News Implied Volatility and Disaster Concerns

Asaf Manela Alan Moreira*

March 21, 2014

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) captures well the disaster concerns of the average investor. NVIX peaks during stock market crashes, times of policy-related uncertainty, world wars and financial crises. We find that periods when people are more concerned with a rare disaster, as proxied by news, are either followed by periods of above average stock returns, or followed by periods of large economic disasters. Concerns related to wars and government policy explain 54% and 21% of the time-variation in risk premia our measure identifies. These findings suggest that time variation in rare disaster risk is an important source of aggregate asset prices fluctuations. We provide parameter values of interest to macro-finance, such as the persistence and volatility of the disaster probability process.

JEL Classification: G12, C82, E44

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

*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, 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 Texas Finance Festival, SFS Cavalcade, AFA meetings, and Rothschild Caesarea Center 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 large rare macroeconomic 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 concerns regarding rare disasters and use this measurement to ask: do time-varying concerns regarding rare disasters drive aggregate stock market returns? If so, which concerns drive variation in risk premia?

To answer these questions, we construct a text-based measure of uncertainty using the frequency of words appearing on the front-page of the *Wall Street Journal*. We use the relatively short sample of options-implied volatility to estimate the relation between the issues the business press writes about and the level of economic uncertainty as measured by volatility implied by option prices (VIX). This empirical strategy measures what the average reader is concerned about when disaster risk is high, and allows us to use the long time-series of news data, which goes back to 1890, to measure disaster concerns over this period. Our empirical strategy relies on two assumptions: (i) the choice of words by the business press is a good and stable reflection of the concerns of the average investor, and (ii) disaster concerns of the average investor are priced in option markets. The first is consistent with the Gentzkow and Shapiro (2006) empirically supported model of news firms. The second is firmly grounded in theory. With respect to disaster risk, standard asset pricing models suggest that option prices are a canary in the coal mine.

We rely on "Big Data" 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 a Support Vector Regression. The key advantage of this method over Ordinary Least Squares is its ability to deal with a large feature space. This quality does not come without pitfalls, and the data intensive nature of the procedure leads us to rely more on out-of-sample validation tests than in-sample formal statical tests as would be standard in a conventional setting. We find that, news implied volatility predicts volatility implied by options out-of-sample well, with an R-squared of 0.34 and root mean squared error of 7.52 percentage points. We also find that our predictive ability over the long sample is quite stable.[1]

News implied volatility (NVIX), the predicted value from this text regression, is a news-based measure of uncertainty, which captures well the disaster concerns of the average investor over more than a century. NVIX peaks during stock market crashes, times of policy-related uncertainty, world wars and financial crises. This measure has two features missing from more standard financial and macro-economic variables used previously. Relative to financial variables it provides insight into the origins of risk premia variation. Relative to macro-variables, it encodes information about concerns that did not materialize in a particular history. These two features of our measure allow us to test whether investor concerns as reflected by business press coverage are consistent with the time-varying disaster risk hypothesis.

We test two key predictions of the time-varying rare disaster risk hypothesis: (i) periods when people are more concerned with a rare disaster are followed by periods of above average returns, or (ii) followed by periods of large economic disasters. We find strong support for both predictions across a variety of sub-samples and for a variety of alternative controls. A one standard deviation increase in NVIX increases annualized excess returns by 4 (3) percentage points over the next three months (year). Consistent with the theory, we find that NVIX also predicts economic disasters. A one standard deviation increase in NVIX leads to an increase in the disaster probability from an unconditional 2% to 3% over the next three months. Our return predictability results are robust and cannot be explained by several plausible alternative explanations. They are not driven by time-variation in stock market volatility or by the truncation induced by the exclusion of disasters from the return forecasting regressions. We find similar and significant return predictability in post Great Depression data (1938-2009) even when including disasters.

We use standard calibrations in the literature to check if our estimates are *quantitatively* consistent with each other. Hypothetically, it could easily be the case that our tests detect predictability both in returns and disasters, but the amount of variation in returns is orders of magnitudes larger than the amount detected in the disaster predictability specification. In this case we would need

[1] We analyze the stability of our measurement error in Section 2.3. One could potentially improve on 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.

an implausibly large disaster size to reconcile the two specifications quantitatively. We find that the point estimates of both the return and the disaster predictability specifications imply together risk-adjusted disaster sizes considered plausible in the literature. Our findings suggest that time variation in rare disaster risk is not only an important source of risk premia variation, but also a quantitatively reasonable one.

While the link between implied volatility and disaster concerns is well-grounded in theory, options implied volatility could reflect additional sources of uncertainty. Examples include variation in normal times volatility and variation in the price of variance risk, which are features of alternative asset pricing models (e.g. Bansal and Yaron, 2004). If one interprets disaster risk as "jump risk," a more targeted measure of disaster risk could focus only on deep out-of-the-money puts. Bollerslev and Todorov (2011) develop such a model free measure of left tail risk (LT). While our results are robust to using LT instead of implied volatility in our estimation, option implied volatility allows a longer sample for estimation, and is consistent with the broader definition of disasters studied in the macro-finance literature.[2] We rely instead on content analysis to sort out which of the option implied volatility components drive risk premia.

Interpretability, a key feature of the text-based approach, allows us to trace back a large part of the variation in risk premia to concerns related to wars (54%) and government policy (21%). Half of the time-series variation in risk premia NVIX identifies is driven by concerns tightly related to war disasters, strongly reinforcing our formal tests of the time-varying rare disaster hypothesis. We find some suggestive evidence that government-related concerns are related to redistribution risk, as our measure traces remarkably well tax-policy changes in the US, though government concerns are more ambiguous than war-related concerns. We decompose NVIX into four additional categories: Stock Markets, Financial Intermediation, Natural Disasters, and a residual component. Of these categories only the residual component reliably predicts future expected returns.

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

[2]Empirically, at the monthly frequency LT is 88 percent correlated with the volatility index we use (VXO). In unreported tests, we created a news-implied left tail index by replacing VXO with LT and found that NVIX and this alternative index are 90 percent correlated and focus on similar words.

such a long sample suggests that its history is special. Barro (2006) and subsequently Barro and Ursua (2008); Barro, Nakamura, Steinsson, and Ursua (2011); 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. The main 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.

Our novel way of measuring ex-ante disaster concerns can shed light on the plausibility of these calibrations. We find that concerns over disasters swing quite a bit, but not quite as persistent as the calibrations in Wachter (2013) and Gourio (2008) assume. Both calibrate the disaster probability process to explain the ability of valuation ratios to predict returns, which means the disaster probability process largely inherits the persistence of valuation ratios. Our results indicate that shocks to the disaster probability process have a half-life between 4 and 8 months, which is fairly persistent but inconsistent with standard calibrations in the literature.

One motivation for our paper is the empirical fact that estimating aggregate risk-return trade-offs is a data intensive procedure. For example, 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.

Our paper is also related to a recent literature that uses asset pricing restrictions to give an interpretation to movements in the VIX. 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. Backus, Chernov, and Martin (2011) challenge the idea that the jumps detected by "overpriced" out of money put options are related to the macroeconomic disasters discussed in the macro-finance literature. Drechsler (2008) interprets abnormal variation in VIX as changes is the degree of ambiguity among investors. Drechsler and Yaron (2011) interpret it as a forward looking measure of risk. Bates (2012) shows that time-changed Lévy processes capture well stochastic volatility and substantial outliers in US stock market returns. Kelly (2012) 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. 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 similarly 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 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.5 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 formally tests the time-variation in disaster risk hypotheses, reports our main results and considers alternative explanations. Section 4 uncovers which concerns drive risk premia to capture the spirit of the times. Section 5 concludes.

[3] Sample 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).

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.[4] Each title and abstract are separately broken into one and two word n-grams using a standard text analysis package that replaces highly frequent words (stop-words) with an underscore, and remove n-grams containing digits.[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 *train* subsample, 1996 to 2009, is used to

[4] We 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'.

[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.

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.[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 $\mathbf{x}_t$, a $K = 374,299$ vector of n-gram frequencies, i.e. $x_{t,i} = \frac{\text{appearences of n-gram } i \text{ in month } t}{\text{total n-grams in month } t}$. 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 *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

$$v_t = w_0 + \mathbf{w} \cdot \mathbf{x}_t + \upsilon_t \qquad\qquad t = 1 \ldots T \tag{1}$$

where $\mathbf{w}$ is a K vector of regression coefficients. Clearly $\mathbf{w}$ cannot be estimated reliably using least squares with a training time series of $T_{train} = 168$ 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.[7] 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

$$H(\mathbf{w}, w_0) = \sum_{t \in train} g_\epsilon (v_t - w_0 - \mathbf{w} \cdot \mathbf{x}_t) + c \|\mathbf{w}\|^2 ,$$

[6] A potential concern is that since the *train* sample period is chronologically after the *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 *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 *train* sample than in earlier times. Such a shift would cause its regression coefficient w_{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.

[7] 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.5.

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

$$\hat{\mathbf{w}}_{SVR} = \sum_{t \in train} (\hat{\alpha}_t^* - \hat{\alpha}_t)\, \mathbf{x}_t \tag{2}$$

where only some of the T_{train} observations' dual weights α_t are non-zero.[8]

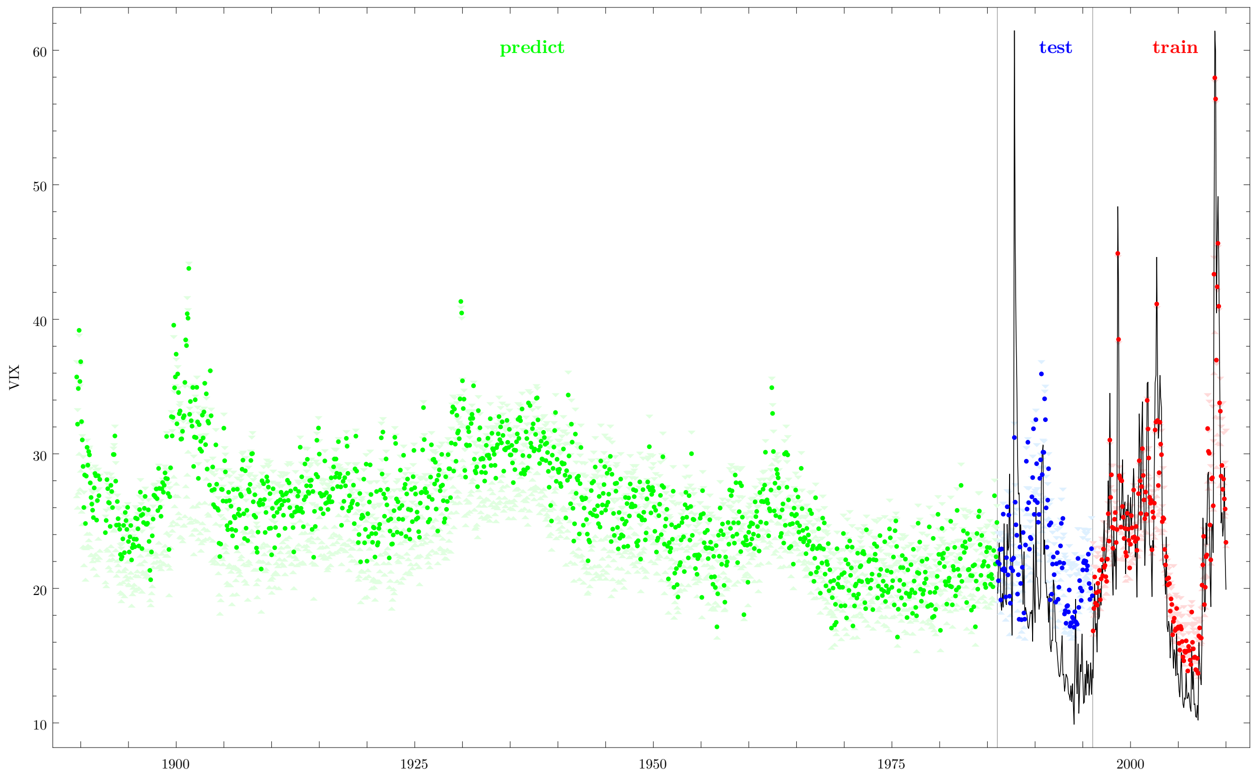
SVR works by carefully selecting a relatively small number of observations called support vectors, and ignoring the rest. The trick is that the restricted form (2) does not consider each of the K linear subspaces separately. By imposing this structure, we reduce an over-determined problem of finding $K \gg T$ coefficients to a feasible linear-quadratic optimization problem with a relatively small number of parameters (picking the T_{train} dual weights α_t). The cost is that SVR cannot adapt itself to concentrate on subspaces of $\mathbf{x}_t$ (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 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 $R^2(train) = \frac{Var(\mathbf{w} \cdot \mathbf{x}_t)}{Var(v_t)} = 0.65$. 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 overfitting.

However, as reported in Table 1, the model's out-of-sample fit over the *test* subsample is quite good, with $R^2(test) = 0.34$ and $RMSE(test) = 7.52$. In addition to these statistics, we also report

[8]SVR estimation requires us to choose two hyper-parameters that control the trade-off between in-sample and out-of-sample fit (the ϵ-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 $\epsilon_{CM2004} = 3.491$. We numerically estimate $\mathbf{w}$ by applying with these parameter values the widely used SVM^{light} package (available at `http://svmlight.joachims.org/`) to our data.

