Internet Appendix

War Discourse and Disaster Premia: 160 Years of Evidence from the Stock Market

David Hirshleifer

Dat Mai

Kuntara Pukthuanthong

Internet Appendix **3**

A Seeded Latent Dirichlet Distribution **3**

B Text Processing Steps **5**

C Additional Figures and Tables for Discourse Topics **7**

C.1 Predict Returns with Small-Sample Bias Correction 7

C.2 *War* versus Conditional Volatility and Skewness 9

C.3 *War* versus Actual Events 9

C.4 Predicting One-Month-Ahead Returns: Controlling for Economic Variables 11

C.5 Predicting One-Month-Ahead Returns: Controlling for Uncertainty and Sentiment Variables 12

C.6 *War* as a Proxy for Time-Varying Risk Aversion 13

C.7 Subperiod Predicting Power 17

C.8 Asset Allocation Implications 18

C.9 *War* and Predictability of Bond Returns 20

C.10 Predicting Returns on Characteristic Portfolios 22

D Robustness Checks **51**

D.1 Robustness Checks: sLDA 51

D.2 Robustness Checks: *War* Innovation 53
D.3 Robustness Checks: Empirical Design 53
D.4 Alternative Bootstrap Approach 57

E Topic Weights Constructed by Counts of Seed Words **77**

F Discourse Topics from the *WSJ* **85**

A Seeded Latent Dirichlet Distribution

In this appendix, we provide more details on the seeded latent Dirichlet distribution model. This paper uses a stochastic topic model to extract latent topic weights from news articles. Topic models are developed based on the core idea that documents are mixtures of topics, where each topic has a probability distribution over words (Blei 2012; Steyvers and Griffiths 2007). Under topic models, we assume that text documents derive from a stochastic generative process. The creation of a new document starts with a document-specific distribution over topics (the document-topic distribution). Each word in the document is chosen first by picking a topic randomly from the document-topic distribution, and then drawing a word from the topic-word distribution for that topic. To model this, every possible word needs to be assigned to a topic.

In this setup, the document-topic distribution for each document and topic-word distribution for each topic (the same across documents) are unobserved parameters that are estimated from the observable word frequencies in the document collection. In other words, we can use standard statistical techniques to estimate the generative process, inferring the topics responsible for generating a collection of documents (Steyvers and Griffiths 2007).

The most widely used topic model is latent Dirichlet allocation (LDA) as introduced by Blei, Ng, and Jordan (2003) and further developed by Griffiths and Steyvers (2004). Under LDA, a document d is generated under the following hierarchical process:

- The word weight vector ω_k of topic k is the vector of probabilities of each word value for the topic k. The prior for these weights is assumed to have a Dirichlet distribution governed by parameter β: $\omega_k \sim \text{Dirichlet}(\beta)$.[1]
- The topic weight in a document d, denoted τ_d, is a vector of topic probabilities, i.e., probabilities that any given word location in the document is about any given topic. The topic weight vector of document d follows a prior Dirichlet distribution governed by parameter α, the same for all documents: $\tau_d \sim \text{Dirichlet}(\alpha)$, the same for all documents.[2]

[1] To illustrate, suppose that topic k has three words: $word_1$, $word_2$, and $word_3$ with respective weights $\omega_k = [w_1, w_2, w_3]$ with $w_1 + w_2 + w_3 = 1$. The model assumes that this ω_k vector follows a Dirichlet distribution.

[2] Similarly, assume document d has four topics $topic_1$, $topic_2$, $topic_3$, $topic_4$ with the weights given to these topics captured by $\tau_d = [\theta_1, \theta_2, \theta_3, \theta_4]$ with $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 1$. The model assumes that this τ_d vector follows a Dirichlet distribution.

- We use v to indicate a word location in a given document, and w to indicate a word value (such as "the" or "cat"). For each word location v in document d, we
  - randomly select a topic from the document-topic distribution:

    $z_{dv} \sim \text{Multinomial}(\tau_d)$ (a distribution which does not depend on v), and then
  - randomly select from a word from that topic:

    $w \sim \text{Multinomial}(\omega_{z_{dv}})$.

In other words, it is the multinomial distribution of word values for the realized topic z_{dv}.

In this setup, the topic-word distribution ω_k and document-topic distribution τ_d are latent parameters that we want to estimate. Estimating these involves a backward inference based on observed word frequencies across documents. The parameters α and β are hyperparameters of the prior distribution whose values are taken from the Latent Dirichlet Distribution topic modelling literature.

The document-topic distribution τ_d is of utmost interest because it summarizes the attention allocated to each topic in each news article. To estimate these parameters using a Bayesian method, Griffiths and Steyvers (2004) specifies that ω_k and τ_d follow two Dirichlet distributions (these two are referred to as the "prior" distribution in Bayesian statistic). From these specifications, we can derive the distribution of the topic assignment z_{dv} conditioned on observed word frequencies (this conditional distribution is referred to as the "posterior" distribution). We then use Gibbs sampling to simulate this posterior distribution and estimate the two hidden model parameters.[3]

Users of the traditional unsupervised LDA developed by Blei, Ng, and Jordan (2003) and Griffiths and Steyvers (2004) only need to prespecify the number of topics K and let the model cluster words into these topics based on word frequencies in a completely unsupervised manner. Specifically, the LDA model is more likely to assign a word w to a topic k in a document d if w has been assigned to k across many different documents and k has been used multiple times in d (Steyvers and Griffiths 2007). The model automatically extracts underlying topics, so users of LDA have no control over topic assignments.

[3]Gibbs sampling is a sampling technique to simulate a high-dimensional distribution by sampling from lower-dimensional subsets of variables where each subset is conditioned on the value of all others. See Griffiths and Steyvers (2004) for details on the implementation of Gibbs sampling in LDA.

Since we are interested in uncovering some specific topics, we employ a recent extension of LDA called seeded LDA (sLDA) developed by Lu et al. (2011). sLDA allows users to regulate topic contents using domain knowledge by injecting seed words (prior knowledge) into the model. Precisely, under sLDA, we specify the topic-word distribution as follows:

$$\omega_k \sim \text{Dirichlet}\left(\beta + C_w\right)_{w \in V}, \tag{A.1}$$

where V is the corpus or text collection, $C_w > 0$ when w is a seed word in topic k and $C_w = 0$ when w is not a seed word. The higher is C_w, the stronger the tilt toward word w appearing in any given topic. Intuitively, sLDA gives preference to seed words w in topic k in the form of pseudo count C_w and clusters words into topics based on their co-occurrences with the seed words. When a seed word is not present in a text collection, it does not enter the sLDA model and has no impact on the estimation process.

Estimation is implemented by the seededla package in R and run on a high-performance computing (HPC) cluster. Following standard practice, we set $\alpha = 50/K$ where K is the number of topics, $\beta = 0.1$, and $C_w = 0.01$ times the number of terms in the corpus.

B Text Processing Steps

Before carrying out text cleaning, we first remove articles with limited contents, i.e., articles containing mostly numbers, names, lists, programs, etc.

We manually check and infer title patterns that indicate limited content. About 1.4 million articles have limited content out of the total 14.7 million articles as shown in Table C.1. List of exclusion patterns are available from authors on requests.

Next, we conduct the following text cleaning steps:

[1] Remove articles with fewer than 100 content words. We consider content words as those outside of the expanded stop word list of 3,346 words developed by Professor Matthew L. Jockers. This list is available at https://www.matthewjockers.net/macroanalysisbook/expanded-stopwords-list/. We append this list with full and abbreviated day and month names (e.g., Monday, Mon, November, Nov, etc.).

[2] Turn all words into lower case and remove Unicode code points, HTML tags, hashtags, URLs, one-letter words, and words containing three or more repeating letters.

[3] Lemmatize texts using part-of-speech tags. Part-of-speech tagging and lemmatization are conducted using the nltk library in Python.

[4] Tokenize texts into unigrams, bigrams, and trigrams within sentence punctuation boundaries. In natural language processing, "tokenize" means breaking documents into words or "tokens." "Unigram" refers to a one-word token, "bigram" a two-word token, and "trigram" a three-word token. Collectively, "ngram"refers to an n-word token. To create sensible ngrams, it is essential to retain punctuations before tokenization. Keeping punctuations and stop words before creating n-grams ensures that our ngrams are present in the corpus. An alternative approach is to tokenize texts after removing punctuations and stop words. However, this approach results in n-grams that do not appear in the documents, thus distorting the original thematic contents of the document.

[5] Remove unigrams of fewer than three letters or being a stop word and bigrams containing stop words. We also remove trigrams containing stopwords unless the stop word is a preposition in the middle position. For example, under within-punctuation boundary tokenization, the sentence "Under current favorable conditions, the revenue of firm A will double next year." is converted into the following unigrams [current, favorable, condition, revenue, firm, double, year], bigrams [current_favorable, favorable_condition], and trigrams [current_favorable_condition, revenue_of_firm] where all stop words and words of less than three characters have been removed. Meaningless ngrams that have stop words on the boundaries such as under_current_favorable (which does not add any additional meaning to current_favorable) have been removed while revenue_of_firm is retained. We also experiment with keeping stop words with future meaning, such as [will, might, could, should, possible, likely, forward, future, pending, etc.], and obtain similar results.

[6] Each month t, with news articles over the past ten years up to and including month t, we create a document-frequency matrix where each row is a document (article), each column is an ngram or token, and each entry is the count of the token in that document. We put all ngrams

into one document-frequency matrix. To mitigate the impact of outliers on document-topic distribution, we remove tokens appearing in fewer than 0.2% and tokens appearing in more than 90% of all documents during each estimation window.

C Additional Figures and Tables for Discourse Topics

This appendix reports additional tables and figures for the discourse topics constructed from the *NYT*. Figure C.1 plots the monthly count and monthly article length of our *NYT* data set. Figure C.2 shows the word clouds for the remaining eight topics, and Figure C.3 shows their time series.

Table C.1 reports the number of *NYT* articles left after each screening step. Table C.2 reports the correlation matrix of the topics and the PLS index.

C.1 Predict Returns with Small-Sample Bias Correction

Table 2 reports an autocorrelation of 0.85 for the war topic index, which is notably high, and 0.70 for the PLS index. This high level of predictability from the previous month's index suggests that the actual changes or innovations in the topic indexes might be more pertinent for predicting market returns than their persistently high autocorrelated levels. Other indexes, such as those for pandemic and confidence, exhibit lower autocorrelations, raising fewer concerns about their predictive validity. Notably, the index producing the most accurate forecasts also shows the highest autocorrelation.

Therefore, we examine whether there is a bias from time series regressions of returns on persistent predictors such as *War* and the PLS index. This is the well-known Stambaugh bias, which applies under conditions outlined in equation (18) of Stambaugh (1999).

Stambaugh describes three key conditions related to bias:

(1) Sample size T: Bias is decreasing with larger T;

(2) Autocorrelation ρ: Bias depends on $1 + 3\rho$, increasing with ρ near 1;

(3) Innovation-residual correlation: Bias depends on p_{uv}/p_v^2 where p_{uv} is the covariance between the regression disturbance and predictor innovation, and p_v^2 is the variance of the predictor innovation.

Our sample size is very large and the correlation between the innovation and residual is small (-0.016 over 1871-2019, compared to -0.96 for dividend-price ratio over the same period). Thus, the Stambaugh bias should be small.

Nevertheless, we address Stambaugh bias with further tests. The typical practice in the return forecasting literature is to use test statistics that directly adjust for Stambaugh bias, rather than replacing predictor variables with their innovations. Intuitively, this approach provides a theoretically motivated test, and avoids throwing away any information.[4]

Furthermore, any concern over spurious in-sample predictability due to high autocorrelation of the regressor is alleviated by the substantial and significant out-of-sample R^2.

We present the results of our tests with the small sample bias correction according to Stambaugh (1999) in Table C.4 and Table C.5 (we use the beta and standard error correction for long-horizon bias in Boudoukh, Israel, and Richardson (2022) which includes the one-period case of Stambaugh (1999)). The significance of the result changes little. So, consistent with expectations, the small sample bias is small in our tests.

As an ancillary analysis, we also perform tests using innovation as a predictor. We measure innovation as residual from an AR(1) process (using ARMA(1,1) or AR(2) yields similar results). The results are consistent for both in-sample and out-of-sample. Over the whole sample, the in- and out-of-sample predictability of *War*'s innovation is comparable to that of *War*'s level. Over the past 20 years, the predictive power of *War*'s innovation is weaker than *War*. However, it is still significant with t-statistic of 1.90. As for the PLS index, the innovation of the PLS is stronger economically and statistically than the *PLS* index in-sample in all periods. However, as with *War*, the PLS index is weaker out-of-sample during the past twenty years. See Section 6.9 and Table D.7.

To sum up, we have addressed concerns over the autocorrelation of *War* and the PLS index in predictive regressions in four ways: (1) On conceptual grounds the bias is expected to be small; (2) we verify the findings using Stambaugh (1999) bias-adjusted statistics, (3) we verify the findings with out-of-sample analysis, and (4) we perform ancillary robustness tests with innovations.

[4]More specifically, according to the rare disaster risk theory, the stock market prices the level of risk of rare events over the long term. Variation in this level can come from both slow-moving and fast-moving components, so it is not just innovations that should be priced.

C.2 *War* versus Conditional Volatility and Skewness

Our *War* index could possibly capture conditional market return, volatility, and skewness that have market return predictability. To investigate this possibility, we re-run our predictive regression and control for variables:

$$R^e_{t+1} = \alpha + \beta War_t + \gamma z_t + \epsilon_{t+1}, \tag{C.1}$$

where z_t is either the current market excess return, conditional volatility, or conditional skewness. We construct the monthly conditional market volatility, $\hat{\sigma}_t$, from daily returns as follows:

$$\hat{\sigma}_t = \sqrt{\frac{1}{n_t - 1}\sum_{\tau=1}^{n_t}(R_\tau - \bar{R}_t)^2}, \tag{C.2}$$

where n_t is the number of trading days in month t, R_τ is the daily return and $\bar{R}_t$ is the average daily returns in month t. Similarly, we construct the monthly conditional skewness as follows:

$$sk_t = \frac{n_t}{(n_t - 1)(n_t - 2)}\sum_{\tau=1}^{n_t}\Big(\frac{R_\tau - \bar{R}_t}{\hat{\sigma}_t}\Big)^3, \tag{C.3}$$

where following standard practice we scale the raw central third moment by the standard deviation. Because daily data on the S&P 500 index becomes available in January 1928, our sample is from January 1928 to October 2019.

We report the results in Table C.6. Panel A reports the results for the whole sample from 1928 to 2019. We find that the predictability of *War* is not affected by any of the return moments. When we control for all three moments in the last column, the predictive power of *War* is still intact. We obtain similar results over two subsamples: 1950-2019 and 2000-2019. This result confirms that the predictability of *War* does not come from other return moments.

C.3 *War* versus Actual Events

As *War* is constructed to be the attention paid to wars and tensions, it is interesting to examine whether *War* has the predictive power beyond the actual rare disasters. To answer this question, we first create indicators for these events reported by GFD:

- Recessions: from NBER;

- Bank failures: if the event is tagged as bank failure, War, or crime;
- Wars: if the event is tagged as war, military, revolution, assassination, rebellion, insurrection, riot, terrorism, battle, or invasion;
- Disasters: if the event is tagged as disaster, earthquake, weather, tornado, hurricane, or typhoon;
- Epidemics: if the event is tagged as epidemic or pandemic;
- All: if the event is tagged with any of the above.

Figure C.4 plots these events over the past 150 years.

We then include these event indicators as controls in the predictive regression:

$$R^e_{t+1} = \alpha + \beta \times War_t + \gamma^j \times D^j_t + \epsilon_{t+1}, \tag{C.4}$$

where D^j_t is a dummy variable for the event j equal to one if there is one event j in month t. If *War* contains additional predictive power, β is expected to be significantly positive.

Panel A of Table C.7 reports the results for the whole sample. Across all events, *War* remains significant as a return predictor. Among the events, only Recessions and Epidemic yield significant prediction coefficients at the 5% and 10% levels, respectively. The prediction slope on Recessions is negative and is thus inconsistent with a risk-based explanation. The results indicate that the actual events themselves, except Recessions, have limited predictive power and therefore cannot be a cause of fluctuations in RRA. This evidence rules out the possibility that *War* only reflects RRA changes triggered by real-world stressful events.

Panel B reports the results in the first half of the sample from 1871 to 1949. During this period, *War* remains significant against Bank Failures, Disasters, and Epidemic. During the second half of the sample, *War* remains significant at least 5% level across all events and drives out the significance of Recessions.

Overall, the findings in this subsection eliminate the alternative explanation that the predictability of *War* from news articles is simply a manifestation of actual events. Indeed, we find that most of the events have no predictive power. Thus, it is undoubtedly the narrative aspects of the events that matter for the stock market.

