected stock market volatility (Merton, 1973), in the variance risk premium (Drechsler, 2008; Drechsler and Yaron, 2011), and in the probability of large disaster events (Gabaix, 2012; Wachter, 2013; Gourio, 2008, 2012). Motivated by this work we study whether fluctuations in NVIX encode information about equity risk premia.

We begin by focusing on the post-war period commonly studied in the literature for which high-quality stock market data is available. We find strong evidence that times of greater investor uncertainty are followed by times of above average stock market returns. A one standard deviation increase in NVIX predicts annualized excess returns higher by 3.4 percentage points over the next year and 3 percentage points annually over the next two years.

We dig deeper into the nature of the uncertainty captured by NVIX and find three pieces of evidence that these return predictability results are driven by variation in investors’ concerns regarding rare disasters as in Gabaix (2012), Wachter (2013) and Gourio (2008, 2012). First, we find that the predictive power of NVIX is orthogonal to risk measures based on contemporaneous or forward-looking measures of stock market volatility. Second, we use alternative option based measures, which are more focused on left tail risk, to estimate their news-based counterparts. Specifically, our news-based extensions of the variance premium (Bollerslev, Tauchen, and Zhou, 2009), the Bollerslev and Todorov (2011) model free measure of left tail risk (LT), and implied volatility slope give similar predictability results. Particularly robust is the predictive ability of LT across multiple horizons.

Interpretability, a key feature of the text-based approach, provides a third piece of evidence by allowing us to trace back a large part of the variation in risk premia to concerns related to wars (47%) and government policy (23%). A substantial part of the time-series variation in risk premia NVIX identifies is driven by concerns tightly related to the type of events discussed in the rare disasters literature (Rietz, 1988; Barro, 2006). We find that government-related concerns are related to redistribution risk, as our measure traces remarkably well tax-policy changes in the US. We decompose NVIX into four additional categories: Stock Markets, Financial Intermediation, Natural Disasters, and a residual component. Of these categories only the residual component this out-of-sample fit using financial variables (e.g. past volatility, default spreads, etc.) at the cost of losing the interpretability of the text-based index, which is central to our analysis.
reliably predicts future expected returns. Interestingly, even though uncertainty regarding the stock market itself—an NVIX component highly correlated with realized volatility—drive a substantial part of the variation in NVIX, this variation is not priced. By contrast, while concerns related to wars or government policy do not drive most of the variation in news implied volatility, they do drive most of its priced variation. These results suggest that time-varying disaster risk in particular is priced in the post-war US stock market.

We then extend our analysis to include the earlier and turbulent 1896–1945 period. This period includes two or three major economic downturns and two World Wars making it well-suited to investigate further if NVIX is related to economic disasters. We start by showing how the predictability result disappears with the inclusion of the Great Depression and the second world war period. From the time-varying disaster risk perspective this result is not surprising, as the positive relation between disaster concerns and future disaster realizations should attenuate the return predictability sample estimates in a sample with large disasters. To evaluate this possibility formally, we adapt the structural approach of Nakamura, Steinsson, Barro, and Ursúa (2013) for identification of economic disasters, and extend it to better identify the correct timing of transition into disaster states.

We identify three different periods as disaster regimes, all in the earlier part of the sample: after World War I, the Great Depression, and the late 1930’s. Our measure of disasters is a disaster probability from an econometrician’s perspective, which also identifies periods of near misses or crises of various sorts: periods such as the financial crises of the early 1980’s and 2008, the oil shock of the early 1970’s, and World War II. Consistent with the notion that NVIX encodes disaster concerns, NVIX predicts innovations in the disaster probability. A one standard deviation increase in NVIX predicts a 2.8% higher probability of disaster within the next year. These results are robust to inclusion of several controls for contemporaneous and forward-looking measures of stock market variance, both qualitatively and quantitatively.

We find that NVIX is abnormally high up to 12 months before a disaster, while conventional measures of financial volatility have virtually no information about impending transitions into a disaster regime. Once we adjust our estimation to take into account the disaster realizations in the pre-war sample, the relationship between NVIX and future returns reemerges, with predictive coefficients that are strikingly similar to the post-war sample estimates.
Our paper fits in a large literature that studies asset pricing consequences of large and rare economic disasters. At least since Rietz (1988), financial economists have been concerned about the pricing consequences of large events that happened not to occur in US data. Brown, Goetzmann, and Ross (1995) argues the fact we can measure the equity premium in the US stock market using such a long sample suggests that its history is special. Barro (2006) and subsequently Barro and Ursúa (2008); Nakamura, Steinsson, Barro, and Ursúa (2013); Barro (2009) show that calibrations consistent with 20th century world history can make quantitative sense of equity premium point estimates in the empirical literature. Gabaix (2012), Wachter (2013), Gourio (2008), and Gourio (2012) further show that calibrations of a time-varying rare disaster risk model can also explain the amount of time-variation in the data.

A major challenge of this literature is whether those calibrations are reasonable. As Gourio (2008) puts it, “this crucial question is hard to answer, since the success of this calibration is solely driven by the large and persistent variation is the disaster probability, which is unobservable.” We bring new data to bear on this question. We find that the overall variation in disaster probabilities used in calibrations such as Wachter (2013) line up well with our estimates. Our estimates, however, suggest substantially lower persistence than previously calibrated by Wachter (2013) and Gourio (2008, 2012). Moreover, we estimate that a 1 percentage point increase in the annual probability of a disaster increases risk premia by 1.14 percentage point. This effect on risk premia is remarkably close to the risk premia disaster sensitivity produced by Wachter (2013), where disaster magnitudes are calibrated to match the distribution of disasters in the Barro and Ursúa (2008) cross county data. We interpret this as evidence that the time-variation in disaster concerns measured by NVIX regards disasters of the same magnitude as studied in the rare disasters literature.

One motivation for our paper is the empirical fact that estimating aggregate risk-return trade-offs is a data intensive procedure. Indeed, Lundblad (2007) shows that the short samples used in the literature is the reason why research on the classic variance-expected return trade-off had been inconclusive. Testing the particular form of risk-return trade-off predicted by the time-varying disaster risk hypothesis is more challenging on two fronts; plausible measures of disaster risk are available for no more than two decades, and validation of these measures is even more challenging, since disasters are rare.

There is a large and fruitful literature that exploits the information embedded in option markets
to learn about the structure of the economy. Drechsler (2008) proposes a theory where the VIX has information about degree of ambiguity aversion among investors. Drechsler and Yaron (2011) interpret it as a forward looking measure of risk. Bollerslev and Todorov (2011) uses a model free approach to back out from option prices a measure of the risk-neutral distribution of jump sizes in the S&P 500 index. Bates (2012) shows that time-changed Lévy processes capture well stochastic volatility and substantial outliers in US stock market returns. Kelly and Jiang (2014) estimates a tail risk measure from a 1963-2010 cross-section of returns and finds it is highly correlated with options-based tail risk measures. Backus, Chernov, and Martin (2011) present an important challenge to the idea that the “overpricing” of out of the money put options can be explained by static rare disaster risk models. Seo and Wachter (2013) show, however, that this apparent inconsistency can be resolved in a model with time-varying disaster risk.

Our paper connects information embedded in VIX with macroeconomic disasters by extending it back a century, and by using cross equation restrictions between disaster and return predictability regressions to estimate disaster probability variance and persistence. Importantly, by decomposing NVIX into word categories we add to this literature interpretable measures of distinct disaster concerns, and gain novel insights about the origins of risk premia variation.3

Broadly, our paper contributes to a growing body of work that applies text-based analysis to fundamental economic questions. Hoberg and Phillips (2010, 2011) use the similarity of company descriptions to determine competitive relationships. Baker, Bloom, and Davis (2013) develop an index of policy-related economic uncertainty using the frequency of newspaper references to policy uncertainty. Tetlock (2007) documents that the fractions of positive and negative words in certain financial columns predict subsequent daily returns on the Dow Jones Industrial Average, and García (2013) shows that this predictability is concentrated in recessions. These effects mostly reverse quickly, which is more consistent with a behavioral investor sentiment explanation than a rational compensation for risk story. By contrast, we examine lower (monthly) frequencies, and find strong return and disaster predictability consistent with a disaster risk premium by funneling front-page appearances of all words through a first-stage text regression to predict economically interpretable VIX. The support vector regression we employ offers substantial benefits over the more common

---

3Sample size is especially important for studying rare events. An alternative approach to our long time-series is to study a large cross-section of countries (e.g. Gao and Song, 2013).
approach of classifying words according to tone (e.g. Loughran and McDonald, 2011). It has been used successfully by Kogan, Routledge, Sagi, and Smith (2010) to predict firm-specific volatility from 10-K filings. We review in Section A.2 alternative text-based methods and explain why the chosen approach is superior for our purposes.

The paper proceeds as follows. Section 2 describes the data and methodology used to construct NVIX. Section 3 tests the hypothesis that time-variation in uncertainty is an important driver of variation in expected returns in post-war US data, reports our main results and identifies time-varying disaster concerns as a likely explanation. Section 4 uncovers which concerns drive risk premia. Section 5 extends our analysis back to 1896. Section 6 concludes.

2 Data and Methodology

We begin by describing the standard asset pricing data we rely on, as well as our unique news dataset and how we use it to predict implied volatility out-of-sample.

We assume throughout that the choice of words by the business press provides a good and stable reflection of the concerns of the average investor. This assumption is quite natural and consistent with a model of a news firm which observes real-world events and then chooses what to emphasize in its report, with the goal of building its reputation. Gentzkow and Shapiro (2006) build a model along these lines and present a variety of empirical evidence consistent with its predictions. The idea that news media reflects the interests of readers is suggested in Tetlock (2007), empirically supported by Manela (2011), and used for structural estimation of the value of information in Manela (2014).

2.1 News Implied Volatility (NVIX)

Our news dataset includes the title and abstract of all front-page articles of the Wall Street Journal from July 1889 to December 2009. We focus on front-page titles and abstracts to make the data collection feasible, and because these are manually edited and corrected following optical character recognition, which improves their earlier sample reliability. We omit titles that appear daily. Each title and abstract are separately broken into one and two word n-grams using a standard text

---

4We omit the following titles keeping their abstracts when available: 'business and finance', 'world wide', 'what’s news', 'table of contents', 'masthead', 'other', 'no title', 'financial diary'.

6
analysis package that replaces highly frequent words (stop-words) with an underscore, and removes n-grams containing digits.\textsuperscript{5}

We combine the news data with our estimation target, the implied volatility indices (VIX and VXO) reported by the Chicago Board Options Exchange. We use the older VXO implied volatility index that is available since 1986 instead of VIX that is only available since 1990 because it grants us more data and the two indices are 0.99 correlated at the monthly frequency.

We break the sample into three subsamples. The \textit{train} subsample, 1996 to 2009, is used to estimate the dependency between news data and implied volatility. The \textit{test} subsample, 1986 to 1995, is used for out-of-sample tests of model fit. The \textit{predict} subsample includes all earlier observations for which options data, and hence VIX is not available.\textsuperscript{6}

We aggregate n-gram counts to the monthly frequency to get a relatively large body of text for each observation. Since there are persistent changes over our sample in the number of words per article, and the number of articles per day, we normalize n-gram counts by the total number of n-grams each month. Each month of text is therefore represented by $x_t$, a $K = 374,299$ vector of n-gram frequencies, i.e. $x_{t,i} = \frac{\text{appearances of n-gram } i \text{ in month } t}{\text{total n-grams in month } t}$. We mark as zero those n-grams appearing less than 3 times in the entire sample, and those n-grams that do not appear in the \textit{predict} subsample. We subtract the mean $\overline{VIX} = 21.42$ to form our target variable $v_t = VIX_t - \overline{VIX}$.

We use n-gram frequencies to predict VIX with a linear regression model

$$v_t = w_0 + w \cdot x_t + v_t \quad t = 1 \ldots T$$

where $w$ is a $K$ vector of regression coefficients. Clearly $w$ cannot be estimated reliably using least

\textsuperscript{5}For example, the sentence “The Olympics Are Coming” results in 1-grams “olympics” and “coming”; and 2-grams “olympics”, “olympics”, and “coming”. We use ShingleAnalyzer and StandardAnalyzer of the open-source Apache Lucene Core project to process the raw text into n-grams. We have experimented with stemming and considering different degree n-grams and found practically identical results, but since this is the procedure we first used, we report its results throughout to get meaningful out-of-sample tests.

\textsuperscript{6}A potential concern is that since the \textit{train} sample period is chronologically after the \textit{predict} subsample, we are using a relationship between news reporting and disaster probabilities that relies on new information, not in the information sets of those who lived during the \textit{predict} subsample, to predict future returns. While theoretically possible, we find this concern empirically implausible because the way we extract information from news is indirect, counting n-gram frequencies. For this mechanism to work, modern newspaper coverage of looming potential disasters would have to use less words that describe old disasters. By contrast, suppose modern journalists now know the stock market crash of 1929 was a precursor for the great depression. As a result, they give more attention to the stock market and the word “stock” gets a higher frequency conditional on the disaster probability in our \textit{train} sample than in earlier times. Such a shift would cause its regression coefficient $w_{\text{stock}}$ to underestimate the importance of the word in earlier times. Such measurement error actually works against us finding return and disaster predictability using our measure.
squares with a training time series of $T_{\text{train}} = 168$ observations.

We overcome this problem using Support Vector Regression (SVR), an estimation procedure shown to perform well for short samples with an extremely large feature space $K$.

While a full treatment of SVR is beyond the scope of this paper, we wish to give an intuitive glimpse into this method, and the structure that it implicitly imposes on the data. SVR minimizes the following objective

$$H(w, w_0) = \sum_{t \in \text{train}} g_{\epsilon} (v_t - w_0 - w \cdot x_t) + c (w \cdot w),$$

where $g_{\epsilon}(e) = \max\{0, |e| - \epsilon\}$ is an “$\epsilon$-insensitive” error measure, ignoring errors of size less than $\epsilon$. The minimizing coefficients vector $w$ is a weighted-average of regressors

$$\hat{w}_{\text{SVR}} = \sum_{t \in \text{train}} (\hat{\alpha}_t^* - \hat{\alpha}_t) x_t$$

(2)

where only some of the $T_{\text{train}}$ observations’ (dual) weights $\alpha_t$ and $\alpha_t^*$ are non-zero.

SVR works by carefully selecting a relatively small number of observations called support vectors, and ignoring the rest. The trick is that the restricted form (2) does not consider each of the $K$ linear subspaces separately. By imposing this structure, we reduce an over-determined problem of finding $K \gg T$ coefficients to a feasible linear-quadratic optimization problem with a relatively small number of parameters (picking the $T_{\text{train}}$ dual weights $\alpha_t$). The cost is that SVR cannot adapt itself to concentrate on subspaces of $x_t$ (Hastie, Tibshirani, and Friedman, 2009). For example, if the word “peace” were to be important for VIX prediction independently of all other words that appeared frequently at the same low VIX months, say “Tolstoy”, SVR would assign the same weight to both. Ultimately, success or failure of SVR must be evaluated in out-of-sample fit which we turn to next.

Figure 1 shows estimation results. Looking at the train subsample, the most noticeable observations are the LTCM crisis in August 1998, September 2002 when the US made it clear an Iraq crisis. The $\epsilon$-insensitive zone and regularization parameter $c$. Rather than make these choices ourselves, we use the procedure suggested by Cherkassky and Ma (2004) which relies only on the train subsample. We first estimate using k-Nearest Neighbor with $k = 5$, that $\sigma_v = 6.664$. We then calculate $c_{CM2004} = 29.405$ and $c_{CM2004} = 3.491$. We numerically estimate $w$ by applying with these parameter values the widely used SVMlight package (available at http://svmlight.joachims.org/) to our data.