Figure 1: News-Implied Volatility 1890-2009



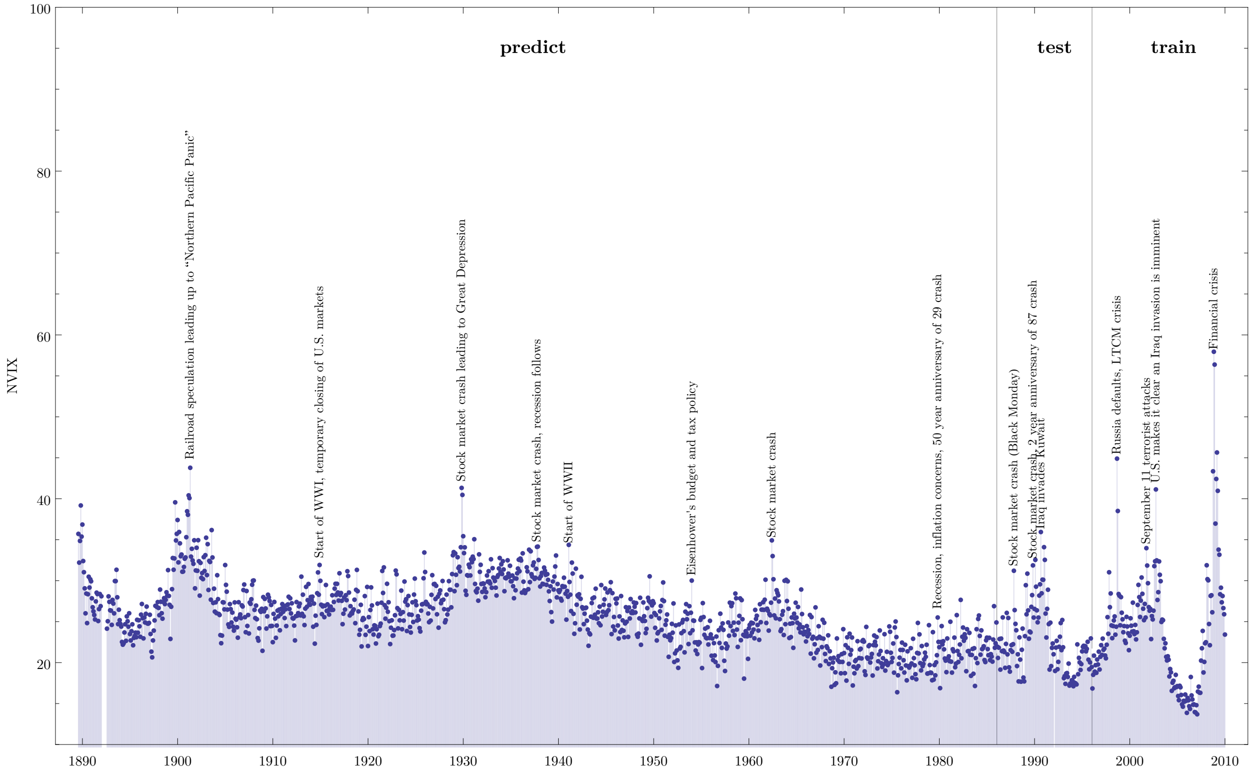
Solid line is end-of-month CBOE volatility implied by options VIX_t. Dots are news implied volatility (NVIX) $\hat{v}_t + \overline{VIX} = w_0 + \overline{VIX} + \mathbf{w} \cdot \mathbf{x}_t$. The *train* subsample, 1996 to 2009, is used to estimate the dependency between news data and implied volatility. 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 $\hat{v}$ using 1000 randomizations. These show the sensitivity of the predicted values to randomizations of the *train* subsample.

Table 1: Out-of-Sample VIX Prediction

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

Reported are out-of-sample model fit statistics using the *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.

Figure 2: News-Implied Volatility Peaks by Decade



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).

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 Disaster Concerns

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 tremulous 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 changes in the disaster probability 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 disaster concerns.

2.3 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.

A specific concern might be 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 first repeat the same estimation procedure over the same *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 *predict* subsample is only slightly higher than in the *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 *test* subsample. We therefore

Table 2: Out-of-Sample Realized Volatility Prediction Using News

Subsample	$RMSE$ SVR	R^2 SVR	$RMSE$ Reg	R^2 Reg	Correlation	Obs.
train	3.35	0.68	2.64	0.93	0.96	168
test	9.60	0.27	9.09	0.20	0.45	119
predict	10.91	0.38	8.49	0.16	0.40	1150

Reported are model fit statistics repeating 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.

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. Our main time-series return data are returns on the Dow Jones index from Global Financial Data, available monthly from July 1871 to December 2010. We refer to this series throughout as "market" returns. Results are similar if we substitute the later part of our sample by returns on the S&P 500 index or total market portfolio index. 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 also use daily return data on the Dow Jones index. This data goes back to January 1896 and is in fact the shortest of our time-series and determines how far back we go in our study. We use this data to construct proxies for realized volatility which is important when we explore alternative stories for our main result. To compute excess returns we use 90 day US government bond yields from Global Financial Data for a measure of the risk free rate that goes back to 1920. For the earlier part of our sample we use yields on 10 year US government bonds. Results do not change if we use long bonds for our entire sample. In addition, 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.

2.5 Implied Volatility, Disaster Probabilities, and Expected Returns

We build on the setup analyzed in Gourio (2008) to motivate our empirical analysis, although our empirical tests rest on the general idea that in periods when agents are more *concerned* with the possibility of a crash, out-of-the-money put options on an index that tracks aggregate wealth should be more expensive as agents try to hedge this risk. We use the term "concerned", because options can be more expensive either because the perceived likelihood of a crash increases or because people become more risk-averse with respect to disaster states. Either way, our goal is to study whether time-variation in disaster *concerns* are behind risk premia fluctuations.

We consider an economy with a representative consumer who has power utility preferences with coefficient of relative risk aversion γ. The endowment and dividend growth are log-normally distributed, but exposed to large and infrequent disasters. We are interested in the connection between disaster probabilities, expected returns and option implied volatility of publicly traded equity in this economy. Proposition 1 shows the tight link between these three quantities and leads to our empirical strategy.

Proposition 1. *Let the joint process of log consumption and public equity dividend growth be*

$$\begin{bmatrix} \Delta c_{t+1} \\ \Delta d_{t+1} \end{bmatrix} = \mu + w_{t+1} + I^D_{t+1} log \begin{bmatrix} 1 - b_c \\ 1 - b_d \end{bmatrix},$$

where w_{t+1} *is a normal random variable with variance* Σ*, and* I^D_{t+1} *is Bernoulli with probability* p_t*. Let* p_{t+1} *follow a positively correlated Markov process that is independent of the* w_{t+1} *and* I^D_{t+1} *realizations,* $\sigma^2_{e,t} = Var\left(R^e_{t+1}|p_t, I^D_t = 0\right)$ *be the variance of public equity returns in periods with no disasters, and* VIX_t *be the CBOE options implied volatility index for period* $t+1$*. It follows that, for small disaster probability* $p \approx 0$*, we can write,*

$$log\frac{E[R^e_{t+1}|p_t]}{E[R^f_{t+1}|p_t]} \approx \gamma\Sigma_{c,d} + p_t\left((1-b_c)^{-\gamma}\, b_d - b_d\right) \tag{3}$$

$$log\frac{E[R^e_{t+1}|p_t, I^D_{t+1}=0]}{E[R^f_{t+1}|p_t, I^D_{t+1}=0]} \approx \gamma\Sigma_{c,d} + p_t(1-b_c)^{-\gamma}b_d \tag{4}$$

$$E_t\left[I^D_{t+1}\right] = p_t \tag{5}$$

$$VIX^2_t \approx p_t(1-b_c)^{-\gamma}\left(A + B\times\sigma^2_{e,t}\right) + \sigma^2_{e,t}. \tag{6}$$

where $B \in [0,1)$ *and* $A = 2(b_d - \log(1-b_d))$.

This result shows that the equity premium's sensitivity to the disaster probability is a function of expected dividend (b^d) and consumption (b^c) drops during a disaster.[9] The term $p_t(1-b_c)^{-\gamma}b_d$ captures ex-ante pricing implications — the risk-premium channel. The term $-b^d p_t$ captures ex-post cash-flow of disasters realizations and drops out if we focus our attention on the behavior of equity premium during normal times as in equation (4). If we had data on p_t, we could use either specification (3) or (4) to test the time-varying disaster risk hypothesis. We have instead options implied volatility, which is an index constructed from short-maturity options on the S&P 500 index.

Proposition 1 tells us that options implied *variance* (VIX^2) is a linear function of p_t, and suggests we can estimate specifications (3) or (4) with VIX^2 instead of p_t. We discuss VIX-based tests in Section 3, but due to the high volatility of stock market returns, these return predictability tests prescribed by theory are data intensive tests, and in the short sample available we lack the power to reject the null of no predictability. Motivated by this empirical challenge we use news-implied volatility, a projection of VIX on news, measurable over a much longer sample.

We focus on estimation of specification (4) using our measure $NVIX^2$ as a proxy for the disaster probability p_t. We focus on specification (4) because (3) is sensitive to the realized severity of disasters in the sample, since it estimates b^d from a relatively small numbers of disasters. To see why omitting disasters increases power in finite samples, consider what we might conclude from a small sample of months, one of which was an economic disaster month that followed an increase in

[9] Note that under the assumptions of Proposition 1, the market drops by the same amount (in percentage terms) as the dividends. This follows from the assumption that innovations to the p_t process are independent of disaster realizations I^D_{t+1}. In a more general framework b^d would be replaced by the total expected market drop, which would include the effect of disasters on disaster probabilities and valuation ratios.

the disaster probability. Confusing realized with expected returns, we would conclude that investors demand a *lower* expected return when the disaster probability is high. This issue is analogous to the result in Broadie, Chernov, and Johannes (2009), which shows that expected returns of deep out of the money put options are statistically challenging to estimate directly.

The more powerful approach of equation (4) avoids the problem of disaster size estimation, but requires identification of disaster events, which we turn to next. This alternative approach of essentially performing return predictability regressions on a truncated distribution of returns also leads to concerns regarding bias in our estimation. We address such concerns in Section 3.2. Equations (4) and (6) also imply that VIX^2 or proxies for it will predict expected returns through the expected variance channel if dividend/consumption volatility is time-varying.[10] We address such concerns in Section 3.2 by controlling for expected variance proxies, and in Section 4.3 by decomposing our predictability results by type of concerns.

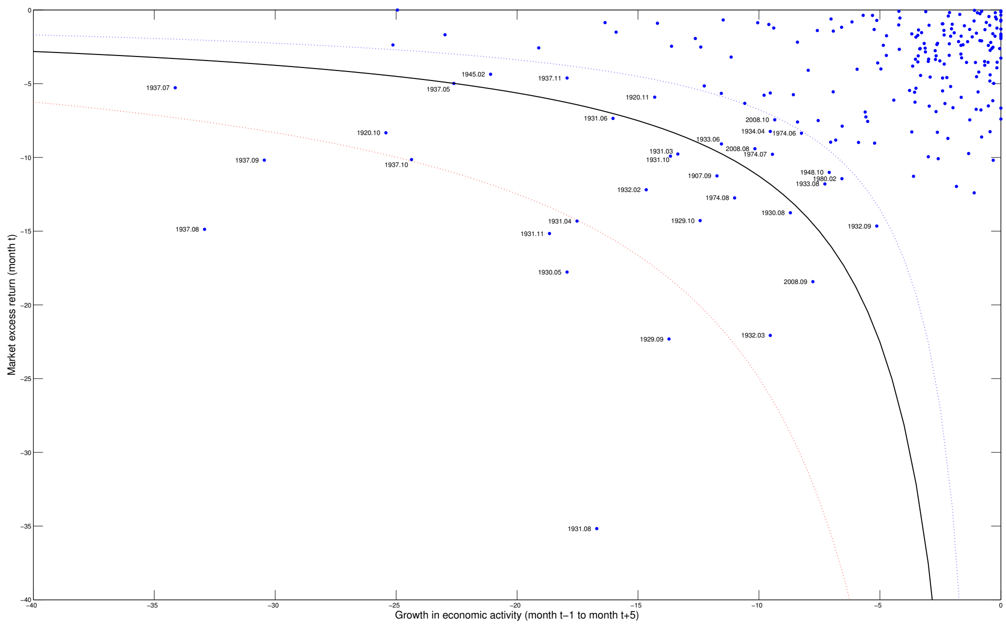
Naturally, a measure of the disaster probability should also predict disasters if investors' beliefs are rational as stated in equation (5). While keeping in mind that such a test suffers from similar pitfalls as testing specification (3), we use it in Section 3.3 as a way to evaluate whether the amount of predictability we detect in expected returns is quantitatively consistent with the amount of disaster predictability in the data.

2.6 Economic Disasters

The time-varying disaster risk hypothesis is really about macroeconomic disasters, measured by sharp drops in consumption. The main challenge to use economic activity data to test the time-varying disaster risk hypothesis is to get the timing of the disaster arrival right. When exactly did investors become aware the Great Depression was happening? One natural answer is "Black Monday". But was the Great Depression really one terrible rare event people learned about on October 29? or was it a sequence of three or four bad events? Information embedded in market prices are a natural channel to detect the timing at which investors realized what was happening, and the extent to which the future drop in economic activity was already anticipated. The cost of focusing on the stock-market to identify disasters is that we distance ourselves from the economic disasters

[10] In our setup returns have time-varying volatility, but p_t innovations are not priced, hence the time-variation in return volatility does not drive equity premium variation. However, this is not be true in general.

Figure 3: Economic Disasters Identification



We call month t a disaster month if $r_t^m < 0$ and $r_t^m \times \Delta y_{t-1,t+5} \geq \kappa$, where r_t^m is the log market return during the month and $\Delta y_{t-1,t+5}$ is log industrial production growth from month $t-1$ to month $t+5$, a six month window including the industrial production growth during month t. The lines depict three different crash identification thresholds, $\{0.5\%, 1.5\%, 2.5\%\}$. The one in the middle (1.5%) is our baseline specification.

that the disaster risk literature has in mind, and shift the focus towards statistical measurement of jumps in asset prices.