C.4 Predicting One-Month-Ahead Returns: Controlling for Economic Variables

In the main text, we found that *War* outperforms 5 economic variables, 13 topics, and numerous crisis event count as indexes in predicting next month market returns. In this subsection, we extend the list of economic variables and consider the following bivariate predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \gamma z_t + \epsilon_{t+1}, \tag{C.5}$$

where z_t is one of the economic predictors. Following Huang, Li, and Wang (2020), we include as economic predictors the 14 variables from Goyal and Welch (2008), the output gap from Cooper and Priestley (2009), and the short interest index from Rapach, Ringgenberg, and Zhou (2016).

The "Univariate" column in Panel A of Table C.9 reports the results of single predictive regressions when each of the 16 economic variables is used alone to predict the next month's excess market return. As shown in Goyal and Welch (2008), most of these variables are not significant as a market predictor. The Treasury Bill rate and short interest are negative predictors, significant at the 5% level. At the same time, the long-term bond return is the only significant positive predictor, marginally significant at the 10% level. The last row reports the prediction results with a PLS index constructed with 16 economic variables (hereafter, the economic PLS index). It significant at the 10% level.

In the "Bivariate" column of Panel A, the topic PLS index is tested against each economic predictor in bivariate regressions. The PLS index is used instead of *War* as the former inherits the latter's features and is a stronger predictor. The PLS index remains significant at the 5% level or stronger against the 16 economic predictors. Finally, when tested against the economic PLS index, the topic PLS index remains significant at the 10% level and drives out the significance of the economic index. Overall, the results indicate that the discourse topics contain substantial information for predicting market returns beyond that in standard economic predictors. Note that we construct the topic and economic PLS indexes using the whole sample: 1871 to 2019 for the topic PLS and 1973 to 2014 for the economic PLS.

C.5 Predicting One-Month-Ahead Returns: Controlling for Uncertainty and Sentiment Variables

In the main text, we document that the topic PLS index contains valuable insights into market returns. We now test whether the topic PLS index has incremental ability to predict returns in comparison with other well-known uncertainty or sentiment variables. To facilitate a fair comparison between the topic PLS indexes and other sentiment and disagreement indexes, we follow Huang et al. (2015) and Huang, Li, and Wang (2020) in using the whole sample to construct the topic PLS index studied in this section.

Recently, ample measures have been introduced into the literature, notably the financial and macro uncertainty indexes from Jurado, Ludvigson, and Ng (2015), the economic policy uncertainty index from Baker, Bloom, and Davis (2016), and the disagreement index from Huang, Li, and Wang (2020). Another commonly used measure of uncertainty is the Chicago Board Options Exchange's Volatility Index (VIX).

Another strand of the predictability literature studies sentiment measures. The most influential of these is the investor sentiment index of Baker and Wurgler (2006) (hereafter BW sentiment), which has been documented to have predictability over small and hard-to-value stocks. Huang et al. (2015) extract the components most relevant to market returns from the BW sentiment using the PLS method to construct a powerful predictor (hereafter PLS sentiment). Jiang et al. (2019) construct manager sentiment from corporate filings to show that manager sentiment has predictability beyond what is captured by investor sentiment. Moreover, Tetlock (2007) and Garcia (2013) find that sentiment extracted from news articles predicts daily market returns. To construct news sentiment from the *NYT*, we compute the difference between the percentages of positive and negative words belonging to the sentiment dictionary developed in Loughran and McDonald (2011) (the most well-known sentiment dictionary in finance research). Finally, we also include the two U.S. stock market confidence indexes introduced by Shiller: the one-year confidence index and the crash confidence index.[5]

[5]These indexes are available at https://som.yale.edu/faculty-research-centers/centers-initiatives/international-center-for-finance/data/stock-market-confidence-indices/united-states-stock-market-confidence-indices.

The "Correlations" column of Panel B of Table C.9 reports the pairwise correlations between the topic index and each uncertainty and sentiment index. The topic PLS index has a significant 26% correlation with the economic policy uncertainty in Baker, Bloom, and Davis (2016) and a significant -19% correlation with the manager sentiment index in Jiang et al. (2019). Furthermore, the topic PLS index is significantly negatively correlated with Shiller's one-year confidence index (correlation -30%).

The "Univariate" column of Panel B reports the univariate prediction for each uncertainty and sentiment variable. We find that the disagreement index, PLS sentiment index, and manager sentiment index show strong prediction results, consistent with previous papers. The financial uncertainty index by Jurado, Ludvigson, and Ng (2015) is a negative predictor, significant at the 5% level. The economic policy uncertainty index of Baker, Bloom, and Davis (2016) is not significant in a one-month regression, consistent with previous studies (see, e.g., Pástor and Veronesi (2013) and Brogaard and Detzel (2015)).

In the "Bivariate" column of Panel B, we test the topic PLS index against the other sentiment and uncertainty variables. The PLS index remains significant in each multivariate predictive regression (at the 10% level or stronger). The last row of Table C.9 reports the results with the PLS index constructed from all the uncertainty and sentiment variables, against which the PLS index remains significant at the 5% level.

The results in Table C.9 indicate that the discourse topic index contains valuable information about market returns after controlling for the strong market predictors recently proposed in the literature.

C.6 *War* as a Proxy for Time-Varying Risk Aversion

In the main text, we show that *War* captures rare disaster probability as *War* is a positive market predictor, and innovations in *War* command a negative risk premium. These results are consistent with the predictions of the rare disaster risk model. In this section, we further show that *War* proxies for time-varying risk aversion, lending empirical support to the ICAPM. We first briefly discuss the ICAPM framework and then present the empirical results.

Before hypothesizing that *War* captures time-varying risk aversion, we first briefly introduce the Merton (1973)'s ICAPM model. In his seminal paper, Merton (1973) derives the following classic risk-return trade-off between the conditional mean of the return on the wealth portfolio, $\mathbb{E}_t[R_{M,t+1} - R_{f,t+1}]$, its conditional volatility, $\sigma^2_{M,t}$, and its conditional covariance with the investment opportunity set, $\sigma_{MF,t}$:

$$\mathbb{E}_t[R_{M,t+1} - R_{f,t+1}] = \left[\frac{-J_{WW}W}{J_W}\right]\sigma^2_{M,t} + \left[\frac{-J_{WF}}{J_W}\right]\sigma_{MF,t}, \tag{C.6}$$

where $J(W(t), F(t))$ is the indirect utility function in wealth, $W(t)$, and any state variables, $F(t)$, describing the evolution of the investment opportunity set over time. The term $\lambda \equiv \left[\frac{-J_{WW}W}{J_W}\right]$ (subscripts denote partial derivatives) is linked to the measurement of relative risk aversion (RRA) and is expected to be positive. Hence, the first term in Equation (C.6) captures the positive risk-return trade-off in which market participants require a higher risk premium on the wealth portfolio when its payoff is expected to be more uncertain. The second term in Equation (C.6) links the risk premium on the wealth portfolio to innovations in the investment opportunity set. Accordingly, investors will demand a higher risk premium on a wealth portfolio that pays off precisely in states where the marginal utility of wealth is low. The converse is true when the wealth portfolio serves as a hedge against investment risks.

Following Lundblad (2007) and the majority of papers in this literature, we consider a univariate version of Equation (C.6):

$$\mathbb{E}_t[R_{M,t+1} - R_{f,t+1}] = \lambda_0 + \lambda_1 \times \sigma^2_{M,t}, \tag{C.7}$$

where we assume that the investment opportunity set is constant or that the representative investor has a log utility function. A natural step then is to empirically test the univariate risk-return trade-off as depicted in Equation (C.7) with the popular GARCH-in-mean framework developed Bollerslev (1986) and Engle and Bollerslev (1986). Specifically, we consider first the following mean equation for the return-volatility trade-off:

$$R_{M,t+1} - R_{f,t+1} = \lambda_0 + \lambda_1 \times \sigma^2_{M,t} + \epsilon_{t+1}, \tag{C.8}$$

where ϵ_{t+1} has a mean of zero with conditional variance $\sigma^2_{M,t}$. Empirical tests of Equation (C.8)

on the U.S. stock market return have yielded mixed results, depending on the sample period and the specification of the volatility equation. Lundblad (2007) reconciles the contradictory findings on the U.S. risk-return trade-off present in the literature. He employs a long sample of U.S. stock market returns and documents a strong positive trade-off. He notes that a weak empirical relation may be an artifact of small samples and hence emphasizes the use of large samples in studying the risk-return relationship.

The specification in Equation (C.7) and Equation (C.8) assumes that the coefficient of relative risk aversion, λ_1, is time-invariant. However, we have no compelling reason to believe this assumption would hold in practice. Indeed, relative risk aversion is modeled as time-varying in several asset pricing models, such as the external habit model by Campbell and Cochrane (1999). If we assume time-varying relative risk aversion, then we can specify the risk-return trade-off as a linear function of some state variable, x_t:

$$R_{M,t+1} - R_{f,t+1} = \lambda_0 + (\lambda_1 + \lambda_2 \times x_t) \times \sigma^2_{M,t} + \epsilon_{t+1}. \tag{C.9}$$

We hypothesize that War proxies for time-varying relative risk aversion, and, thus, we replace the state variable, x_t, with War in Equation (C.9). Hence, $\lambda_t = \lambda_1 + \lambda_2 \times War_t$. If this hypothesis holds with real-world data, then we expect (1) the adjusted R^2 of Equation (C.9) to be higher than that of (C.8), as the former is a more proper representation of the risk-return trade-off, and (2) the coefficient λ_2 in Equation (C.9) to be significantly positive as risk aversion is expected to rise when War is high.

To complete the GARCH-M framework, we need a specification for the conditional volatility equation. Following Lundblad (2007), we consider four different volatility specifications, namely, GARCH (Bollerslev, 1986), IGARCH (Engle and Bollerslev, 1986), TGARCH (Zakoian, 1994), and

EGARCH (Nelson, 1991):

$$\begin{aligned}
\text{GARCH}(1,1): \quad & \sigma^2_{M,t} = \delta_0 + \delta_1 \epsilon^2_t + \delta_2 \sigma^2_{M,t-1} \\
\text{IGARCH}(1,1): \quad & \sigma^2_{M,t} = \delta_0 + \delta_1 \epsilon^2_t + (1-\delta_1) \sigma^2_{M,t-1} \\
\text{TGARCH}(1,1): \quad & \sigma^2_{M,t} = \delta_0 + \delta_1 \epsilon^2_t + \delta_3 D_t \epsilon^2_t + \delta_2 \sigma^2_{M,t-1} \\
\text{EGARCH}(1,1): \quad & ln(\sigma^2_{M,t}) = \delta_0 + \delta_1 \left(\frac{|\epsilon_t|}{\sigma_{M,t}} \right) - \delta_3 \left(\frac{\epsilon_t}{\sigma_{M,t}} \right) + \delta_2 ln(\sigma^2_{M,t-1}),
\end{aligned} \tag{C.10}$$

where D_t is an indicator equal to one when ϵ_t is negative and zero otherwise.

Panel A of Table C.8 reports the results using the standard GARCH(1,1) model. Over the whole 150-year sample, the coefficient of RRA, λ_1, is 2.17, significant at the 1% level. Hence, we observe the positive risk-return trade-off with a large sample size. However, the adjusted R^2 is negative at -0.38% as the conditional volatility is very smooth, failing to explain the variations in realized returns. These results are consistent with those of Lundblad (2007). Moving on to the time-varying RRA specification, if *War* proxies for time-varying RRA, we expect the interaction term λ_2 to be significantly positive and the conditional volatility to have higher explanatory power for return variations. The empirical results in Panel A confirm these conjectures. Specifically, λ_2 is 2.05, significant at the 1% level, and the adjusted R^2 jumps from -0.38% to 0.27%, indicating a better fit. Notably, the coefficient capturing constant RRA, λ_1, collapses toward zero.

We obtain similar results when decomposing the whole 150-year sample into two subsamples as in the previous tests of return predictability. In the first half of the sample, the time-varying RRA specification yields a better model fit as measured by R^2, and the coefficient λ_2 is significant at the 10% level. In the second subsample, R^2 jumps more than eight times, and λ_2 is significant at the 5% level under the time-varying RRA model.

Panels B, C, and D of Table C.8 report the results with different specifications for the volatility equation. We obtain consistent results across both the models and sample periods, except for EGARCH in the whole sample, confirming that *War* captures risk aversion, enhancing the risk-return relationship.

C.7 Subperiod Predicting Power

This subsection investigates the predictive power of discourse topics during different subsamples: expansion versus recession and high versus low sentiment. The literature seems to have reached a consensus that sentiment indexes can better predict the market during recessionary times (see, e.g., Garcia (2013), Huang et al. (2015), Jiang et al. (2019), among others). The intuition underlying this view is that the fear and anxiety investors feel related to the economic hardships during recessions increase their sensitivity to sentiment (Garcia, 2013).

The literature also shows that sentiment indexes have stronger predictability during high sentiment periods when mispricings are likely to occur because of short-sale constraints (Huang et al., 2015; Jiang et al., 2019; Stambaugh, Yu, and Yuan, 2012). Huang, Li, and Wang (2020) find that their disagreement index yields stronger predictability when sentiment is high: high disagreement leads to higher average bias and more overvaluation. This effect is stronger when investors are more optimistic (Huang, Li, and Wang, 2020). While these observations lean toward the behavioral channel, the predictability of our topics is more risk-based, so whether we can observe similar subsample concentrations in predictability remains unclear.

To examine the above question, we follow Rapach, Strauss, and Zhou (2010) and Huang et al. (2015), among others. We compute the subsample R^2 as follows:

$$R_c^2 = 1 - \frac{\sum_{t=1}^{T} I_t^c \left(\hat{\epsilon}_t\right)^2}{\sum_{t=1}^{T} I_t^c \left(R_t^e - \bar{R}^e\right)^2}, \quad c = exp,\ \ rec,\ \ high,\ \ low, \tag{C.11}$$

where I_t^c is an indicator that takes a value of one when month t is an expansion (recession) period or high (low) sentiment period; $\hat{\epsilon}_t$ is the fitted residual based on the in-sample predictive regression (2); $\bar{R}^e$ is the full sample mean of the excess market return; and T is the number of observations for the full sample of 1871–2019. We classify months into expansions and recessions based on the National Bureau of Economic Research (NBER) business cycles. For sentiment periods, we follow Stambaugh, Yu, and Yuan (2012) and Huang et al. (2015) and classify a month as high (low) sentiment if the Baker and Wurgler (2006) investor sentiment level in the previous month is above (below) is median value for the sample. Unlike the full sample R^2, the subsample R^2 can be positive or negative.

In the same spirit as Equation (C.11), we compute the out-of-sample R^2_{OS} for each period. Similar to the previous out-of-sample analyses, we use the expanding estimation window, and the evaluation period began in January 1891.

Panel A of Table C.10 reports the results with the in-sample R^2. For the in-sample results, we use the topic PLS index constructed using the whole sample. Accordingly, *War* and the PLS index yield higher R^2's during recessions (0.91% in recessions vs. 0.06% in expansions for *War*, and 1.80% in recessions vs. 0.58% in expansions for PLS). These results are consistent with the observation of concentrated predictive power during recessions documented in the literature. However, the out-of-sample R^2 with an expanding window in Panel B suggests both *War* and the PLS index have stronger predictive power in expansions (0.69% in expansions vs. -0.46% in recessions for *War*, and 0.10% in expansions vs. -0.31% in recessions for PLS). In sum, whether topics have stronger prediction power in recessions remains inconclusive.

We consistently find that discourse topics can better predict the market during low sentiment periods for both in-sample and out-of-sample analyses. For example, the in-sample R^2 for *War* is 0.50% during low sentiment periods versus 0.01% during high sentiment periods, while the figure for PLS is 1.90% versus -0.16%, respectively. For out-of-sample prediction, *War* yields an R^2_{OS} of 0.47% during low sentiment months versus -0.02% during high sentiment months, while the numbers for PLS are 0.85% and -0.66%. While this result is contradictory to the sentiment literature, it is intuitive. When people are in a bad mood, they are more receptive to stressful news.

In short, while we do not find evidence of different predicting powers of topics across the business cycles as commonly documented in the literature, we note that topics can better predict the market during low sentiment periods. This result is opposite to the sentiment literature. This further indicates that economic topics predict market outcomes via a different channel from sentiment.

C.8 Asset Allocation Implications

In this subsection, we examine the economic value of news topics from an asset allocation perspective. Following Campbell and Thompson (2008), we compute the certainty equivalent return (CER) gain and Sharpe ratio for a mean-variance investor who optimally allocates her portfolio

between the stock market and the riskfree asset using out-of-sample return forecasts.