\footnote{See Kogan, Levin, Routledge, Sagi, and Smith (2009); Kogan, Routledge, Sagi, and Smith (2010) for an application in finance or Vapnik (2000) for a thorough discussion of theory and evidence. We discuss alternative approaches in Section A.2.}

\footnote{SVR estimation requires us to choose two hyper-parameters that control the trade-off between in-sample and out-of-sample fit (the $\epsilon$-insensitive zone and regularization parameter $c$). Rather than make these choices ourselves, we use the procedure suggested by Cherkassky and Ma (2004) which relies only on the train subsample. We first estimate using k-Nearest Neighbor with $k = 5$, that $\sigma_v = 6.664$. We then calculate $c_{CM2004} = 29.405$ and $c_{CM2004} = 3.491$. We numerically estimate $w$ by applying with these parameter values the widely used SVMlight\textsuperscript{\textregistered} package (available at http://svmlight.joachims.org/) to our data.}
invasion is imminent, the abnormally low VIX from 2005 to 2007, and the financial crisis in the fall of 2008. In-sample fit is quite good, with an $R^2 (\text{train}) = \frac{\text{Var}(w'x_t)}{\text{Var}(v_t)} = 0.65$. The tight confidence interval around $\hat{v}_t$ suggests that the estimation method is not sensitive to randomizations (with replacement) of the train subsample. This gives us confidence that the methodology uncovers a fairly stable mapping between word frequencies and VIX, but with such a large feature space, one must worry about over-fitting.

However, as reported in Table 1, the model’s out-of-sample fit over the test subsample is quite good, with $R^2 (\text{test}) = 0.34$ and $RMSE (\text{test}) = 7.52$. In addition to these statistics, we also report results from a regression of test subsample actual VIX values on news-based values. We find that NVIX is a statistically powerful predictor of actual VIX. The coefficient on $\hat{v}_t$ is statistically greater than zero ($t = 3.99$) and no different from one ($t = -1.33$), which supports our use of NVIX to extend VIX to the longer sample.

### 2.2 NVIX is a Reasonable Proxy for Uncertainty

NVIX captures well the fears of the average investor over this long history. Noteworthy peaks in NVIX include the stock market crash of October and November 1929 and other tumultuous periods which we annotate in Figure 2. Stock market crashes, wars and financial crises seem to play an important role in shaping NVIX. Noteworthy in its absence is the “burst” of the tech bubble in March 2000, thus not all market crashes indicate rising concerns about future disasters. Our model produces a spike in October 1987 when the stock market crashed and a peak in August 1990 when Iraq invaded Kuwait and ignited the first Gulf War. This exercise gives us confidence in using the model to predict VIX over the entire predict subsample, when options were hardly traded, and actual VIX is unavailable.

We find it quite plausible that spikes in uncertainty perceived by the average investor would coincide with stock market crashes, world wars and financial crises. Since these are exactly the times when NVIX spikes due to each of these concerns, we find it is a plausible proxy for investor uncertainty.

It is perhaps surprising that NVIX is relatively smooth over the predict sample. For example, the Great Depression is widely regarded as the most severe realized economic disaster the US experienced over the past century, but during that time period, NVIX increases from about 25% to
30%, peaking at 41% on October 1929. We note, however, that like options-implied volatility, NVIX is a forward-looking measure of uncertainty, in contrast with backward-looking realized volatility, which mechanically spikes during disaster realizations. Alternatively, this could happen because measurement attenuates NVIX, a concern we turn to next.

2.3 Word-choice Stability and Measurement Error

We assume throughout that the choice of words by the business press provides a good and stable reflection of the concerns of the average investor. Otherwise, the type of “Big Data” techniques we use to interpret text data would produce noisy estimates of implied volatility. Such measurement error would bias our predictability results toward zero.

One concern is that the issues worrying investors change over time. For example, the “Dust Bowl” was a uniquely salient feature of the 1930s, which featured severe dust storms, drought, and agricultural damage. Since it is unlikely to concern modern day investors enough to make front page news during our training sample, we might measure with error the actual uncertainty that prevailed during the thirties. Technically, to estimate reliably the relationship between specific sources of aggregate uncertainty and word usage of the business press, we require variation in both during our train subsample. We choose to train on the recent sample, and test on the earlier one, so we can get a sense of out-of-sample fit when we go even further back in time. This choice is not innocuous. If we were to reverse the order and train on the earlier sample, our text regression would miss important variation due to the financial crisis of 2008, and instead focus on the stock market crash of 1987. We are therefore (ironically) fortunate to have experienced the financial crisis during the sample where options-implied volatility is available.

A related concern is that the meaning of certain words or phrases used by the business press has changed considerably over our long sample. For example, the mapping from the 2-gram “Japanese navy” to investor concerns about disaster risks in the 1940s is likely different than in the 2000s. Ideally, we would only consider more common phrases with a stable meaning, such as “war”. The techniques we use are, however, designed to avoid such overfitting pitfalls, and proved successful in related settings (Antweiler and Frank, 2004; Kogan, Routledge, Sagi, and Smith, 2010).

Nonetheless, we wish to quantify how measurement error changes when moving from the test subsample to the predict subsample, but VIX is not available during this earlier period. Instead,
we use realized volatility, a closely related variable highly correlated with VIX. We repeat the same estimation procedure over the same \textit{train} subsample as before, only replacing VIX with realized volatility as the dependent variable of the SVR (1).

We find that our predictive ability over the long sample is quite stable. Table 2 reports several different measures of realized volatility fit to news data over the three subsamples. The most natural measure of fit is root mean squared error of the text regression ($RMSE_{SVR}$), according to which, measurement error in the \textit{predict} subsample is only slightly higher than in the \textit{test} subsample. RMSE increases from 9.6 percent to 10.9 percent annualized volatility. R-squared measures of fit are higher or comparable in the predict subsample relative to the \textit{test} subsample. We therefore expect only a modest increase in measurement error of NVIX as we extend VIX further back to times the index did not exist.

### 2.4 Asset Pricing Data

We use two different data sources for our stock market data. We use CRSP total market portfolio for the period from 1926 to 2009 and the Dow Jones index from Global Financial Data, available monthly from July 1896 to 1926. We refer to this series as “market” returns. Results are similar if we use the Dow Jones index throughout. We also use Robert Shiller’s time series of aggregate S&P 500 earnings from his website. We chose to use this data to run our predictability tests because this index is representative of the overall economy and goes back a long way. We use daily return data on the CRSP total market portfolio and the Dow Jones index to construct proxies for realized volatility, which is important when we explore alternative explanations of our main result. To compute excess returns we use the one-month t-bill rate as a measure of the risk free rate. When this data is not available we use yields on 10 year US government bonds from Global Financial Data. We use the difference between Moody’s Baa and Aaa yields as our measures of credit spreads. This data is only available after 1919. We use the VXO and VIX indices from the CBOE. They are implied volatility indices derived from a basket of option prices on the S&P 500 (VIX) and S&P 100 (VXO) indices. The VIX time series starts in January 1990 and VXO starts in January 1986. The LT measure of Bollerslev and Todorov (2011) was kindly provided to us by the authors. We use Option Metrics data to construct a measure of the slope of the implied volatility curve for the S&P 500 index.
3 Post-War Compensation for Risks Measured by NVIX

In this section we test the hypothesis that time-variation in uncertainty is an important driver of variation in expected returns on US equity over the post-World War II, 1945 to 2009 sample. High-quality stock market data is available for this period. During this period, commonly studied in the literature, the US experienced no economic disasters of the magnitude of a Great Depression or a World War. We start with our main findings that NVIX predicts returns. We then show that stochastic volatility is not behind this result and that we get even stronger results with alternative uncertainty measures more focused on tail risk.

3.1 NVIX Predicts Returns

Asset pricing models with time-varying risk premia predict that times when risk is relatively high would be followed by above average returns on the aggregate market portfolio. For example, the dynamic risk-return tradeoff in Merton (1973) predicts a linear relation between the conditional expected excess return on the market to its conditional variance, as well as its conditional covariance with other priced risk factors. The more recent time-varying rare disaster models, predict a linear relationship between expected excess returns and the variance premium, which is linear in the time-varying probability of a rare disaster (e.g. Gabaix, 2012). Therefore, our main tests try to explain future excess returns on the market portfolio at various horizons with lagged forward-looking measures of risk as measured by NVIX squared. We place our measure in variance space because in all the above-mentioned models, risk premia are linear in variances as opposed to standard deviations.9 To alleviate any concerns about news-based measures that rely on weekend news coverage not yet priced in the stock market, we skip a month to err on the side of caution. Because our forecasts use overlapping monthly data we adjust standard errors to reflect the dependence that this introduces into forecast errors using four different ways: Newey and West (1987), Hansen and Hodrick (1980) , Hodrick (1992), and Bootstrap. For the first three standard errors use the same number of lags as the forecasting window. In our empirical analysis, results for all of these test statistics are similar, and robust to the use of somewhat longer lags. We report Newey and West (1987) standard error trough out.

9The results are very similar in terms of statistical significance and economic magnitude if we use NVIX instead.
The last two columns of Table 3 show that in the short-sample for which option prices are available, the ability of VIX to predict returns is statistically rather weak. In the sample for which VIX is available, the implied volatility index predicts excess returns in the six months to twelve months horizons. If we consider a slightly longer sample for which the VXO implied volatility index on the S&P 100 is available, the evidence for return predictability becomes weaker. Would these results change if we had a longer sample of such forward-looking measures of uncertainty?

While we do not have new options data to bring to bear, we use NVIX to extrapolate these options-based measures of uncertainty back in time. NVIX largely inherits the behavior of VIX and VXO in sample periods where both are available. Point estimates and standard errors are quite similar, especially for the VIX sample. This is hardly surprising, because NVIX was constructed to fit these implied volatility indices, though we only use post 1995 data for NVIX estimation.

The advantage of using NVIX, however, is the ability to consider much longer samples. The first two columns of Table 3 reports our main results for two alternative extended sample periods. In the first column we see that return predictability for the entire post-war period going from 1945 to 2009 is well estimated with larger point estimates relative to the VIX sample. From six months to twenty-four months horizons the coefficients are statistically significant at the 1 percent level, unlike for the VIX sample. The second column reports results for the sample period for which we did not use any in-sample option price data. Out-of-sample, estimates are even larger and statistically significant at one to twenty-four months horizons.

Comparing columns (3) and (5), which respectively use NVIX and VXO over the sample 1986–2009 sample shows that the point estimates are rather similar, but statistically indistinguishable from zero. By contrast, the full pre-war sample in column (1) and even just the pre-1995 which we did not use to fit the model, grants us the statistical power to reject zero at usual significance levels.

We interpret the extended sample results as strong evidence for the joint hypothesis that NVIX measures investors’ uncertainty and that time-variation in uncertainty drives expected returns. The coefficient estimates imply substantial predictability with a one standard deviation increase in NVIX$^2$ leading to $\sigma_{NVIX^2} \times \beta_1 = 21.11 \times 0.16 = 3.4\%$ higher excess returns in the following year. Unsurprisingly, R-squares are small and attempts to exploit this relationship carry large risks even for a long-run investor. Annualized forecasting coefficients are fairly stable across forecasting
horizons. We next dig deeper into the nature of this priced uncertainty captured by NVIX.

### 3.2 Stochastic Volatility Does Not Explain These Results

One potential explanation for the ability of NVIX to predict returns is that NVIX measures variation in current stock market volatility (Merton, 1973). According to this story NVIX predicts returns because investors demand higher expected returns during more volatile periods.

We test this story using lagged realized variance as well as four alternative variance forecasting models, gradually adding additional predictors, such as additional realized variance lags, price to earnings ratio, NVIX and credit spread. The bottom line of Table 4 compares the ability of the alternative variance forecasting models to predict future variance. We use these models to control for fluctuation in stock market risk in our return predictability specification.

Table 4 shows that the coefficient on NVIX is about the same and that standard errors either decrease or only slightly increase when we control for realized or expected variance. Note that coefficient do not change even after we add NVIX to the variance forecasting model (model 4). This establishes that NVIX embeds priced information that is largely orthogonal to any information NVIX or other standard predictor variables might contain regarding future volatility.

### 3.3 Horse Races

We horse race NVIX directly against different predictors. If the concerns encoded in NVIX are the same concerns reflected in the other predictors, then the predictor measured with more noise should be driven out of the regression, and if not driven completely out we would expect the coefficient magnitude to decrease.

The results in Table 5 show remarkably stable coefficients across specifications, suggesting that NVIX captures additional information relative to what is reflected in a variance based measured of VIX, the credit spread, or the price to earnings ratio. Comparing R-squared across horizons we see that the predictive power of NVIX and the other variables roughly adds up. At the yearly horizon, the NVIX has a (univariate) R-squared of 3.5% and a marginal contribution of 2.6% (column 4 less column 5 is 8.8%–6.2%). All the other variables together have an R-squared of 6.2% with marginal contribution of 5.3% (8.8%–3.5%). This pattern appears across alternative sample periods, and strongly suggests that these variables measure different things.
In the Appendix A.2 we horse race NVIX against other text based measures that have been shown to be successful in aggregating information from text. NVIX remains a significant return across all specifications.

3.4 Alternative Measures of Uncertainty Focused on Tail Risk

In Table 6 we replicate our analysis using alternative measures of uncertainty, which are more focused on left tail risk. For each of these measures we reproduce the methodology we applied to VIX. The first column reproduces our main results. In the second column we have VIX premium \( = VIX_t^2 - E_t[Var(R_{t+1})] \), where \( E_t[Var(R_{t+1})] \) is constructed using an AR(1) for realized variance (Bollerslev, Tauchen, and Zhou, 2009). In the third column we have the options-based and model free Left-Tail (LT) measure of Bollerslev and Todorov (2011). In the fourth column we have the slope of option implied volatility curve, constructed using 30-day options from Option metrics. We report in the raw correlations across these different measures of tail risk in the appendix.

Results for these alternative measures of news implied uncertainty give similar and even stronger predictability results. Particularly robust is the predictive ability of the Bollerslev and Todorov (2011) model-free measure of left tail risk (LT). The direction of return predictability is consistent with the hypothesis that the predictability is driven by time-varying disaster risk concerns. When tail-risk is high, as measured by any of the four alternative measures, average returns are higher going forward.

All of these measures in one way or another take higher values when options that pay off in bad states of the world are relatively expensive. These options can be expensive because investors' attitudes towards these states take a turn for the worse, as in the time-varying Knightian uncertainty model of Drechsler (2008), or because the objective probability that these states occur increases, as in time-varying rare disaster models (Gourio, 2008, 2012; Gabaix, 2012; Wachter, 2013). In either case, NVIX appears to capture concerns about large and infrequent macroeconomic events.

4 Origins of Uncertainty Fluctuations

In this section, we lever the text-based feature of our uncertainty measure to gain novel insights into the origins of uncertainty fluctuations. The results in Section 3 imply that priced variation
in NVIX is unrelated with standard measures of stock market risk, and likely to be related to
fluctuations in tail risk. Guided by this evidence, we decompose our uncertainty measure into
five interpretable categories meant to capture different types of shocks: Government, Financial
Intermediation, Natural Disasters, Stock Markets and War. We find that a substantial amount of
risk premia variation is driven by concerns about war and expropriation-related disasters.

4.1 Important Words

We begin by calculating the fraction of NVIX variance that each word drives over the predict
subsample. Define \( \hat{v}_t(i) \equiv x_{t,i}w_i \) as the value of VIX predicted only by n-gram \( i \in \{1..K\} \). We construct
\[
h(i) \equiv \frac{\text{Var}(\hat{v}_t(i))}{\sum_{j \in K} \text{Var}(\hat{v}_t(j))}
\]
as a measure of the n-gram specific variance of NVIX.\(^{10}\) Table 7 reports \( h(i) \) for the top vari-
ance driving n-grams and the regression coefficient \( w_i \) from the model (1) for the top variance
n-grams. Note that the magnitude of \( w_i \) does not completely determine \( h(i) \) since the frequency
of appearances in the news interacts with \( w \) in (3).