Our approach to identify disasters uses both pieces of information, requiring that big stock market drops be followed by large drops in economic activity. This approach has the advantage of keeping the focus on macroeconomic disasters, while using stock markets to identify the timing of economic disasters. We construct a "disaster-index" $Z_t = \min\{e^{r_t^m} - 1, 0\} \times (e^{y_{t+5} - y_{t-1}} - 1)$, where r_t^m is the log market return during month t and $y_{t+5} - y_{t-1}$ is log industrial production growth from month $t-1$ to month $t+5$, a six month window including the industrial production growth during month t. We classify month t as an economic disaster if month t disaster index is in the top κ percentile among all months in our full sample $(1896 - 2009)$, that is $I_t^D = 1 \times (Z_t \geq Z_\kappa)$. Our baseline case is $\kappa = 1.5$, but we consider also $\kappa's$ from 0 to 3. We assume that a disaster event lasts for three months, $I_t^{D,r} = I_{t+1}^{D,r} = I_{t+2}^{D,r} = 1$ if $I_t^D = 1$, and construct $I_{t \to t+\tau}^{D,r} = 1 - \prod_{j=1}^{\tau}\left(1 - I_{t+j}^{D,r}\right)$, which turns on as long there is at least one disaster month during the window τ. This approach follows work by Schularick and Taylor (2009) and Krishnamurthy and Vissing-Jorgensen (2012).

When predicting disasters we want to focus on the initial disaster month. We classify as a disaster $I_t^{D,D} = I_t^{D,r}$ if $I_t^{D,r} + I_{t-1}^{D,r} + I_{t-2}^{D,r} = 1$. That is, if month t is a disaster month, but the previous two periods were not. We construct $I_{t\to t+\tau}^{D,D} = \frac{12}{\tau}\sum_{j=1}^{\tau} I_{t+j}^{D,D}$, which is the annualized disaster realization rate during this period. The dummy $I_{t\to t+\tau}^{D,r}$ is the right construction if we are interested in identifying disaster periods, while $I_{t\to t+\tau}^{D,D}$ is the right construction if we are interested in measuring disaster probabilities (see Appendix). We use $I_{t\to t+\tau}^{D,r}$ in the return predictability regressions, and $I_{t\to t+\tau}^{D,D}$ in the disaster predictability specifications.

We report in Figure 3 months that this approach identifies as disasters ($I_t^D = 1$). The different lines depict disaster regions identified by different thresholds κ. A month is characterized as a disaster according to a threshold if it falls below the line. This approach requires negative signals from both economic activity and market returns to classify a month as a disaster.

We use industrial production because it is the only measure of aggregate activity available during our entire sample. It is available monthly since 1919 and annually since the beginning of the sample. This leaves us with about 25 years (1896-1919) where we do not have a monthly measure of economic activity. In this case we use gross domestic product which is measured quarterly by the NBER, and interpolate linearly to measure monthly growth in economic activity.

3 Time-Varying Disaster Concerns

In this section we formally test the hypothesis that time-variation in disaster risk is an important driver of variation in 1986 to 2009 expected returns on US equity. We start with our main findings. We then explore and rule out alternative stories. Finally, we discuss the quantitative implications of our regression results for the persistence and volatility of the disaster probability process, and the size of disaster concerns.

3.1 Results

The time-varying disaster risk explanation of asset pricing puzzles has two key empirical predictions regarding asset prices: (i) periods of high rare disaster concerns are periods when put options on the market portfolio are abnormally expensive, and (ii) these periods are followed either by economic disasters or above average excess returns on the market portfolio. Since disaster concerns are unobservable, we test these two predictions jointly. Specifically we test if periods of high option prices are followed by either disasters or periods of above average excess returns.

The formal motivation for our empirical strategy described in Section 2.5 boils down to testing if NVIX predicts future returns in paths without disasters, and if NVIX predicts disasters. To implement these two tests we rely on disaster indicators $I^{D,r}_{t\to t+\tau}$, which are 1 if there is a month classified as a disaster during the period t to $t+\tau$, not including t.

Because our forecasts use overlapping monthly data we adjust standard errors to reflect the dependence that this introduces into forecast errors using three different ways. First, we calculate Newey and West (1987) standard errors. As a second alternative, we calculate Hansen and Hodrick (1980) standard errors. As a third alternative we calculate standard error using Bootstrap. Both Newey and West and Hansen and Hodrick 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.

The last two columns of Table 3 show that in the short-sample for which option prices are available the results are weak. In the sample for which VIX is available, the implied volatility index predicts excess returns in the three months to six months horizons. However, VIX was abnormally low just before the sole economic disaster in 2008, what does not reject the time-varying disaster risk story but suggests return predictability is the result of other economic forces. 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, with the six months horizon coefficient losing its statistical significance.

While we do not have new options data to bring to bear, we use NVIX to extrapolate investors' disaster concerns. Our test of the time-varying disaster concern hypothesis is a joint test that NVIX measures investors disaster concerns *and* that disaster concerns drive expected returns on the market index. Hence, a failure to reject the null can either mean NVIX does not accurately measures disaster concerns or disaster concerns do not drive expected returns. NVIX largely inherits the behavior of VIX and VXO in sample periods where both are available. Point estimates are very similar, especially for the VIX sample, but the predictability coefficient on NVIX is estimated less precisely. To some extent this should not be surprising as NVIX was constructed to fit these

Table 3: Return Predictability

$$r^e_{t\to t+\tau} = \beta_0^R + \beta_1^R X_t^2 + \epsilon_{t+\tau} \quad \text{if } I^{D,r}_{t\to t+\tau} = 0$$

X_t	$NVIX_t$								VXO_t		VIX_t	
	1896-2009		1896-1995		1986-2009		1990-2009		1986-2009		1990-2009	
τ	β_1^R $t(\beta_1)$	R^2 T	β_1^R $t(\beta_1)$	R^2 T	β_1^R $t(\beta_1)$	R^2 T	β_1^R $t(\beta_1)$	R^2 T	β_1^R $t(\beta_1)$	R^2 T	β_1^R $t(\beta_1)$	R^2 T
1	0.19**	0.41	0.19*	0.36	0.10	0.21	0.13	0.48	0.03	0.04	0.16	0.83
	[2.04]	1322	[1.90]	1145	[0.55]	284	[0.70]	236	[0.22]	284	[0.85]	236
3	0.21***	1.55	0.20**	1.14	0.20*	2.68	0.22**	4.16	0.14**	2.85	0.23***	5.87
	[3.13]	1306	[2.45]	1131	[1.93]	282	[2.09]	234	[2.21]	282	[2.7]	234
6	0.16***	1.89	0.16**	1.55	0.16*	3.51	0.16*	4.19	0.09	2.41	0.17**	6.29
	[2.71]	1284	[2.19]	1112	[1.77]	279	[1.71]	231	[1.49]	279	[2.24]	231
12	0.14**	2.55	0.15**	2.49	0.08	1.66	0.10	2.84	0.05	1.35	0.09	3.05
	[2.34]	1248	[2.01]	1082	[0.89]	273	[1.11]	225	[0.9]	273	[1.21]	225
24	0.08*	2.12	0.09	2.24	0.04	0.93	0.04	0.80	0.02	0.34	0.02	0.38
	[1.70]	1176	[1.49]	1022	[0.59]	261	[0.53]	213	[0.35]	261	[0.32]	213

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 sample excludes any period with an economic disaster ($I^{D,r}_{t\to t+\tau} = 1$). The dependent variables are annualized log excess returns on the market index. The first and third columns report results for the sample period for which VXO is available, while the second and fourth columns are for the sample period for which VIX is available. t-statistics are Newey-West corrected with number of lags/leads equal to the size of the return forecasting window.

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 sample going from 1896 to 2009 is well estimated with a point estimate similar to the VIX sample, and a t-stat over 2 from one month to twelve months horizons. Coefficients are statistically significant up to 24 months, contrasting with a much shorter horizon for the VIX sample. The second column reports results for the sample period where no option prices are available, and the third column for the sample period for which we did not use any in-sample option price data. Estimates are similar across different samples.

We interpret the extended sample results as strong evidence for the joint hypothesis that NVIX measures disaster concerns and time-variation in disaster concerns drives expected returns. The coefficient estimates imply substantial predictability with a one standard deviation increase in $NVIX^2$ leading to $\sigma_{NVIX^2} \times \beta_1 = 21.66 \times 0.19 = 4.12\%$ higher annualized excess returns in

Table 4: Disaster Predictability: Extended Sample Tests

$I^{D,D}_{t\to t+\tau} = \beta^D_0 + \beta^D_1 NVIX^2_t + \epsilon_t$				
	1896-2009		1938-2009	
τ	$\beta^D_1(\times 100)$ $t(\beta^D_1)$	R^2 T	$\beta^D_1(\times 100)$ $t(\beta^D_1)$	R^2 T
1	0.35**	0.67	0.29	1.11
	[1.97]	1367	[1.10]	863
3	0.23**	0.89	0.10	0.37
	[2.09]	1367	[0.85]	863
6	0.21*	1.46	0.04	0.10
	[1.75]	1367	[0.51]	863
12	0.21	2.77	0.01	0.03
	[1.58]	1367	[0.27]	863
24	0.14	2.26	-0.05	0.72
	[1.10]	1367	[1.04]	863
N_D	8		2	

Reported are monthly return predictability regressions based on news implied volatility (NVIX). The dependent variable is the dummy variable $I^{D,D}_{t\to t+\tau}$ that turns on if there was an economic disaster between months t (excluding) and $t+\tau$. t-statistics are Newey-West corrected with number of lags/leads equal to the size of the disaster forecasting window.

the following month. At the annual frequency excess returns are 3.03% higher. Unsurprisingly, R-squares are small and attempts to exploit this relationship carry large risks even in samples without economic disasters. Forecasting coefficients are monotonically decreasing in the forecasting horizon, consistent with the fact that disaster concerns are persistent but substantially less persistent than alternative return predictors such as dividend yields and equivalents. Disaster concerns are 0.79 autocorrelated at the monthly frequency compared to 0.98 for the the dividend yield.

A second prediction of the time-varying disaster concerns hypothesis is that disaster concerns should be abnormally high before disasters. This prediction does not say economic disasters are predictable, but rather that in a long enough sample, disasters should happen more often when disaster concerns are high. This relationship is challenging to estimate as rare disasters are rare by definition. As we argued before in the sample for which we have option prices available, option market did not reveal an abnormally high concern with an economic disaster on the eve of the 2008-2009 Great Recession. Implied volatilities were running below realized volatility in the months preceding the stock market crash. We test this disaster predictability hypothesis using a simple linear probability model.

Table 4 reports disaster predictability regression results for the extended sample that relies on NVIX. We find that, in the full sample, NVIX is high just before disaster events. In the entire sample under the baseline specification for identifying disasters we identify eight disasters, which results in a 1.75% per month probability of a disaster event.[11] When NVIX is one standard deviation above its mean this probability increases from 1.75% to 3%. These are large numbers in terms of economic significance. It is important to note that these results rely heavily on the pre-war sample as the majority of the disasters that our criteria identify are in the earlier part of the century. In the second column of Table 4, we see that when we focus on the post-Great Depression sample the coefficients remain positive, indicating that NVIX is typically high before crashes but we cannot reject the null, what is not surprising since we only identify two disasters during this sample. The extended NVIX sample seems necessary to draw strong conclusions about disaster predictability.

3.2 Robustness and Alternative Explanations

We next explore alternative explanations for our main results. One possibility is that NVIX is not measuring variation in disaster concerns but rather current stock market volatility. According to this story NVIX predicts returns because investors demand higher expected returns during more volatile periods. We test this story by using two different proxies for expected volatility. We use past realized stock market realized variance and predicted stock market realized variance. We use an AR(4) model to predict variance. The results can be seen in Table 5. The coefficient on NVIX is slightly reduced and statistical significance is also reduced, suggesting that at least a piece of the information embedded in NVIX is related to current stock market volatility. However, the estimates show that NVIX has substantial additional new information relative to current stock market volatility.

A second concern we have is that excluding disasters could mechanically generate predictability in a world of time-varying volatility. The argument is as follows: suppose stock-market ex-ante volatility is moving around in a predictable fashion. Our strategy to identify disasters is more likely to identify disasters in periods of high ex-ante volatility. Suppose NVIX has information about future volatility. Since disaster months are excluded from the regression, we are truncating

[11] We discuss how we classify a period as an economic disaster in Section 4.3 and show how the results change if we change our disaster classification.

Table 5: Alternative Explanations

$r^e_{t\to t+\tau} = \beta_0^R + \beta_1^R NVIX_t^2 + \beta_2 X_t + \epsilon_t$ if $I^{D,r}_{t\to t+\tau} = 0$						
	Realized Variance		Expected Variance		Truncation	
τ	β_1^R $t(\beta_1^R)$	R^2 T	β_1^R $t(\beta_1^R)$	R^2 T	$\beta_1^R - \gamma$ $t(\beta_1^R - \gamma)$	R^2 T
1	0.14	0.68	0.11	0.79	0.15*	0.41
	[1.33]	1322	[1.13]	1319	[1.65]	1322
3	0.17***	1.99	0.17**	1.97	0.18***	1.55
	[2.58]	1306	[2.42]	1303	[2.66]	1306
6	0.15**	2.05	0.13**	2.31	0.13**	1.89
	[2.36]	1284	[2.12]	1281	[2.24]	1284
12	0.11*	3.11	0.10*	3.57	0.11*	2.55
	[1.93]	1248	[1.69]	1245	[1.89]	1248
24	0.08	2.24	0.07	2.37	0.06	2.12
	[1.53]	1176	[1.42]	1173	[1.22]	1176

This table presents return predictability regressions based on our constructed NVIX series and realized stock stock market variance. We measure stock market realized variance using daily returns on the Dow Jones index within the relevant month. The dependent variable are market annualized log excess returns. Each row and each column represents a different regression. Rows show different forecasting horizons. The first column shows predictability coefficients of NVIX squared on future returns controlling for past realized variance. In the second column we show the same predictability coefficient after we subtract γ, the predictability coefficient implied by the time-varying truncation that our procedure of excluding disasters induces (for a full discussion see Section A.3). The third column controls simultaneously for both realized variance and truncation. t-statistics are Newey-West corrected with number of lags/leads equal to the size of the disaster forecasting window.

the left tail of the distribution exactly when volatility is higher. This mechanism would make our proxy for volatility artificially predict returns. This story calls for a selection adjustment, which we derive explicitly in Section A.3.