At the end of period t, the investor optimally allocates

$$w_{t+1} = \frac{1}{\gamma} \frac{\hat{R}^e_{t+1}}{\hat{\sigma}^2_{t+1}} \tag{C.12}$$

of the portfolio to equities during period $t+1$, where risk aversion coefficient γ is set to three following Huang, Li, and Wang (2020),[6] $\hat{R}^e_{t+1}$ is the predicted excess return, and $\hat{\sigma}^2_{t+1}$ is the variance forecast. The investor then allocates $1-w_{t+1}$ of the portfolio to the riskfree asset. The $t+1$ realized portfolio return is

$$R^p_{t+1} = w_{t+1} R^e_{t+1} + R^f_{t+1}, \tag{C.13}$$

where R^f_{t+1} is the riskfree return. Following Campbell and Thompson (2008), we use a rolling window of 60 months to estimate the variance forecast of the excess market return, constrain w_t to be between 0 and 1.5 to exclude short sales, and allow a maximum 50% leverage.

The CER of the portfolio is

$$CER_p = \hat{\mu}_p - 0.5\gamma\hat{\sigma}^2_p, \tag{C.14}$$

where $\hat{\mu}_p$ and $\hat{\sigma}^2_p$ are the samples mean and variance, respectively, for the realized portfolio returns over the evaluation period. The CER gain is the difference between the CER for an investor who uses a forecasting model to predict the excess market return and the CER for an investor who uses the historical mean forecast. We annualize the CER gain by multiplying by 12 so that it can be interpreted as the maximum annual management fee the investor is willing to pay to gain access to the predictive forecasts. In addition to the CER gain, we compute the annualized monthly Sharpe ratios of the portfolio's realized returns. We test the statistical significance of the CER gain and the Sharpe ratios (against the historical mean benchmark) using the test statistics in DeMiguel, Garlappi, and Uppal (2009).

Panel A of Table C.11 reports the asset allocation results when individual predictors are used to make return forecasts. Treasury Bill produces the highest utility gains among all predictors during

[6] To conserve space, the results with a risk aversion coefficient of five are not reported but are similar to the reported results.

1881–2019 at 1.31%, while *Money* comes in second at 1.03% and *War* comes third at 0.86%. While Treasury Bill continues to deliver utility gains over each subsample, over the past 20 years, EP (the earnings price ratio) and *War* produce better allocation performance at 3.88% and 2.01%, respectively.

Panel B of Table C.11 shows that over 2000-2019, results via PLS yield a utility gain of 4.11%, the highest among all setups considered. Consistent with the OOS R^2 results, using all 14 topics yields superior performance to utilizing only narratives from Shiller (2019).

The right panel of Table C.11 shows the results for the annualized Sharpe ratio. Among all the individual predictors in Panel A, Treasury Bill yields the best results for the whole sample, followed by *Money* and *War*. Combining all economic predictors or topics via PLS only delivers significant results from 2000-2019. As a benchmark, the last row reports the annualized monthly Sharpe ratio from buying and holding the S&P 500 index in the corresponding periods. Allocations using the combination of topics outperform the buy-and-hold strategy in general.

Overall, the allocation results suggest that using return forecasts from *War* or the combination of discourse topics via PLS offers real-time economic values to investors. The economic gains increase over time, consistent with the R^2_{OOS} results.

C.9 *War* and Predictability of Bond Returns

We have shown that *War* is a positive stock market predictor. This is consistent with the disaster risk model (Barro 2006, 2009), or with a behavioral model in which rare risks are overestimated or overweighted in investors' expected utility functions. Gabaix (2012) theoretically shows that disaster risks should also affect bond risk premia. Specifically, disaster probabilities should increase risk premia on risky bonds such as long-term high-yield corporate bonds and decrease risk premia on safe bonds such as short-term government and investment-grade corporate bonds. A behavioral setting in which investors overweight rare risks suggests a similar prediction. We now test these predictions.

Specifically, we run the following predictive regression:

$$R^e_{t+1\to t+h} = \alpha + \beta War_t + \epsilon_{t+1\to t+h}, \tag{C.15}$$

where $R^e_{t+1\to t+h}$ is the annualized excess returns over the next h months on Treasury bond indexes, investment grade, and high-yield corporate bond indexes, and War_t, as before, is *War* standardized to zero mean and unit variance. We consider the prediction horizons of 1, 3, 6, and 12 months ahead. The Treasury bond indexes are from Datastream and available from December 1988 to October 2019; the corporate bond indexes are from S&P and available from January 1993 to October 2019. To measure the statistical strength of β, we report the Newey and West (1987) standard error with h lags.

Panel A of Table C.12 reports the results for Treasury bonds. We consider first the whole sample 1988-2019 over which the data for Treasury bond indexes are available. During this period, *War* does not have a discernible impact on government bonds' excess returns. We next consider the post-2000 period, in which *War* displays the most robust predictive power for stocks. During this sample, *War* negatively predicts excess returns over the next 1 up to 12 months ahead on short-term Treasury bonds having a maturity of up to 7 years. For longer-maturity Treasury bonds, *War* has no predictive power.

For investment grade corporate bonds in Panel B, over 1993-2019, *War* is a negative predictor of excess returns of bonds maturing in under one year, marginally significant at the 10% level. From 2000-2019, the impact is more substantial and powerful than in the longer sample period, for bonds with a maturity of 3 years or less.

For high-yield corporate bonds in Panel C, *War* positively predicts excess returns on bonds having 3 years or more until maturity over the whole sample 1993-2019. The most robust predictability is found in 5-7 year high yield bonds, significant at the 5% level for the next month's return. For the 2000-2019 sample, *War* is still a positive predictor of subsequent one-month excess returns on high-yield corporate bonds ranging from 3 to 10 years to maturity.

We also perform the out-of-sample predictions of bond excess returns using *War* and report the results in Table C.13. To facilitate comparison among different types of bond indexes, we limit

the sample to the range of January 1993 to October 2019. We employ an expanding estimation window using the first 10 years of the sample as the initial estimation window. We find, consistent with the in-sample results, that *War* has out-of-sample predictive power for returns of short-term government bonds (1-3 years to maturity) and investment-grade corporate bonds (0-1 years to maturity). *War* also yields significant R^2_{OS}'s for following one-month returns on high-yield corporate bonds having 3 to 10 years of maturity.

Overall, our bond prediction results are consistent with the predictions of the rare disaster risk model and with behavioral models in which rare risks are overweighted by investors. While the empirical disaster probability captured by *War* is a negative predictor of safe assets such as short-term government bonds and investment grade corporate bonds, it is associated with an increase in the return premia of risky investments such as stocks and mid- to long-term high-yield corporate bonds. The absolute coefficients on high-yield bonds are about seven to 15 times larger than those of investment-grade bonds with the same maturity.

C.10 Predicting Returns on Characteristic Portfolios

In the main text, we document that *War* and the discourse topic index predict market returns. In this appendix, we investigate whether the return predictability of topics holds at the individual portfolio level. Following Huang et al. (2015), we consider 40 characteristics-sorted portfolios, including 10 industry portfolios, 10 size portfolios, 10 book-to-market (BM) portfolios, and 10 momentum portfolios. The sample period for this analysis is from January 1927 to October 2019.

To examine the predictability of topics over the risk premium on the characteristics portfolios, we run the following predictive regression:

$$R^e_{i,t+1} = \alpha_i + \beta_i x_t + \epsilon_{i,t+1}, \quad i = 1, \ldots, 40 \tag{C.16}$$

where $R^e_{i,t+1}$ is the excess return on portfolio i, and x_t is either *War* or the PLS index.

Panel A of Table C.14 reports results with 10 industry portfolios. Both *War* and the PLS index yield positive slope coefficients across industries, although most of the prediction coefficients on *War* are insignificant. On the other hand, the PLS index can significantly predict returns on all

industries, with the strongest predicting powers found in Durable.

The rest of Table C.14 reports results with the size, BM, and momentum portfolios. Both *War* and the PLS index yield positive slopes for these portfolios, but the prediction coefficients on *War* are not as strong as those on the PLS index. The slopes on the ten-size portfolios increase monotonically from the large to small portfolios for both *War* and the PLS index. The topics also better predict value (high BM) and past loser stocks. Thus, returns on small, distressed (high BM) and, recently, underperforming stocks are more sensitive to War.

Figure C.1. *NYT* Article Count and Length

This figure plots the time series of the monthly total count and the monthly average length of articles in the *NYT*. Article length is measured as the sum of unigrams (one-word terms), bigrams (two-word terms), and trigrams (three-word terms) of each article. The sample period is from January 1871 to October 2019. Articles with limited content have been removed.

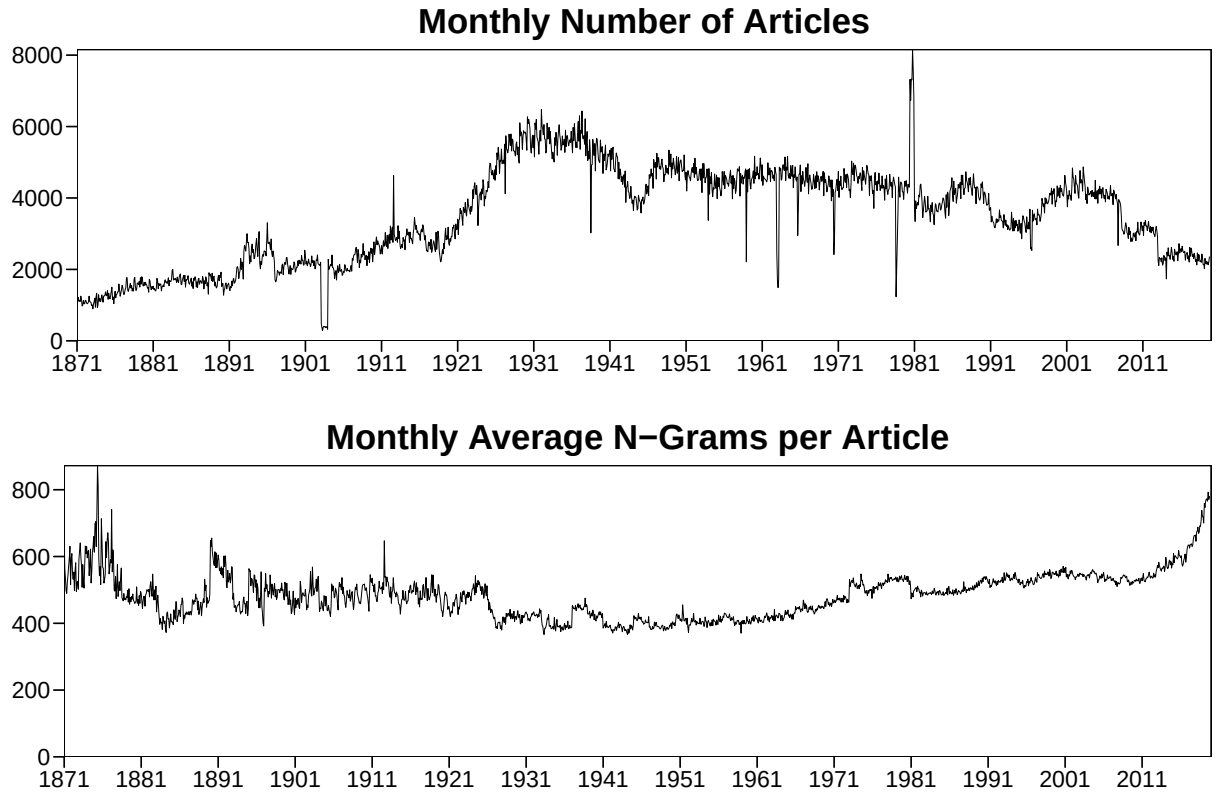


Figure C.2. Discourse Topic Contents

This figure plots the frequencies of n-grams per topic over time. Frequencies are constructed according to the sLDA model described in Section 3, and the size of each n-gram indicates its frequency. The sample period is from January 1871 to October 2019.

Confidence

Savings

Conspicuous Consumption

Technology Replacing Jobs

Real Estate Crashes

Stock Market Bubbles

Stock Market Crashes

Wage and Labor Unions

Figure C.3. Time Series of Topic Weights

This figure plots the time series of monthly topic weights constructed according to the sLDA model as described in Section 3. Topi weights are demeaned for ease of visualization. The shades indicate NBER-dated recessions. The sample period is from January 1871 to October 2019.

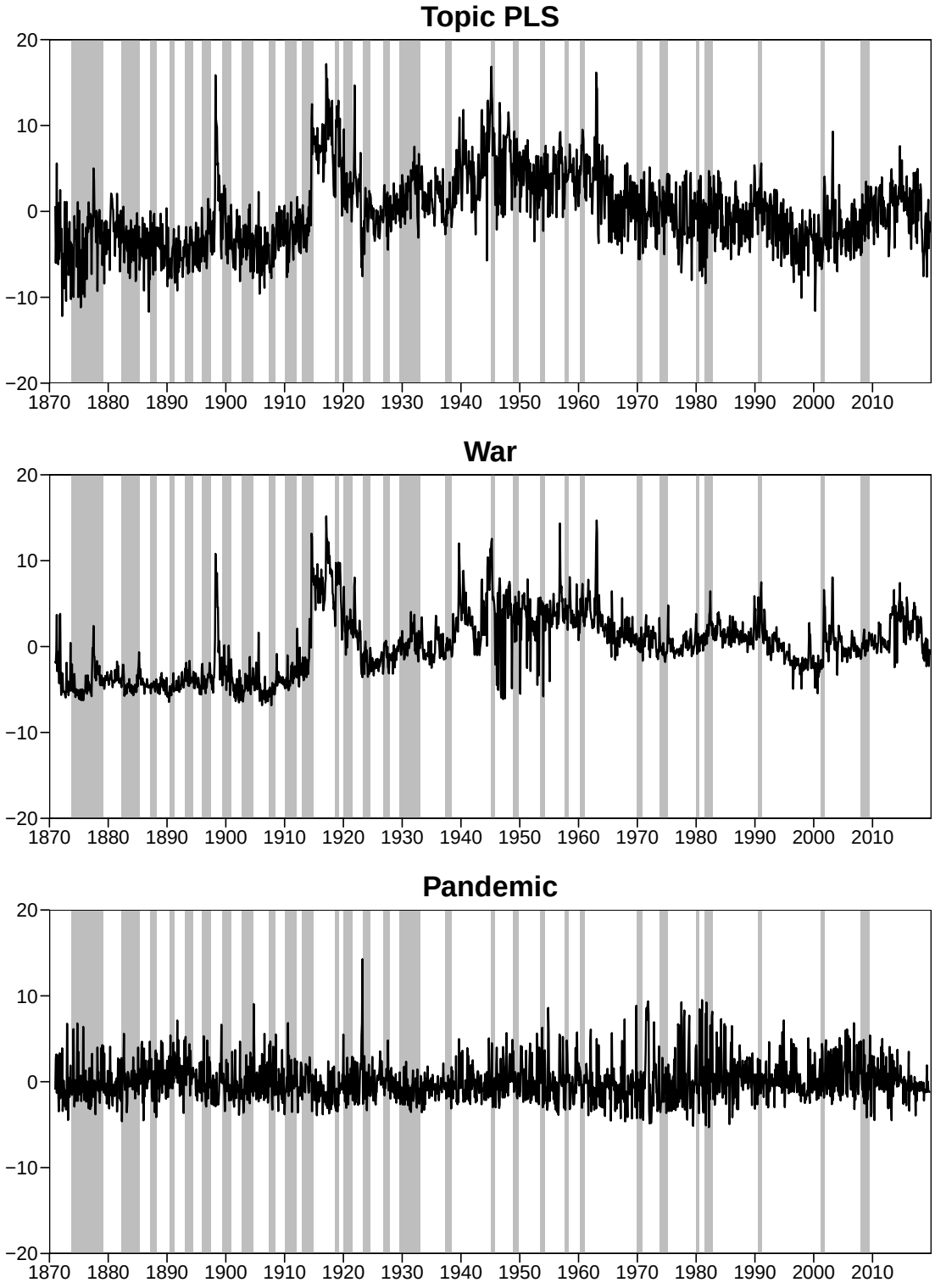


Figure C.3. Time Series of Topic Weights (Cont.)