Clearly, when the stock market makes an unusually high fraction of front page news it is a
strong indication of high implied volatility. The word “stock” alone accounts for 37 percent of NVIX
variance. Examining the rest of the list, we find that stock market-related words are important as
well. This should not be surprising since when risk increases substantially, stock market prices tend
to fall and make headlines. War is the fourth most important word and accounts for 6 percent.

4.2 Word Categorization

We rely on the widely used WordNet and WordNet::Similarity projects to classify words.\(^{11}\) WordNet
is a large lexical database where nouns, verbs, adjectives and adverbs are grouped into sets of
cognitive synonyms (synsets), each expressing a distinct concept. We select a number of root
synsets for each of our categories, and then expand this set to a set of similar words which have a
path-based WordNet::Similarity of at least 0.5.

\(^{10}\) Note that in general \( \text{Var}(\hat{v}_t) \neq \sum_{j \in K} \text{Var}(\hat{v}_t(j)) \) due to covariance terms.
\(^{11}\) WordNet (Miller, 1995) is available at http://wordnet.princeton.edu. WordNet::Similarity (Pedersen, Patward-
Table 8 reports the percentage of NVIX variance \(= \sum_{i \in C} h(i)\) that each n-gram category drives over the predict subsample. Stock market related words explain over half the variation in NVIX. War-related words explain 6 percent. Unclassified words explain 37 percent of the variation. Clearly there are important features of the data, among the 374,299 n-grams that the automated SVR regression picks up. While these words are harder to interpret, they seem to be important in explaining VIX behavior in-sample, and predicting it out-of-sample.

Each NVIX component can be interpreted as a distinct type of disaster concerns. Figure 3 plots each of the four NVIX categories responsible for more of its variation to provide some insight into their interpretation. We omit the easily interpretable Natural Disasters category because it generates a negligible amount of NVIX variation.

The NVIX Stock Markets component has a lot to do with stock market volatility as shown in Figure 3a. Attention to the stock market as measured by this component seems to spike at market crashes and persist even when stock market volatility declines. This component likely captures proximate concerns about the stock market that have other ultimate causes, but can also capture concerns with the market itself.

Wars are clearly a plausible driver of disaster risk because they can potentially destroy a large amount of both human and physical capital and redirect resources. Figure 3b plots the NVIX War component over time. The index captures well the ascent into and fall out of the front-page of the *Journal* of important conflicts which involved the US to various degrees. A common feature of both world wars is an initial spike in NVIX when war in Europe starts, a decline, and finally a spike when the US becomes involved.

The most striking pattern is the sharp spike in NVIX in the days leading up to US involvement in WWII. The newspaper was mostly covering the US defensive buildup until the Japanese Navy’s surprise attack at Pearl Harbor on December 1941. Following the attack, the US actively joined the ongoing War. NVIX[War] jumps from 0.87 in November to 2.86 in December and mostly keeps rising. The highest point in the graph is the Normandy invasion on June 1944 with the index reaching 4.45. The *Journal* writes on June 7, 1944, the day following the invasion: “Invasion of the continent of Europe signals the beginning of the end of America’s wartime way of economic life.” Clearly a time of elevated disaster concerns. Thus, NVIX captures well not only whether the US was engaged in war, but also the degree of concern about the future prevalent at the time.
Policy-related uncertainty as captured by our Government component tracks well changes in the average marginal tax rate on dividends as shown in Figure 3c. An important potential disaster from a stock market investor perspective is expropriation of ownership rights through taxation. While in retrospect, a socialist revolution did not occur in the US over this period, its probability could have been elevated at times.

Financial Intermediation-related NVIX spikes when expected, mostly during financial crises. Figure 3d shows that the Intermediation component is high during banking crises identified by Reinhart and Rogoff (2011), but also during other periods when bank failures were high, such as the late 1930s and early 1970s. Apparent in the figure are the panic of 1907, the Great Depression of the 1930s, the Savings & Loans crisis of the 1980s and the Great Recession of 2008.

4.3 Which Concerns Drive Risk Premia Variation?

We report a text-based decomposition of risk premia variation in Table 9. The shares of risk premia variation due to each of the categories is in parentheses, which add up to more than one because the different categories are not mutually orthogonal. At the yearly horizon, Government (57%) and War (14%) related concerns capture the bulk of the post-war variation in risk premia. Both categories have a statistically reliable relation with future market excess returns. Concerns related to Financial Intermediation (1.5%), Stock Markets (0%), and Natural disasters (6%), account for some of the variation in expected returns, but the relationship is statistically unreliable. The harder to interpret orthogonal residual accounts for 20% of the variation.

During the post-war sample, war-related concerns explain a substantial part of the variation in risk premia. This is somewhat surprising to a 21st century economist who knows that the US economy did not contract sharply during any of its 1945–2009 military conflicts. We stress again, however, that NVIX captures the concerns prevalent at the time, without the benefit of hindsight. During the 1896–1945 period, which included two world wars, war-related concerns explain a much larger 60 percent share of this variation, or 47 percent in the full sample. A substantial part of the variation in risk premia is therefore unequivocally related to disaster concerns.

Government-related concerns allows for a wider range of potential interpretations. Work by Pastor and Veronesi (2012), Croce, Nguyen, and Schmid (2012), and Baker, Bloom, and Davis (2013) emphasizes the role of policy-related uncertainty in inducing volatility and reducing asset
prices in the recent period. We find that policy-related uncertainty explains a substantial part of risk-premia variation, but not in the early sample. This finding is consistent with an increasing role for the government in the aftermath of the Great Depression and World-War II.

One might argue that policy-related uncertainty is a very different type of risk than the the rare disaster risk that the macro-finance literature has in mind. However, we find the tight relation between our government concerns measure and the evolution of US capital taxation shown in Figure 3c suggestive that our measure captures concerns related to expropriation risk. Not the typical cash-flow shock we use to model risk, but from the average capital holder perspective, a sudden sharp rise in taxes is a disaster. These results suggest that we may need to go beyond representative agent models to fully account for variation in risk premia.

Just as important, this result does not say that most of the variation in news implied volatility is related to disaster concerns, but shows that most of variation in news implied volatility that is *priced* in the stock market is due to disaster concerns. The fact that a substantial fraction of the variation in risk premia over the last century is due to concerns related to wars and taxation, strongly suggests that risk premia estimates likely reflect the special realization of history the US happened to experience during this period (Brown, Goetzmann, and Ross, 1995).

Stock Markets-related concerns are not reliably related to future returns. Figure 3a shows that these concerns track well the time-series of realized volatility. Common sense and theory predicts that investors pay more attention to the stock market in periods of high volatility (Abel, Eberly, and Panageas, 2007; Huang and Liu, 2007). While these concerns of the stock market with itself explain about half the variation in NVIX, we find that this variation is not priced.

We were surprised to find that Financial Intermediation does not account for much of the time-variation in risk premia in our data. This was puzzling to us because the largest event in the sample we estimate NVIX is the 2007–2008 financial crisis. We think there are different possible conclusions from the empirical evidence: it could be that our measure of uncertainty fails to pick up concerns related to the intermediary sector appropriately. However, Figure 3d strongly suggests that our measure gets the timing of the major financial events right. For example, note that during the great depression the intermediation measure peaks in 1933, three years after NVIX peaks. This timing lines up exactly with the the declaration of a national banking holiday and with the peak in bank failures (Cole and Ohanian, 1999). A second possibility is that financial
crises are intrinsically different since they are liability crises, essentially credit booms gone bust (Schularick and Taylor, 2009; Krishnamurthy and Vissing-Jorgensen, 2012). Reinhart and Rogoff (2011) suggests that financial crises are the result of complacency and a sense of “this time is different” among investors, what would suggest that financial crises happen only when investors are not concerned about financial intermediaries. Moreira and Savov (2013) build a macroeconomic model consistent with the notion that financial crises happen when investors perceive risk to be low, and predict that times of high intermediary activity are periods of low risk premia. The fact that Financial Intermediation does not account for much of the time-variation in risk premia in our data is consistent with our measure picking up financial intermediary activity during normal times, and concerns related to financial intermediaries during financial crises.

Our fifth category, Natural Disasters, also fails to predict returns. This is somewhat expected because we perceive as unlikely that there is time-variation in the likelihood of natural disasters at the frequencies we examine, in particular regarding natural disasters that are large enough to impact aggregate wealth. However, a behavioral story of overreaction to past disasters could generate such a link. Nonetheless, we find no such link in the data, at least not when one focuses on the entire stock market. Even though a large fraction of NVIX variation is not interpretable, as the overwhelming majority of words are unclassified, this residual component explains 1–19% of the variation in risk premia at yearly frequencies. Our ex-ante chosen categories seem to do a good job of capturing the concerns that impact risk premia, but there is still a non-trivial fraction of risk premia left unexplained.

Taken together these results paint a novel picture of the origins of aggregate fluctuations. Of the roughly 4% a year variation in risk premia news implied volatility can measure (Table 9, column 6), 49 percent is driven by war concerns, tightly related to the type of disasters that motivates the rare disaster literature. An additional 23 percent of this variation is plausibly related to expropriation risk, which is quite different from the cash-flow shocks usually studied in rare disaster models.

5 A Century of Disaster Concerns: Structural Analysis

We extend our analysis to include the earlier sample (1896–1945) to further evaluate the time-varying disaster risk hypothesis. The occurrence of the Great Depression, two world wars, and
several near misses allows us to test if NVIX has information regarding the likelihood of a disaster, and how disaster realizations impact the predictability pattern documented above.

Figure 4 shows that the inclusion of either the Great Depression or World War II has a large impact on our estimates. The figure depicts how the return predictability estimates evolve over our sample, with the date on the x-axis denoting the beginning of the estimation sample. Once each of these rare events drops out of the estimation sample, the coefficient increases sharply, implying that the relation between NVIX and future returns is quite different in the earlier sample.

Two plausible channels can reconcile this result with our findings about the post-war sample: (i) disaster realizations could statistically attenuate the return predictability relation if NVIX successfully predicts disasters; (ii) long-lasting disaster periods, a salient feature of the Nakamura, Steinsson, Barro, and Ursúa (2013) cross-country data, could have a similar effect.

We next investigate whether the large macroeconomic events of the early sample are indeed behind the sharp break in return predictability. We start by developing a formal methodology to empirically identify the occurrence and the precise timing of disasters, and use this measure to test if NVIX encodes forward-looking information regarding disaster realizations.

Consistent with the two channels, we find that NVIX is abnormally high up to 12 months before a disaster. Moreover, once we adjust our estimation to take into account the disaster realizations in the pre-war sample, and the persistence of disasters, the relationship between NVIX and future returns reemerges, with predictive coefficients that are strikingly similar to the post-war sample estimates.

5.1 Measuring Rare Disasters

Before we can say anything about the ability of NVIX to predict disasters, we need to identify disasters and their exact timing. We formally measure disasters using a statistical model of rare disasters in the spirit of Nakamura, Steinsson, Barro, and Ursúa (2013), who use a Bayesian framework to statistically distinguish disaster periods from normal periods by measuring the behavior of the annual consumption series for a large cross-section of countries. We extend their framework to include stock market returns as an additional signal about the state of the economy. We deliberately focus on a framework where the probability of a disaster transition is constant. This choice may appear puzzling and at odds with the time-varying disaster risk hypothesis, but it allows us
a conservative measurement of disasters. This approach identifies disasters exclusively from the ex-post behavior of consumption and stock-market data. By contrast, had we allowed the disaster probability to vary over time, the Bayesian filter applied to the data would mechanically infer a high disaster probability just before disaster realizations, or conversely, if we introduced a disaster probability signal (NVIX) in the estimation, the Bayesian procedure would be more likely to find disasters in periods the disaster signal was high.

5.1.1 State Dynamics

Consumption of the representative agent follows a two state Markov chain. Let states be \( s_t \in \{0, 1\} \), where \( s_t = 1 \) denotes a disaster state. Then log consumption and dividend growth follow,

\[
\begin{align*}
\Delta c_{t+1} &= \mu_c(s_t) + \sigma_c \epsilon_{c,t+1}^c, \\
\Delta d_{t+1} &= \mu_d(s_t) + \sigma_d \epsilon_{d,t+1}^d,
\end{align*}
\]

with state transitions \( Pr(s_{t+1} = 1|s_t = 0) = p < Pr(s_{t+1} = 1|s_t = 1) = q \). Disasters are persistent. Once in a disaster, the economy is likely to stay in a disaster for a while. For simplicity, we rule out any high-frequency co-variation between dividends and consumption. For our purposes this assumption is harmless.

In addition to the state \( s_t \) which is the only driver of time-variation in consumption growth, the economy features variation in non-priced shocks to dividend volatility \( \sigma_{d,t} \) that evolves,

\[
\sigma_{d,t+1}^2 = (1 - \rho_v) \sigma_d^2 + \rho_v \sigma_{d,t}^2 + \sigma_v \epsilon_v^{v,t+1}.
\]

The role of stochastic volatility is to provide a competing mechanism for the model to fit large return surprises.

We are interested in estimating \( I_t^D \equiv Pr(s_t = 1|y^T) \), the probability the economy was in a disaster regime at time \( t \), and \( I_{t+1}^{N \rightarrow D} \equiv Pr(s_{t+1} = 1, s_t = 0|y^T) \), the probability that economy transitioned into a disaster regime from time \( t \) to \( t + 1 \).
5.1.2 Asset Pricing Framework

We assume that the representative consumer in our model has a stochastic discount factor given by \( M_{t+1} = M(\Delta c_{t+1}, s_{t+1}, s_t) \).\(^{12}\) This specification implies that the price-dividend ratio of a dividend stream that satisfies the dynamics in (4) can be written as a function of the disaster state \( \pi(s_t) = \pi e^{\psi(s_t)} \), with \( \psi(0) = 0 \). Substituting this expression in the log return of the dividend claim we obtain the realized return dynamics. Using the standard log-linearization around the average price-dividend ratio we obtain,

\[
\log(R^e_{t+1}) = \mu_d(s_t) + \sigma_{d,t} \epsilon_{t+1}^d + \kappa_1 \psi(s_{t+1}) - \psi(s_t) + \kappa_0,
\]

where \( \kappa_1 \) and \( \kappa_0 \) are log-linearization constants. Realized returns reflect permanent shocks to dividends \( \sigma_{d,t} \epsilon_{t+1}^d \) and transitions into and out of a disaster state \( \kappa_1 \psi(s_{t+1}) - \psi(s_t) \).

Realized returns are informative about regime transitions to the extent that the price-dividend ratio sensitivity to the disaster state \( \psi(1) - \psi(0) \) is large relative to the volatility \( \sigma_{d,t} \). The model interprets large negative returns as more likely to be a transition into a disaster, if volatility has been previously low, and if future periods exhibit a substantial and persistent reduction in consumption growth. Large negative returns that are not followed by drops in economic activity are interpreted as a mix of increases in volatility and unusually large negative dividend innovations \( \sigma_{d,t} \epsilon_{t+1}^d \).

5.1.3 Measurement

We use a “mixed-frequency” approach adapted from Schorfheide, Song, and Yaron (2013) to simultaneously use economic data measured at different frequencies. This allow us to use the best consumption growth data that is available in each sample period. We model the true monthly consumption growth as hidden to the econometrician, and use annual consumption growth (Barro and Ursua (2008), 1896–1959) as signals. Whenever data on monthly consumption growth (NIPA, 1960–2009) is available we assume it measures consumption perfectly.