The intuition is analogous to studies where we only observe a selected sample. The standard procedure is to control for this selection effect.[12] In our exercise we know the model and the selection criteria under the null, so there is no need for an instrument. In the first stage we estimate a selection equation, where we estimate the ability of NVIX to predict the probability that a given period is a disaster under the null where all that is happening is time-variation in volatility. This specification gives a benchmark coefficient which is the coefficient that a regression of future returns on NVIX should have if predictability was the result of this truncation story. In short, instead of the null hypothesis being a zero coefficient, it is a new adjusted coefficient.

[12] For example, the Heckman selection model is a popular example of such selection corrections.

We develop this analysis in full in the appendix, but it is convenient to inspect an equation to grasp the intuition of our test. We classify a month as a disaster if returns in the month are lower than a threshold $\underline{r}_{t+1}$. Expected returns for a period that was not classified as a disaster are higher than average because of the truncation our classification imposes,

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

where $\lambda(\cdot)$ is commonly know as the mills ratio. According to the truncation rationale NVIX will predict returns in paths that were not classified as disasters to the extent it predicts $\sigma_{t+1}\lambda(\underline{r}_{t+1})$. Under the null that the return predictability is the result of truncation we have:

$$E[\lambda(\underline{r}_t)\sigma_{t+1}|NVIX_t^2] = \gamma_0 + \gamma NVIX_t^2 \tag{7}$$

So what this selection problem prescribes is to run our main forecasting regression, and test if β_1 is different from γ, which is the coefficient of the regression of the truncated mean return on NVIX. If we could not reject equality of coefficients, then the time-varying truncation hypothesis would be consistent with our results.

Results with the adjusted coefficients and t-stats are in Table 5. Both the statistical and economic significance of the results survive once we adjust for this mechanical selection effect. Yet one might still be concerned that the parametric structure that we impose on the return distribution might not be doing a good job of capturing this truncation effect. One way to alleviate these concerns is to focus on a sub-sample where there were fewer disasters but not exclude the disasters from the sample. Since the majority of our disasters happen during the Great Depression, we focus on a sub-sample starting in 1938, when the NBER officially declared the end of the Great Depression. We run the return forecasting regression in the post depression sample *without* excluding disasters, where only two economic disasters happened according to our baseline criteria. Since we are not excluding disasters truncation concerns become mute.

We report in Table 6 similar and significant return predictability in post Great Depression data (1938-2009) even when including disasters. The coefficients for horizons 6, 12, and 24 months do not change both in magnitude or in statistical significance. Coefficients for one and three months

Table 6: Post Depression Sample *including* Disasters

$r^e_{t\to t+\tau} = \beta_0 + \beta_1^R NVIX_t^2 + \beta_2 X_t + \epsilon_t$								
Sample:	1938-2009				1896-2009			
Disasters	Excluded		Included		Excluded		Included	
τ	β_1^R / $t(\beta_1^R)$	R^2 / T	β_1^R / $t(\beta_1^R)$	R^2 / T	β_1^R / $t(\beta_1^R)$	R^2 / T	β_1^R / $t(\beta_1^R)$	R^2 / T
1	0.20	0.46	0.10	0.48	0.14	0.68	0.03	0.39
	[1.52]	857	[0.80]	863	[1.33]	1322	[0.26]	1367
3	0.23***	2.18	0.08	0.22	0.17***	1.99	0.03	0.32
	[2.81]	853	[0.65]	863	[2.58]	1306	[0.35]	1367
6	0.21***	3.16	0.13*	1.10	0.15**	2.05	0.05	0.48
	[3.02]	847	[1.72]	863	[2.36]	1284	[0.77]	1367
12	0.15**	2.84	0.12**	1.85	0.11*	3.11	0.02	0.15
	[2.22]	835	[2.15]	863	[1.93]	1248	[0.34]	1367
24	0.10*	2.43	0.12**	3.59	0.08	2.24	0.01	0.08
	[1.74]	811	[2.43]	863	[1.53]	1176	[0.17]	1367

This table presents return predictability regressions based on our constructed NVIX series and realized stock stock market variance for two different subsamples. We measure stock market realized variance using daily returns on the Dow Jones index within the relevant month. Each row and each column represents a different regression. Different forecasting horizons are in rows. The table shows predictability coefficients of NVIX squared on future returns controlling for past realized variance. t-statistics are Newey-West corrected with number of lags/leads equal to the size of the disaster forecasting window.

are slashed by roughly half and lose their statistical significance. We will see in Table 7 that these short horizons are also more sensitive to our criteria to identify economic disasters, but longer horizons estimates remain significant.

Table 7 examines the sensitivity of our main return predictability results to different disaster thresholds. Our baseline specification ($\kappa = 1.5\%$) identifies twenty-one disasters in our sample. The results change somewhat depending on the criteria used, with evidence for return predictability becoming weaker as we use a more strict definitions of a disaster. This works exactly as expected as including disaster months in the return predictability regression biases the coefficient downward.

3.3 Volatility, Persistence and Size of Disaster Concerns

In this section we connect the coefficient estimates from our two predictability specifications to structural parameters of interest in the macro-finance literature. In particular, our regressions can recover the volatility and persistence of the disaster probability process, and the risk neutral disaster size. We report our estimates for these quantities in Table 8. We explain formally where our estimates come from and provide an economic interpretation.

Table 7: Disaster Threshold Sensitivity

$$r^e_{t\to t+\tau} = \beta_0 + \beta_1^R NVIX_t^2 + \epsilon_t \text{ if } I^{D,r}_{t\to t+\tau} = 0$$

Threshold	$\kappa = 0\%$		$\kappa = 0.5\%$		$\kappa = 1\%$		$\kappa = 1.5\%$		$\kappa = 2\%$		$\kappa = 2.5\%$		$\kappa = 3\%$	
τ	β_1^R $t(\beta_1^R)$	R^2 T	β_1^R $t(\beta_1^R)$	R^2 T	β_1^R $t(\beta_1^R)$	R^2 T	β_1^R $t(\beta_1^R)$	R^2 T	β_1^R $t(\beta_1^R)$	R^2 T	β_1^R $t(\beta_1^R)$	R^2 T	β_1^R $t(\beta_1^R)$	R^2 T
1	-0.05	0.04	0.05	0.04	0.16*	0.32	0.19**	0.41	0.19**	0.42	0.21**	0.55	0.23***	0.66
	[0.52]	1367	[0.52]	1348	[1.77]	1335	[2.04]	1322	[2.04]	1311	[2.39]	1296	[2.73]	1280
3	-0.02	0.01	0.08	0.23	0.19***	1.30	0.21***	1.55	0.20***	1.48	0.23***	1.96	0.22***	1.68
	[0.18]	1367	[0.81]	1339	[2.88]	1322	[3.13]	1306	[3.04]	1291	[3.42]	1268	[3.18]	1248
6	0.01	0.01	0.11*	0.95	0.16***	1.79	0.16***	1.89	0.14**	1.52	0.15**	1.75	0.14**	1.41
	[0.15]	1367	[1.83]	1327	[2.67]	1305	[2.71]	1284	[2.44]	1263	[2.52]	1230	[2.26]	1205
12	0.01	0.00	0.12**	1.92	0.14**	2.60	0.14**	2.55	0.12**	2.00	0.12**	2.06	0.11*	1.65
	[0.08]	1367	[2.44]	1312	[2.40]	1278	[2.34]	1248	[2.11]	1215	[2.04]	1175	[1.79]	1144
24	0.00	0.00	0.10**	3.29	0.10*	2.70	0.08*	2.12	0.07	1.68	0.06	1.25	0.05	0.84
	[0.04]	1367	[2.44]	1288	[1.95]	1230	[1.70]	1176	[1.48]	1130	[1.27]	1067	[1.03]	1024
N_D	0		7		14		21		27		34		41	

This table presents return predictability regressions based on our constructed NVIX series for different disaster thresholds and different horizons. The sample excludes any period with an economic disaster. The dependent variable are annualized stock market log excess returns. t-statistics are Newey-West corrected. The sample period is 1896 to 2009.

As shown in Proposition 1, p_t is linearly related to VIX_t^2. A linear approximation of equation (5) that VIX movements are driven by disaster concerns ($\hat{p}_t$) and normal times risk ($\widehat{\Sigma_t^e}$),

$$\widehat{VIX_t^2} = \widehat{\Sigma_t^e} + \phi \hat{p}_t$$

where $\widehat{x}$ denotes deviations from the unconditional mean and ϕ is an approximating constant. Under the conditions discussed in Section 2.5, we can use our news-based measure of implied volatility, NVIX, and recover estimates of β_1^D and β_1^R. We show next how to recover additional information about disaster risk. In particular, we are interested in the volatility, persistence and the expected disaster size.

If the disaster probability process follows an AR(1) process with persistence parameter ρ_p, an OLS regression of the (annualized) average number of disaster $\frac{12}{\tau}\sum_{s=0}^{\tau-1} I_{t+s\to t+s+1}^D$ on $\widehat{VIX_t^2}$ recovers $plim(\beta_{1.\tau}^D) = \frac{\phi VAR(\widehat{p_t})\sum_{j=0}^{\tau-1}\rho_p^j}{Var(\widehat{VIX_t^2})}$, which implies that the ratio between coefficients of two different horizons τ_L and τ_S recovers the disaster probability persistence ρ_p:

$$plim\left(\frac{\tau_L}{\tau_S}\frac{\beta_{\tau_L}^D}{\beta_{\tau_S}^D}\right) = \frac{1-\rho_p^{\tau_L}}{1-\rho_p^{\tau_S}}. \tag{8}$$

We follow a similar strategy with the return predictability specification, but in this case a similar computation only recovers the overall persistence of expected returns ρ_R:

$$plim\left(\frac{\tau_L\beta_{\tau_L}^R}{\tau_S\beta_{\tau_S}^R}\right) = \frac{1-\rho_R^{\tau_L}}{1-\rho_R^{\tau_S}}. \tag{9}$$

If disaster probability variation is the only driver of expected returns, i.e. either there is no variation in normal times risk or normal times risk does not impact risk premia, then this persistence should match the one we recover from the disaster predictability specification ($\rho_p = \rho_R$). In general, however, expected returns are also driven by time-variation from other sources of uncertainty. If such movements are priced, they would also lead to variation in risk premia. Evidence presented in Section 4.3 strongly suggests that the bulk of time variation in risk premia our measure identifies is directly related to disasters. In this section we reinforce that result by evaluating quantitatively whether the amount of predictability and persistence in expected returns is consistent with the amount of predictability and persistence in disasters.

Table 8: Disaster Risk: Persistence, Volatility and Size

(a) Disaster Probability Persistence Implied by Return Predictability Regressions

$$\rho_R = \left\{ x \left| \frac{\tau_L \beta^R_{\tau_L}}{\tau_S \beta^R_{\tau_S}} = \frac{1-x^{\tau_L}}{1-x^{\tau_S}} \right. \right\}$$

$\tau_S \backslash \tau_L$	3	6	12	24
1	1.12	0.95	0.94	0.92
3		0.82	0.90	0.89
6			0.94	0.91
12				0.88

(b) Disaster Probability Persistence Implied by Disaster Predictability Regressions

$$\rho_p = \left\{ x \left| \frac{\beta^D_{\tau_L}}{\beta^D_{\tau_S}} = \frac{1-x^{\tau_L}}{1-x^{\tau_S}} \right. \right\}$$

$\tau_S \backslash \tau_L$	3	6	12	24
1	0.61	0.78	0.90	0.91
3		0.93	0.98	0.95
6			1.01	0.96
12				0.92

(c) Risk Neutral Disaster Size

τ	$\left\lvert b^d(1-b^c)^{-\gamma}\right\rvert = \frac{\beta^R_\tau}{\beta^D_\tau}$
1	0.53
3	0.91
6	0.79
12	0.65
24	0.57

(d) Disaster Probability and Expected Return Volatility

τ	$\sigma(\frac{12}{\tau}E[\sum_{s=0}^{\tau} p_{t,t+1} \lvert NVIX_t^2])$	$\sigma(\frac{12}{\tau}E^{ND}[r^e_{t\to t+\tau} \lvert NVIX_t^2])$
1	7.52 %	4.01%
3	4.96 %	4.50%
6	4.46 %	3.51%
12	4.53 %	2.93%
24	3.13 %	1.79%

Table 8 panel (a) and (b) reports persistence estimates. We find a tight range of persistence estimates implied by return predictability specifications, except perhaps for the results that rely on very short horizons. Disaster probability specifications suggest similar magnitudes. While the estimates suggest disaster concerns are persistent, they are smaller than the numbers used in the literature. For example, Gourio (2012) uses approximately $\rho_p = 0.96$ at the monthly frequency, and Wachter (2013) chooses $\rho_p = 0.99$ to match the persistence of valuation ratios.

We exploit the cross-equation restriction between the return predictability and the disaster predictability regressions imposed by the rare disaster risk model to evaluate whether implied disaster sizes are quantitatively reasonable. Under the hypothesis that all time-variation in expected returns detected by NVIX is driven by variation in the disaster probability, the ratio between return and disaster predictability regression coefficients recovers the risk-adjusted disaster size:

$$plim\left(\frac{\beta^R_\tau}{\beta^D_\tau}\right) = -b^d\left(1-b^c\right)^{-\gamma}. \tag{10}$$

Our point estimates for this quantity range from 53% to 91%, with a median point estimate of 65% (Table 8(c)). Are these disaster sizes reasonable? We can compare these estimates with the Barro and Ursua (2008) calibration based on a large sample of countries some of which experienced economic disasters. In their calibration, agents have a coefficient of relative risk aversion of 3.5, the average rare disaster exhibits a 22% consumption drop and a 32% stock market decline. In this case we have $\left|b^d\left(1-b^c\right)^{-\gamma}\right| \approx \left|(-0.32)(0.78)^{-3.5}\right| = 76\%$. This number is in the same ballpark as our estimates. This indicates that the amount of predictability we detect in expected returns is consistent with the amount of predictability given a reasonable calibration of rare disasters.