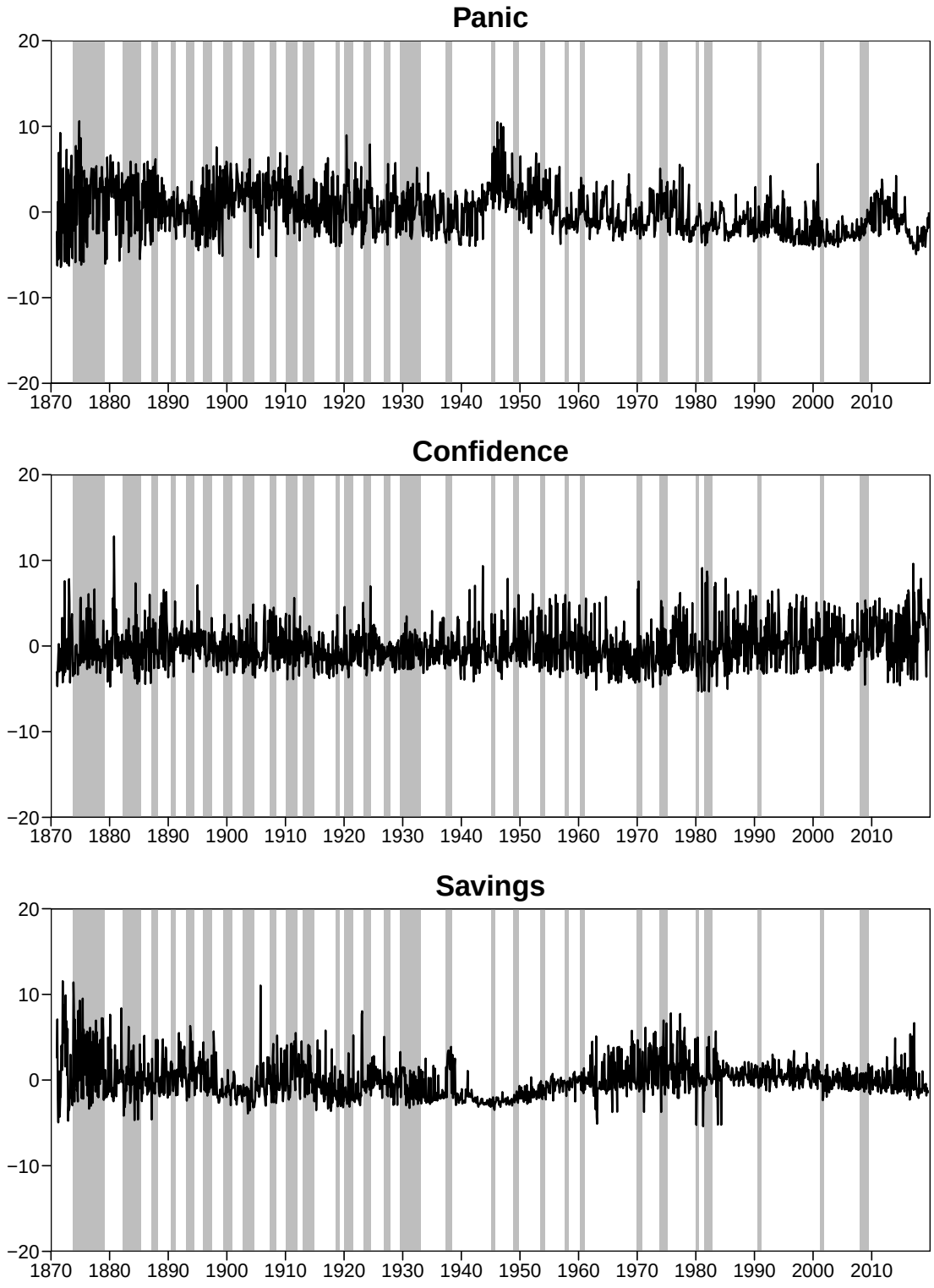


Figure C.3. Time Series of Topic Weights (Cont.)

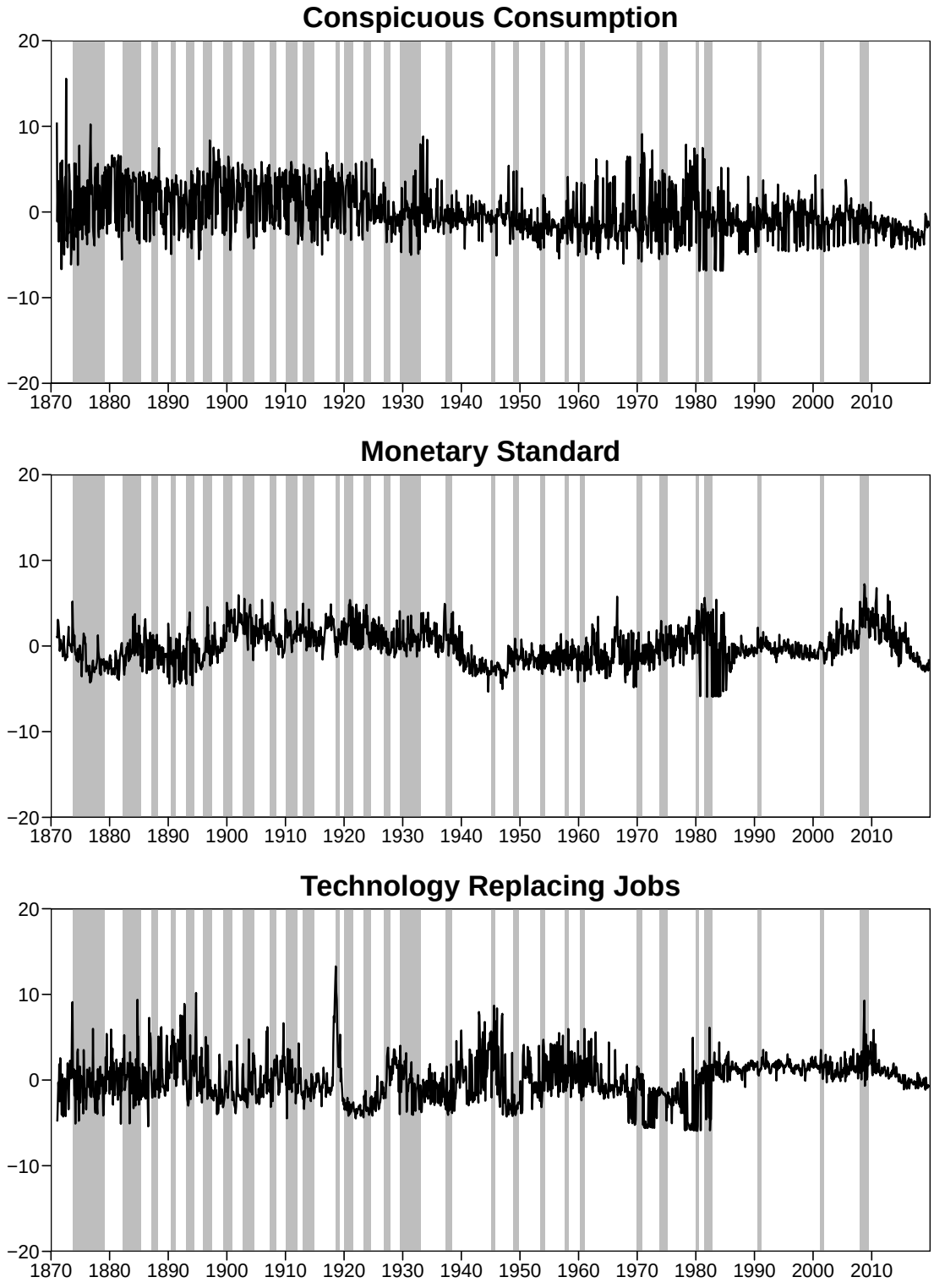


Figure C.3. Time Series of Topic Weights (Cont.)

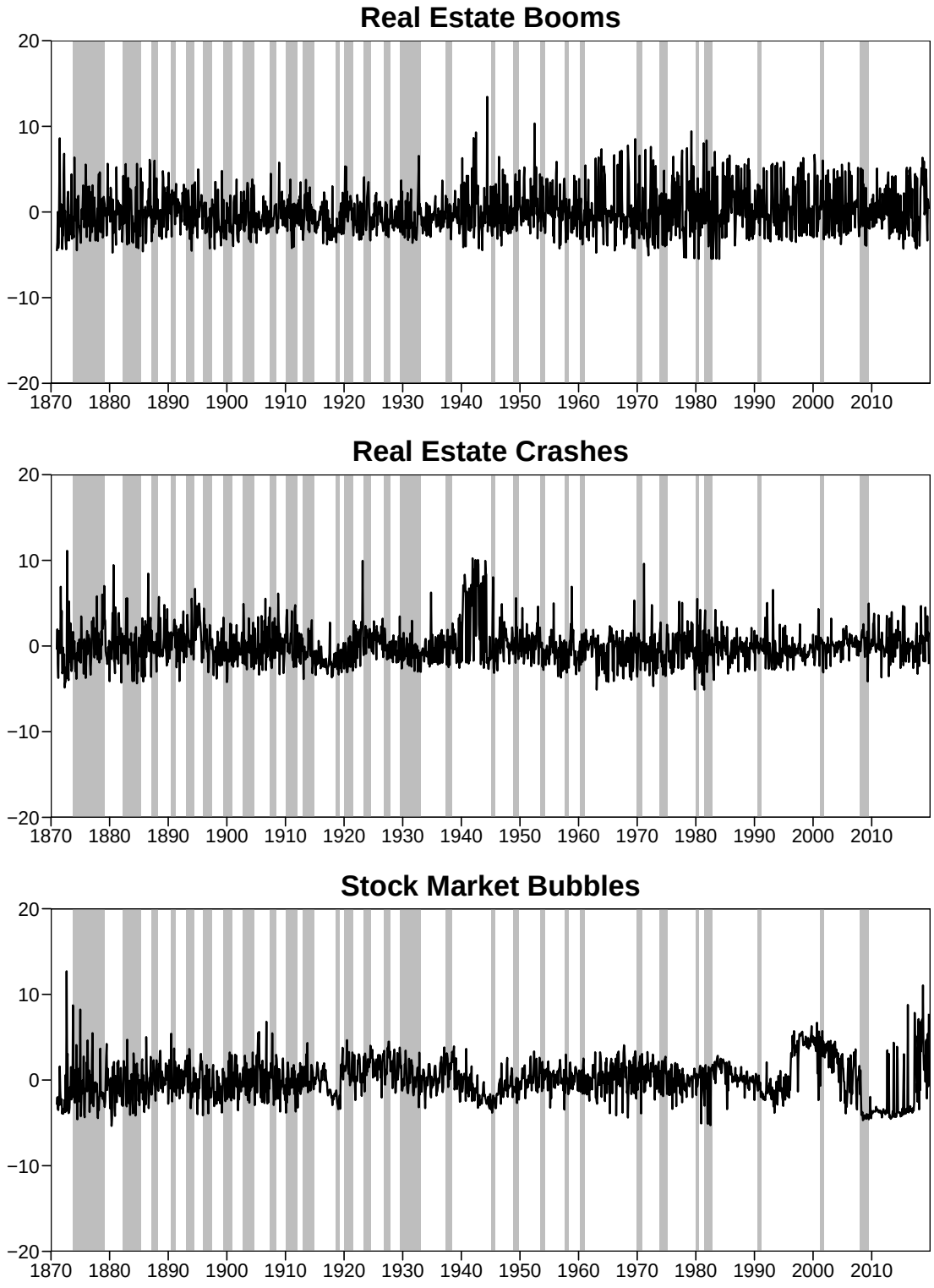


Figure C.3. Time Series of Topic Weights (Cont.)

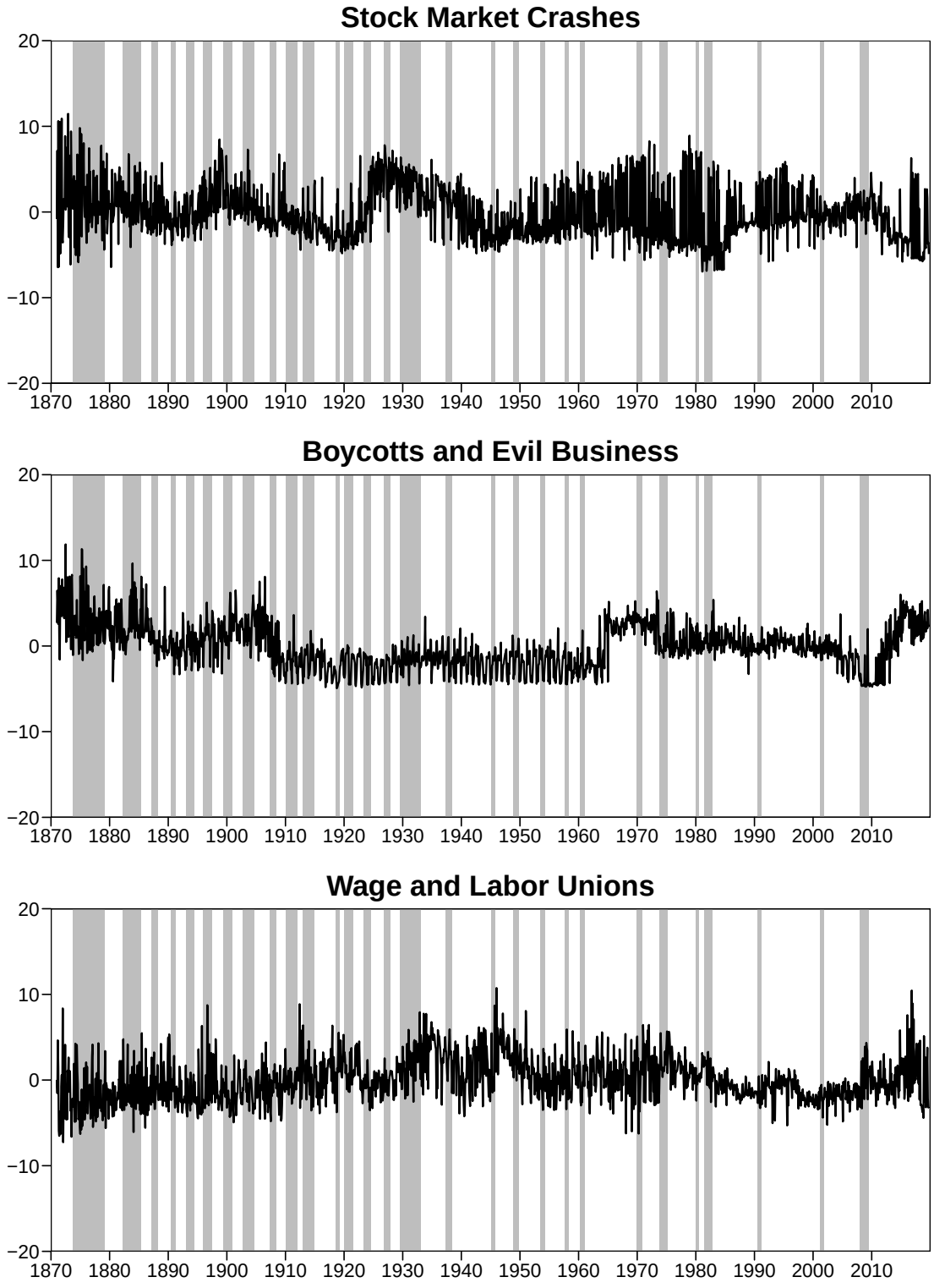


Figure C.4. Historical Events

This figure plots the time series of monthly historical U.S. events. The horizontal gray bar indicates at least one labeled event in that month. Recession dates are from NBER, while other event dates are from Global Financial Data. The sample period is from January 1871 to October 2019.

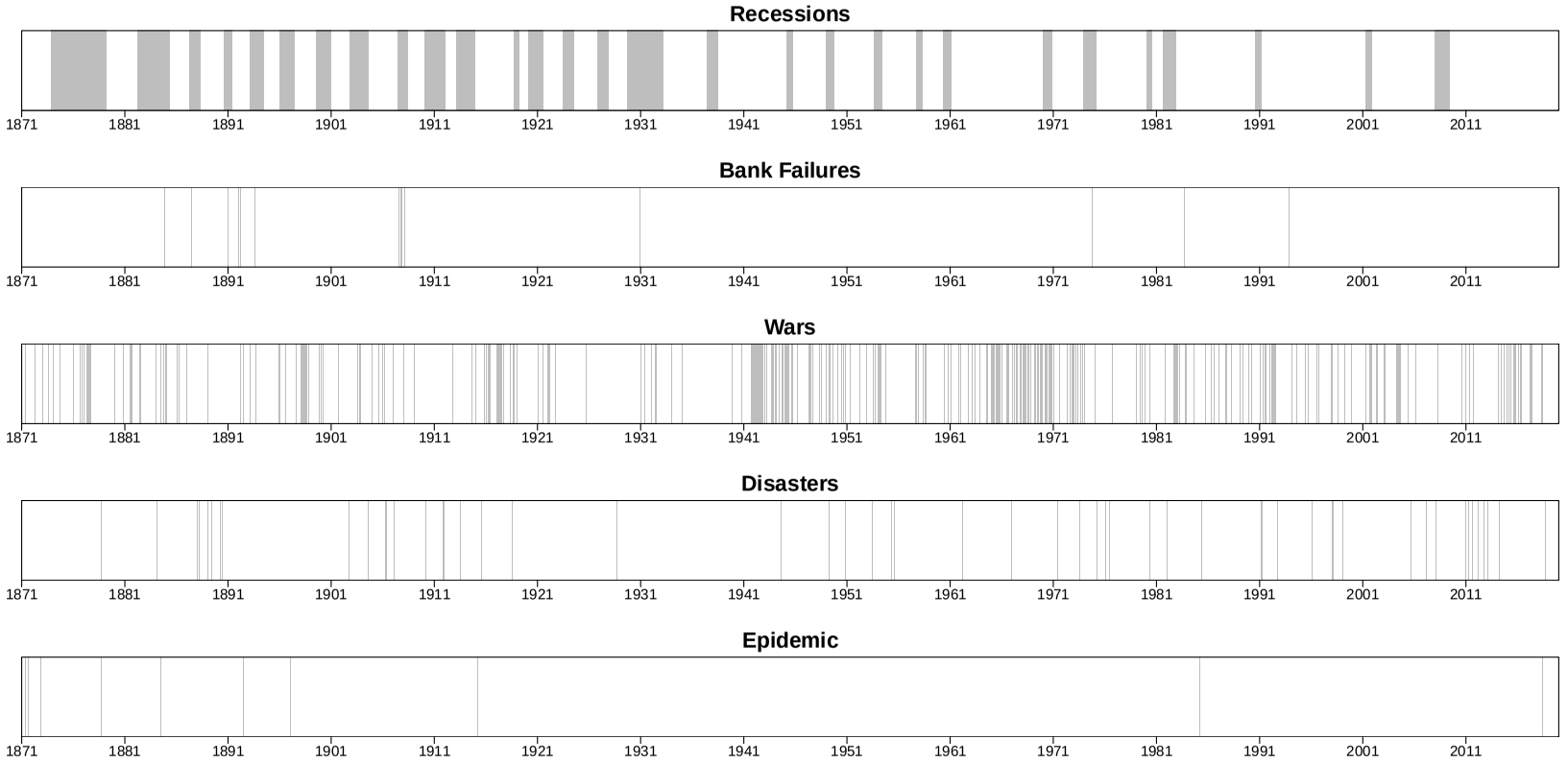


Table C.1
Data Screening

This table reports the number of *NYT* articles after each cleaning step. The whole sample is from January 1871 to October 2019.