We represent monthly time subscript \( t \) as \( t = 12(j-1) + m \), where \( m = 1, \ldots, 12 \), \( j \) indexes the year and \( m \) the month within the year. Annual consumption is the sum of monthly consumption

---

\(^{12}\)This would be consistent with both CRRA-type preferences and with recursive preferences developed by Epstein and Zin (1989) and Weil (1990).
over the span of a calendar year, \( C_{(j)} = \sum_{m=1}^{12} C_{12(j-1)+m} \). Following Schorfheide, Song, and Yaron (2013) we represent annual consumption growth rates as a function of monthly growth rates. We first log-linearize this relationship around a monthly value \( C^* \) and define \( c \) as the percent deviations from \( C^* \):

\[
c_{(j)} = \frac{1}{12} \sum_{m=1}^{12} c_{12(j-1)+m}.
\]

(7)

Because monthly consumption growth can be written as \( g_{c,t} = c_t - c_{t-1} \), annual growth rates are

\[
g_{a_{(j)}} = c_{(j)} - c_{(j-1)} = \sum_{\tau=1}^{23} \left( \frac{12 - |\tau - 12|}{12} \right) g_{c,12j-\tau+1}.
\]

(8)

We measure realized variance using daily stock market return data within the month \( t \), which satisfies,

\[
rvar_t = \sigma^2_{d,t} + \sigma_{rvar} w_{rvar}^t,
\]

(9)

where \( w_{rvar} \) represents measurement error.

### 5.1.4 State Space Representation

We now construct the system state evolution and measurement equations. Let us define consumption growth shocks as deviations from the conditional (on the economic regime \( s \)) expected growth rate, \( \epsilon_{c,t+1} = g_{c,t+1} - \mu_c(s_t) \), and define the hidden state \( x_t \) as

\[
x_t = \begin{bmatrix} \sigma_{d,t}^2 - \bar{v}^2 \\ \epsilon_{t}^c \\ \epsilon_{t-1}^c \\ \vdots \\ \epsilon_{t-22}^c \end{bmatrix}.
\]

(10)

Its evolution can be represented as an auto-regressive process given by

\[
x_{t+1} = Ax_t + C\epsilon_{t+1},
\]

(11)

where \( \epsilon = [\epsilon^c, \epsilon^d, \epsilon^v] \). The measurement vector
can be represented as a function of the hidden states and the hidden disaster regimes as follows

\[ y_{t+1} = \begin{bmatrix} \log(R_{t+1}^e) - \log(R_{t+1}^f) \\ rvars_{t+1} \\ \Delta c^m_{t+1} \\ \Delta c^0_{t+1} \end{bmatrix} \]  

(12)

\[ H_{t+1} \times y_{t+1} = H_{t+1} \times F \left( \{ s_{t-j} \}_{j=0}^{11} \right) + Gx_t + B(x_t)\epsilon_{t+1} + Dw_{t+1}, \]  

(13)

where the matrix \( H_{t+1} \) selects the components of the measurement vector that are observed in a particular period. For example, annual consumption growth is only observed in the end of the year, so \( H_{t+1} \) selects the fourth row only when \( t + 1 \) is a December month and the annual consumption growth data is available. Monthly consumption is only available after 1959, so the matrix \( H_{t+1} \) selects the third row if \( t + 1 \) is in a year after 1959. The vector \( F(\{ s_{t-j} \}_{j=0}^{11}) \) adds the expected value of each of the measurement variables as function of the hidden economic states \( s_t \). The matrix \( G \) maps hidden state variables into the observable variables, the vector \( \epsilon \) groups economic shocks, and the vector \( w_{t+1} \) groups measurement errors.

5.1.5 Bayesian Filtering

Our goal is to filter the time-series of realized disasters. We keep the estimation simple by calibrating the parameters and using a Bayesian approach to infer state transitions. Given a set of calibrated parameters \( \Theta \) and the observed data \( Y = \{ y_t \}_{t=1}^{T} \), we estimate the most likely trajectory of the disaster state \( S = \{ s_t \}_{t=1}^{T} \) and the hidden variables \( X = \{ x_t \}_{t=1}^{T} \),

\[ p(S, X|Y) \propto p(Y|X, S)p(X|S)p(S) \]  

(14)

Bayesian inference requires the specification of a prior distribution \( p(Z) \) which we choose consistent with the 2% per year probability of a disaster event estimated by Barro and Ursua (2008) using cross-country data.

We use a Gibbs sampler to construct the posterior by repeating the following two steps:
1. Draw $S^{(i)} \sim p(S \mid X^{(i-1)}, Y)$

2. Draw $X^{(i)} \sim p(X \mid S^{(i)}, Y)$

Where we construct $p(X \mid S^{(i)}, Y)$ using a Kalman smoother. The Gibbs sampler generates a sequence of random variables which converges to the posterior $p(S, X \mid Y)$.

### 5.1.6 Calibration

Table 10 summarizes the calibrated parameters. Most of these are easily estimated from the data. We estimate the parameters driving the hidden volatility process by first fitting an AR(3) to realized variance and then estimating an AR(1) on the one step ahead variance predictor. The realized variance measurement error is constructed from the forecasting error of this specification. Consumption growth is calibrated to have annual volatility of 2%, and annual growth rate of 3.5% in good times and -2% during disasters. Disasters are assumed to start with a 2% probability per year, and economies transition out of disaster within a year with a 10% probability (this number pins down $q$). The log-linearization constant $\kappa_1$ is constructed using the average price-dividend ratio in the post-war sample. We set $\psi(s = 1)$ to be consistent with a stock market drop on a normal times to disaster transition of -25%, and set the quantity $\kappa_0 + \mu_d(s_t = 0) - \log(R_{ft})$ to fit the equity premium during the post-war period. The change in dividend growth is chosen so that $\mu_{d,t}(s = 1) - \mu_{d,t}(s_t = 0)$ lines up with the consumption drop during a disaster.

### 5.2 Results

#### 5.2.1 Filtered Disasters

Figure 5 shows the posterior probability that the economy is in a disaster state from the econometrician’s perspective. We identify three distinct disaster periods: the two disasters during the period known as the Great Depression, October 1929 to January 1933, and then June 1937 to 1939, as well as a four year period that starts with the US entry into the first world war in 1917 and lasts until the end of the depression of 1920–1921. Other periods stand out as near misses, such as the 2007–2009 financial crisis, the Volcker recession during the early 1980’s, the oil shock of the 1970’s, and the US entry into the second world war.
An important feature of our approach is that disasters are identified from the behavior of consumption data, but the exact timing is identified from stock market return data. This allows us to pinpoint precisely the timing of the regime change even in the earlier part of our sample, when consumption is only available annually. For example, while annual consumption growth in 1929 was a healthy 3%, the sharp drop in the stock market in October 1929 implies a high likelihood that the economy transitioned to a disaster during this month. Note that other periods of sharp return drops that are not followed by a large contraction in economic activity, such as the crash of 1987 are not identified as disasters. Allowing for stochastic volatility lets the estimation neatly separate the identification of economic disasters from the variation in volatility of stock market returns that is unrelated to macroeconomic developments.

5.2.2 Disaster Predictability

An important prediction of a rational time-varying disaster risk theory is that disaster concerns are abnormally high before disasters. This prediction does not say economic disasters are fully predictable, but rather that in a long enough sample, disasters occur more often when disaster concerns are high. We test whether our proxy for the disaster predictability, NVIX, predicts disasters using a simple linear probability model.

We construct a disaster transition probability

\[
I_{t \rightarrow t+\tau}^{N \rightarrow D} = \bigcup_{s=1}^{\tau} Pr(s_{t+s} = 1 | s_t = 0, y^T),
\]

which is the probability that the economy transitions into a disaster between \( t \) and \( t + \tau \).\(^{13}\) The fact that this measure is continuous, allows us to improve inference as it relies not only on clear transitions into a disaster regime, but also on the less sharp cases, such as the 2007–2009 period.

Our main specification, tests if NVIX predicts a disaster transition as proxied by \( I_{t \rightarrow t+\tau}^{N \rightarrow D} \), controlling for the contemporaneous disaster state \( s_t = I_t^D > 0.5 \) and the interaction of the current disaster state and NVIX. Intuitively, the interaction controls for the mechanical effect that NVIX cannot predict a transition into disaster when the economy is already in a persistent disaster state. We control for expected stock market variance and its interaction with the current disaster state, using the same variance models of Section 3.2.

Table 11 reports the coefficient on NVIX for different horizons and subject to alternative controls.

\(^{13}\)This expression can be written as

\[
I_{t \rightarrow t+\tau}^{N \rightarrow D} = Pr(s_{t+1} = 1 | s_t = 0, y^T) + \prod_{j=1}^{\tau-1}(1 - Pr(s_{t+j} = 1 | s_{t+j-1} = 0, y^T))Pr(s_{t+j+1} = 1 | s_{t+j} = 0, y^T) = I_{t+1}^{N \rightarrow D} + \prod_{j=1}^{\tau-1}(1 - I_{t+j}^{N \rightarrow D})I_{t+j+1}^{N \rightarrow D}.
\]
for expected stock market risk. As in the return predictability regressions, we run the regression in
variance space to be consistent with theory (e.g. Gabaix, 2012). We find that, in the full sample,
NVIX is high just before disaster transitions. When the filtered disaster probability is zero and
$NVIX^2$ is one standard deviation above its mean, the probability of a disaster within the next
twelve months increases by 2.7% (column 3 times the NVIX standard deviation). Columns 4 to
8 use various models to control for expected stock market variance. Coefficients and statistical
significance are stable across specifications. Column 7 (model 4) is of special interest as it uses
NVIX in the variance forecasting model. Similar to what we found for returns in Table 4, we find
that the disaster forecasting ability of NVIX is orthogonal to its ability to forecast variance. Only
when we add $NVIX^2$ and credit spreads (model 5), the coefficients become less precisely estimated
and lose their statistical significance. The coefficients themselves do not change.

These results show that disaster risk is quite different from volatility risk. Even though disasters
are periods of elevated volatility, financial volatility has little forecasting power about transitions
into an economic disaster. This feature is especially evident at medium term horizons, 3 to 24
months, where volatility forecasts barely impact the regression R-squared. Only when we add
as control another disaster sensitive measure, the credit spread, do we experience an increase in
R-squared of similar magnitude (column 8).

Figure 6 illustrates this predictability result by showing the average behavior of NVIX and a
realized variance-based measure of VIX around disasters. We see that up to 12 months before a
disaster NVIX is consistently above its long-run mean, while the variance-based measure remains
close to its long-run mean. After a disaster, the variance measure mechanically spikes up as
expected, and as time passes differences in their behavior disappears.

Overall these results reinforce the hypothesis that NVIX captures concerns about disaster risk.
These results also tell us that these concerns are rational in the weak sense that disaster concerns are
associated with future transitions into a disaster regime. The magnitude of disaster risk variation
is also reasonable; our estimates imply that the probability of the economy transitioning into a
disaster within a year has a standard deviation of 2.8% (Table 11, average across specifications,
$\sigma(E[N_{t+1} \rightarrow D_{t+12} | NVIX^2_{t-1}]) = \beta_1 \times \sigma(NVIX^2) = 0.13 \times 21.66 = 2.8\%$). Since in our sample the annual
unconditional probability of transition into a disaster is 3.78%, our estimates imply that that the
annual probability of a disaster is under 9.5% more than 95% of the time. The probability of being
in a disaster state is substantially higher, because some disasters are quite persistent.

5.2.3 Return Predictability

If the probability of a disaster per period is low enough, a time-varying rare-disaster model predicts that realized excess returns can be written in terms of the expected probability of a disaster event and the actual disaster realizations as follows,

\[ r_{t \rightarrow t+\tau}^e = \beta_0 + \beta_1 E_t[I_{t \rightarrow t+\tau}^N] + \left( \beta_2 + \epsilon_{t \rightarrow t+\tau}^D \right) I_{t \rightarrow t+\tau}^D + \epsilon_{t \rightarrow t+\tau}^N, \]

(15)

where \( \beta_2 = E_t[r_{t \rightarrow t+\tau}^e | I_{t \rightarrow t+\tau}^N = 1] \) is the expected excess return conditional on a disaster event, a large negative number. In models like Gabaix (2012), \( \beta_1 \) is the expected disaster loss under the risk-neutral measure. If \( M_{t,t+\tau} \) is the stochastic discount factor that prices cash-flows between \( t \) and \( t+1 \), \( \beta_1 = -E_t[M_{t,t+\tau} (r_{t \rightarrow t+\tau}^e) | I_{t \rightarrow t+\tau}^N = 1] \). In richer models, such as Wachter (2013) and Gourio (2012), \( \beta_1 \) also includes the risk premia associated with disaster probability risk. If there is no risk premia associated with disaster or disaster-probability risks, then \( \beta_1 = -\beta_2 \). In general we expect \( \beta_1 > -\beta_2 \) if investors require a premium to be exposed to disaster risk.

In samples without disasters, a univariate regression of excess returns on the disaster probability recovers a consistent estimate of \( \beta_{1,\tau} \), and that is how we interpret our post-war results. When there are disaster realizations in the sample, however, the recovered estimates depend on the number of disasters in the sample, which renders the coefficient estimates not directly interpretable.

A regression of realized returns on the disaster probability that excludes disaster realizations allows us to recover consistent estimates of \( \beta_{1,\tau} \) as long \( E\left[I_{t \rightarrow t+\tau}^N \epsilon_{t \rightarrow t+\tau}^N \right] = 0 \). This condition is satisfied in the time-varying rare disaster model, but it is not satisfied under the plausible alternative model that there are no disasters, but stock market variance is variable and predictable. We pursue here the strategy of excluding disasters and delay a full discussion and adjustment for this potential bias to the last two paragraphs of this section.

Formally, we follow Schularick and Taylor (2009) and Krishnamurthy and Vissing-Jorgensen (2012) and construct \( I_{t \rightarrow t+\tau}^R = \left( I_{t \rightarrow t+\tau}^N > 0.5 \right) \), which is an indicator variable that turns on whenever the probability of a normal-times to disaster transition in the forecasting window is higher than 50%. We exclude from our estimated sample any month for which \( I_{t \rightarrow t+\tau}^R \) indicates a transition into
a disaster within the forecasting window. This procedure removes a very small number of months; for one-month (twelve-month) ahead forecast it excludes 3 (46) observations.

We add controls for the contemporaneous disaster state \( s_t = I_t^D > 0.5 \) and interactions of the state with NVIX. This specification mirrors the disaster predictability specification of Table 11, and is prescribed by models that feature persistent disaster states (e.g. Gourio, 2012). In such model, variation in option implied volatility reflects variation is disaster probabilities during normal times, while during disasters it reflects only variation in stock market variance.

Table 12 reports the normal times predictability coefficient of news-implied variance, its t-statistic, and the R-squared for alternative horizons and controls. Column 1 presents full sample estimates. Consistent with Figure 4, the positive relationship between NVIX and future returns is statistically weak and only statistically significant at the 12 month horizon. Column 2 excludes periods where the forecasting window has a disaster transition (\( I_{t-\tau+1}^R = 1 \)). Consistent with the results for the post-war sample, the predictability coefficient \( \beta_1 \) is positive and statistically significant at the 6, 12, and 24 months.

Columns 3 to 8 add controls for alternative measures of expected variance. Results are robust across all specifications. Note that in the columns where we use the credit spread in the variance model, the sample starts in 1919, and as a result the coefficients increase quite a bit and become better measured. The most conservative specification is in column 6, which includes NVIX in the variance model, and controls for any variance forecasting ability NVIX might have.

Overall these results reinforce the time-varying rare disaster risk interpretation of our findings. We document in the full sample a strikingly similar relation between NVIX and future returns to the one we found in the post-war sample. Consistent with this interpretation, the relationship between NVIX and future returns is mostly present during normal times. During normal times, time-variation in disaster concerns implies a large amount of risk premia variation in frequencies from 6 months up to two years.