Let us now turn to the volatility of the disaster probability process. To estimate this volatility we follow a similar approach and identify variation in disaster probabilities from the disaster predictability specification:

$$\sigma_p^2 = Var(E[p_{t,\tau}|\widehat{NVIX}_t^2]) = (\beta_\tau^D)^2 Var(\widehat{NVIX}_t^2).$$

A similar computation backs out the amount of expected return variation detected by our measure. In panel (d) we report our estimates for the (annualized) volatility of disaster probability shocks and expected returns shocks detected by NVIX. It is again useful to contrast these quantities with parameter choices currently used in the literature. Wachter (2013) calibrates a continuous time model to produce an unconditional standard deviation of $\sigma_p \approx 2.9\%$ per year in disaster probability. Gourio (2012) calibrates a discrete time model with an unconditional annual volatility of $\sigma_p \approx 2.3\%$. We estimate this same quantity to be 4.53%, at least 50% larger. Relative to the calibrations in the literature we find disaster concerns to be more volatile and less persistent. Note these results are not a mechanical consequence of NVIX being a noisy proxy for the true disaster probability process. Estimated volatility is recovered from the ability of NVIX to predict disasters at different horizons and estimated persistence is recovered from the amount of predictability detected at different horizons.

4 The Origins of Disaster Concerns

What drives investors' concerns at different periods? Are these concerns reasonable? The NVIX index we constructed relies on the relative frequency of words during each month in the sample. In this section we investigate which words play an important role and try to describe the zeitgeist - the spirit of the times. We gain novel insights about the origins of disaster concerns by utilizing our text-based measure of uncertainty.

Table 9: Top Variance Driving n-grams

n-gram	Variance Share, %	Weight, %	n-gram	Variance Share, %	Weight, %
stock	37.28	0.10	oil	1.39	-0.03
market	6.74	0.06	banks	1.36	0.06
stocks	6.53	0.08	financial	1.32	0.11
war	6.16	0.04	__ u.s	0.88	0.05
u.s	3.62	0.06	bonds	0.81	0.04
tax	2.01	0.04	__ stock	0.80	0.03
washington	1.78	0.02	house	0.77	0.05
gold	1.46	-0.04	billion	0.67	0.06
special	1.44	0.02	economic	0.64	0.05
treasury	1.43	0.06	like	0.59	-0.05

We report the fraction of NVIX variance $h(i)$ that each n-gram drives over the *predict* subsample as defined in (11), and the regression coefficient w_i from (1), for the top 20 n-grams.

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

$$h(i) \equiv \frac{Var(\hat{v}_t(i))}{\sum_{j \in K} Var(\hat{v}_t(j))} \tag{11}$$

as a measure of the n-gram specific variance of NVIX.[13] Table 9 reports $h(i)$ for the top variance 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 $\mathbf{w}$ in (11).

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.

[13]Note that in general $Var(\hat{v}_t) \neq \sum_{j \in K} Var(\hat{v}_t(j))$ due to covariance terms.

Table 10: Categories Total Variance Share

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

We report the percentage of NVIX variance ($= \sum_{i \in C} h(i)$) that each n-gram category C drives over the *predict* subsample.

4.2 Word Categorization

To study the principal word categories driving this variation, we classify n-grams into five broad categories of words: Government, Intermediation, Natural Disasters, Stock Markets and War. We rely on the widely used WordNet and WordNet::Similarity projects to classify words.[14] 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.

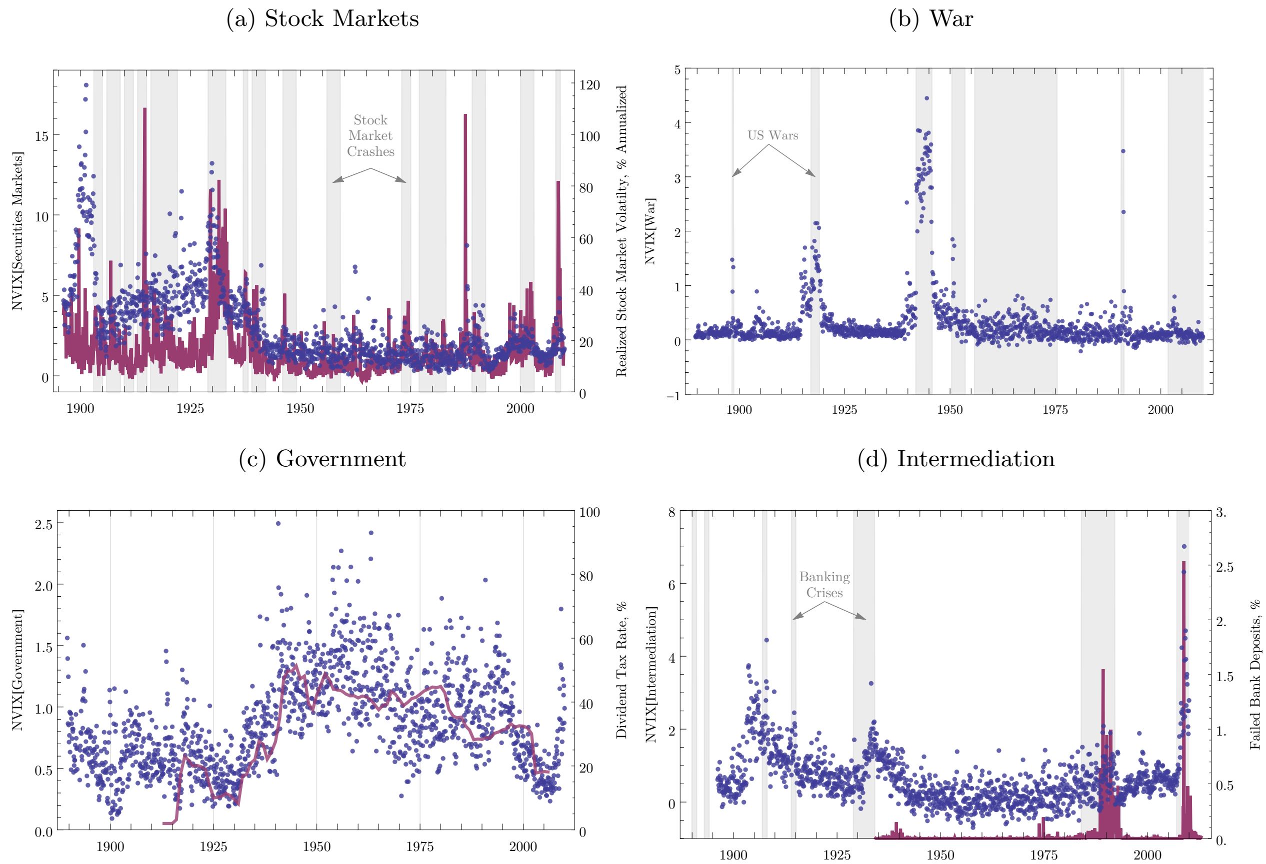
Table 10 reports the percentage of NVIX variance ($= \sum_{i \in C} h(i)$) that each n-gram category drives over the *predict* subsample. Stock market related words explain over half the variation in NVIX. War-related words explain 6 percent. Unclassified words explain 37 percent of the variation. Clearly 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 4 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 4a. Attention to the stock market as measured by this component seems to spike at market

[14] WordNet (Miller, 1995) is available at http://wordnet.princeton.edu. WordNet::Similarity (Pedersen, Patwardhan, and Michelizzi, 2004) is available at http://lincoln.d.umn.edu/WordNet-Pairs.

Figure 4: News Implied Volatility due to Different Word Categories



In all panels dots are monthly NVIX due only to category C-related words $\hat{v}_t(C) = \mathbf{x}_t \cdot \mathbf{w}(C)$. 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).

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 4b 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 4c. 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 4d 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?

Our approach relies on interpretable news to measure uncertainty. This feature allows us to investigate which types of concerns are responsible for time-variation in expected returns. We use the word classification to decompose variation in risk premia, and report the results in Table 11.

At the one-year horizon, War (54%) and Government (21%) related concerns capture the bulk of the variation in risk premia. Both categories have a statistically reliable relation with future market excess returns. Concerns related to Financial Intermediation (13%), Stock Markets (10%), and Natural disasters (2%), account for some of the variation in expected returns, but the relationship

Table 11: Risk premia decomposition

Panel A - Regression Coefficients: $r^e_{t\to t+\tau} = \beta^R_0 + \sum_{j=1}^{N} \beta^R_j X^j_t + \epsilon_{t+\tau}$ if $I^{D,r}_{t\to t+\tau} = 0$							
τ	Gov't	War	Intermediation	Stock Markets	Natural Disasters	Residual	R^2/T
1	1.62	2.90**	1.69	2.41	-0.72	2.80	0.65
	[0.99]	[2.40]	[0.68]	[1.48]	[0.63]	[1.37]	1323
3	2.78**	2.76***	2.32	2.65*	-0.25	3.02**	2.41
	[2.41]	[2.84]	[1.30]	[1.92]	[0.32]	[2.28]	1307
6	2.12**	2.99***	1.78	1.49	0.19	2.73**	3.96
	[1.99]	[3.27]	[1.27]	[1.13]	[0.29]	[2.39]	1285
12	1.77*	2.77***	1.54	1.30	0.50	1.92*	5.81
	[1.69]	[3.74]	[1.47]	[1.05]	[0.90]	[1.67]	1249
24	2.16**	1.56**	1.28	1.39	-0.18	0.59	6.62
	[2.29]	[2.33]	[1.62]	[1.31]	[0.46]	[0.65]	1177

Panel B - Risk premia variation share by concern $\left(\frac{(\beta^R_j)^2 \times \hat{\sigma}^2(X^j_t)}{\hat{\sigma}^2(\sum_{j=1}^{N} \beta^R_j X^j_t \mid \beta^R)}\right)$						
τ	Gov't	War	Intermediation	Stock Markets	Natural Disasters	Residual
1	0.13	0.42	0.12	0.27	0.03	0.31
3	0.31	0.32	0.19	0.26	0.00	0.29
6	0.22	0.46	0.13	0.10	0.00	0.30
12	0.21	0.54	0.13	0.10	0.02	0.19
24	0.61	0.35	0.17	0.23	0.00	0.03

Reported are monthly return predictability regressions based on the six word categories constructed from news implied volatility (NVIX). The sample excludes any period with an economic disaster ($I^{D,r}_{t\to t+\tau} = 1$). The dependent variables are annualized log excess returns on the market index. All six variables are normalized to have a standard deviation equal to one. Panel B reports the share of risk premia variation due to each of the categories. Note that they add up to more than one as the different categories are not orthogonal to each other. t-statistics are Newey-West corrected with number of lags/leads equal to the size of the return forecasting window.

is statistically unreliable. The harder to interpret orthogonal residual accounts for 19%.[15]

This decomposition strongly supports the time-varying rare disaster risk model. About half of the variation in risk premia is unequivocally related to disaster fears. War concerns is the main driver of our return predictability results. This result shows that the relationship between news implied volatility and expected returns documented in Section 3 are mostly driven by disaster concerns. 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

[15]Note that shares add up to slightly more than one due to fact that the shocks are not orthogonal to each other.

of the variation in risk premia in the last century is due to concerns related to wars, strongly suggests that risk premia estimates likely reflect the special realization of history the US happened to experience during this period.[16]

Government-related concerns allows for a considerably wider range of possible interpretations. Work by Baker, Bloom, and Davis (2013), Pastor and Veronesi (2012), and Croce, Nguyen, and Schmid (2012) emphasizes the role of policy-related uncertainty in inducing volatility and reducing asset prices in the recent period. 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 4c 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.

While Stock Markets-related concerns are not reliably related to future returns at all frequencies, at higher frequencies (3 months) it predicts returns with the same economic and statistical magnitude as Government and War. Stock market related concerns spike following periods of high stock market volatility. Consistent with these regression results, Figure 4a shows that stock market concerns tracks well the time-series of realized volatility. This result is not surprising, 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). Interestingly, but nevertheless hard to interpret, is that this attention persists longer than the shocks to volatility themselves and that attention to the stock market predict returns, while realized volatility does not. We think there are several plausible stories that fit this empirical pattern. One story could be a high-frequency version of the depression babies narrative (Malmendier and Nagel, 2011), where investors fear risk for a long time after a period of high volatility. A slightly different interpretation of the data is that periods of high volatility incite crash fears in investors. Our disaster predictability estimates (see Table 8) imply that such persistent disaster fears would not be implausible, as disaster probabilities are substantially more persistent than stock market volatility. A third possibility is that this measure tracks broadly defined variation in uncertainty, for example due to long-run risk concerns.

[16]See Brown, Goetzmann, and Ross (1995) for this argument.

While we cannot sort out these different explanations, we think our Stock Markets category can be useful for future studies to study theories of optimal inattentive behavior.

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 4d 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 perception of risks are low [17], 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 should be expected as we perceive as unlikely that there is time-variation in the likelihood of natural disasters at the frequencies that we are looking at, 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 not 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 19% of the variation in risk premia at yearly frequencies and up to 31% at monthly frequencies. Our ex-ante

[17] That is, periods when they perceive risk to be low.

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. In particular, we find that a substantial amount of risk premia variation is driven by disaster concerns. From the roughly 4% a year variation in risk premia news implied volatility can measure, at least 54% is driven by war concerns, tightly related to the type of disasters that motivates the rare disaster literature. An additional 21% of this variation is plausibly related to expropriation risk, which is quite different from the types of cash-flow shocks we are used to study in rare disaster models.

5 Conclusion

We use a text-based method to extend options-implied measures of disaster concerns back to the end of the 19th century, bringing new data to bear on the time-varying rare disaster asset pricing model. We show that our news-based measure of implied volatility, NVIX, predicts returns at frequencies up to 24 months. Consistent with the time-varying disaster risk view our measure also predicts disasters. Importantly, the amounts of predictability detected in stock returns and disasters are quantitatively consistent with each other, leading to implied disaster sizes that are in the same ball park as the ones estimated by Barro and Ursua (2008) using cross sectional data. Using content analysis, we find that at least half of the predictability in excess returns detected by news implied volatility is driven by concerns related to wars, which is strong evidence for the view that time-varying disaster concerns is an important driver of asset prices.