Screening Steps	Number of Articles (Millions)
Original Sample	14.73
After dropping articles whose title indicates limited content	13.41
After further dropping articles having fewer than 100 content words	6.89

Table C.2
Topic Correlations

This table presents the pairwise correlations among 14 monthly topic weights constructed according to the sLDA model described in Section 3 and the PLS index. All numbers (except sample size) are expressed as percentages. The sample period is from January 1871 to October 2019.

	War	Pandemic	Panic	Confidence	Saving	Consumption	Money	Tech	Real Estate Boom	Real Estate Crash	Stock Bubble	Stock Crash	Boycott	Wage	PLS
War		-6.5	-17.1	-7.9	-22.2	-26.2	-10.0	-1.5	-1.7	-10.1	-10.1	-19.9	-32.3	22.7	82.1
Pandemic	-6.5		-6.5	-10.7	-2.2	-8.8	-2.3	-4.5	-7.3	-1.2	-3.7	-13.9	-2.2	-9.0	-28.5
Panic	-17.1	-6.5		-6.2	-14.9	5.3	-3.5	-8.2	-6.1	-3.0	-15.3	-14.1	-3.4	0.7	15.7
Confidence	-7.9	-10.7	-6.2		-3.3	-10.0	-8.0	6.6	-8.5	-1.6	-5.9	-20.0	-0.1	-11.5	-0.5
Saving	-22.2	-2.2	-14.9	-3.3		-5.8	-4.9	-6.8	-0.9	-7.3	-0.8	3.1	16.0	-22.3	-34.1
Consumption	-26.2	-8.8	5.3	-10.0	-5.8		10.4	-14.4	-17.9	-4.3	-7.0	-1.2	2.7	-13.5	-4.6
Money	-10.0	-2.3	-3.5	-8.0	-4.9	10.4		-14.6	-9.3	-12.4	-10.8	1.8	-13.5	5.9	-15.3
Tech	-1.5	-4.5	-8.2	6.6	-6.8	-14.4	-14.6		-4.1	-2.8	-8.6	-7.0	-14.2	-22.4	-8.1
Real Estate Boom	-1.7	-7.3	-6.1	-8.5	-0.9	-17.9	-9.3	-4.1		-1.8	-0.9	-17.4	-2.6	-6.7	-32.7
Real Estate Crash	-10.1	-1.2	-3.0	-1.6	-7.3	-4.3	-12.4	-2.8	-1.8		-7.7	-6.0	-10.0	-6.2	-1.1
Stock Bubble	-10.1	-3.7	-15.3	-5.9	-0.8	-7.0	-10.8	-8.6	-0.9	-7.7		-0.0	-0.8	-13.8	-14.7
Stock Crash	-19.9	-13.9	-14.1	-20.0	3.1	-1.2	1.8	-7.0	-17.4	-6.0	-0.0		-0.3	-5.6	-2.3
Boycott	-32.3	-2.2	-3.4	-0.1	16.0	2.7	-13.5	-14.2	-2.6	-10.0	-0.8	-0.3		-24.4	-41.7
Wage	22.7	-9.0	0.7	-11.5	-22.3	-13.5	5.9	-22.4	-6.7	-6.2	-13.8	-5.6	-24.4		37.3
PLS	82.1	-28.5	15.7	-0.5	-34.1	-4.6	-15.3	-8.1	-32.7	-1.1	-14.7	-2.3	-41.7	37.3	

Table C.3
Correlations between *War* and other Variables

This table presents the correlation between *War* and other variables including the uncertainty variables (financial and macro uncertainty indexes from Jurado, Ludvigson, and Ng (2015), economic policy uncertainty index from Baker, Bloom, and Davis (2016), disagreement index from Huang, Li, and Wang (2020), implied volatility (VIX), and news implied volatility (NVIX) from Manela and Moreira (2017)), sentiment variables (news sentiment, investor sentiment from Baker and Wurgler (2006), PLS sentiment from Huang et al. (2015), and manager sentiment from Jiang et al. (2019)), Shiller's confidence indexes (one-year confidence index and crash confidence index), log growth of trading volume of S&P 500, log growth of Google search volume, and log growth of RMI S&P 500 buzz words. Correlations are expressed as percentages. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

	Correlation with *War*	Sample Period
Financial uncertainty	-16.48 ***	196007-201910
Macro uncertainty	-8.04 **	196007-201910
Economic policy uncertainty	23.44 ***	198501-201910
Implied volatility (VIX)	-11.66 **	199001-201910
News implied volatility (NVIX)	-6.67 ***	188907-201603
Disagreement	-9.05 **	196912-201812
News sentiment	-4.00 *	186612-201910
Investor sentiment (BW)	-4.11	196507-201812
Investor sentiment (PLS)	-4.24	196507-201812
Manager sentiment	-28.00 ***	200301-201712
Shiller's one-year confidence index	-26.04 ***	200107-201910
Shiller's crash confidence index	-6.53	200107-201910
Trading volume growth	1.25	188802-201910
Google search volume growth	-4.33	200402-201910
SP500 buzz growth	-3.91	199802-201910

Table C.4
Predicting One-Month Market Returns With Correction for Small-Sample Bias

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month, x_t is one of the discourse topics or the PLS indexes, and β, the coefficient of interest, measures the strength of predictability. "Shiller PLS" uses only the topics from Shiller (2019), excluding *War* and Pandemic. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Reported are the estimates of β and t-statistics corrected for small-sample bias in Boudoukh, Israel, and Richardson (2022). The sample period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	1871-2019	1871-1949	1950-2019	2000-2019
War	3.80 ***	3.46 *	4.11 **	9.97 ***
t-stat	(2.83)	(1.72)	(2.41)	(3.14)
Pandemic	-2.60 *	-3.97 *	-1.53	-2.21
t-stat	(-1.93)	(-1.94)	(-0.90)	(-0.68)
Panic	2.21	3.08	2.55	3.26
t-stat	(1.63)	(1.51)	(1.49)	(1.00)
Confidence	0.63	-0.12	1.09	0.65
t-stat	(0.47)	(-0.06)	(0.64)	(0.20)
Saving	-1.36	-1.62	-1.41	-0.84
t-stat	(-1.01)	(-0.79)	(-0.82)	(-0.26)
Consumption	0.85	0.64	2.59	0.06
t-stat	(0.63)	(0.31)	(1.51)	(0.02)
Money	-2.33 *	-1.22	-3.41 **	-0.60
t-stat	(-1.73)	(-0.60)	(-2.00)	(-0.19)
Tech	-0.84	0.57	-3.08 *	-14.19 ***
t-stat	(-0.62)	(0.28)	(-1.80)	(-4.54)
Real Estate Boom	-3.04 **	-3.58 *	-3.02 *	-6.59 **
t-stat	(-2.25)	(-1.75)	(-1.76)	(-2.03)
Real Estate Crash	0.61	0.59	0.98	-2.36
t-stat	(0.45)	(0.29)	(0.57)	(-0.72)
Stock Bubble	-0.54	-1.66	0.28	-5.00
t-stat	(-0.40)	(-0.81)	(0.16)	(-1.56)
Stock Crash	0.78	2.09	-0.18	-0.62
t-stat	(0.58)	(1.02)	(-0.10)	(-0.19)
Boycott	-1.54	-2.34	-0.47	5.17
t-stat	(-1.14)	(-1.15)	(-0.28)	(1.62)
Wage	1.15	0.69	1.93	8.85 ***
t-stat	(0.85)	(0.34)	(1.13)	(2.75)
PLS	4.66 ***	3.89 *	5.37 ***	9.77 ***
t-stat	(3.29)	(1.73)	(3.15)	(3.05)
Shiller PLS	2.79 **	2.36	3.70 **	2.85
t-stat	(1.96)	(1.04)	(2.16)	(0.87)

Table C.5
Predicting Long-Horizon Market Returns With Correction for Small-Sample Bias

This table presents the results of the following predictive regression:

$$R^e_{t+1\to t+h} = \alpha + \beta x_t + \epsilon_{t+1\to t+h},$$

where $R^e_{t+1\to t+h}$ is the excess market return over the next h months, x_t is either *War* or the PLS indexes constructed from 14 topics, and β, the coefficient of interest, measures the strength of predictability. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Reported are the estimates of β and t-statistics corrected for small-sample bias in Boudoukh, Israel, and Richardson (2022). The sample period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	War	t-stat	PLS	t-stat
		1871-2019		
$h = 1$	3.80 ***	(2.83)	4.66 ***	(3.29)
$h = 3$	2.87 **	(2.30)	3.72 ***	(2.95)
$h = 6$	3.03 ***	(2.62)	3.28 ***	(2.93)
$h = 12$	2.78 ***	(2.74)	2.77 ***	(2.99)
$h = 24$	2.09 **	(2.52)	2.39 ***	(3.36)
$h = 36$	2.27 ***	(3.19)	3.06 ***	(5.11)
		1871-1949		
$h = 1$	3.46 *	(1.72)	3.89 *	(1.73)
$h = 3$	2.49	(1.31)	2.99	(1.42)
$h = 6$	2.86	(1.60)	2.98	(1.52)
$h = 12$	2.84 *	(1.77)	2.73	(1.58)
$h = 24$	1.87	(1.40)	2.51 *	(1.78)
$h = 36$	2.07 *	(1.79)	3.29 ***	(2.72)
		1950-2019		
$h = 1$	4.11 **	(2.41)	5.37 ***	(3.15)
$h = 3$	3.20 **	(2.20)	4.42 ***	(3.17)
$h = 6$	2.92 **	(2.36)	3.40 ***	(2.96)
$h = 12$	2.22 **	(2.26)	2.57 ***	(2.90)
$h = 24$	1.85 **	(2.53)	2.00 ***	(3.06)
$h = 36$	1.95 ***	(3.22)	2.50 ***	(4.67)
		2000-2019		
$h = 1$	9.97 ***	(3.14)	9.77 ***	(3.05)
$h = 3$	7.73 ***	(2.80)	5.34 **	(2.28)
$h = 6$	6.05 **	(2.53)	4.06 **	(2.27)
$h = 12$	5.44 ***	(2.84)	3.15 **	(2.41)
$h = 24$	4.79 ***	(3.34)	2.51 ***	(2.67)
$h = 36$	4.56 ***	(3.81)	3.33 ***	(4.44)

Table C.6
Predicting One-Month Market Returns: *War* versus Return Moments

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta War_t + \gamma z_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month, War_t is the *NYT War* index, z_t is one of the current excess market return (MKT), conditional volatility (VOL), or conditional skewness (SK), and β measures the strength of predictability. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. The whole sample is from January 1928 to October 2019.

Panel A: 1928-2019				
War	3.40 **	3.71 **	3.48 **	3.61 **
	(2.13)	(2.27)	(2.16)	(2.27)
MKT	5.28			6.22
	(1.39)			(1.46)
VOL		0.68		2.31
		(0.14)		(0.44)
SK			-2.33	-3.27
			(-0.89)	(-1.37)
R^2	0.78	0.13	0.25	0.93
Panel B: 1950-2019				
War	3.97 **	3.71 **	4.00 **	3.63 **
	(2.54)	(2.33)	(2.53)	(2.30)
MKT	1.44			1.10
	(0.66)			(0.53)
VOL		-1.92		-1.62
		(-0.67)		(-0.57)
SK			-1.29	-1.45
			(-0.74)	(-0.81)
R^2	0.52	0.58	0.50	0.46
Panel C: 2000-2019				
War	9.49 ***	8.81 ***	9.75 ***	8.71 ***
	(3.31)	(2.92)	(3.39)	(2.86)
MKT	2.27			0.36
	(0.55)			(0.09)
VOL		-5.10		-4.94
		(-0.95)		(-0.86)
SK			-1.01	-0.99
			(-0.38)	(-0.37)
R^2	3.18	3.96	3.02	3.18

Table C.7
War versus Real Events

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta \times War_t + \gamma^j \times D^j_t + \epsilon_{t+1}$$

where R^e_{t+1} is the excess market return over the next month, D^j_t is a dummy variable for event j equal to one if there is one event j in month t. Returns are expressed as annualized percentages, and *War* is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Recessions	Bank Failures	Wars	Disasters	Epidemic	All
Panel A: 1871-2019						
War	3.02 **	3.82 ***	3.58 ***	3.75 ***	3.84 ***	3.55 ***
	(2.52)	(3.37)	(3.06)	(3.32)	(3.40)	(3.14)
Event	-8.33 **	3.36	2.90	-7.06	18.32 *	-8.81 ***
	(-2.08)	(0.30)	(0.72)	(-1.10)	(1.66)	(-3.07)
R^2(%)	0.74	0.33	0.37	0.38	0.39	0.92
Panel B: 1871-1949						
War	2.54	3.62 **	2.68	3.42 **	3.54 **	3.35 *
	(1.37)	(2.09)	(1.45)	(1.98)	(2.05)	(1.93)
Event	-11.24 **	12.69	9.36	-8.11	20.43	-11.65 ***
	(-2.29)	(1.07)	(1.32)	(-1.15)	(1.57)	(-2.85)
R^2(%)	0.85	0.14	0.38	0.14	0.19	0.95
Panel C: 1950-2019						
War	4.12 **	4.02 **	4.16 ***	4.01 **	4.07 **	4.26 ***
	(2.58)	(2.53)	(2.62)	(2.53)	(2.57)	(2.66)
Event	-4.55	-32.15 *	-2.44	-6.22	9.74	-6.19
	(-0.64)	(-1.80)	(-0.56)	(-0.64)	(0.51)	(-1.57)
R^2(%)	0.53	0.58	0.48	0.50	0.44	0.79

Table C.8
Risk-Return Trade-Off

This table presents the results of the GARCH-M framework with the constant relative risk aversion specification (constant RRA):

$$R_{M,t+1} - R_{f,t+1} = \lambda_0 + \lambda_1 \times \sigma^2_{M,t} + \epsilon_{t+1},$$

and the time-varying RRA specification (varying RRA):

$$R_{M,t+1} - R_{f,t+1} = \lambda_0 + (\lambda_1 + \lambda_2 \times War_t) \times \sigma^2_{M,t} + \epsilon_{t+1},$$

in the *mean* equation. Panels A–D report the results with different specifications for the *volatility* equation, namely, GARCH (Bollerslev, 1986), IGARCH (Engle and Bollerslev, 1986), TGARCH (Zakoian, 1994), and EGARCH (Nelson, 1991):

$$\begin{aligned}
\text{GARCH}(1,1): &\quad \sigma^2_{M,t} = \delta_0 + \delta_1 \epsilon^2_t + \delta_2 \sigma^2_{M,t-1}, \\
\text{IGARCH}(1,1): &\quad \sigma^2_{M,t} = \delta_0 + \delta_1 \epsilon^2_t + (1-\delta_1) \sigma^2_{M,t-1}, \\
\text{TGARCH}(1,1): &\quad \sigma^2_{M,t} = \delta_0 + \delta_1 \epsilon^2_t + \delta_3 D_t \epsilon^2_t + \delta_2 \sigma^2_{M,t-1}, \\
\text{EGARCH}(1,1): &\quad ln(\sigma^2_{M,t}) = \delta_0 + \delta_1 \left(\frac{|\epsilon_t|}{\sigma_{M,t}} \right) - \delta_3 \left(\frac{\epsilon_t}{\sigma_{M,t}} \right) + \delta_2 ln(\sigma^2_{M,t-1}),
\end{aligned}$$

where D_t is an indicator equal to one when ϵ_t is negative and zero otherwise. The coefficient of interest λ_2, which measures the sensitivity of RRA to *War*, is in bold. The whole sample is from January 1871 to October 2019.