Quantitatively the coefficients are in line with the post-war results, with a one standard deviation increase in \( NVIX^2 \) leading to \( \sigma_{NVIX^2} \times \beta_1 \in [3.2\%, 4.8\%] \) higher excess returns in the following year depending on the model we use to control for stock market risk. This compares with \( \sigma_{NVIX^2} \times \beta_1 \in [3.4\%, 4\%] \) in the post-war sample.

The strategy of removing disasters of the estimation can lead to biases if our predictor forecasts
stock market variance. Because the timing of disaster realizations are greatly influenced by the realization of abnormally low returns, it is entirely plausible that the true model features time-varying volatility (at least in the pre-war sample), but not time-varying disaster risk. Under this alternative model, our procedure would be classifying as disasters, periods of high variance that turn out to have low return realizations, and a variable that predicts variance (but not returns), could show up as predicting returns if we exclude “disaster periods” from the regressions.

We develop fully this discussion in the appendix, but according to this truncation mechanism it is enough to control for the forecast of the Mills ratio \( \sigma_{t+1} \lambda(\frac{\bar{r} - \mu_r}{\sigma_{t+1}}) \), the truncated mean of the returns distribution.\(^\text{14}\) If return predictability excluding “disasters” is only a result of time-varying truncation, a predictor of the Mills ratio should completely drive out NVIX. To be consistent with the specification we used last section when forecasting disasters, our specification controls for the predicted Mills ratio using several alternative variables. Table 13 in the Appendix presents the results. Neither the coefficients or their statistical significance are impacted by including the relevant Mills ratio forecast. As before, the most conservative specification is model (4), which includes NVIX, price to earning ratio and three lags of realized variance. Including credit spreads make the results stronger as before, but only because it restricts the sample (the credit spread variable is only available after 1919) and places more weight on the the post war sample when return predictability results are stronger.

5.2.4 A Quantitative Check

Time-varying rare disaster risk models were developed as a candidate explanation for the excess volatility puzzle. Their quantitative success as an explanation for the puzzle hinges on the pattern of time-variation in disaster risk, and the sensitivity of excess returns to disaster probability risk. Structural models such as Gourio (2012) and Wachter (2013) use cross country studies such as Barro and Ursua (2008) to determine how bad disasters are in terms of drops in consumption and losses in the financial claims of interest. Together with assumptions about investors preferences, this disciplines the model implied relationship between time-variation disaster probability and risk premia. However, in the absence on direct measurements of the disaster probability process, the modeler is free to pick a disaster probability process that fits the pattern of return predictability.

\(^{14}\)Where \( \lambda(z) = E[x|x \geq z] = \int_z^{\infty} x \phi(x) dx \)
and excess volatility observed in the data. Our disaster probability estimates can be useful to this literature by providing discipline on the statistical features of the disaster probability process can be considered plausible given the data. Our return probability estimates are also useful in providing a check if the cross country data of disaster sizes extrapolates well to the US. In particular, by analyzing the sensitivity of excess returns to disaster probability shocks we can evaluate if the disaster concerns that we measure is related to events of the same magnitude as the ones implied by the cross country data.

We use Wachter (2013) as a comparison because her study provides direct counterparts to the quantities we are interested in. In her model, the disaster probability has an unconditional distribution where the disaster probability spends 95% of the time in values lower than 10%. This lines up surprisingly well with our estimates, where the disaster probability spends 95% of time in values below 9.5% (see Section 5.2.2). So in terms of overall variation in the disaster probability, Wachter (2013) calibration is in line with our estimates. However, in her calibration this disaster probability distribution is achieved by considering a highly persistent disaster probability process. In her calibration, the (annualized) disaster probability has a persistence of 0.9934 at the monthly frequency, and a standard deviation of 0.36%. Our estimates indicate a substantially different pattern. We estimate a 1.55% standard deviation and 0.8 persistence at the same frequency. At the yearly frequency, her model implies a volatility of disaster probability of 1.21% and persistence of 0.92, while our estimates indicate a standard deviation of 2.26% and persistence of 0.5. The lower persistence of the disaster predictability detected in the data implies it can explain much less of the very long-run movements observed in risk premia and price-dividend ratios.

Her setup produces a sensitivity of risk-premium to disaster 1.8 in the baseline case which features recursive utility with coefficient of relative risk aversion of 3 and inter-temporal elasticity of substitution of 1. For example, if the instantaneous disaster probability increases by one percentage point, the instantaneous risk premium increases by 1.8 percentage points. In her Constant Relative Risk Aversion utility specification, the sensitivity is slightly above 1, with a 1 pp increase in the probability of a disaster mapping into a 1 pp higher risk premia, so the recursive utility kernel amplifies the risk premia effect substantially. Basically agents become averse not only to disaster shocks, but also disaster probability shocks.

\[15\text{This quantity can be computed directly from Wachter (2013), Figure 3}\]
We estimate that a one-standard deviation movement in NVIX increases the probability of a disaster by 2.8% and expected returns by 3.2%, implying a risk premia sensitivity to disaster of 1.14, very close to the Wachter (2013) CRRA specification. This indicates that the type of disaster risk that our measure is capturing is related to events of similar severity as the ones implied by the cross country data.

Our estimates suggest that the disaster probability process and the risk premia variation it induces are consistent with a state of the art calibration of the rare disaster risk model. While our estimates and the Wachter (2013) calibration agree on the overall unconditional distribution of disaster risk shocks, our estimates point to shocks (to the disaster probability) that are substantially larger, but much less persistent than assumed in her calibration. Overall disaster concerns can produce large, but relatively short-lived spikes in risk premia.

6 Conclusion

We use a text-based method to extend options-implied measures of uncertainty back to the end of the 19th century. We find that our news-based measure of implied volatility, NVIX, predicts returns at frequencies from 6 up to 24 months. Four pieces of evidence suggest that these return predictability results are driven by variation in investors’ concerns regarding rare disasters. First, we find that the predictive power of NVIX is orthogonal to contemporaneous or forward-looking measures of stock market volatility. Second, we use alternative options-based measures, which are more focused on left tail risk, to estimate their news-based counterparts and find similar return predictability results. Third, using content analysis we trace a large part of the variation in risk premia to concerns about wars and government policy, which are tightly related to the type of events discussed in the rare disasters literature. Lastly, we show that our measure predicts disasters, even after controlling for stock market volatility. Importantly, the amounts of predictability detected in stock returns and disasters are quantitatively consistent with disasters of the same magnitude as documented by Barro and Ursua (2008) using cross sectional data.


A Appendix

A.1 Truncation

A concern that we have regarding our approach to identify disasters is that the predictability results in Table 12 might be a mechanical artifact of truncating the left tail of the distribution during periods of higher volatility. If NVIX is a proxy for future stock-market volatility, periods of higher volatility where no disaster was observed will also be periods of artificially higher than average returns. This will be a mechanical artifact of the exclusion of very low return months from the regression analysis. Correcting for this mechanism is fairly straightforward. Consider the following model featuring time-varying volatility and constant expected returns,

\[
\sigma_{t+1}^2 = \mu_\sigma + \rho_\sigma \sigma_t^2 + \omega \sqrt{\sigma_t^2} H_\sigma w_{t+1}
\]

\[
r_{t+1} = \mu_r + \sigma_{t+1} H_r w_{t+1}
\]

In this counter-factual economy there is no predictability, and there is no sense that very low returns are special. But suppose in this environment we use threshold $r$ to split the sample in disaster periods and normal times. In this case we would have average returns in normal-periods periods given by:

\[
E[r_{t+1} | r_{t+1} \geq r, \sigma_{t+1}] = \mu_r + \sigma_{t+1} E[w_{r,t+1} | w_{r,t+1} \geq \frac{r - \mu_r}{\sigma_{t+1}}] = \mu_r + \sigma_{t+1} \lambda \left( \frac{r - \mu_r}{\sigma_{t+1}} \right),
\]

where $\lambda(x)$ is known as the Mills ratio. In the context of our exercise we know exactly the threshold $r$. If $NVIX_t$ predicts future volatility the truncation effect will lead us to find that $NVIX$ predicts returns when in fact it does not. In this case, conditional expectations are given by:

\[
E[r_{t+1} | r_{t+1} \geq r | NVIX_t^2, \sigma_t^2] = \mu_r + E[\sigma_{t+1} \lambda \left( \frac{r - \mu_r}{\sigma_{t+1}} \right) | NVIX_t^2, \sigma_t]
\]

The above expression tells us that in order to test the time-varying rare disaster story against the truncation story it suffices to control for the best predictor of the quantity $\sigma_{t+1} \lambda \left( \frac{r - \mu_r}{\sigma_{t+1}} \right)$. Note that under the null NVIX is allowed to predict “disasters”, between quotes because there are no disasters under the null, only periods that are classified as disasters. The essence of this test is
the restriction imposed by the truncation hypothesis that any return predictability has to happen through the prediction of the Mills ratio multiplied by the volatility. That is, we first estimate,

$$\sigma_{t+1} \lambda \left( \frac{r_{t+1} - \mu_{r}}{\sigma_{t+1}} \right) = \Gamma X_t + \epsilon_t,$$

where $X_t$ is a set of predictors (NVIX inclusive) and the constant, we then run

$$r_{t+1} = \beta_0 + \beta_1 NVIX^2 + \beta_2 \hat{\Gamma} X_t.$$

Under the null that all predictability is driven by truncation, we have $\beta_1 = 0$ and $\beta_2 = 1$.

For multi-period return forecasts, a observation is excluded as long there is at least one disaster transition in the forecasting window. To derive the truncation bias formally write multi-period expected returns as,

$$E\left[ \sum_{i=1}^{\tau} r_{t+i} \right] | \{ r_{t+z} \geq \tau | \{ 1 \leq z \leq \tau \}, X_t \} = \frac{1}{\tau} \sum_{i=1}^{\tau} E[ E[r_{t+i} | r_{t+i} \geq \tau] | X_t] = \frac{1}{\tau} \sum_{i=1}^{\tau} E[ \sigma_{t+i} \lambda \left( \frac{r_{t+i} - \mu_{r}}{\sigma_{t+i}} \right) | X_t].$$

We implement this by constructing multi-period forecasts of the Mills ratio,

$$\frac{1}{\tau} \sum_{i=1}^{\tau} \sigma_{t+i} \lambda \left( \frac{r_{t+i} - \mu_{r}}{\sigma_{t+i}} \right) = \Gamma_{\tau} X_t + \epsilon_{t+\tau},$$

and using $EMILLS_{t-1, \tau} = \hat{\Gamma}_{\tau} X_{t-1}$ as a control variable.

A.2 Alternative Text-based Analysis Approaches

We estimate the relationship between news and volatility, disaster concerns and returns in our dataset using support vector regression (1). SVR overcomes the main challenge, which is the large dimensionality of the feature space (number of unique n-grams). Our approach lets the data speak without much human interaction. Two alternative approaches have been suggested by previous literature.

The first approach, creates a topic-specific compound full-text search statement and counts the resulting number of articles normalized by a measure of normal word count. The result is a univariate time-series that can be used in a least squares regression. An advantage of this approach
is that resulting articles are highly likely to be related to the specific topic, resulting in a fine-grained index that is easily interpretable. However, it requires a very large body of text every period and ignores many other articles that also relate to the same topic. Furthermore, unlike in our approach which relies on an objective measure of success (VIX) in constructing the uncertainty index, this approach relies on the econometrician’s judgment. Since out-of-sample fit is paramount in our paper, we find the text regression superior for our purposes.

A leading example of this approach is the news-based economic policy uncertainty index suggested in Baker, Bloom, and Davis (2013). It searches for articles containing the term 'uncertainty' or 'uncertain', the terms 'economic' or 'economy' and one or more policy terms such as 'policy', 'tax', etc. Our attempt to apply the Baker, Bloom, and Davis (2013) methodology to our dataset, classified as discussing economic policy uncertainty only 47 out of 320000 articles, or 43 out of 1439 months. We found no return predictability using this index.

A second approach, classifies words into dictionaries or word lists that share a common tone. One then counts all occurrences of words in the text belonging to a particular word list, again normalized by a measure of normal word count. An advantage of this approach is that it reduces the feature space from the number of n-grams to the number of word lists. One disadvantage is that all words within a word list are equally-weighted. Thus the words ‘war’ and ‘yawn’ would count the same, even though the importance of their appearance on the front page of a newspaper is quite different.

A recent contribution by Loughran and McDonald (2011) develops a negative word list, along with five other word lists, that reflect tone in financial text better than the widely used Harvard Dictionary and relate them to 10-K filing returns. We applied the Loughran and McDonald (2011) methodology to our sample of articles. We tried both tf (proportional weights) and tf.idf weights of words appearing in their Negative, Positive, Uncertainty, Modal Strong, and Modal Weak word lists. Unlike NVIX, the ten time-series do not appear to capture important historical events. We then run return predictability regressions on the scores of each word list separately and together with NVIX. The intermediate step of regressing VIX on the scores is unnecessary here because the predicted value of VIX would just be a constant multiplying the raw word list score. Most of

---

16 Examples of this approach can be found in Antweiler and Frank (2004), Tetlock (2007), Engelberg (2008), and Tetlock, Saar-Tsechansky, and Macskassy (2008).
the lists have no predictive power. Only Uncertainty and Modal Weak using proportional weights
are significant but do not drive out NVIX. We therefore conclude that support vector regression is
better suited to our purposes given our data.

Table 14 repeats our analysis but this time includes also the tone scores as a second independent
variable in addition to NVIX. Both tables show that NVIX remains a significant return predictor
throughout.

A.3 Relationship Between Alternative Measures of Tail Risk

Panel A in Table 15 reports the raw correlations between the option constructed measures of tail
risk. Panel B in Table 15 reports the raw correlations between their news based counterparts which
we use as predictors in Table 6.
References


Drechsler, I., 2008, Uncertainty, time-varying fear, and asset prices, in *AFA 2010 Atlanta Meetings Paper*.


Gao, George P., and Zhaogang Song, 2013, Rare disaster concerns everywhere, Working paper.


Moreira, Alan, and Alexi Savov, 2013, The macroeconomics of shadow banking, *Available at SSRN 2310361*.