A Appendix

A.1 Proposition 1

We borrow heavily from Gourio (2008) .Consider a Lucas type economy with a representative consumer with coefficient of relative risk-aversion of γ and impatience ρ. We are interested in pricing cash flows and derivatives of an arbitrary security in this economy. Let the joint stochastic process of consumption and dividend growth be

$$\begin{bmatrix} \Delta c_{t+1} \\ \Delta d_{t+1} \end{bmatrix} = \begin{bmatrix} \mu_c \\ \mu_d \end{bmatrix} + w_{t+1} + I^D_{t+1} log \begin{bmatrix} 1-b_c \\ 1-b_d \end{bmatrix},$$

where w_{t+1} is a normal random variable with variance $\Sigma = \begin{bmatrix} \sigma_c^2 & \sigma_{c,d} \\ \sigma_{c,d} & \sigma_d^2 \end{bmatrix}$, and I^D_{t+1} is Bernoulli with probability $p_t = Prob_t(I^D_{t+1} = 1)$, and b_c and b_d are percentage drops in consumption and cash flows.The disaster probability process follows a Markov process such that p_{t+1} $F(p_{t+1}|p_t)$, with $p_{t+1} \perp (I^D_t, \epsilon_t, \epsilon_{t+1})\Big| p_t$ and $p_t \in [\underline{p}, \bar{p}]$.

The stochastic discount factor can be written as follows,

$$m_{t+1} = e^{-\rho}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} = e^{-\rho-\gamma\mu_c-\gamma\sigma_c w_{t+1}} \begin{cases} 1 & I^D_{t+1} = 0 \\ (1-b_c)^{-\gamma} & I^D_{t+1} = 1. \end{cases}$$

A.1.1 Return predictability

The risk-free rate can be written as

$$log(R^f_{t+1}) = log(E_t[m_{t+1}]^{-1}) = \rho + \gamma\mu_c - \frac{\gamma^2\sigma_c^2}{2} - log\left(1 - p_t + p_t(1-b_c)^{-\gamma}\right).$$

The price of the cash-flow claim has to respect a similar asset pricing restriction,

$$\frac{P_t}{D_t} = E_t\left[e^{-\rho}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\frac{D_{t+1}}{D_t}\left(\frac{P_{t+1}}{D_{t+1}} + 1\right)\right].$$

Exploring conditional independence of the disaster probability process and the fact that the disaster probability is the only source of persistence in this economy we can write the price dividend ratio as a function of the disaster probability, $\frac{P_t}{D_t} = q(p_t)$, where $q(\cdot)$ satisfies the following recursion,

$$q(p) = e^{-\rho-\gamma\mu_c+\frac{\gamma^2\sigma_c^2}{2}+\mu_c+\frac{\sigma_d^2}{2}-\gamma\sigma_{c,d}} \left(1-p_t+p_t(1-b_c)^{-\gamma}(1-b_d)\right) \int_{\underline{p}}^{\bar{p}} \left(q(p')+1\right) dF(p'|p).$$

See Gourio (2008) for a proof of existence and uniqueness of the function $q(\cdot)$. Armed with a construction for the valuation ratio we can write the returns on the cash flow claim as,

$$\begin{aligned} E_t[R^d_{t+1}] &= E_t\left[\frac{D_{t+1}}{D_t}\frac{q(p_{t+1})+1}{q(p_t)}\right] = E_t\left[\frac{D_{t+1}}{D_t}\right] E_t\left[\frac{q(p_{t+1})+1}{q(p_t)}\right] \\ &= e^{\rho+\gamma\mu_c-\frac{\gamma^2\sigma_c^2}{2}+\gamma\sigma_{c,d}} \left(1-p_t+p_t(1-b_d)\right)\left(1-p_t+p_t(1-b_c)^{-\gamma}(1-b_d)\right)^{-1}. \end{aligned} \tag{12}$$

We are now ready to obtain equation (3) by computing the implied equity premium,

$$log\frac{E[R^d_{t+1}|p_t]}{E[R^f_{t+1}|p_t]} = \gamma\sigma_{c,d} + \log\left(\frac{(1-b_d)\left(1-p_t+p_t(1-b_c)^{-\gamma}\right)}{(1-p_t+p_t(1-b_c)^{-\gamma}(1-b_d)}\right).$$

Since we are interested int he case of rare disasters it is useful to consider an approximation as $p \to 0$. At this limit the equity premium can be written as

$$\begin{aligned} log\frac{E[R^d_{t+1}|p]}{E[R^f_{t+1}|p]} &= \gamma\sigma_{c,d} + p_t\frac{\partial}{\partial p}log\frac{E[R^d_{t+1}|p]}{E[R^f_{t+1}|p]}\Bigg|_{p=0} = \gamma\sigma_{c,d} + p_t\left(b_d(1-b_c)^{-\gamma}-b_d\right), \\ log\frac{E[R^d_{t+1}|p, I^D_{t+1}=0]}{E[R^f_{t+1}|p, I^D_{t+1}=0]} &= \gamma\sigma_{c,d} + p_t\frac{\partial}{\partial p}log\frac{E[R^d_{t+1}|p, I^D_{t+1}=0]}{E[R^f_{t+1}|p, I^D_{t+1}=0]}\Bigg|_{p=0} = \gamma\sigma_{c,d} + p_t b_d(1-b_c)^{-\gamma}. \end{aligned}$$

Where the second line follows from evaluating equity expected returns in paths with no disasters, what comes down to replacing $(1-p_t+p_t(1-b_d))$ for 1 in the dividend growth expectation term in expression (11). These results follow almost immediately from Gourio (2008). We now extend his framework to derive an explicit link between option implied volatility as computed by the CBOE and disaster probabilities.

A.1.2 Option implied volatility

The idealized VIX index can be constructed as an weighted average of options of different moneyness k as follows

$$VIX^2 = 2e^{r_{f,T}} \left[\int_0^{F_T} \left(\frac{1}{k^2} \right) Put_T(k)dk + \int_{F_T}^{\infty} \left(\frac{1}{k^2} \right) Call_T(k)dk \right],$$

where F_t is the normalized forward price for the S&P 500 index, the forward and options are for a maturity of one month, and t$r_{f,t}$ is the one-month continuously compounded risk-free rate. In practice the formula applied by the CBOE it an approximation of the above expressions as one does not have a continuous of strike prices or options with exactly thirty days maturity.Given our asset pricing framework we can write the price of a put options maturity as,

$$\begin{aligned} Put_T(k) &= pE\left[(1-b_c)^{-\gamma} e^{-\gamma\mu - \gamma\sigma_c w_c} \left[k - \frac{q(p')}{q(p)} (1-b_d) e^{\mu + \sigma_d w_d} \right]^+ \right] + (1-p)\, E\left[e^{-\gamma\mu - \gamma\sigma_c w_c} \left[k - \frac{q(p')}{q(p)} e^{\mu + \sigma_d w_d} \right]^+ \right] \\ &= p(1-b_c)^{-\gamma}(1-b_d) Put_T^{ND}\left((1-b_d)^{-1} k \right) + pPut_T^{ND}(k). \end{aligned}$$

where $Put_T^{ND}(k) \equiv E\left[e^{-\gamma\mu - \gamma\sigma_c w_c} \left[k - \frac{q(p')}{q(p)} e^{\mu + \sigma_d w_d} \right]^+ \right]$, and $\frac{q(p')}{q(p)} e^{\mu + \sigma_d w_d} = R_{e,t}$, the simple index return in paths without crashes. Applying analogous decomposition to the call price and applying to the VIX formula we obtain,

$$\begin{aligned} VIX^2 = 2e^{r_{f,T}} \Bigg[& p(1-b_c)^{-\gamma}(1-b_d) \left(\int_0^{F_T} \left(\frac{1}{k^2} \right) Put_T^{ND}((1-b_d)^{-1}k)dk + \int_{F_T}^{\infty} \left(\frac{1}{k^2} \right) Call_T^{ND}((1-b_d)^{-1}k)dk \right) \\ & + (1-p) \left(\int_0^{F_T} \left(\frac{1}{k^2} \right) Put_T^{ND}(k)dk + \int_{F_T}^{\infty} \left(\frac{1}{k^2} \right) Call_T^{ND}(k)dk \right) \Bigg]. \end{aligned}$$

We now apply Result 1 in Martin (2011). Consider a continuous time limit for the return process. In particular we can write

$$\frac{q(p')}{q(p)} e^{\mu + \sigma_d w_d} = e^{\int_0^1 dR_t^e},$$

where log equity returns follow $dR_t^e = \mu_t dt + \sigma_t dB_t$. Under this return process result 1 in Martin (2011), implies this particular weighted sum of calls and puts converge to the expected integral of the quadratic return variation under the risk-neutral measure $(*)$ from 0 to T,

$$
\begin{aligned}
2e^{r_{f,T}}\left[\int_0^{F_T}\left(\frac{1}{k^2}\right)Put_T^{ND}(k)dk+\int_{F_T}^{\infty}\left(\frac{1}{k^2}\right)Call_T^{ND}(k)dk\right] &= E^*\left[\int_0^T\sigma_t^2dt\right] \\
&\approx Var^*(R_T^e) \\
&\approx Var(R_T^e)
\end{aligned}
$$

The second line applies Result 2 in Martin (2011) that this quantity is approximately equal to the variance of realized equity return at the horizon T. The third line follows from the fact that under the assumption that p_t innovations are independent of shocks to the stochastic discount factor. This implies that variance risk is not priced, hence expected risk-neutral variance equal to expected variance. Turning to the first term, a change of variables, $\tilde{k}=(1-b_d)^{-1}k$ gives

$$
\begin{aligned}
&p_t(1-b_c)^{-\gamma}\left(\int_0^{(1-b_d)^{-1}F_T}\left(\frac{1}{\tilde{k}^2}\right)Put_T^{ND}(\tilde{k})d\tilde{k}+\int_{(1-b_d)^{-1}F_T}^{\infty}\left(\frac{1}{\tilde{k}^2}\right)Call_T^{ND}(\tilde{k})d\tilde{k}\right) \\
&= p(1-b_c)^{-\gamma}[\int_0^{F_T}\left(\frac{1}{\tilde{k}^2}\right)Put_T^{ND}(\tilde{k})d\tilde{k} \\
&+\int_{F_T}^{\infty}\left(\frac{1}{\tilde{k}^2}\right)Call_T^{ND}(\tilde{k})d\tilde{k}+\int_{F_T}^{F_T(1-b_d)^{-1}}\left(\frac{1}{\tilde{k}^2}\right)\left[Put_T^{ND}(\tilde{k})d\tilde{k}-Call_T^{ND}(\tilde{k})\right]d\tilde{k}].
\end{aligned}
$$

Our previous analysis applies to the term in the first line. We can apply put-call parity to the third line to obtain it as a function of the risk-free rate,

$$
\int_{F_T}^{F_T(1-b_d)^{-1}}\left(\frac{1}{\tilde{k}^2}\right)\left[Put_T^{ND}(\tilde{k})d\tilde{k}-Call_T^{ND}(\tilde{k})\right]d\tilde{k}=-e^{-r_{f,T}}\left(\log\left(1-b_d\right)-b_d\right).
$$

Putting it all together we get,

$$
\begin{aligned}
VIX^2 &= (1-p)Var\left(R_T^e\right)+p(1-b_c)^{-\gamma}\left(Var\left(R_T^e\right)+2\left(b_d-\log\left(1-b_d\right)\right)\right) \\
&= Var\left(R_T^e\right)+p(1-b_c)^{-\gamma}\left(\frac{(1-b_c)^{-\gamma}-1}{(1-b_c)^{-\gamma}}Var\left(R_T^e\right)+2\left(b_d-\log\left(1-b_d\right)\right)\right).
\end{aligned}
$$

Note that disaster probability impact option implied volatility through two distinct channels. The first term is the risk-aversion channel , $\left((1-b_c)^{-\gamma}-1\right)Var\left(R_T^e\right)$. VIX is higher because the price of risk is higher during disasters, and not because the market drops during a disaster. The second channel is due to the actual expected market drop during a disaster, the term $2\left(b_d - \log\left(1-b_d\right)\right)$, which also is weighted by the marginal utility in the disaster state. To gauge the relative magnitudes of these two terms, note that $\frac{(1-b_c)^{-\gamma}-1}{(1-b_c)^{-\gamma}} \in [0,1)$, and realized variance averages 0.16^2, peaking at 0.8^2 during periods of extreme volatility, implying $Var\left(R_T^e\right)$ is on average a number like 0.0021, reaching numbers like 0.06 during very extreme circumstances. The second term on the other hand is about $2\left(b_d - \log\left(1-b_d\right)\right) \approx 1.3$ for a market drop of 30% during a disaster event.

$$VIX^2 = Var\left(R_T^e\right) + p(1-b_c)^{-\gamma}\left(\frac{(1-b_c)^{-\gamma}-1}{(1-b_c)^{-\gamma}}Var\left(R_T^e\right) + 2\left(b_d - \log\left(1-b_d\right)\right)\right).$$

A.2 Multi period regressions

We have shown that one-month ahead return forecasts and disaster forecasts can be written as linear functions of disaster probabilities,

$$\begin{aligned} E_t^{ND}[R_{t+1}^e] &\approx E^{ND}[R_{t+1}^e] - \frac{E^D[m_{t+1}R_{t+1}^e]}{E^{ND}[m_{t+1}]}p_t, \\ E_t[I_{t\to t+1}^D] &= p_t. \end{aligned}$$

Multi-period forecasts are useful to investigate the persistence of the disaster probability process. Assume p_t follows an AR(1),

$$p_{t+1} = \mu_p + \rho_p p_t + \sigma_p \epsilon_{t+1}.$$

In this case, we should have that the decay rate of predicted number of disasters over the predictability horizon τ should be informative about the persistence of the disaster probability process. In annualized terms we have

$$E_t\left[\frac{12}{\tau}\sum_{s=0}^{\tau-1} I_{t+s\to t+s+1}^D\right] = \frac{12}{\tau}\sum_{s=0}^{\tau-1}\rho_p^s p_t = p_t\frac{12}{\tau}\frac{1-\rho_p^\tau}{1-\rho_p}.$$

To link return predictability regressions across periods we impose a conditional log-normal structure for returns in periods without disasters. Let r be the log market return

$$r_{t+1} = \mu_{r,t} + \sigma_r\epsilon_{t+1} + \zeta_r I_{t+1}^D.$$

The expected return in paths without disasters $\mu_{r,t}$ is our quantity of interest. Analogously, the log stochastic discount factor growth is,

$$\tilde{m}_{t+1} = \mu_m + \sigma_m \epsilon_{t+1} + \zeta_m I_{t+1}^D.$$

No arbitrage implies $E_t \left[e^{\tilde{m}_{t+1} + r_{t+1}} \right] = 1$. Exploiting conditional log-normality we get

$$E_t^{ND}[r_{t+1} - r_{f,t+1}] = -\sigma_m \sigma_r + p_t e^{\zeta_m} \left(e^{\zeta_r} - 1 \right) + \frac{1}{2} \sigma_r^2.$$

Cumulative annualized excess returns for τ horizon are then,

$$\begin{aligned}
\frac{12}{\tau} E_t^{ND}[\sum_{s=0}^{\tau-1} r_{t+s+1} - r_{f,t+s+1}] &= -12\sigma_m\sigma_r + 12\frac{1}{2}\sigma_r^2 + e^{\zeta_m}\left(e^{\zeta_r} - 1\right)\frac{12}{\tau}\sum_{s=0}^{\tau-1}\rho_p^s p_t \\
&= -12\sigma_m\sigma_r + 12\frac{1}{2}\sigma_r^2 + e^{\zeta_m}\left(e^{\zeta_r} - 1\right)\frac{12}{\tau}\frac{1-\rho_p^\tau}{1-\rho_p} p_t.
\end{aligned}$$

where the ratio between disaster and return predictability identifies risk-neutral disaster sizes, $e^{\zeta_m} \left(e^{\zeta_r} - 1 \right)$.