	1871-2019				1871-1949				1950-2019			
	Constant RRA		Varying RRA		Constant RRA		Varying RRA		Constant RRA		Varying RRA	
	Coef.	*t*-stat	Coef.	*t*-stat	Coef.	*t*-stat	Coef.	*t*-stat	Coef.	*t*-stat	Coef.	*t*-stat
Panel A: GARCH												
λ_0	0.00	1.47	0.00	-0.85	0.00	0.37	0.00	-0.72	0.00	1.09	0.00	-0.83
λ_1	2.17	2.62	0.22	0.21	2.20	2.42	0.81	0.68	2.58	1.32	-0.78	-0.33
λ_2			**2.05**	**3.20**			**1.56**	**1.89**			**3.10**	**2.46**
δ_0	0.00	3.56	0.00	3.64	0.00	2.67	0.00	2.73	0.00	2.47	0.00	2.45
δ_1	0.14	6.42	0.14	6.60	0.16	4.92	0.17	5.08	0.12	3.75	0.12	3.72
δ_2	0.82	31.85	0.82	32.67	0.81	20.95	0.81	21.66	0.83	24.50	0.83	23.73
Adj. R^2(%)	-0.38		0.27		-0.73		-0.30		0.10		0.84	
Panel B: IGARCH												
λ_0	0.00	2.17	0.00	-0.73	0.00	0.68	0.00	-0.65	0.00	1.95	0.00	-0.64
λ_1	1.69	2.84	-0.03	-0.03	1.81	2.41	0.56	0.56	1.88	1.80	-1.18	-0.77
λ_2			**1.95**	**4.08**			**1.50**	**2.48**			**2.79**	**2.57**
δ_0	0.00	3.60	0.00	3.63	0.00	2.57	0.00	2.62	0.00	2.93	0.00	2.88
δ_1	0.18	6.65	0.18	6.76	0.20	4.65	0.20	4.85	0.16	5.50	0.16	5.36
δ_2	0.82		0.82		0.80		0.80		0.84		0.84	
Adj. R^2(%)	-0.32		0.29		-0.61		-0.21		0.11		0.76	
Panel C: TGARCH												
λ_0	0.00	2.11	0.00	-1.15	0.00	0.43	0.00	-2.05	0.01	4.19	0.00	-0.19
λ_1	2.33	10.18	0.07	0.09	2.09	2.02	1.06	1.32	0.31	0.33	-2.09	-0.96
λ_2			**2.17**	**9.34**			**1.77**	**6.80**			**2.55**	**4.82**
δ_0	0.00	6.48	0.00	4.93	0.00	4.01	0.00	2.81	0.01	1.33	0.01	1.65
δ_1	0.14	11.79	0.14	13.85	0.14	8.72	0.15	10.77	0.12	4.52	0.12	4.95
δ_2	0.84	805.33	0.84	75.72	0.85	616.62	0.84	43.90	0.76	7.04	0.76	9.01
δ_3	0.26	3.38	0.28	3.24	0.20	2.29	0.21	2.22	0.76	1.71	0.68	2.03
Adj. R^2(%)	-0.26		0.13		-0.35		-1.07		-0.11		0.42	
Panel D: EGARCH												
λ_0	0.00	1.17	0.00	-0.12	0.00	-0.95	0.00	-1.18	0.01	1.62	0.00	-2.02
λ_1	2.88	43.43	0.23	0.03	2.81	16.15	0.77	4.50	0.89	0.44	-1.92	-1.98
λ_2			**2.03**	**1.15**			**1.47**	**3.47**			**2.69**	**5.58**
δ_0	-0.28	-43.00	-0.28	-0.97	-0.21	-23.38	-0.19	-3.45	-0.81	-2.77	-0.80	-8.26
δ_1	-0.06	-3.75	-0.06	-0.72	-0.05	-2.38	-0.05	-2.40	-0.14	-2.59	-0.13	-4.25
δ_2	0.96	65420.25	0.95	20.30	0.97	4136.73	0.97	112.70	0.88	19.74	0.88	59.15
δ_3	0.26	8.27	0.26	2.35	0.26	5.47	0.25	5.69	0.22	5.46	0.22	5.71
Adj. R^2(%)	-0.56		0.20		-1.17		-0.17		-0.13		0.50	

Table C.9
Predicting One-Month Market Returns: Discourse Topics versus Economic, Sentiment, and Uncertainty Predictors

This table presents the results of the following univariate predictive regression:

$$R^e_{t+1} = \alpha + \gamma z_t + \epsilon_{t+1},$$

and the following bivariate predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \gamma z_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month and x_t is the discourse topic PLS index. In Panel A, z_t is one of the 14 economic predictors from Goyal and Welch (2008), output gap from Cooper and Priestley (2009), and short interest from Rapach, Ringgenberg, and Zhou (2016); in Panel B, z_t is one of the uncertainty variables (financial and macro uncertainty indexes from Jurado, Ludvigson, and Ng (2015), economic policy uncertainty index from Baker, Bloom, and Davis (2016), disagreement index from Huang, Li, and Wang (2020), implied volatility (VIX), and news implied volatility (NVIX) from Manela and Moreira (2017)), sentiment variables (news sentiment, investor sentiment from Baker and Wurgler (2006), aligned sentiment from Huang et al. (2015), and manager sentiment from Jiang et al. (2019)), or Shiller's confidence indexes: one-year confidence index and crash confidence index. The last row of each panel reports the results using the PLS index constructed from all predictors in that panel. Returns and correlations are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table C.9
Predicting One-Month Market Returns:
Discourse Topics versus Economic, Sentiment, and Uncertainty Predictors (Cont.)

Panel A: Economic Predictors

	Univariate		Bivariate			
Economic Predictor	γ	R^2	β	γ	R^2	Period
Dividend-price ratio (DP)	1.39	0.00	6.00 ***	0.40	1.02	187101-201910
Dividend yield (DY)	2.03	0.07	5.89 ***	1.03	1.05	187102-201910
Earnings-price ratio (EP)	2.46	0.13	5.80 ***	1.23	1.06	187101-201910
Dividend payout ratio (DE)	-1.00	-0.03	6.05 ***	-0.87	1.04	187101-201910
Stock variance (SVAR)	-0.08	-0.06	6.08 ***	-0.37	1.02	187101-201910
Book-to-market Ratio (BM)	5.19	0.58	5.09 ***	3.49	1.06	192103-201910
Net equity expansion (NTIS)	-4.11	0.31	6.87 ***	-4.83	1.33	192612-201910
Treasury bill rate (TBL)	-3.68 **	0.36	5.46 ***	-2.11	1.14	187101-201910
Long term bond yield (LTY)	-2.82	0.11	6.25 ***	-0.60	0.87	191901-201910
Long term bond return (LTR)	3.36 *	0.18	6.62 ***	3.88 **	1.13	192601-201910
Term spread (TMS)	-2.70	0.10	6.30 ***	-0.46	0.87	191901-201910
Default yield spread (DFY)	2.87	0.12	6.37 ***	2.65	1.03	191901-201910
Default return spread (DFR)	2.30	0.04	6.29 ***	2.21	0.89	192601-201910
Inflation (INFL)	-3.32	0.20	5.70 ***	-3.60	0.95	191302-201910
Output Gap (OG)	-3.39	0.20	6.33 ***	-3.24	1.10	191902-201910
Short Interest (SI)	-5.70 **	0.94	4.98 **	-5.27 **	1.60	197301-201412
Economic PLS	4.40 *	0.48	4.53 *	3.05	0.93	197301-201412

Table C.9
Predicting One-Month Market Returns:
Discourse Topics versus Economic, Sentiment, and Uncertainty Predictors (Cont.)

Panel B: Uncertainty and Sentiment Predictors

	Correlations	Univariate		Bivariate			
Economic Predictor	Corr. with PLS	γ	R^2	β	γ	R^2	Period
Financial uncertainty	-5.77	-5.75 **	1.15	4.95 ***	-5.47 *	1.96	196007-201910
Macro uncertainty	-2.43	-4.30	0.58	5.17 ***	-4.17	1.48	196007-201910
Economic policy uncertainty	25.51 ***	4.03	0.38	4.02 *	3.01	0.73	198501-201910
Implied volatility (VIX)	-1.62	0.40	-0.27	6.32 **	0.50	1.11	199001-201910
News implied volatility (NVIX)	6.55 **	0.03	-0.07	5.80 ***	-0.35	0.80	188907-201603
Disagreement	-8.12 **	-8.43 ***	2.43	4.98 ***	-8.02 ***	3.17	196912-201812
News sentiment	14.71 ***	-0.52	-0.05	6.28 ***	-1.44	1.08	186612-201910
Investor sentiment (BW)	-13.83 ***	-2.50	0.08	4.68 **	-1.86	0.73	196507-201812
Investor sentiment (PLS)	-2.54	-7.32 ***	1.86	4.75 **	-7.20 ***	2.56	196507-201812
Manager sentiment	-19.31 ***	-9.06 **	3.32	6.87 ***	-7.74 **	4.93	200301-201712
Shiller's one-year confidence index	-29.87 ***	-4.77	0.48	10.48 ***	-1.64	4.17	200107-201910
Shiller's crash confidence index	-22.47 ***	-2.07	-0.28	11.06 ***	0.41	4.07	200107-201910
Uncertainty PLS	18.43 **	2.38	-0.39	8.84 ***	0.75	2.22	200301-201603

Table C.10
Subperiod R^2

This table reports the R^2 statistic as a percentage computed over different subperiods: expansion (exp) versus recession (rec) and high sentiment versus low sentiment. Expansions and recessions are based on the National Bureau of Economic Research (NBER) business cycles. A month is classified as high (low) sentiment if the Baker and Wurgler (2006) investor sentiment level in the previous month is above (below) the median value for the sample. Panel A reports the results for the in-sample analysis, and the entire sample period is January 1871 to October 2019. Panel B reports the results for the out-of-sample analysis with an expanding estimation window, and the evaluation period begins in January 1891.

	R^2	R^2_{exp}	R^2_{rec}	R^2_{high}	R^2_{low}
Panel A: In Sample					
War	0.39	0.06	0.91	0.01	0.50
PLS	1.07	0.58	1.80	-0.16	1.90
Panel B: Out of Sample					
War	0.17	0.69	-0.46	-0.02	0.47
PLS	-0.08	0.10	-0.31	-0.66	0.85

Table C.11
Asset Allocation Results

This table reports the annualized certainty equivalent returns (utility) gains as percentages and the annualized monthly Sharpe ratio for a mean-variance trading strategy. The strategy uses 6 economic predictors or 14 discourse topics to make return forecasts compared to historical mean returns. "Shiller Topics" uses only the topics from Shiller (2019), excluding *War* and Pandemic. Panels A and B report the results using OLS and PLS, respectively. The last row reports the annualized monthly Sharp ratio of the S&P 500 index. All out-of-sample forecasts are estimated recursively using the data available in the expanding estimation window. The evaluation period begins in January 1881, and the whole sample is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ (based on the test statistics in DeMiguel, Garlappi, and Uppal (2009)).

	Utility Gain (%)				Sharpe Ratio			
	1881-2019	1881-1949	1950-2019	2000-2019	1881-2019	1881-1949	1950-2019	2000-2019
Panel A: OLS								
Dividend-price ratio (DP)	-0.31	0.07	-0.70	0.24	0.36	0.25	0.48	0.21
Dividend yield (DY)	-0.50	-0.08	-0.94	1.02	0.35	0.23	0.50	0.30
Earnings-price ratio (EP)	0.59	0.53	0.64	3.88 *	0.44	0.28	0.56	0.65 ***
Dividend payout ratio (DE)	0.42	0.92	-0.08	-0.10	0.42	0.32 *	0.49	0.26
Stock variance (SVAR)	-0.44	-0.36	-0.53	-0.44	0.35	0.19	0.46	0.23
Treasury bill rate (TBL)	1.31 **	0.89	1.72 *	1.93 *	0.48 **	0.32 **	0.62 **	0.40 **
War	0.86 **	0.87	0.85 **	2.01 ***	0.45 **	0.32 *	0.55 **	0.38 ***
Pandemic	-0.09	-0.19	0.02	1.18	0.38	0.21	0.50	0.32
Panic	0.37	1.04 *	-0.31	-0.18	0.42	0.33 *	0.49	0.21
Confidence	-0.02	0.01	-0.05	-0.13	0.38	0.23	0.49	0.25
Saving	-0.04	-0.09	0.01	-0.21	0.38	0.21	0.50	0.24
Consumption	0.08	0.13	0.03	-0.04	0.39	0.24	0.50	0.24
Money	1.03 ***	0.66	1.40 **	1.86	0.47 ***	0.30	0.59 **	0.36
Tech	-0.53	-1.09	0.04	0.33	0.34	0.13	0.50	0.27
Real estate boom	0.24	0.15	0.32	1.89	0.40	0.24	0.52	0.37
Real estate bust	-0.06	-0.17	0.05	-0.08	0.38	0.20	0.50	0.25
Stock bubble	-0.14	0.06	-0.34	-0.25	0.37	0.23	0.47	0.24
Stock crash	-0.09	0.16	-0.35	-0.03	0.38	0.24	0.47	0.25
Boycott	0.20	0.02	0.38	-0.73	0.40	0.23	0.52	0.22
Wage	-0.17	-0.31	-0.04	0.10	0.37	0.19	0.49	0.26
Panel B: PLS								
Economic	0.44	0.70	0.16	3.08	0.42	0.31	0.53	0.65 **
Narratives	0.69	0.66	0.70	4.11 **	0.43	0.31 *	0.55	0.54 **
Shiller Narratives	0.46	0.50	0.40	1.05	0.43	0.29	0.56	0.29
Buy and Hold					0.39	0.28	0.55	0.35

Table C.12
Predicting Bond Returns

This table presents the results of the following predictive regression:

$$R^e_{t+1\to t+h} = \alpha + \beta War_t + \epsilon_{t+1\to t+h},$$

where $R^e_{t+1\to t+h}$ is the excess returns over the next h months on Datastream Treasury bond indexes (Panel A), S&P investment grade corporate bond (Panel B), and S&P high yield corporate bond (Panel C). Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors using the corresponding h lags. The sample is from December 1988 to October 2019. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Panel A: Government Bond Indexes

	h=1	h=3	h=6	h=12
	1988-2019			
U.S. Goverment Bond 1-3 Years	-0.290	-0.225	-0.192	-0.133
(t-stat)	(-1.41)	(-1.13)	(-1.00)	(-0.52)
R^2	0.12	0.25	0.41	0.22
U.S. Goverment Bond 3-5 Years	-0.57	-0.44	-0.33	-0.16
(t-stat)	(-1.03)	(-0.99)	(-0.94)	(-0.38)
R^2	-0.04	0.07	0.11	-0.11
U.S. Goverment Bond 5-7 Years	-0.73	-0.60	-0.42	-0.16
(t-stat)	(-0.91)	(-0.98)	(-0.88)	(-0.33)
R^2	-0.07	0.09	0.07	-0.17
U.S. Goverment Bond 7-10 Years	-0.91	-0.81	-0.50	-0.13
(t-stat)	(-0.86)	(-1.01)	(-0.80)	(-0.24)
R^2	-0.09	0.14	0.03	-0.23
U.S. Goverment Bond 10+ Years	-0.94	-0.78	-0.22	0.16
(t-stat)	(-0.54)	(-0.57)	(-0.20)	(0.21)
R^2	-0.20	-0.13	-0.25	-0.25
	2000-2019			
U.S. Goverment Bond 1-3 Years	-0.591 **	-0.529 **	-0.586 ***	-0.590 **
(t-stat)	(-2.43)	(-2.41)	(-2.64)	(-2.21)
R^2	1.58	3.30	7.07	10.98
U.S. Goverment Bond 3-5 Years	-1.07	-0.89 *	-0.93 **	-0.88 **
(t-stat)	(-1.53)	(-1.86)	(-2.39)	(-2.08)
R^2	0.44	1.26	2.98	5.92
U.S. Goverment Bond 5-7 Years	-1.31	-1.11	-1.06 *	-0.91 *
(t-stat)	(-1.27)	(-1.60)	(-1.95)	(-1.77)
R^2	0.20	0.93	1.95	3.60
U.S. Goverment Bond 7-10 Years	-1.52	-1.31	-1.07	-0.77
(t-stat)	(-1.09)	(-1.40)	(-1.46)	(-1.35)
R^2	0.05	0.69	1.04	1.47
U.S. Goverment Bond 10+ Years	-1.40	-0.86	-0.29	0.00
(t-stat)	(-0.59)	(-0.51)	(-0.20)	(0.00)
R^2	-0.29	-0.27	-0.40	-0.44

Table C.12
Predicting Bond Returns (Cont.)