Figure 1: News-Implied Volatility 1890–2009

Solid line is end-of-month CBOE volatility implied by options $VIX_t$. Dots are news implied volatility (NVIX) $\hat{\omega}_t + \nabla VIX = w_0 + VIX + w \cdot x_t$. The *train* subsample, 1996 to 2009, is used to estimate the dependency between news data and implied volatility. The *test* subsample, 1986 to 1995, is used for out-of-sample tests of model fit. The *predict* subsample includes all earlier observations for which options data, and hence VIX is not available. Light-colored triangles indicate a nonparametric bootstrap 95% confidence interval around $\hat{\omega}$ using 1000 randomizations. These show the sensitivity of the predicted values to randomizations of the *train* subsample.
We describe NVIX peak months each decade by reading the front page articles of *The Wall Street Journal* and cross-referencing with secondary sources when needed. Many of the market crashes are described in Mishkin and White (2002). See also Noyes (1909) and Shiller and Feltus (1989).
Figure 3: News Implied Volatility due to Different Word Categories

(a) Stock Markets

(b) War

(c) Government

(d) Intermediation

In all panels dots are monthly NVIX due only to category $C$-related words $\tilde{v}_t(C) = x_t \cdot w(C)$. Panel (a): Solid line is annualized realized stock market volatility. Shaded regions indicate stock market crashes identified by Reinhart and Rogoff (2011). Panel (b): Shaded regions are US wars, specifically the American-Spanish, WWI, WWII, Korea, Vietnam, Gulf, Afghanistan, and Iraq wars. Panel (c): Solid line is the annual average marginal tax rate on dividends from Sialm (2009). Panel (d): Solid line is percent of total insured deposits held by US banks that failed each month, from the FDIC starting April 1934. Shaded regions indicate banking crises identified by Reinhart and Rogoff (2011).
Rolling window coefficient $\beta_1^R$ estimates for the regression $r_{t+\tau} = \beta_0 + \beta_1^R NVIX_{t-1}^2 + \epsilon_t$ at the $\tau = 12$ months horizon. Years on the x-axis represent the start of the estimation window. All windows run until 2009, the end of our sample. Shaded region represents 95% confidence interval.
Figure 5: Filtered Probability that the Economy is in a Disaster State, $I_t^D = \text{Prob}(s_t = 1 | y^T)$
Gray dots are a realized variance-based forecast of VIX (Model 6 in Table 11). Black dots are the component of $NVIX^2$ orthogonal to the variance-based forecast. Both measures are demeaned and standardized using their sample standard deviation. Reported are averages across disaster transitions, where a disaster transition is defined as a month $t$ where the disaster transition probability is higher than 0.1 ($I_{t-1}^{N \rightarrow D_t} > 0.1$).
Table 1: Out-of-Sample VIX Prediction

<table>
<thead>
<tr>
<th>Panel (a) Out-of-Sample Fit</th>
<th>Panel (b) Out-of-Sample OLS Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( v_t = a + b\hat{v}_t + e_t, \quad t \in \text{test} )</td>
</tr>
<tr>
<td>( R^2 (\text{test}) = \frac{\text{Var}(\hat{v}_t)}{\text{Var}(v_t)} )</td>
<td>0.34</td>
</tr>
<tr>
<td>( \text{RMSE (test)} = \sqrt{\frac{1}{T_{\text{test}}} \sum_{t \in \text{test}} (v_t - \hat{v}_t)^2} )</td>
<td>7.52</td>
</tr>
<tr>
<td>( T_{\text{test}} )</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>( a \quad -3.55^{***} \ (0.51) )</td>
</tr>
<tr>
<td></td>
<td>( b \quad 0.75^{***} \ (0.19) )</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Reported are out-of-sample model fit statistics using the \( \text{test} \) subsample. Panel (a) reports variance of the predicted value (NVIX) as a fraction of actual VIX variance, and the root mean squared error. Panel (b) reports a univariate OLS regression of actual VIX on NVIX. In parenthesis are robust standard errors. *** indicates 1% significance.
Table 2: Out-of-Sample Realized Volatility Prediction Using News

<table>
<thead>
<tr>
<th>Subsample</th>
<th>RMSE SVR</th>
<th>$R^2$ SVR</th>
<th>RMSE Reg</th>
<th>$R^2$ Reg</th>
<th>Correlation</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>train</td>
<td>3.35</td>
<td>0.68</td>
<td>2.64</td>
<td>0.93</td>
<td>0.96</td>
<td>168</td>
</tr>
<tr>
<td>test</td>
<td>9.60</td>
<td>0.27</td>
<td>9.09</td>
<td>0.20</td>
<td>0.45</td>
<td>119</td>
</tr>
<tr>
<td>predict</td>
<td>10.91</td>
<td>0.38</td>
<td>8.49</td>
<td>0.16</td>
<td>0.40</td>
<td>1150</td>
</tr>
</tbody>
</table>

Reported are model fit statistics repeating the estimation procedure over the same *train* subsample as before, only replacing VIX with realized volatility as the dependent variable of the SVR (1). The *train* subsample, 1996 to 2009, is used to estimate the dependency between monthly news data and realized volatility. The *test* subsample, 1986 to 1995, is used for out-of-sample tests of model fit. The *predict* subsample includes all earlier observations for which options data, and hence VIX is not available. RMSE SVR is root mean squared error of the SVR. $R^2$ SVR is the variance of the predicted value as a fraction of actual realized volatility’s variance. RMSE Reg and $R^2$ Reg are the variance of the predicted value as a fraction of actual realized volatility’s variance, and the root mean squared error from a subsequent univariate OLS regression of actual realized volatility on realized volatility implied by news.
Table 3: NVIX Predicts Post-war Stock Market Returns

\[ r_{t+\tau}^r = \beta_0 + \beta_1 X_{t-1}^2 + \epsilon_{t+\tau} \]

<table>
<thead>
<tr>
<th>X:</th>
<th>NVIX</th>
<th>VXO</th>
<th>VIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau )</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>1</td>
<td>( \beta_1 )</td>
<td>0.15</td>
<td>0.33**</td>
</tr>
<tr>
<td>( t(\beta_1) )</td>
<td>1.04</td>
<td>2.21</td>
<td>0.58</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.37</td>
<td>0.74</td>
<td>0.28</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_1 )</td>
<td>0.12</td>
<td>0.41***</td>
</tr>
<tr>
<td>( t(\beta_1) )</td>
<td>0.87</td>
<td>3.60</td>
<td>0.27</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.64</td>
<td>2.98</td>
<td>0.13</td>
</tr>
<tr>
<td>6</td>
<td>( \beta_1 )</td>
<td>0.18***</td>
<td>0.39***</td>
</tr>
<tr>
<td>( t(\beta_1) )</td>
<td>2.59</td>
<td>3.72</td>
<td>1.44</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>2.56</td>
<td>4.91</td>
<td>1.93</td>
</tr>
<tr>
<td>12</td>
<td>( \beta_1 )</td>
<td>0.16***</td>
<td>0.28***</td>
</tr>
<tr>
<td>( t(\beta_1) )</td>
<td>3.27</td>
<td>2.79</td>
<td>1.64</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>3.50</td>
<td>4.78</td>
<td>2.99</td>
</tr>
<tr>
<td>24</td>
<td>( \beta_1 )</td>
<td>0.14***</td>
<td>0.19**</td>
</tr>
<tr>
<td>( t(\beta_1) )</td>
<td>3.55</td>
<td>2.17</td>
<td>2.13</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>5.12</td>
<td>4.26</td>
<td>6.13</td>
</tr>
</tbody>
</table>

Reported are monthly return predictability regressions based on news implied volatility (NVIX), S&P 100 options implied volatility (VXO), and S&P 500 options implied volatility (VIX). The dependent variables are annualized log excess returns on the market index. The first column examines the entire post-war period, while the second focuses on a sample that was not used in fitting NVIX to options implied volatility. The third and fourth columns report results for the sample period for which VXO and VIX are available. t-statistics are Newey-West corrected with number of lags/leads equal to the return forecasting horizon.
Table 4: Stochastic Volatility Does Not Explain the Return Predictability Results

\[ r_{t\rightarrow t+\tau}^e = \beta_0 + \beta_1 NVIX_{t-1}^2 + \beta_2 EVAR_{t-1} + \epsilon_t \]

| \( \tau \) | \( \beta_1 \) | \( t(\beta_1) \) | \( R^2 \) | \( \beta_1 \) | \( t(\beta_1) \) | \( R^2 \) | \( \beta_1 \) | \( t(\beta_1) \) | \( R^2 \) | \( \beta_1 \) | \( t(\beta_1) \) | \( R^2 \) | \( \beta_1 \) | \( t(\beta_1) \) | \( R^2 \) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.21 | [1.59] | 0.55 | 0.14 | [1.05] | 0.70 | 0.19** | [2.51] | 2.57 | 0.17*** | [3.15] | 3.56 | 0.15*** | [3.32] | 5.18 | 9.21 |
| 3 | 0.23 | [1.6] | 0.46 | 0.16 | [1.34] | 0.75 | 0.22*** | [2.64] | 2.75 | 0.21*** | [2.77] | 3.75 | 0.17*** | [2.79] | 5.51 | 25.53 |
| 6 | 0.21 | [1.64] | 0.52 | 0.18 | [1.51] | 0.89 | 0.24*** | [2.91] | 3.01 | 0.20*** | [2.98] | 4.19 | 0.21*** | [2.8] | 6.27 | 25.87 |
| 12 | 0.23*** | [1.62] | 0.49 | 0.17 | [1.53] | 0.85 | 0.23*** | [2.93] | 2.87 | 0.20*** | [2.92] | 4.14 | 0.27** | [2.39] | 7.35 | 28.22 |
| 24 | 0.26** | [1.67] | 0.48 | 0.17 | [1.67] | 0.89 | 0.27** | [2.44] | 2.94 | 0.26** | [2.39] | 4.36 | 0.30*** | [2.67] | 8.67 | 31.83 |

EVAR Model \( R^2 \) \( 9.21 \) \( 25.53 \) \( 25.87 \) \( 28.22 \) \( 31.83 \)

Return predictability regressions controlling for expected variance, where the dependent variables are market annualized log excess returns, over the post-war 1945–2009 period. Each row and each column represents a different regression. Rows show different forecasting horizons. EVAR is predicted variance using the following variables: model 1 uses lagged lagged variance, model 2 uses an three lags of realized variance, model 3 adds price to earning ratio to model 2, model 4 adds \( NVIX^2 \) to model 3, and model 5 adds credit spread to model 4. We measure stock market realized variance using daily returns on the Dow Jones index within the relevant month. The last row reports the percent R-squared from the variance predictability regression used to estimate EVAR. t-statistics are Newey-West corrected with number of lags/leads equal to the return forecasting horizon.
Table 5: Horse Races with Financial Predictors

\[ r_{t+\tau} = \beta_0 + \beta_1 \text{NVIX}_{t-1}^2 + \sum_{j=2}^{N} \beta_j X_{j,t-1} + \epsilon_{t+\tau} \]

<table>
<thead>
<tr>
<th>( \tau )</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \beta_1 )</td>
<td>0.15</td>
<td>0.20</td>
<td>0.21</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[1.04]</td>
<td>[1.45]</td>
<td>[1.43]</td>
<td>[1.32]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.37</td>
<td>0.45</td>
<td>0.51</td>
<td>0.85</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_1 )</td>
<td>0.12</td>
<td>0.16</td>
<td>0.16</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[0.87]</td>
<td>[1.29]</td>
<td>[1.29]</td>
<td>[1.15]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.64</td>
<td>0.72</td>
<td>0.99</td>
<td>1.92</td>
</tr>
<tr>
<td>6</td>
<td>( \beta_1 )</td>
<td>0.18***</td>
<td>0.22***</td>
<td>0.22***</td>
<td>0.21**</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[2.59]</td>
<td>[2.64]</td>
<td>[2.63]</td>
<td>[2.42]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>2.56</td>
<td>2.73</td>
<td>3.51</td>
<td>5.34</td>
</tr>
<tr>
<td>12</td>
<td>( \beta_1 )</td>
<td>0.16***</td>
<td>0.19***</td>
<td>0.19***</td>
<td>0.18***</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[3.27]</td>
<td>[2.78]</td>
<td>[2.79]</td>
<td>[2.62]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>3.50</td>
<td>3.72</td>
<td>4.47</td>
<td>8.80</td>
</tr>
<tr>
<td>24</td>
<td>( \beta_1 )</td>
<td>0.14***</td>
<td>0.17***</td>
<td>0.17***</td>
<td>0.15***</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[3.55]</td>
<td>[2.82]</td>
<td>[2.82]</td>
<td>[3.01]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>5.12</td>
<td>5.49</td>
<td>5.49</td>
<td>16.46</td>
</tr>
</tbody>
</table>

Controls

| \( \text{NVIX}_{t-1}^2 \) | yes | yes | yes | yes | no |
| \( E[VIX_{t-1}^2|VAR] \) | no | yes | yes | yes | yes |
| \( Creditspread_{t-1} \) | no | no | yes | yes | yes |
| \( (P/E)_{t-1} \) | no | no | no | yes | yes |

Reported are monthly return predictability regressions based on news implied volatility (NVIX) and controls. \( E[VIX_{t-1}^2|VAR] \) is the variance based VIX, the predicted value of \( VIX^2 \) using the contemporaneous variance plus two additional lags. The model is estimated in the sample where VIX is available (1990–2009). The dependent variables are annualized log excess returns on the market index. Each row and each column represents a different regression. Rows show different forecasting horizons. The sample is Jan/1945 to Dec/2009. t-statistics are Newey-West corrected with number of lags/leads equal to the return forecasting horizon.
Table 6: Alternative Measures of Uncertainty Focused on Tail Risk

\[ r_{t-t+\tau}^e = \beta_0 + \beta_1 \tilde{X}_{t-1} + \beta_2 EVAR_{t-1} + \epsilon_{t+\tau} \]

<table>
<thead>
<tr>
<th>( \tau )</th>
<th>( X : )</th>
<th>VIX(^2)</th>
<th>VIX premium</th>
<th>LT</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \beta_1 )</td>
<td>0.21</td>
<td>0.42***</td>
<td>1.39*</td>
<td>128.21*</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[1.47]</td>
<td>[2.62]</td>
<td>[1.82]</td>
<td>[1.93]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.46</td>
<td>1.34</td>
<td>0.43</td>
<td>0.53</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_1 )</td>
<td>0.16</td>
<td>0.18</td>
<td>1.54**</td>
<td>98.49*</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[1.34]</td>
<td>[1.46]</td>
<td>[2.53]</td>
<td>[1.83]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.75</td>
<td>0.70</td>
<td>1.37</td>
<td>0.85</td>
</tr>
<tr>
<td>6</td>
<td>( \beta_1 )</td>
<td>0.22***</td>
<td>0.18**</td>
<td>1.33**</td>
<td>80.13**</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[2.64]</td>
<td>[2.14]</td>
<td>[2.02]</td>
<td>[1.98]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>2.75</td>
<td>1.60</td>
<td>2.22</td>
<td>1.40</td>
</tr>
<tr>
<td>12</td>
<td>( \beta_1 )</td>
<td>0.19***</td>
<td>0.12*</td>
<td>1.26**</td>
<td>57.19*</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[2.77]</td>
<td>[1.87]</td>
<td>[2.45]</td>
<td>[1.73]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>3.75</td>
<td>1.67</td>
<td>3.51</td>
<td>1.53</td>
</tr>
<tr>
<td>24</td>
<td>( \beta_1 )</td>
<td>0.17***</td>
<td>0.11**</td>
<td>0.82*</td>
<td>54.65**</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[2.79]</td>
<td>[2.2]</td>
<td>[1.7]</td>
<td>[2.33]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>5.51</td>
<td>2.34</td>
<td>3.15</td>
<td>2.39</td>
</tr>
</tbody>
</table>

This table replicates our main results of Table 3 for alternative tail risk measures, over the post-war 1945–2009 period. For each of these measures we reproduce the methodology we applied to VIX. The symbol \( \tilde{X}_{t} \) denotes the text based estimator of variable \( X_{t} \). The first column reproduces our main results. In the second column we have VIX premium (\( = VIX^2_t - E_t[Var(R_{t+1})] \)), where \( E_t[Var(R_{t+1})] \) is constructed using an AR(1). In the third column we have the Left-Tail (LT) measure from Bollerslev and Todorov (2011). In the fourth column we have the slope of option implied volatility curve, constructed using the 30-day volatility curve from Option metrics. We use puts with delta of -0.5 and -0.8 to compute the slope. The variable \( EVAR_{t-1} \) controls for expected future variance using an AR(3) model (Model 2 of Table 4).
<table>
<thead>
<tr>
<th>n-gram</th>
<th>Variance Share, %</th>
<th>Weight, %</th>
<th>n-gram</th>
<th>Variance Share, %</th>
<th>Weight, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>stock</td>
<td>37.28</td>
<td>0.10</td>
<td>oil</td>
<td>1.39</td>
<td>-0.03</td>
</tr>
<tr>
<td>market</td>
<td>6.74</td>
<td>0.06</td>
<td>banks</td>
<td>1.36</td>
<td>0.06</td>
</tr>
<tr>
<td>stocks</td>
<td>6.53</td>
<td>0.08</td>
<td>financial</td>
<td>1.32</td>
<td>0.11</td>
</tr>
<tr>
<td>war</td>
<td>6.16</td>
<td>0.04</td>
<td>u.s</td>
<td>0.88</td>
<td>0.05</td>
</tr>
<tr>
<td>u.s</td>
<td>3.62</td>
<td>0.06</td>
<td>bonds</td>
<td>0.81</td>
<td>0.04</td>
</tr>
<tr>
<td>tax</td>
<td>2.01</td>
<td>0.04</td>
<td>stock</td>
<td>0.80</td>
<td>0.03</td>
</tr>
<tr>
<td>washington</td>
<td>1.78</td>
<td>0.02</td>
<td>house</td>
<td>0.77</td>
<td>0.05</td>
</tr>
<tr>
<td>gold</td>
<td>1.46</td>
<td>-0.04</td>
<td>billion</td>
<td>0.67</td>
<td>0.06</td>
</tr>
<tr>
<td>special</td>
<td>1.44</td>
<td>0.02</td>
<td>economic</td>
<td>0.64</td>
<td>0.05</td>
</tr>
<tr>
<td>treasury</td>
<td>1.43</td>
<td>0.06</td>
<td>like</td>
<td>0.59</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