A.3 Truncation

A concern that we have regarding our approach to identify disasters is that the predictability results in non-crash periods 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 be period of artificially higher returns. This will me a mechanical of artifact of the truncation. Testing this hypothesis is fairly straightforward. Consider the following model featuring time-varying volatility ,

$$\begin{aligned}
\sigma_{t+1}^2 &= \mu_\sigma + \rho_\sigma \sigma_t^2 + \omega \sqrt{\sigma_t^2} H_\sigma W_{t+1} \\
r_{t+1} &= \mu_r + \sigma_{t+1} H_r W_{t+1} \\
\Delta y_{t+1} &= \mu_y + \rho_y \Delta y_t + \sigma_{t+1} H_y W_{t+1}
\end{aligned}$$

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

$$
\begin{aligned}
E[r_{t+1}|r_{t+1} \geq \underline{r}(\sum_{j=0}^{dw-1} \Delta y_{t+j}), \sigma_{t+1}] &= \mu_r + \sigma_{t+1} E\left[W_{r,t+1} | W_{r,t+1} \geq \frac{\underline{r}(\sum_{j=0}^{dw} \Delta y_{t+j}) - \mu_r}{\sigma_{t+1}} \right] \\
&= \mu_t + \phi \sigma_t^2 + \sigma_{t+1} \lambda(\underline{r}(\sum_{j=0}^{dw} \Delta y_{t+j})).
\end{aligned}
$$

$\lambda(\underline{r})$ is the well known Mills ratio, which is the mean of a truncated variable, and dw is the disaster window we use to select disasters ($dw = 6$). In the context of our example we know exactly how months were selected as disasters, so we know the threshold $\underline{r}(\sum_{j=0}^{dw} \Delta y_{t+j})$. If $NVIX_t$ predicts future volatility the truncation will lead us to find predictability that NVIX predicts returns when in fact it does not. In this case conditional expectations are given by:

$$
E[r_{t+1}|r_{t+1} \geq \underline{r}(\sum_{j=0}^{dw} \Delta y_{t+j}), NVIX_t^2, \sigma_t^2] = \mu_r + E[\sigma_{t+1} \lambda(\underline{r}(\sum_{j=0}^{dw} \Delta y_{t+j}))|NVIX_t^2, \sigma_t].
$$

Focusing on linear expectations we can write this as:

$$
\begin{aligned}
E[\lambda(\underline{r}(\sum_{j=0}^{dw} \Delta y_{t+j}))\sigma_{t+1}|NVIX_t^2, \sigma_t^2] &= \gamma_0 + \gamma_1 NVIX_t^2 + \gamma_2 \sigma_t^2, \\
E[r_{t+1}|r_{t+1} \geq \underline{r}(\sum_{j=0}^{dw} \Delta y_{t+j}), NVIX_t^2, \sigma_t^2] &= \mu_r + \gamma_0 + \gamma_1 NVIX_t^2 + \gamma_2 \sigma_t^2.
\end{aligned}
$$

The above expression tells us that in order to test the time-varying rare disaster story against the truncation story it suffices to test the NVIX coefficient against γ_1 instead of zero. If the estimated coefficient is larger than γ_1 we can reject the null that the predictability during periods without disasters is induced by this truncation effect. 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 to compare the amount of return predictability we detect in the data with what one would expect if expected returns were exclusively driven by time-varying volatility, but truncation was inducing a spurious correlation between NVIX and future returns.

For multi-period return forecasts, a observation is excluded as long there is at least on disasters in the forecasting window, implying that the bias will be less severe for longer horizons, since the independent variable is an average of truncated and non truncated months. To derive the truncation bias formally let $X_t = \begin{bmatrix} NVIX_t^2 & \sigma_t^2 \end{bmatrix}'$, we can write multi-period expected returns as,

$$\begin{aligned} E[\frac{\sum_{i=1}^{\tau} r_{t+i}}{\tau}|\{r_{t+z} \geq \underline{r}(\sum_{j=0}^{dw} \Delta y_{t+j})|1 \leq z \leq \tau\}, X_t] &= \frac{1}{\tau}\sum_{i=1}^{\tau} E[E[r_{t+i}|r_{t+i} \geq \underline{r}(\sum_{j=0}^{dw} \Delta y_{t+i+j}), X_{t+i-1}]|X_t] \\ &= \frac{1}{\tau}\sum_{i=1}^{\tau} E[\mu_r + \gamma_0 + \gamma_1 NVIX^2_{t+i-1} + \gamma_2 \sigma^2_{t+i-1}|X_t] \\ &= \mu_r + \gamma_0 + \frac{1}{\tau}\left(\gamma_1 1'_{NVIX^2} + \gamma_2 1'_{\sigma^2}\right)\sum_{i=0}^{\tau-1} \Gamma^i X_t, \end{aligned}$$

where 1_x denotes a column vector with 1 on the position of variable x. Under the null, the multi-period regression coefficient should be,

$$\beta^R_{1,null} = \frac{1}{\tau}\left(\gamma_1 1_{NVIX^2} + \gamma_2 1_{\sigma^2}\right)\sum_{i=1}^{\tau} \Gamma^i 1_{NVIX^2}.$$

A.4 Time-Varying Expected Returns: are Rare Disasters the Whole Story?

Proponents of the rare disaster explanation suggest it can explain one of the key facts regarding the time-series properties of stock-market returns, that the dividend yield on the market portfolio predicts future returns far into the future. Our results suggest NVIX captures variation at different frequencies than the dividend yield, but it seems only natural to horse-race them. If the concerns encoded in NVIX are the same concerns reflected in dividend yields one of the variable should drive the other out of the regression. We would expect that the variable measured with more more noise to be driven out of the regression. And if not driven completely out we would expect the coefficient magnitude to decrease.

Table 12 shows that if we focus on the whole sample this is approximately what happens, with the coefficients on both NVIX and price to earnings ratios decreasing in the multivariate specification. It is reassuring that the NVIX coefficient is always estimated more reliably than the price to earnings ratio coefficient. With the difference being specially meaningful for shorter horizons. However, comparing R^2 across horizons we see that the predictive power of the two

Table 12: Price-to-Earnings Ratio Predictability

$r^e_{t\to t+\tau} = \beta_0 + \beta_1 NVIX^2_t + \beta_2(\frac{P}{E})_t + \epsilon_t$ if $I^D_{t\to t+\tau} = 0$						
Sample Period	1896-2010			1896-1994		
Dependent Variable	β_1 $t(\beta_1)$	β_2 $t(\beta_2)$	R^2 T	β_1 $t(\beta_1)$	β_2 $t(\beta_2)$	R^2 T
$r^e_{t\to t+1}$	0.18**		0.41	0.19*		0.36
	[2.03]		1323	[1.9]		1146
		-0.49*	0.28		-0.66	0.27
		[1.73]	1323		[1.49]	1146
	0.18**	-0.47*	0.67	0.22**	-0.80*	0.75
	[1.97]	[1.67]	1323	[2.07]	[1.72]	1146
$r^e_{t\to t+3}$	0.21***		1.56	0.20**		1.14
	[3.13]		1307	[2.45]		1132
		-0.44**	0.68		-0.56	0.58
		[1.99]	1307		[1.6]	1132
	0.20***	-0.41*	2.15	0.22***	-0.69*	2.00
	[3.02]	[1.87]	1307	[2.67]	[1.94]	1132
$r^e_{t\to t+6}$	0.16***		1.89	0.16**		1.55
	[2.71]		1285	[2.19]		1113
		-0.44**	1.40		-0.59*	1.24
		[2.06]	1285		[1.76]	1113
	0.15**	-0.42*	3.11	0.18**	-0.68**	3.18
	[2.55]	[1.93]	1285	[2.47]	[2.08]	1113
$r^e_{t\to t+12}$	0.14**		2.55	0.15**		2.49
	[2.34]		1249	[2.01]		1083
		-0.53**	3.88		-0.76**	3.88
		[2.39]	1249		[2.47]	1083
	0.12**	-0.50**	5.98	0.17**	-0.82***	7.02
	[2.09]	[2.24]	1249	[2.3]	[2.81]	1083
$r^e_{t\to t+24}$	0.08*		2.12	0.09		2.24
	[1.7]		1177	[1.49]		1023
		-0.50**	7.60		-0.65***	6.47
		[2.2]	1177		[2.61]	1023
	0.07	-0.48**	9.19	0.10*	-0.68***	9.30
	[1.48]	[2.1]	1177	[1.76]	[2.78]	1023

This table presents return predictability regressions based on our constructed NVIX series and price-to-earning ratios. The sample excludes any period with an economic disaster. The dependent variable are annualized log excess returns on the market index. Price-to-earnings ratios are from Shiller, where earnings are 10 years averages of S&P 500 earnings.t-statistics are Newey-West corrected with number of lags/leads equal to the size of the disaster forecasting window. The first column reports the results for our entire sample period, and the second column for the sample period for which we did not use any in sample option price data.

different variables roughly adds up. This pattern is replicated across alternative sample periods, and strongly suggests that these variables are measuring different things.

We interpret these results as saying disaster concerns, at least the ones we can measure though NVIX, are not likely to be the whole explanation behind time-variation in expected returns. A possibility put forth by Wachter (2013) is that different disaster concerns might move at different frequencies, generating return predictability at different frequencies. Under this story VIX and price to earnings ratio would put different weights in these different concerns.

A.5 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.[18] 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 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.

Tables 13 and 14 repeat our analysis but this time include also the tone scores as a second independent variable in addition to NVIX. Both tables show that NVIX remains a significant return and disaster predictor throughout.

[18] Examples of this approach can be found in Antweiler and Frank (2004), Tetlock (2007), Engelberg (2008), and Tetlock, Saar-Tsechansky, and Macskassy (2008).

Table 13: Return Predictability Horse races: NVIX vs Tone Word Lists

$$r^e_{t\to t+\tau} = \beta_0 + \beta_1 NVIX^2_t + \beta_2 Dictionary + \epsilon_t \text{ if } I^D_{t\to t+\tau} = 0$$