Panel B: Investment Grade Corporate Bond Indexes

	h=1	h=3	h=6	h=12
	1993-2019			
Corporate Bond 0-1 Years	-0.13	-0.13 *	-0.11 *	-0.13 *
(t-stat)	(-1.28)	(-1.70)	(-1.72)	(-1.83)
R^2	-0.04	0.34	0.55	1.42
Corporate Bond 1-3 Years	-0.34	-0.35	-0.30	-0.24
(t-stat)	(-1.02)	(-1.35)	(-1.41)	(-1.10)
R^2	-0.06	0.30	0.50	0.60
Corporate Bond 3-5 Years	-0.16	-0.26	-0.21	-0.13
(t-stat)	(-0.27)	(-0.61)	(-0.61)	(-0.38)
R^2	-0.29	-0.19	-0.17	-0.23
Corporate Bond 5-7 Years	0.02	-0.27	-0.19	-0.00
(t-stat)	(0.03)	(-0.44)	(-0.38)	(-0.01)
R^2	-0.31	-0.24	-0.26	-0.32
Corporate Bond 7-10 Years	0.22	-0.23	-0.11	0.09
(t-stat)	(0.21)	(-0.30)	(-0.17)	(0.16)
R^2	-0.30	-0.28	-0.30	-0.31
Corporate Bond 10+ Years	0.62	-0.14	0.15	0.34
(t-stat)	(0.39)	(-0.12)	(0.14)	(0.44)
R^2	-0.27	-0.31	-0.30	-0.19
	2000-2019			
Corporate Bond 0-1 Years	-0.26 *	-0.26 **	-0.25 **	-0.27 **
(t-stat)	(-1.81)	(-2.14)	(-2.31)	(-2.58)
R^2	0.42	1.63	2.85	5.29
Corporate Bond 1-3 Years	-0.54	-0.58 *	-0.58 *	-0.47
(t-stat)	(-1.30)	(-1.82)	(-1.91)	(-1.51)
R^2	0.13	1.07	2.20	2.59
Corporate Bond 3-5 Years	-0.45	-0.64	-0.69	-0.53
(t-stat)	(-0.64)	(-1.28)	(-1.53)	(-1.15)
R^2	-0.28	0.29	1.02	1.03
Corporate Bond 5-7 Years	-0.34	-0.75	-0.80	-0.51
(t-stat)	(-0.35)	(-1.09)	(-1.26)	(-0.79)
R^2	-0.38	0.08	0.53	0.21
Corporate Bond 7-10 Years	-0.22	-0.79	-0.77	-0.45
(t-stat)	(-0.17)	(-0.90)	(-0.99)	(-0.62)
R^2	-0.41	-0.04	0.25	-0.02
Corporate Bond 10+ Years	0.17	-0.61	-0.46	-0.18
(t-stat)	(0.09)	(-0.46)	(-0.38)	(-0.19)
R^2	-0.42	-0.32	-0.32	-0.41

Table C.12
Predicting Bond Returns (Cont.)

Panel C: High Yield Corporate Bond Indexes

	h=1	h=3	h=6	h=12
	1993-2019			
Corporate Bond 0-1 Years	0.56	0.22	-0.04	-0.14
(t-stat)	(0.81)	(0.54)	(-0.12)	(-0.35)
R^2	-0.18	-0.25	-0.31	-0.25
Corporate Bond 1-3 Years	1.34	0.81	0.53	0.30
(t-stat)	(1.42)	(1.05)	(0.74)	(0.43)
R^2	0.10	0.02	-0.11	-0.20
Corporate Bond 3-5 Years	2.68 **	1.69 *	1.30	1.14
(t-stat)	(2.21)	(1.69)	(1.35)	(1.05)
R^2	1.02	0.70	0.57	0.86
Corporate Bond 5-7 Years	3.48 **	2.03 *	1.77 *	1.65
(t-stat)	(2.49)	(1.87)	(1.72)	(1.43)
R^2	1.42	0.87	1.02	1.74
Corporate Bond 7-10 Years	3.33 **	1.73	1.54	1.62
(t-stat)	(2.16)	(1.54)	(1.46)	(1.34)
R^2	0.91	0.46	0.68	1.66
Corporate Bond 10+ Years	2.73 *	1.50	1.03	0.88
(t-stat)	(1.70)	(1.19)	(0.83)	(0.70)
R^2	0.42	0.10	-0.04	0.01
	2000-2019			
Corporate Bond 0-1 Years	0.32	-0.11	-0.49	-0.65
(t-stat)	(0.30)	(-0.20)	(-1.02)	(-1.37)
R^2	-0.39	-0.41	-0.05	0.88
Corporate Bond 1-3 Years	1.19	0.29	-0.20	-0.54
(t-stat)	(0.87)	(0.29)	(-0.20)	(-0.55)
R^2	-0.17	-0.39	-0.41	-0.13
Corporate Bond 3-5 Years	2.98 *	1.60	0.91	0.41
(t-stat)	(1.89)	(1.37)	(0.77)	(0.30)
R^2	0.90	0.31	-0.08	-0.32
Corporate Bond 5-7 Years	3.95 **	1.89	1.37	0.94
(t-stat)	(2.12)	(1.41)	(1.01)	(0.58)
R^2	1.35	0.39	0.20	0.10
Corporate Bond 7-10 Years	4.48 **	2.22	1.73	1.55
(t-stat)	(2.23)	(1.64)	(1.23)	(0.88)
R^2	1.38	0.60	0.57	1.01
Corporate Bond 10+ Years	2.89	1.15	0.13	-0.20
(t-stat)	(1.39)	(0.78)	(0.08)	(-0.12)
R^2	0.25	-0.23	-0.43	-0.43

Table C.13
Predicting Bond Returns: Out-Of-Sample R^2

This table reports the out-of-sample R^2 (R^2_{OS}) statistic (Campbell and Thompson 2008) in predicting the excess returns over the next h months on Datastream Treasury bond indexes (Panel A), S&P investment grade corporate bond (Panel B), and S&P high yield corporate bond (Panel C) using *NYT War*. All the out-of-sample forecasts are estimated recursively using the data available in the expanding estimation window. All numbers are expressed as percentages. The evaluation period begins in January 2003, and the whole sample is from January 1993 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ (based on the Clark and West (2007) MSFE-adjusted statistic).

	h=1	h=3	h=6	h=12
Panel A: Government Bond Indexes				
1-3 Years	0.21 **	0.18 *	3.38 **	7.31 **
3-5 Years	-1.39	-1.16	-0.76	1.23
5-7 Years	-1.57	-1.73	-1.90	-0.88
7-10 Years	-1.50	-1.87	-3.07	-2.15
10+ Years	-0.86	-1.73	-3.65	-2.08
Panel B: Investment Grade Corporate Bond Indexes				
0-1 Years	-0.27	0.29 *	0.46 *	0.65
1-3 Years	-0.84	-0.72	0.49	0.05
3-5 Years	-1.29	-1.28	-0.61	-1.54
5-7 Years	-1.19	-1.49	-0.94	-1.52
7-10 Years	-0.94	-1.43	-1.45	-1.40
10+ Years	-0.62	-1.31	-1.95	-1.27
Panel C: High Yield Corporate Bond Indexes				
0-1 Years	-0.07	-1.16	-0.79	-1.16
1-3 Years	0.02	-1.45	-1.54	-1.88
3-5 Years	0.30 *	-1.82	-3.06	-5.86
5-7 Years	0.44 **	-1.99	-3.36	-5.93
7-10 Years	0.58 **	-1.31	-1.64	-3.10
10+ Years	0.36 *	-1.04	-1.23	-1.16

Table C.14
Predicting Returns of Characteristics Portfolios

This table presents the results of the following predictive regression:

$$R^e_{i,t+1} = \alpha_i + \beta_i x_t + \epsilon_{i,t+1}, \quad i = 1, \ldots, 40$$

where $R^e_{i,t+1}$ is the excess return on portfolio i over the next month, x_t is either War or the PLS index, and β_i, the coefficient of interest, measures the strength of predictability. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. The sample period is January 1927 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	War(%)	t-stat	R^2 (%)	PLS (%)	t-stat	R^2 (%)
Panel A: Industry Portfolios						
Nondurable	2.53 *	(1.80)	0.12	4.07 ***	(2.64)	0.46
Durable	3.29	(1.42)	0.04	6.96 ***	(2.82)	0.48
Manufacture	2.16	(1.21)	-0.01	5.25 ***	(2.58)	0.40
Energy	1.69	(0.89)	-0.04	4.75 **	(2.30)	0.33
Technology	3.62	(1.57)	0.09	6.54 ***	(2.61)	0.48
Telecom	0.98	(0.70)	-0.06	2.43	(1.57)	0.11
Shop	3.18 *	(1.77)	0.12	6.01 ***	(3.34)	0.66
Health	2.17	(1.24)	0.02	4.95 ***	(2.65)	0.46
Utility	2.23	(1.35)	0.02	4.19 **	(2.42)	0.32
Other	3.75 **	(1.98)	0.15	6.38 ***	(2.92)	0.60
Panel B: Size Portfolios						
Small	6.17 **	(1.97)	0.18	10.67 ***	(2.83)	0.72
2	4.99 *	(1.88)	0.14	8.76 ***	(2.85)	0.62
3	5.32 **	(2.22)	0.23	8.39 ***	(3.02)	0.69
4	4.39 **	(1.97)	0.16	7.93 ***	(3.13)	0.71
5	4.20 **	(1.98)	0.16	7.06 ***	(2.99)	0.62
6	3.63 *	(1.78)	0.11	6.90 ***	(3.03)	0.64
7	3.91 **	(1.98)	0.17	6.71 ***	(3.11)	0.68
8	3.67 **	(2.01)	0.16	6.33 ***	(3.08)	0.66
9	3.13 *	(1.83)	0.11	5.93 ***	(3.10)	0.65
Large	2.54 *	(1.69)	0.09	4.90 ***	(2.93)	0.57
Panel C: Book-to-market Portfolios						
Growth	2.33	(1.37)	0.03	5.35 ***	(2.80)	0.53
2	2.49	(1.50)	0.06	5.11 ***	(2.88)	0.56
3	2.36	(1.48)	0.04	4.75 ***	(2.76)	0.46
4	2.01	(1.23)	-0.01	4.83 **	(2.50)	0.38
5	3.15 *	(1.90)	0.13	5.88 ***	(3.15)	0.67
6	2.30	(1.34)	0.01	5.15 **	(2.54)	0.42
7	3.47 *	(1.86)	0.11	5.89 ***	(2.79)	0.50
8	3.78 *	(1.92)	0.13	7.02 ***	(3.16)	0.68
9	4.54 **	(2.04)	0.16	7.77 ***	(3.09)	0.63
Value	5.77 **	(2.08)	0.19	9.08 ***	(2.88)	0.60
Panel D: Momentum Portfolios						
Losers	7.10 **	(2.41)	0.28	10.43 ***	(3.24)	0.71
2	3.26	(1.39)	0.02	6.29 **	(2.35)	0.34
3	2.99	(1.48)	0.04	5.32 **	(2.37)	0.32
4	2.56	(1.42)	0.02	5.34 ***	(2.65)	0.41
5	2.81 *	(1.71)	0.07	5.39 ***	(2.80)	0.49
6	2.73 *	(1.67)	0.07	5.45 ***	(2.93)	0.54
7	2.81 *	(1.79)	0.10	5.46 ***	(3.14)	0.61
8	2.24	(1.40)	0.04	5.44 ***	(3.13)	0.65
9	3.70 **	(2.18)	0.22	6.37 ***	(3.47)	0.83
Winners	3.92 *	(1.82)	0.17	6.43 ***	(2.83)	0.61

Table C.15
Predicting Expected and Unexpected Returns, and Cash Flow and Discount Rate News with Topic PLS Index

This table reports results from the following univariate predictive regressions:

$$
\begin{aligned}
r_{t+1} &= \alpha + \beta PLS_t + \epsilon_{t+1}, \\
\mathbb{E}_t r_{t+1} &= \alpha^E + \beta^E PLS_t + \epsilon^E_{t+1}, \\
CF_{t+1} &= \alpha^{CF} + \beta^{CF} PLS_t + \epsilon^{CF}_{t+1}, \quad \text{and} \\
DR_{t+1} &= \alpha^{DR} + \beta^{DR} PLS_t + \epsilon^{DR}_{t+1},
\end{aligned}
$$

where r_{t+1} is log market return, $\mathbb{E}_t r_{t+1}$ is expected return, CF_{t+1} is cash flow news, DR_{t+1} is discount rate news, and $r_{t+1} = \mathbb{E}_t r_{t+1} + CF_{t+1} + DR_{t+1}$. $\mathbb{E}_t r_{t+1}$, CF_{t+1}, and DR_{t+1} are estimated via a VAR(1) model consisting of log return, log dividend-price (DP), log earnings-price (EP), Treasury bill (TBL), and stock variance (SVAR). Each row reports results for a different combination of the return predictors. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. t-statistics are computed with Newey and West (1987) standard errors. The sample period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	r_{t+1}	$E_t r_{t+1}$	CF_{t+1}	DR_{t+1}
R, DP	3.25 **	0.17	3.53 **	0.47 **
	(2.44)	(1.09)	(2.36)	(2.08)
R, DP, EP	3.25 **	0.29 *	3.76 **	0.81 ***
	(2.44)	(1.67)	(2.53)	(3.12)
R, DP, SVAR	3.25 **	0.15	3.53 **	0.44 *
	(2.44)	(0.92)	(2.36)	(1.95)
R, DP, TBL	3.25 **	0.16	3.53 **	0.45 **
	(2.44)	(1.01)	(2.36)	(2.01)
R, DP, EP, SVAR, TBL	3.25 **	0.45 ***	3.52 **	0.73 ***
	(2.44)	(2.60)	(2.34)	(2.64)

D Robustness Checks

In this appendix, we report robustness checks of the main prediction results.

D.1 Robustness Checks: sLDA

We perform a battery of robustness checks in this section. The number and type of topics and seed words are key inputs to our technique. To examine robustness of these inputs, we first examine the strategy of using a very large number of topics. In such a case, the weights of seeded topics can be approximated by the frequency of the seed words in the corpus. Hence, we investigate this case by constructing topic weights as the counts of seed words scaled by the article length and present the results in Appendix E. Frequency-based topic weights still yield results consistent with the sLDA ones, but their out-of-sample performance is weaker. We also experiment with choices of number of topics and seed words.

In this paper, since we are interested in studying 2 disaster-focused topics and 12 non-disaster topics from Shiller (2019), we only include 15 topics in sLDA (14 seeded plus 1 unseeded topic). Our exogenous specification of the number of topics is consistent with the view in Gentzkow, Kelly, and Taddy (2019) that in many applications of topic models, the goal is to provide an intuitive description of text rather than infer underlying "true" parameters. Our out-of-sample return predictability tests provide a validation that our specification of the number of topics generates meaningful results.

We also experiment with choices of number of topics and seed words. First, we add another disaster-related topic—*Natural Disaster*, which includes seed words such as "earthquake, flood," and "hurricane," while enhancing our *Pandemic* topic by increasing the number of seed words from 2 ("epidemic" and "pandemic") to 12 and keeping the seed words for other topics unchanged. The updated list of seed words for this specification is reported in Table D.1. This could affect our results if these "disaster-like" topics use words that are correlated with *War*. As reported in Table D.1, the results for *War* overall are similar (slightly weaker in the first half of the sample and much stronger in the second half). The signs of *Pandemic* and *Natural Disaster* coefficients are unstable

with *Pandemic* being significant only during the past two decades. The result suggests that *War* has the most predictive power, as it is the most important extreme disaster. This conclusion is consistent with the findings of Barro (2006, 2009), who used historical data from wars to estimate disaster probabilities. They used this to explain equity premia and other asset pricing phenomena.

We next experiment with various modifications of seed words and number of topics: increasing the number of seed words for *War* while maintaining the other topics in Table D.1 (see Table D.2); modifying the seed words for *War* and removing duplicates in seed words within and across topics (see Table D.3); and increasing the number of unseeded topics from 1 to 50 (see Table D.4). We find the results for *War* remain robust.

At the extreme, in Table D.5, we only include 1 seeded topic for *War* having only 1 seed word "war" with 50 unseeded topics. Under this specification, the results for *War* become much stronger in the first half and slightly weaker in the second half, but over the whole sample, the results are still significant at the 1% level for both in- and out-of-sample tests. All of our experiments show that the predictive power of *War* is not sensitive to the specifications of sLDA.