We report the fraction of NVIX variance $h(i)$ that each n-gram drives over the predict subsample as defined in (3), and the regression coefficient $w_i$ from (1), for the top 20 n-grams.
Table 8: Categories Total Variance Share

<table>
<thead>
<tr>
<th>Category</th>
<th>Variance Share, %</th>
<th>n-grams</th>
<th>Top n-grams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government</td>
<td>2.59</td>
<td>83</td>
<td>tax, money, rates, government, plan</td>
</tr>
<tr>
<td>Intermediation</td>
<td>2.24</td>
<td>70</td>
<td>financial, business, bank, credit, loan</td>
</tr>
<tr>
<td>Natural Disaster</td>
<td>0.01</td>
<td>63</td>
<td>fire, storm, aids, happening, shock</td>
</tr>
<tr>
<td>Stock Markets</td>
<td>51.67</td>
<td>59</td>
<td>stock, market, stocks, industry, markets</td>
</tr>
<tr>
<td>War</td>
<td>6.22</td>
<td>46</td>
<td>war, military, action, world war, violence</td>
</tr>
<tr>
<td>Unclassified</td>
<td>37.30</td>
<td>373988</td>
<td>u.s, washington, gold, special, treasury</td>
</tr>
</tbody>
</table>

We report the percentage of NVIX variance \(= \sum_{i \in C} h(i)\) that each n-gram category \(C\) drives over the predict subsample.
Table 9: Risk Premia Decomposition

\[ r_{t \rightarrow t+\tau} = \beta_0 + \sum_{j=1}^{N} \beta_j X_{t-1}^j + \epsilon_{t+\tau} \]

<table>
<thead>
<tr>
<th>Horizon ( \tau ):</th>
<th>6</th>
<th>12</th>
<th>6</th>
<th>12</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government</td>
<td>4.34***</td>
<td>4.22***</td>
<td>-0.45</td>
<td>-0.57</td>
<td>2.65**</td>
<td>2.54**</td>
</tr>
<tr>
<td></td>
<td>[3.18]</td>
<td>[2.90]</td>
<td>[0.17]</td>
<td>[0.26]</td>
<td>[2.29]</td>
<td>[2.12]</td>
</tr>
<tr>
<td></td>
<td>(44.89)</td>
<td>(57.18)</td>
<td>(0.38)</td>
<td>(0.57)</td>
<td>(23.7)</td>
<td>(23.19)</td>
</tr>
<tr>
<td>War</td>
<td>4.78***</td>
<td>3.03**</td>
<td>3.09*</td>
<td>3.76***</td>
<td>3.77***</td>
<td>3.63***</td>
</tr>
<tr>
<td></td>
<td>[3.26]</td>
<td>[2.32]</td>
<td>[1.95]</td>
<td>[2.65]</td>
<td>[3.71]</td>
<td>[4.37]</td>
</tr>
<tr>
<td></td>
<td>(25.02)</td>
<td>(3.54)</td>
<td>(45.1)</td>
<td>(59.99)</td>
<td>(47.68)</td>
<td>(47.45)</td>
</tr>
<tr>
<td>Intermediation</td>
<td>0.31</td>
<td>0.70</td>
<td>0.24</td>
<td>1.19</td>
<td>0.98</td>
<td>1.38</td>
</tr>
<tr>
<td></td>
<td>[0.12]</td>
<td>[0.40]</td>
<td>[0.09]</td>
<td>[0.52]</td>
<td>[0.57]</td>
<td>[0.97]</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(1.49)</td>
<td>(0.13)</td>
<td>(3.09)</td>
<td>(3.2)</td>
<td>(6.8)</td>
</tr>
<tr>
<td>Stock Markets</td>
<td>0.97</td>
<td>-0.73</td>
<td>-2.91</td>
<td>-2.78</td>
<td>-0.87</td>
<td>-1.07</td>
</tr>
<tr>
<td></td>
<td>[0.26]</td>
<td>[0.24]</td>
<td>[1.35]</td>
<td>[1.09]</td>
<td>[0.54]</td>
<td>[0.58]</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.16)</td>
<td>(28.42)</td>
<td>(23.44)</td>
<td>(2.54)</td>
<td>(4.09)</td>
</tr>
<tr>
<td>Natural Disaster</td>
<td>0.99</td>
<td>1.08*</td>
<td>-2.16</td>
<td>-0.28</td>
<td>0.69</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>[1.12]</td>
<td>[1.70]</td>
<td>[1.09]</td>
<td>[0.15]</td>
<td>[0.8]</td>
<td>[1.54]</td>
</tr>
<tr>
<td></td>
<td>(3.68)</td>
<td>(5.88)</td>
<td>(3.61)</td>
<td>(0.05)</td>
<td>(1.61)</td>
<td>(3.87)</td>
</tr>
<tr>
<td>Residual</td>
<td>2.39**</td>
<td>2.13*</td>
<td>0.47</td>
<td>0.59</td>
<td>1.88*</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>[2.06]</td>
<td>[1.87]</td>
<td>[0.14]</td>
<td>[0.19]</td>
<td>[1.70]</td>
<td>[1.38]</td>
</tr>
<tr>
<td></td>
<td>(18.81)</td>
<td>(20.11)</td>
<td>(0.39)</td>
<td>(0.54)</td>
<td>(11.8)</td>
<td>(7.95)</td>
</tr>
</tbody>
</table>

\( R^2 \) 6.78 9.12 3.63 6.52 3.98 6.33

Obs 779 779 588 588 1367 1367

Reported are monthly return predictability regressions based on six word categories constructed from news implied volatility (NVIX). The dependent variables are annualized log excess returns on the market index. All six variables are in variance space and normalized to have unit standard deviation over the entire sample. Newey-West corrected t-statistics with number of lags/leads equal to the size of the return forecasting window are in brackets. The shares of risk premia variation due to each of the categories is in parentheses, which add up to more than one because the different categories are not mutually orthogonal.
Table 10: Filtering Disasters: Calibration Parameters

<table>
<thead>
<tr>
<th>$\bar{v}^2$</th>
<th>$\sigma_v$</th>
<th>$\rho_v$</th>
<th>$\sigma_{rvar}$</th>
<th>$\kappa_0$</th>
<th>$\kappa_0 + \mu_d(0) - \log(R_t^f)$</th>
<th>$\mu_d(1) - \mu_d(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0025</td>
<td>0.0018</td>
<td>0.7300</td>
<td>0.0041</td>
<td>0.9948</td>
<td>0.0050</td>
<td>-0.0058</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\mu_c(0)$</th>
<th>$\mu_c(1) - \mu_c(0)$</th>
<th>$\sigma_c$</th>
<th>$\psi(1)$</th>
<th>$p$</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0029</td>
<td>-0.0058</td>
<td>0.0058</td>
<td>-0.2500</td>
<td>0.0017</td>
<td>0.9715</td>
</tr>
</tbody>
</table>

All quantities are at the monthly frequency.
Table 11: NVIX Predicts Disasters

\[ I_{t\rightarrow t+\tau}^{N\rightarrow D} = \beta_0 + \beta_1 NVIX^2_{t-1} + \beta_2 s_{t-1} + \beta_3 s_{t-1} \times NVIX^2_{t-1} + \beta_4 EVAR_{t-1} + \beta_5 s_{t-1} \times EVAR_{t-1} + \epsilon_{t+\tau} \]

<table>
<thead>
<tr>
<th>(\tau)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(\beta_1\times100)</td>
<td>0.01*</td>
<td>0.01*</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>(t(\beta_1))</td>
<td>-</td>
<td>[1.79]</td>
<td>[1.83]</td>
<td>[1.60]</td>
<td>[1.45]</td>
<td>[1.46]</td>
<td>[1.14]</td>
<td>[0.67]</td>
<td>[1.56]</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.01</td>
<td>0.27</td>
<td>0.31</td>
<td>0.33</td>
<td>0.33</td>
<td>0.51</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>(\beta_1\times100)</td>
<td>0.06**</td>
<td>0.07*</td>
<td>0.07*</td>
<td>0.07*</td>
<td>0.07*</td>
<td>0.06</td>
<td>0.07*</td>
<td></td>
</tr>
<tr>
<td>(t(\beta_1))</td>
<td>-</td>
<td>[1.97]</td>
<td>[1.96]</td>
<td>[1.96]</td>
<td>[1.90]</td>
<td>[1.90]</td>
<td>[1.68]</td>
<td>[1.21]</td>
<td>[1.86]</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.00</td>
<td>2.03</td>
<td>2.23</td>
<td>2.29</td>
<td>2.29</td>
<td>3.87</td>
<td>2.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>(\beta_1\times100)</td>
<td>0.12*</td>
<td>0.12*</td>
<td>0.14**</td>
<td>0.14*</td>
<td>0.15*</td>
<td>0.14</td>
<td>0.14*</td>
<td></td>
</tr>
<tr>
<td>(t(\beta_1))</td>
<td>-</td>
<td>[1.90]</td>
<td>[1.86]</td>
<td>[1.97]</td>
<td>[1.95]</td>
<td>[1.94]</td>
<td>[1.88]</td>
<td>[1.51]</td>
<td>[1.95]</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.01</td>
<td>3.52</td>
<td>3.68</td>
<td>3.93</td>
<td>3.90</td>
<td>3.89</td>
<td>6.11</td>
<td>3.98</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>(\beta_1\times100)</td>
<td>0.17*</td>
<td>0.18</td>
<td>0.21*</td>
<td>0.21*</td>
<td>0.23*</td>
<td>0.22</td>
<td>0.22*</td>
<td></td>
</tr>
<tr>
<td>(t(\beta_1))</td>
<td>-</td>
<td>[1.65]</td>
<td>[1.62]</td>
<td>[1.79]</td>
<td>[1.79]</td>
<td>[1.78]</td>
<td>[1.47]</td>
<td>[1.80]</td>
<td></td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.14</td>
<td>3.92</td>
<td>4.06</td>
<td>4.49</td>
<td>4.49</td>
<td>4.39</td>
<td>4.37</td>
<td>6.99</td>
<td>4.52</td>
</tr>
</tbody>
</table>

Obs 1368 1367 1367 1362 1360 1360 1359 1090 1361

Reported are monthly disaster predictability regressions based on news implied volatility (NVIX). The dependent variable is \(I_{t\rightarrow t+\tau}^{N\rightarrow D}\), which reflects the probability that the economy transitions into a disaster between \(t\) and \(t + \tau\). Each row and each column represents a different regression. Rows show different forecasting horizons. Column (1) shows disaster predictability only controlling for the regime of the economy \(s_t = I^{D} > 0.5\). Column (2) adds NVIX\(^2\), and column (3) adds the interaction NVIX\(^2\times s_t\). Columns (4)-(8) control for alternative measures of expected stock market variance: (1) past realized variance; (2) AR(3) forecasting model; (3) adds price to earnings ratio to model (2); (4) adds NVIX\(^2\) to model (3); (5) adds credit spread to model (4); model (6) use as a variance proxy \(E[VIX^2_{t-1}|VAR]\) from Table 5, the forecast of VIX\(^2\) using contemporaneous and two lags of realized variance (estimated in the Jan/1990-Dec/2009 sample). The sample goes from Jan/1896 to Dec/2009 for columns 1-7 and 9, and Jan/1919 to Dec/2009 for column 8. t-statistics are Newey-West corrected with number of lags/leads equal to the size of the disaster forecasting window.
Table 12: Return Predictability in the Full Sample

\[ r_{t+\tau} = \beta_0 + \beta_1 NVIX^2_{t-1} + \beta_2 s_{t-1} + \beta_3 s_{t-1} \times NVIX^2_{t-1} + \beta_4 EVAR_{t-1} + \beta_5 s_{t-1} \times EVAR_{t-1} + \epsilon_{t+\tau} \]

<table>
<thead>
<tr>
<th>( \tau )</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \beta_1 )</td>
<td>0.12</td>
<td>0.14</td>
<td>0.13</td>
<td>0.13</td>
<td>0.14</td>
<td>0.13</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[1.30]</td>
<td>[1.55]</td>
<td>[1.33]</td>
<td>[1.32]</td>
<td>[1.40]</td>
<td>[1.06]</td>
<td>[1.1]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.50</td>
<td>0.63</td>
<td>0.68</td>
<td>0.69</td>
<td>0.67</td>
<td>0.67</td>
<td>1.23</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_1 )</td>
<td>0.05</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.14</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[0.64]</td>
<td>[0.99]</td>
<td>[0.96]</td>
<td>[0.87]</td>
<td>[0.96]</td>
<td>[0.85]</td>
<td>[1.53]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.48</td>
<td>0.67</td>
<td>0.67</td>
<td>0.71</td>
<td>0.69</td>
<td>0.69</td>
<td>2.13</td>
</tr>
<tr>
<td>Excl./Obs</td>
<td>0/1367</td>
<td>9/1358</td>
<td>9/1353</td>
<td>9/1351</td>
<td>9/1351</td>
<td>9/1351</td>
<td>8/1082</td>
<td>8/1352</td>
</tr>
<tr>
<td>6</td>
<td>( \beta_1 )</td>
<td>0.08</td>
<td>0.11**</td>
<td>0.12*</td>
<td>0.11*</td>
<td>0.13*</td>
<td>0.13*</td>
<td>0.24***</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[1.41]</td>
<td>[2.02]</td>
<td>[1.78]</td>
<td>[1.70]</td>
<td>[1.81]</td>
<td>[1.68]</td>
<td>[3.05]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>1.37</td>
<td>1.86</td>
<td>2.30</td>
<td>2.74</td>
<td>2.70</td>
<td>2.66</td>
<td>5.71</td>
</tr>
<tr>
<td>Excl./Obs</td>
<td>0/1367</td>
<td>20/1347</td>
<td>20/1342</td>
<td>20/1340</td>
<td>20/1340</td>
<td>20/1340</td>
<td>20/1073</td>
<td>20/1341</td>
</tr>
<tr>
<td>12</td>
<td>( \beta_1 )</td>
<td>0.09*</td>
<td>0.13***</td>
<td>0.14**</td>
<td>0.14**</td>
<td>0.15*</td>
<td>0.27***</td>
<td>0.13**</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[1.68]</td>
<td>[2.69]</td>
<td>[2.17]</td>
<td>[2.07]</td>
<td>[2.18]</td>
<td>[1.95]</td>
<td>[3.49]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>2.37</td>
<td>3.64</td>
<td>5.00</td>
<td>5.78</td>
<td>5.69</td>
<td>5.62</td>
<td>13.99</td>
</tr>
<tr>
<td>Excl./Obs</td>
<td>0/1367</td>
<td>46/1321</td>
<td>46/1316</td>
<td>46/1314</td>
<td>46/1314</td>
<td>46/1314</td>
<td>46/1055</td>
<td>46/1315</td>
</tr>
<tr>
<td>24</td>
<td>( \beta_1 )</td>
<td>0.08</td>
<td>0.14***</td>
<td>0.15**</td>
<td>0.14**</td>
<td>0.15**</td>
<td>0.16**</td>
<td>0.29***</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[1.47]</td>
<td>[2.78]</td>
<td>[2.35]</td>
<td>[2.28]</td>
<td>[2.35]</td>
<td>[2.14]</td>
<td>[4.07]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>3.60</td>
<td>6.87</td>
<td>9.19</td>
<td>9.97</td>
<td>9.95</td>
<td>9.88</td>
<td>22.05</td>
</tr>
<tr>
<td>Excl./Obs</td>
<td>0/1367</td>
<td>94/1273</td>
<td>94/1268</td>
<td>94/1266</td>
<td>94/1266</td>
<td>94/1266</td>
<td>94/1019</td>
<td>94/1267</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controls</th>
<th>yes</th>
<th>yes</th>
<th>yes</th>
<th>yes</th>
<th>yes</th>
<th>yes</th>
<th>yes</th>
<th>yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_{t-1} \times NVIX^2_{t-1} )</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>( EVAR_{t-1} )</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>( s_{t-1} \times EVAR_{t-1} )</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Variance model</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
</tr>
</tbody>
</table>