τ	1			3			6			12			24		
Dictionary	β_1 $t(\beta_1)$	β_2 $t(\beta_{12})$	R^2	β_1 $t(\beta_1)$	β_2 $t(\beta_{12})$	R^2	β_1 $t(\beta_1)$	β_2 $t(\beta_{12})$	R^2	β_1 $t(\beta_1)$	β_2 $t(\beta_{12})$	R^2	β_1 $t(\beta_1)$	β_2 $t(\beta_{12})$	R^2
Negative_tf	0.17*	-0.00	0.45	0.21***	0.00	1.57	0.17***	0.00	1.93	0.14**	0.00	2.58	0.08*	-0.00	2.15
	[1.76]	[0.7]	0.00	[3.12]	[0.31]	0.00	[2.78]	[0.42]	0.00	[2.49]	[0.26]	0.00	[1.8]	[0.17]	0.00
Negative_tfidf	0.19*	0.00	0.41	0.21***	0.00	1.58	0.17***	0.00	1.93	0.14**	0.00	2.61	0.08*	0.00	2.19
	[1.95]	[0.08]	0.00	[3.08]	[0.36]	0.00	[2.68]	[0.36]	0.00	[2.35]	[0.28]	0.00	[1.79]	[0.18]	0.00
Uncertainty_tf	0.21**	-0.00	0.46	0.26***	-0.00**	2.18	0.21***	-0.00**	2.95	0.19***	-0.00**	4.76	0.13***	-0.00**	6.21
	[2.02]	[0.85]	0.00	[3.82]	[2.54]	0.00	[3.75]	[2.29]	0.00	[3.6]	[2.35]	0.00	[3.25]	[2.44]	0.00
Uncertainty_tfidf	0.18**	0.04	0.44	0.21***	0.04	1.61	0.16***	0.03	1.94	0.13**	0.01	2.57	0.08*	0.00	2.12
	[2.04]	[0.6]	0.00	[3.16]	[0.6]	0.00	[2.74]	[0.44]	0.00	[2.36]	[0.18]	0.00	[1.69]	[0.03]	0.00
Positive_tf	0.21**	-0.00	0.45	0.23***	-0.00	1.63	0.17***	-0.00	1.94	0.14**	-0.00	2.61	0.10**	-0.00	2.56
	[1.99]	[0.7]	0.00	[3.17]	[0.77]	0.00	[2.81]	[0.49]	0.00	[2.47]	[0.36]	0.00	[2.16]	[0.68]	0.00
Positive_tfidf	0.19**	0.02	0.42	0.21***	0.02	1.58	0.16***	0.01	1.93	0.14**	0.01	2.64	0.08*	-0.00	2.12
	[2.02]	[0.36]	0.00	[3.12]	[0.44]	0.00	[2.71]	[0.39]	0.00	[2.36]	[0.37]	0.00	[1.71]	[0.01]	0.00
ModalStrong_tf	0.19**	-0.00	0.44	0.22***	-0.00	1.73	0.17***	-0.00	2.07	0.14**	-0.00	2.64	0.09*	-0.00	2.54
	[2.02]	[0.64]	0.00	[3.32]	[1.25]	0.00	[2.98]	[0.97]	0.00	[2.5]	[0.59]	0.00	[1.93]	[1.15]	0.00
ModalStrong_tfidf	0.18**	0.11	0.41	0.20***	0.28	1.59	0.15***	0.42	2.05	0.13**	0.37	2.80	0.08*	0.01	2.12
	[2.02]	[0.15]	0.00	[3.15]	[0.55]	0.00	[2.71]	[0.9]	0.00	[2.29]	[0.83]	0.00	[1.67]	[0.03]	0.00
ModalWeak_tf	0.19**	-0.00	0.52	0.22***	-0.00**	1.98	0.17***	-0.00**	3.00	0.15***	-0.00***	4.46	0.09**	-0.00***	5.90
	[2.08]	[1.22]	0.00	[3.28]	[2]	0.00	[3]	[2.46]	0.00	[2.7]	[2.77]	0.00	[2.23]	[3.61]	0.00
ModalWeak_tfidf	0.19**	0.31	0.45	0.21***	0.16	1.59	0.16***	0.03	1.89	0.14**	-0.01	2.55	0.08*	-0.04	2.14
	[2.05]	[0.78]	0.00	[3.14]	[0.51]	0.00	[2.7]	[0.09]	0.00	[2.33]	[0.03]	0.00	[1.7]	[0.12]	0.00

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 1896 to 2009.

Table 14: Disaster Predictability Horse races: NVIX vs Tone Word Lists

$$I^D_{t\to t+\tau} = \beta_0 + \beta_1 NVIX^2_t + \beta_2 Dictionary + \epsilon_t$$

τ	1			3			6			12			24		
Dictionary	β_1 $t(\beta_1)$	β_2 $t(\beta_2)$	R^2	β_1 $t(\beta_1)$	β_2 $t(\beta_2)$	R^2	β_1 $t(\beta_1)$	β_2 $t(\beta_2)$	R^2	β_1 $t(\beta_1)$	β_2 $t(\beta_2)$	R^2	β_1 $t(\beta_1)$	β_2 $t(\beta_2)$	R^2
Negative_tf	0.04	0.00	1.04	0.08**	0.00	1.32	0.11*	0.00	1.46	0.18*	0.00	2.01	0.13	-0.00	0.72
	[1.61]	[0.79]	0.00	[2.06]	[0.42]	0.00	[1.84]	[0.26]	0.00	[1.89]	[0.47]	0.00	[0.92]	[0.23]	0.00
Negative_tfidf	0.04	0.00	0.94	0.08**	0.00	1.30	0.11*	-0.00	1.44	0.17*	0.00	1.91	0.15	0.01	0.83
	[1.62]	[0.52]	0.00	[2]	[0.15]	0.00	[1.72]	[0.02]	0.00	[1.67]	[0.26]	0.00	[0.95]	[0.42]	0.00
Uncertainty_tf	0.04	-0.00	0.97	0.07*	0.00	1.32	0.09	0.00	1.68	0.13	0.00	2.54	0.04	0.00*	2.65
	[1.51]	[0.59]	0.00	[1.7]	[0.42]	0.00	[1.36]	[1.07]	0.00	[1.23]	[1.32]	0.00	[0.27]	[1.81]	0.00
Uncertainty_tfidf	0.04*	-0.00	0.94	0.08**	0.01	1.32	0.11*	0.02	1.51	0.17*	0.04	2.06	0.13	0.18	2.05
	[1.66]	[0.33]	0.00	[2.06]	[0.97]	0.00	[1.76]	[1.06]	0.00	[1.69]	[0.97]	0.00	[0.82]	[1.29]	0.00
Positive_tf	0.04	0.00	0.95	0.06	0.00	1.78	0.08	0.00	2.20	0.10	0.00	3.50	-0.01	0.00**	5.10
	[1.29]	[0.25]	0.00	[1.5]	[1.41]	0.00	[1.32]	[1.41]	0.00	[1.18]	[1.55]	0.00	[0.08]	[2.16]	0.00
Positive_tfidf	0.04*	0.00	0.94	0.08**	0.01	1.37	0.11*	0.02	1.55	0.17*	0.04	2.14	0.15	0.10	1.72
	[1.66]	[0.61]	0.00	[2.05]	[1.18]	0.00	[1.76]	[1]	0.00	[1.7]	[1.02]	0.00	[0.95]	[1.11]	0.00
ModalStrong_tf	0.04	-0.00	0.94	0.08*	0.00	1.30	0.10	0.00	1.53	0.15	0.00	2.40	0.09	0.01*	2.90
	[1.55]	[0.02]	0.00	[1.94]	[0.04]	0.00	[1.62]	[0.61]	0.00	[1.54]	[1.06]	0.00	[0.6]	[1.91]	0.00
ModalStrong_tfidf	0.04	-0.04	0.95	0.07**	0.18	1.38	0.10*	0.32	1.59	0.16*	0.42	2.05	0.11	1.31	1.49
	[1.64]	[0.62]	0.00	[2.01]	[0.87]	0.00	[1.72]	[0.98]	0.00	[1.67]	[0.87]	0.00	[0.75]	[1.26]	0.00
ModalWeak_tf	0.04*	-0.00	0.96	0.08**	-0.00	1.32	0.11*	-0.00	1.47	0.17	0.00	1.98	0.13	0.00	0.88
	[1.65]	[0.72]	0.00	[2.03]	[0.48]	0.00	[1.74]	[0.4]	0.00	[1.63]	[0.43]	0.00	[0.82]	[0.57]	0.00
ModalWeak_tfidf	0.04*	0.05*	0.98	0.08**	0.06	1.32	0.11*	0.11	1.48	0.17*	0.26	2.05	0.14	0.85	1.58
	[1.67]	[1.93]	0.00	[2.06]	[0.82]	0.00	[1.77]	[0.85]	0.00	[1.7]	[1.19]	0.00	[0.91]	[1.29]	0.00

This table presents disaster predictability regressions based on our constructed NVIX series and the different "language tone" dictionaries developed by Loughran and McDonald (2011). The dependent variable is the dummy variable $I^D_{t\to t+\tau}$ that turns if there was an economic disaster between months t (excluding) and $t+\tau$ on sample excludes any period with an economic disaster ($I^D_{t\to t+\tau}=1$). Month t is classified as an economic disaster if the crash-index of month t is large than the crash index of 98.5% of the months in the period $1896-2009$. The crash index is described in Section 2.6, and is the product of market return in the month t and economic growth in a six month window succeeding month t for months which the market return is negative. t-statistics are Newey-West corrected with number of lags/leads equal to the size of the disaster forecasting window.

References

Abel, Andrew B, Janice C Eberly, and Stavros Panageas, 2007, Optimal inattention to the stock market, *American economic review* 97, 244–249.

Antweiler, Werner, and Murray Z. Frank, 2004, Is all that talk just noise? the information content of internet stock message boards, *Journal of Finance* 59, 1259–1293.

Backus, D., M. Chernov, and I. Martin, 2011, Disasters implied by equity index options, *Journal of Finance* 66, 1969–2012.

Baker, Scott, Nicholas Bloom, and Steven Davis, 2013, Measuring economic policy uncertainty, Working Paper 13-02 Chicago Booth.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Barro, R.J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823–866.

——— , E. Nakamura, J. Steinsson, and J.F. Ursua, 2011, Crises and recoveries in an empirical model of consumption disasters, Working paper.

Barro, R.J., and J.F. Ursua, 2008, Consumption disasters in the twentieth century, *American Economic Review* 98, 58–63.

Barro, Robert J., 2009, Rare disasters, asset prices, and welfare costs, *American Economic Review* 99, pp. 243–264.

Bates, David S., 2012, U.s. stock market crash risk, 1926–2010, *Journal of Financial Economics* 105, 229–259.

Bollerslev, T., and V. Todorov, 2011, Tails, fears, and risk premia, *Journal of Finance* 66, 2165–2211.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2009, Understanding index option returns, *Review of Financial Studies* 22, 4493–4529.

Brown, Stephen J, William N Goetzmann, and Stephen A Ross, 1995, Survival, *Journal of Finance* 50, 853–873.

Cherkassky, V., and Y. Ma, 2004, Practical selection of svm parameters and noise estimation for svm regression, *Neural networks* 17, 113–126.

Cole, Harold L, and Lee E Ohanian, 1999, The great depression in the united states from a neoclassical perspective, *Federal Reserve Bank of Minneapolis Quarterly Review* 23, 2–24.

Croce, Mariano M., Thien T. Nguyen, and Lukas Schmid, 2012, The market price of fiscal uncertainty, *Journal of Monetary Economics* 59, 401 – 416.

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

——— , and A. Yaron, 2011, What's vol got to do with it, *Review of Financial Studies* 24, 1–45.

Engelberg, Joseph, 2008, Costly information processing: Evidence from earnings announcements, *Working paper.*

Gabaix, X., 2012, Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645–700.

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

García, Diego, 2013, Sentiment during recessions, *Journal of Finance* 68, 1267–1300.

Gentzkow, Matthew, and Jesse M. Shapiro, 2006, Media bias and reputation, *Journal of Political Economy* 114, pp. 280–316.

Gourio, Francois, 2008, Time-series predictability in the disaster model, *Finance Research Letters* 5, 191–203.

———, 2012, Disaster risk and business cycles, *American Economic Review* 102, 2734–2766.

Hansen, Lars Peter, and Robert J Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *The Journal of Political Economy* pp. 829–853.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, 2009, *The elements of statistical learning* (Springer) second edition edn.

Hoberg, Gerard, and Gordon Phillips, 2010, Product market synergies and competition in mergers and acquisitions: A text-based analysis, *Review of Financial Studies* 23, 3773–3811.

———, 2011, Text-based network industries and endogenous product differentiation, *Working paper.*

Huang, Lixin, and Hong Liu, 2007, Rational inattention and portfolio selection, *Journal of Finance* 62, 1999–2040.

Kelly, Bryan T., 2012, Tail risk and asset prices, *Working Paper.*

Kogan, S., D. Levin, B.R. Routledge, J.S. Sagi, and N.A. Smith, 2009, Predicting risk from financial reports with regression, in *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics* pp. 272–280. Association for Computational Linguistics.

Kogan, S., B. Routledge, J. Sagi, and N. Smith, 2010, Information content of public firm disclosures and the sarbanes-oxley act, *Working paper.*

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen, 2012, Short-term debt and financial crises: What we can learn from us treasury supply, *Unpublished Working Paper.*

Loughran, T., and B. McDonald, 2011, When is a liability not a liability? textual analysis, dictionaries, and 10-ks, *Journal of Finance* 66, 35–65.

Lundblad, Christian, 2007, The risk return tradeoff in the long run: 1836–2003, *Journal of Financial Economics* 85, 123 – 150.

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression babies: Do macroeconomic experiences affect risk taking?, *Quarterly Journal of Economics* 126, 373–416.

Manela, Asaf, 2011, Spreading information and media coverage: Theory and evidence from drug approvals, Ph.D. thesis University of Chicago.

——— , 2014, The value of diffusing information, *Journal of Financial Economics* 111, 181–199.

Martin, Ian, 2011, Simple variance swaps, Discussion paper National Bureau of Economic Research.

Miller, G.A., 1995, Wordnet: a lexical database for english, *Communications of the ACM* 38, 39–41.

Mishkin, F.S., and E.N. White, 2002, Us stock market crashes and their aftermath: implications for monetary policy, *NBER Working paper.*

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

Newey, WK, and KD West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Noyes, A.D., 1909, A year after the panic of 1907, *Quarterly Journal of Economics* 23, 185–212.

Pastor, Lubos, and Pietro Veronesi, 2012, Uncertainty about government policy and stock prices, *Journal of Finance* 67, 1219–1264.

Pedersen, Ted, Siddharth Patwardhan, and Jason Michelizzi, 2004, Wordnet::similarity - measuring the relatedness of concepts, in Daniel Marcu Susan Dumais, and Salim Roukos, ed.: *HLT-NAACL 2004: Demonstration Papers* pp. 38–41 Boston, Massachusetts, USA. Association for Computational Linguistics.

Reinhart, Carmen M., and Kenneth S. Rogoff, 2011, From financial crash to debt crisis, *American Economic Review* 101, pp. 1676–1706.

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

Schularick, Moritz, and Alan M Taylor, 2009, Credit booms gone bust: monetary policy, leverage cycles and financial crises, 1870–2008, Discussion paper National Bureau of Economic Research.

Shiller, R., and W. Feltus, 1989, Fear of a crash caused the crash, *The New York Times* p. F3.

Sialm, Clemens, 2009, Tax changes and asset pricing, *American Economic Review* 99, pp. 1356–1383.

Tetlock, Paul, Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: Quantifying language to measure firms' fundamentals, *Journal of Finance* 63, 1437–1467.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62, 1139–1168.

Vapnik, N. Vladimir, 2000, *The Nature of Statistical Learning Theory* (Springer-Verlag, New York.).

Wachter, Jessica A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *Journal of Finance* 68, 987–1035.