Reported are monthly return predictability regressions based on news implied volatility (NVIX^2). The dependent variables are annualized log excess returns on the market index. Each row and each column represents a different regression. Rows show different forecasting horizons. Column (1) controls for the state of the economy \( s_t = D^P > 0.5 \), and the interaction with NVIX^2. Column (2) excludes disasters. Column (3) controls for disasters. Columns (4-9) exclude disasters and control for alternative measures of expected stock market variance: (1) past realized variance; (2) AR(3) forecasting model; (3) adds price to earnings ratio to model (2); (4) adds NVIX^2 to model (3); (5) adds credit spread to model (4); model (6) uses as a variance proxy \( E[VIX^2_{t-1}|VAR] \) from Table 5, the forecast of VIX^2 using contemporaneous and two lags of realized variance (estimated in the Jan/1990-Dec/2009 sample). The sample goes from Jan/1896 to Dec/2009 for columns 1-7 and 9, and Jan/1919 to Dec/2009 for column 8. t-statistics are Newey-West corrected with number of lags/leads equal to the size of the disaster forecasting window. A observation \( t \) is excluded as a disaster month if \( I_{t+\tau}^R = I_{t+\tau}^{D} > 0.5 = 1 \), that is, if the filtered probability implies a higher than 50% probability that the economy transitioned into a disaster during the return forecasting window.
Table 13: Return Predictability: Controlling for Truncation Effects

\[ r_{t\rightarrow t+\tau}^e = \beta_0 + \beta_1 NVIX^2_{t-1} + \beta_2 s_{t-1} + \beta_3 s_{t-1} \times NVIX^2_{t-1} + \beta_4 EVAR_{t-1} + \beta_5 s_{t-1} \times EVAR_{t-1} + \beta_6 EMILLS_{t-1,\tau} + \epsilon_{t+\tau} \]

<table>
<thead>
<tr>
<th>( \tau )</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \beta_1 )</td>
<td>0.12</td>
<td>0.14</td>
<td>0.13</td>
<td>0.13</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>&amp; ( t(\beta_1) )</td>
<td>[1.3]</td>
<td>[1.55]</td>
<td>[1.29]</td>
<td>[1.40]</td>
<td>[1.50]</td>
<td>[1.10]</td>
<td>[1.06]</td>
</tr>
<tr>
<td>&amp; ( R^2 )</td>
<td>0.50</td>
<td>0.63</td>
<td>0.69</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>1.24</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_1 )</td>
<td>0.05</td>
<td>0.08</td>
<td>0.07</td>
<td>0.07</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>&amp; ( t(\beta_1) )</td>
<td>[0.64]</td>
<td>[0.99]</td>
<td>[0.84]</td>
<td>[0.83]</td>
<td>[0.96]</td>
<td>[0.83]</td>
<td>[1.16]</td>
</tr>
<tr>
<td>&amp; ( R^2 )</td>
<td>0.48</td>
<td>0.67</td>
<td>0.85</td>
<td>0.84</td>
<td>0.69</td>
<td>0.69</td>
<td>3.31</td>
</tr>
<tr>
<td>Excl./Obs</td>
<td>0/1367</td>
<td>9/1358</td>
<td>9/1353</td>
<td>9/1351</td>
<td>9/1351</td>
<td>9/1351</td>
<td>9/1082</td>
</tr>
<tr>
<td>6</td>
<td>( \beta_1 )</td>
<td>0.08</td>
<td>0.11***</td>
<td>0.11*</td>
<td>0.10</td>
<td>0.12*</td>
<td>0.14*</td>
</tr>
<tr>
<td>&amp; ( t(\beta_1) )</td>
<td>[1.41]</td>
<td>[2.02]</td>
<td>[1.71]</td>
<td>[1.56]</td>
<td>[1.77]</td>
<td>[1.76]</td>
<td>[3.02]</td>
</tr>
<tr>
<td>&amp; ( R^2 )</td>
<td>1.37</td>
<td>1.86</td>
<td>2.35</td>
<td>3.13</td>
<td>2.76</td>
<td>2.74</td>
<td>6.17</td>
</tr>
<tr>
<td>Excl./Obs</td>
<td>0/1367</td>
<td>20/1347</td>
<td>20/1342</td>
<td>20/1340</td>
<td>20/1340</td>
<td>20/1340</td>
<td>20/1073</td>
</tr>
<tr>
<td>12</td>
<td>( \beta_1 )</td>
<td>0.09*</td>
<td>0.13***</td>
<td>0.13**</td>
<td>0.13**</td>
<td>0.14**</td>
<td>0.15*</td>
</tr>
<tr>
<td>&amp; ( t(\beta_1) )</td>
<td>[1.68]</td>
<td>[2.69]</td>
<td>[2.04]</td>
<td>[1.97]</td>
<td>[2.19]</td>
<td>[1.96]</td>
<td>[3.43]</td>
</tr>
<tr>
<td>&amp; ( R^2 )</td>
<td>2.37</td>
<td>3.64</td>
<td>5.18</td>
<td>6.02</td>
<td>5.73</td>
<td>5.64</td>
<td>15.02</td>
</tr>
<tr>
<td>Excl./Obs</td>
<td>0/1367</td>
<td>46/1321</td>
<td>46/1316</td>
<td>46/1314</td>
<td>46/1314</td>
<td>46/1314</td>
<td>46/1055</td>
</tr>
<tr>
<td>24</td>
<td>( \beta_1 )</td>
<td>0.08</td>
<td>0.14***</td>
<td>0.14**</td>
<td>0.14**</td>
<td>0.15**</td>
<td>0.15***</td>
</tr>
<tr>
<td>&amp; ( t(\beta_1) )</td>
<td>[1.47]</td>
<td>[2.78]</td>
<td>[2.25]</td>
<td>[2.18]</td>
<td>[2.41]</td>
<td>[2.12]</td>
<td>[4.17]</td>
</tr>
<tr>
<td>&amp; ( R^2 )</td>
<td>3.60</td>
<td>6.87</td>
<td>9.58</td>
<td>10.23</td>
<td>10.69</td>
<td>10.51</td>
<td>22.33</td>
</tr>
<tr>
<td>Excl./Obs</td>
<td>0/1367</td>
<td>94/1273</td>
<td>94/1268</td>
<td>94/1266</td>
<td>94/1266</td>
<td>94/1266</td>
<td>94/1019</td>
</tr>
</tbody>
</table>

**Controls**

| & yes | yes | yes | yes | yes | yes | yes |
|---|---|---|---|---|---|---|---|
| s\( t-1 \) | yes | yes | yes | yes | yes | yes | yes |
| s\( t-1 \) \times NVIX^2_{t-1} | yes | yes | yes | yes | yes | yes | yes |
| EVAR_{t-1} | no | no | yes | yes | yes | yes | yes |
| s\( t-1 \) \times EVAR_{t-1} | no | no | yes | yes | yes | yes | yes |
| EMILLS_{t-1,\tau} | no | no | yes | yes | yes | yes | yes |
| Mills\| Variance model | (1) | (2) | (3) | (4) | (5) | (6) | (7) |

Reported are monthly return predictability regressions based on news implied volatility (\( NVIX^2 \)). The dependent variables are annualized log excess returns on the market index. Each row and each column represents a different regression. Rows show different forecasting horizons. Column (1) controls for the state of the economy \( s_t = I^{ND}_t > 0.5 \), and the interaction with \( NVIX^2 \). Column (2) excludes disasters. Columns (3-7) control for the predicted stock market variance and the predicted mills ratio using the following variables: (1) past realized variance; (2) three lags of realized variance; (3) adds price to earnings ratio to model (2); (4) adds \( NVIX^2 \) to model (3); (5) adds credit spread to model (4). The sample goes from Jan/1896 to Dec/2009 for columns 1-6, and Jan/1919 to Dec/2009 for columns 7. \( \text{t-statistics are Newey-West corrected with number of lags/leads equal to the size of the disaster forecasting window. Observation } t \text{ is excluded as a disaster month if } I^{ND}_{t-1,\tau} = 1 = (I^{ND}_{t-1,\tau} > 0.5), \text{i.e. if the filtered probability implies a higher than 50}\% \text{ probability that the economy transitioned into a disaster during the return forecasting window.} \)
Table 14: Return Predictability Horse races: NVIX vs Tone Word Lists

\[ r_{t \rightarrow t+\tau}^f = \beta_0 + \beta_1 NVIX_{t-1}^2 + \beta_2 X_{t-1} + \epsilon_{t+\tau} \]

<table>
<thead>
<tr>
<th>( \tau ) (months)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \beta_1 )</td>
<td>0.14</td>
<td>0.15</td>
<td>0.14</td>
<td>0.15</td>
<td>0.15</td>
<td>0.16</td>
<td>0.15</td>
<td>0.13</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[0.98]</td>
<td>[1.03]</td>
<td>[0.99]</td>
<td>[1.03]</td>
<td>[1.01]</td>
<td>[0.99]</td>
<td>[1.08]</td>
<td>[1.04]</td>
<td>[0.93]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.47</td>
<td>0.38</td>
<td>0.88</td>
<td>0.38</td>
<td>0.42</td>
<td>0.42</td>
<td>0.45</td>
<td>0.41</td>
<td>0.93</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_1 )</td>
<td>0.11</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[0.83]</td>
<td>[0.87]</td>
<td>[0.85]</td>
<td>[0.87]</td>
<td>[0.88]</td>
<td>[0.86]</td>
<td>[0.87]</td>
<td>[0.87]</td>
<td>[0.83]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.75</td>
<td>0.65</td>
<td>1.18</td>
<td>0.65</td>
<td>0.64</td>
<td>0.65</td>
<td>0.65</td>
<td>0.66</td>
<td>0.98</td>
</tr>
<tr>
<td>6</td>
<td>( \beta_1 )</td>
<td>0.18***</td>
<td>0.18**</td>
<td>0.18**</td>
<td>0.18***</td>
<td>0.18**</td>
<td>0.18***</td>
<td>0.18**</td>
<td>0.17***</td>
<td>0.18**</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[2.69]</td>
<td>[2.55]</td>
<td>[2.57]</td>
<td>[2.55]</td>
<td>[2.68]</td>
<td>[2.52]</td>
<td>[2.66]</td>
<td>[2.57]</td>
<td>[2.58]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>2.57</td>
<td>2.57</td>
<td>3.12</td>
<td>2.56</td>
<td>2.57</td>
<td>2.56</td>
<td>2.57</td>
<td>2.60</td>
<td>2.95</td>
</tr>
<tr>
<td>12</td>
<td>( \beta_1 )</td>
<td>0.16***</td>
<td>0.16***</td>
<td>0.15***</td>
<td>0.16***</td>
<td>0.16***</td>
<td>0.16***</td>
<td>0.16***</td>
<td>0.15***</td>
<td>0.15***</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[3.23]</td>
<td>[3.2]</td>
<td>[3.29]</td>
<td>[3.19]</td>
<td>[3.36]</td>
<td>[3.10]</td>
<td>[3.36]</td>
<td>[3.27]</td>
<td>[3.27]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>3.50</td>
<td>3.50</td>
<td>4.61</td>
<td>3.53</td>
<td>3.50</td>
<td>3.52</td>
<td>3.51</td>
<td>3.64</td>
<td>4.30</td>
</tr>
<tr>
<td>24</td>
<td>( \beta_1 )</td>
<td>0.14***</td>
<td>0.14***</td>
<td>0.14***</td>
<td>0.14***</td>
<td>0.14***</td>
<td>0.14***</td>
<td>0.14***</td>
<td>0.14***</td>
<td>0.14***</td>
</tr>
<tr>
<td></td>
<td>( t(\beta_1) )</td>
<td>[3.01]</td>
<td>[3.63]</td>
<td>[3.59]</td>
<td>[3.61]</td>
<td>[3.48]</td>
<td>[3.60]</td>
<td>[3.47]</td>
<td>[3.61]</td>
<td>[3.51]</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>5.12</td>
<td>5.12</td>
<td>7.12</td>
<td>5.21</td>
<td>5.56</td>
<td>5.32</td>
<td>5.64</td>
<td>5.15</td>
<td>6.31</td>
</tr>
</tbody>
</table>

Controls

<table>
<thead>
<tr>
<th>Text measure</th>
<th>Negative</th>
<th>Uncertainty</th>
<th>Positive</th>
<th>Modal Strong</th>
<th>Modal Weak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight scheme</td>
<td>tf</td>
<td>tf.idf</td>
<td>tf</td>
<td>tf.idf</td>
<td>tf</td>
</tr>
</tbody>
</table>

This table presents return predictability regressions based on our constructed NVIX series and the different “language tone” dictionaries developed by Loughran and McDonald (2011). The sample excludes any period with an economic disaster. The dependent variables are annualized log excess returns on the market index. t-statistics are Newey-West corrected with leads/lags equal to the size of the disaster forecasting window. The sample period is 1945 to 2009.
### Table 15: Correlations Between Alternative Measures of Tail Risk

#### Panel A: Option based measures

<table>
<thead>
<tr>
<th></th>
<th>VIX²</th>
<th>VIX premium</th>
<th>LT</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIX²</td>
<td>1.0000</td>
<td>0.9457</td>
<td>0.9787</td>
<td>0.8478</td>
</tr>
<tr>
<td>VIX premium</td>
<td>1.0000</td>
<td>0.9077</td>
<td>0.8641</td>
<td>0.8626</td>
</tr>
<tr>
<td>LT</td>
<td>1.0000</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td></td>
<td>1.0000</td>
<td></td>
</tr>
</tbody>
</table>

The options-based measured correlations are for the period Jan/1996 to Dec/2008 for which we have all four quantities.

#### Panel B: News Based Measures

<table>
<thead>
<tr>
<th></th>
<th>NVIX²</th>
<th>VIX premium</th>
<th>LT</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIX²</td>
<td>1.0000</td>
<td>0.7790</td>
<td>0.6209</td>
<td>0.8261</td>
</tr>
<tr>
<td>VIX premium</td>
<td>1.0000</td>
<td>0.6858</td>
<td>0.8795</td>
<td></td>
</tr>
<tr>
<td>LT</td>
<td>1.0000</td>
<td>1.0000</td>
<td></td>
<td>0.7526</td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
</tr>
</tbody>
</table>

The options-based measured correlations are for the period Jan/1996 to Dec/2008 for which we have all four quantities. The News based measure correlations are for the full sample, Jan/1896 to Dec/2009.