In choosing the number of topics, it would not be appropriate to optimize using the entire sample period to maximize ex-post predictive power, as this would expose the study to look-ahead bias. Another possibility might be to use statistical methods such as Bayes factors or cross-validation to select the optimal number of topics during each monthly estimation. However, under such an approach, the optimal numbers of topics can change from month to month. As a result, monthly changes in topic weights would not be attributable to the shifts in public attention to different topics. Such shifts in public attention are central to our approach. Alternatively, Lu et al. (2011) suggest using a number of topics equal to the number of seeded topics plus one.[7] Our robustness checks indicate that the results are not sensitive to this choice.

As detailed in Internet Appendix B, we create n-grams within punctuation boundaries before removing stop words. We then remove bigrams containing stop words because these bigrams add

[7]The rationale for adding an unseeded topic is to allow the model to discover and consider an additional topic that may not be captured by the seeded topics provided. By using the number of seeded topics plus one, the model can better accommodate any unforeseen or unanticipated topics that are relevant to the data but were not part of the initial seeded topics. This approach helps strike a balance between incorporating prior knowledge through seeded topics and remaining flexible enough to account for any new information or patterns that may arise from the data during analysis.

no additional value to topic estimation beyond the unigrams contained therein. We also remove trigrams containing stop words unless the stop word is a preposition in the middle position. Our n-gram creation method ensures we only consider meaningful terms in the text data. However, we also experiment with removing stop words *before* creating n-grams and retaining all the resulting n-grams and find consistent results for *War*.

When cleaning the texts, we remove articles containing mostly numbers, names, and lists (i.e., articles having limited content). As a robustness check, we keep all news articles in our data set and find *War* still has the same predictive power.

D.2 Robustness Checks: *War* Innovation

In Table D.7, we use innovation instead of level of 14 topics reported in the main text to predict next month market returns. We estimate innovation as residuals of an AR(1) process for each topic.[8] The PLS indexes are constructed from these innovations. We find that the predicting power of *War* innovation is consistent with *War* level over the whole sample and weaker than *War* level over the second half of the sample.

D.3 Robustness Checks: Empirical Design

To examine the robustness of our choice of five seed words for *War* in our baseline results (Table 1), we implement a non-parametric bootstrap. Accordingly, we generate 1,000 null words that are expected to not predict returns. In each iteration from 1 to 10,000, we

- randomly select 5 words from the list of 1,000 null words,
- compute the frequency of these 5 selected words in 150 years of *NYT* text data to generate a time series of null topic weight,
- use the topic weight to run in- and out-of-sample return prediction tests,
- and store the Newey-West adjusted t-statistic of the in-sample predictive regression, in-sample R^2, and out-of-sample R^2.

After 10,000 iterations, we have the empirical null distributions of the in-sample and out-of-sample

[8] Using an AR(2) process yields similar results.

test statistics. We then compute the percentages of iterations when these bootstrap statistics are greater than the corresponding statistics generated by the frequency of our five *War* seed words for reported in Table E.2 and Table E.4. These form the empirical p-values of our prediction statistics under the null of no predictability. Note that we decide to run the bootstrap with frequency count of seed words instead of running sLDA because running 10,000 sLDA models over our 160 years of text data would require at least 10,000 days to complete on a high computing cluster.

Table D.6 reports the empirical p-values generated by the bootstrap. Column p_t reports the percentage of iterations that yield a Newey-West t-statistic positively larger than the t-statistic of our War seed words in the corresponding sample period (1.85% for 1871-2019, 0.21% for 1950-2019, and 1.48% for 2000-2019). Note that we care about the sign of the coefficient because War is expected to positively predict returns. Column p_{IS-R^2} reports the percentage of iterations that yield an in-sample R^2 larger than the one generated by our War seed words in the corresponding sample period (2.96% for 1871-2019, 0.51% for 1950-2019, and 4.37% for 2000-2019). Column p_{OS-R^2} reports the percentage of iterations that yield an out-of-sample R^2 larger than the one generated by our War seed words in the corresponding sample period (3.20% for 1881-2019, 3.98% for 1950-2019, and 0.99% for 2000-2019). The last column of Table D.6 reports the percentage of iterations that simultaneously beat all three statistics of *War* seed words in the corresponding sample period. Overall, we can conclude that under 1% of the times can five randomly selected words match the prediction performance of our *War* seed words.

Here is the list of 1,000 null words:[9] [abaft, abandon, abhorrent, abject, ablaze, abound, absorb, accessible, acidic, acoustic, acrid, actor, addict, adjustment, advertisement, advice, advise, afford, afternoon, afterthought, agonize, air, ajar, alive, allow, aloof, ambiguous, amuck, amuse, analyze, angle, angry, animal, answer, ant, apparel, appear, argue, argument, arithmetic, aromatic, arrange, attraction, automatic, average, badge, bait, bake, balance, ban, bang, base, baseball, bashful, basket, basketball, bathe, beautiful, bed, bedroom, bee, behave, belief, believe, bell, bend, berry, berserk, big, bike, bird, bite, bizarre, black, blade, bleach, bless, blind, blink, bloody, blot, blush, board, boil, bolt, bone, book, boot, boundary, boy, brainy, brash, brass, breakable, breath, breathe,

[9] We use this website (https://www.randomlists.com/random-words) to generate the list and manually check them.

brick, bridge, brief, brown, bruise, brush, bulb, bump, bumpy, burly, burn, burst, butter, cable, cactus, calculator, call, callous, calm, camera, can, capricious, car, careful, cart, cast, cat, cattle, cause, cave, ceaseless, cemetery, certain, chance, change, changeable, channel, charge, chase, cheap, check, cheer, cheese, cherry, chew, childlike, chunky, church, clammy, classy, clever, clip, close, cloudy, clover, club, clumsy, coach, coast, cobweb, coil, collar, colossal, colour, comb, common, communicate, compete, competition, complete, complex, concentrate, concern, condition, contain, control, cool, coordinate, copper, corn, count, courageous, cow, cowardly, crate, crazy, cream, creepy, crib, crime, crook, crow, crowd, crown, cruel, cup, curl, curtain, cut, cycle, dam, damage, damp, dapper, dark, daughter, day, debonair, debt, decision, decisive, decorate, decorous, deeply, defective, defiant, degree, delight, delirious, deliver, demonic, derange, descriptive, desert, design, desk, destroy, detail, develop, different, dime, dinner, dinosaur, dirty, disarm, discreet, disgust, dislike, distance, distinct, dock, doctor, dog, door, doubtful, downtown, drab, draconian, drawer, dreary, dress, drive, drop, drown, dull, dynamic, earsplitting, earth, educate, effect, egg, elate, elbow, electric, elegant, elfin, eminent, employ, entertain, enthusiastic, equal, erect, error, ethereal, evasive, even, event, example, excite, exercise, exist, existence, exotic, expansion, expect, expensive, experience, expert, extend, extra_large, extra_small, exultant, eye, fabulous, fail, fair, fairy, false, fancy, fantastic, far, fast, fasten, fearless, fence, fertile, fierce, file, filthy, find, finger, first, fish, five, fix, flag, flagrant, flaky, flame, flash, flashy, flavor, flawless, flight, flock, floor, flowery, fluffy, flutter, fog, fold, follow, fool, force, forego, forgetful, fork, fragile, frame, frantic, fresh, friction, friend, friendly, frighten, frog, fry, fumble, functional, furniture, furry, gainful, gape, gate, gather, gaudy, gaze, general, ghost, giant, gigantic, giraffe, girl, glamorous, gleam, glib, glisten, glue, good, grab, grandfather, grape, grass, grateful, grease, greedy, grey, grin, grind, groan, grotesque, grubby, grumpy, guarantee, guard, guess, guide, guitar, habitual, haircut, hall, hammer, hand, handsome, handy, hang, hapless, harass, harbor, hard_to_find, harm, harmonious, harmony, harsh, hat, health, heap, heartbreaking, helpful, hideous, high_pitch, hill, hiss, hole, holiday, homeless, homely, hop, hope, horn, horrible, hose, hospitable, hour, hover, hug, hulking, hum, humor, humorous, hungry, hunt, hurry, hydrant, ice, ignorant, ill_fate, imaginary, immense, imminent, impartial, imperfect, important, impress, improve, inconclusive, increase, inform, inject, ink, innate, inquisitive, insidious, instinctive, intend,

interest, interrupt, invite, iron, irritate, itch, jam, jeans, jelly, jellyfish, jobless, jog, join, jolly, joyous, judicious, juice, jumble, jumpy, kaput, keen, key, kick, kind, kiss, kitty, knit, knot, knowledge, knowledgeable, label, laborer, lace, lackadaisical, ladybug, lamp, languid, last, lean, learn, leather, leg, letter, lettuce, library, lick, light, lighten, like, limit, line, list, literate, little, long, lopsided, loss, loud, lowly, lucky, ludicrous, lumber, lunch, luxuriant, madden, madly, magenta, magic, maid, mailbox, malicious, man, manage, march, marry, massive, match, measure, meat, meddle, medical, mellow, memorize, merciful, mere, metal, milk, milky, minor, mint, minute, mix, morning, motion, murky, mushy, mute, mysterious, nail, name, nappy, nasty, natural, naughty, near, nebulous, need, needle, needless, neighborly, nerve, new, nifty, nine, nod, noise, noisy, nonstop, note, notice, numb, numerous, nut, nutritious, oafish, oatmeal, obeisant, obey, object, obnoxious, obscene, obsequious, obtain, obtainable, offbeat, offend, office, omniscient, open, orange, ossify, oven, overflow, own, paddle, page, painful, paint, paper, parallel, parch, partner, party, passenger, pastoral, pause, pear, pedal, pencil, perfect, perform, perpetual, pest, pet, physical, pick, pickle, pie, pink, pipe, place, plain, plan, plane, plant, plastic, plate, plausible, play, pleasant, please, pleasure, plot, point, poise, polish, pollution, possess, possible, post, pot, potato, powerful, pray, preach, precede, prepare, prevent, prickly, produce, profuse, program, promise, prose, protect, protective, psychedelic, psychotic, pump, punch, puncture, punishment, purple, purpose, purr, puzzle, quarter, quartz, queen, question, queue, quiet, quilt, quiver, quizzical, rabbit, race, rag, rain, rainy, raise, rambunctious, raspy, rate, ratty, ray, reaction, real, receipt, receive, receptive, recognise, recondite, record, reflective, refuse, regret, reject, rejoice, release, religion, remind, reminiscent, replace, reply, request, rescue, return, rhyme, rhythm, rice, rich, riddle, right, ring, risk, ritzy, robust, rock, roll, romantic, room, root, royal, rub, run, rural, rustic, sable, sack, sad, sail, same, sand, savory, scarce, scare, scarf, scary, scent, scintillate, scissor, scrape, scream, screw, scribble, sea, seal, seat, secretive, sedate, selection, self, sense, separate, serve, shaggy, shaky, shallow, shame, shape, ship, shirt, shoe, shrill, shrug, shy, sick, sign, signal, sin, sink, sip, sister, six, skillful, slap, sleep, slippery, sloppy, slow, smart, smell, smelly, smile, smooth, snail, snatch, sneaky, snobbish, snore, soap, society, sock, soda, soft, soggy, song, sophisticate, sound, sour, spade, sparkle, special, spectacular, spicy, spider, spill, spiritual, spiteful, spoil, spoon, spotty, spray, spring, sprout, square, squash, squeak, squeal,

squeeze, squirrel, stain, stake, stamp, stand, stay, steady, stem, step, stick, stir, stitch, stomach, store, stormy, strange, straw, string, stroke, stuff, stupendous, subdue, subtract, succeed, successful, sudden, suffer, suggest, suit, sun, superb, superficial, suppose, surprise, surround, swanky, sweater, swelter, system, taboo, tacit, tail, talented, tame, tan, tap, tasteless, tawdry, teeny_tiny, telephone, tell, temper, tent, tenuous, texture, thank, therapeutic, thick, thing, thirsty, thoughtless, threaten, three, throne, thumb, thunder, tidy, tie, time, tire, toe, toothbrush, toothpaste, top, tower, toy, trap, tree, tremendous, tricky, trite, trouble, trouser, trust, try, tug, turn, two, type, typical, ugly, unable, unadvised, uncover, undesirable, uneven, unfasten, unhealthy, uninterested, unique, unite, unkempt, unlock, unwieldy, unwritten, uppity, upset, uptight, use, useful, vacuous, vague, various, veil, vein, vengeful, venomous, verse, vigorous, violent, violet, voice, voiceless, wacky, waggish, wail, wait, wakeful, walk, wander, want, warm, warn, wasteful, watch, water, wave, wax, way, weak, weary, welcome, whimsical, whine, whistle, wicked, wilderness, wind, wine, wire, wise, wish, wistful, wobble, woman, wood, wooden, wool, woozy, word, worry, worthless, wreck, wretched, wriggle, yard, yawn, yell, yield, yoke, youthful, zany, zephyr, zesty, zinc, zippy, zonked, zoo].

D.4 Alternative Bootstrap Approach

The previous bootstrap exercise is designed to evaluate the likelihood of randomly discovering any 5 seed words that predict returns as effectively as *War*. However, this approach might be unconvincing because selecting random, unrelated words typically generates time-series predictors that lack persistence and do not correlate with the unobserved, persistent expected return process. The issue is that, by design, we calculate 14 non-random time series that inherently possess some degree of persistence, potentially leading to spurious correlations with the unobserved (persistent) expected return process.

To address this issue, we perform a simpler exercise for bootstrapping the likelihood of the observed return predictability. Inspired by Ferson, Sarkissian, and Simin (2003), we simulate 14 AR(1) predictors with autocorrelations matching the estimated autocorrelations for the 14 discourse variables. In each simulation, we run separate time-series predictive regressions for all 14 and record the t-stat and R^2 for the single best of the 14 bootstrapped predictors. We compare these

maximum-order bootstrapped statistics to the observed values.

We have performed the bootstrap as follows:

[1] Compute the first-order autocorrelation ρ for each topic.

[2] For each topic, simulate a time series: $x_{t+1} = \rho \cdot x_t + e_{t+1}$, where $x_0 = 0$ and e_{t+1} is randomly sampled from a standard normal distribution $N(0, 1)$. Each time series has a length of $n+500$ where n is the number of time observations in our sample from 1871 to 2019, and the first 500 generated observations are discarded to remove the dependency on the starting value x_0.

[3] Use the 14 generated series to run 14 univariate predictive regressions of next month's returns and store the highest t-statistic and adjusted R^2 among these 14 regressions (t and R^2 can come from two different regressions).

[4] Repeat steps 1-3 10,000 times to generate the distributions of t and R^2.

[5] Compute the proportion of times these values are greater than the t and R^2 of *War* as in Table 3.

	$\boldsymbol{p_t}$	$\boldsymbol{p_{R^2}}$	$\boldsymbol{p}_{\text{Both}}$
Value	0.0073	0.0973	0.0073

Bootstrap results

The results above present p-values obtained from the bootstrapped distribution of our test statistics: the highest t-statistic ($\boldsymbol{p_t}$), the highest adjusted R^2 ($\boldsymbol{p_{R^2}}$), and the joint consideration of both statistics ($\boldsymbol{p}_{Both}$). These results support the predictive power of *War* on stock market returns, as summarized below:

[1] $\boldsymbol{p_t} = 0.0073$ indicating that only 0.73% of the bootstrapped predictive regressions generates a t-statistic as high as or higher than the one observed for *War*. This strongly suggests that the association between *War* and stock returns is not due to random chance but is likely a real predictive relationship.

[2] $\boldsymbol{p_{R^2}} = 0.0973$, falling just shy of the traditional significance threshold $p < 0.10$. It nonetheless indicates that fewer than 10% of the simulations produced an adjusted R^2 as high as that observed for *War*, lending moderate support to the predictive relevance of this variable.

[3] $\boldsymbol{p}_{Both} = 0.0073$ presenting the joint consideration of both the highest t-statistic and adjusted

R^2. This result further reinforces the significant predictive power of *War*, mirroring the individual t-statistic p-value. This underscores the robustness of *War* as a predictor, not just in terms of its statistical significance but also in terms of the strength of the relationship.

These results affirm our premise that discourse surrounding war predicts subsequent stock returns significantly.

Table D.1
Predicting One-Month Market Returns: More Seed Words for *Pandemic* and *Natural Disasters*

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month, x_t is one of the discourse topics or the PLS indexes, and β, the coefficient of interest, measures the strength of predictability. "Shiller PLS" uses only the 12 topics from Shiller (2019). Seed words for 15 topics are listed below. Panel A reports the in-sample results. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. Panel B reports the out-of-sample R^2 in percentages. All out-of-sample forecasts are estimated recursively using the data available in the expanding estimation window. Out-of-sample R^2 is tested using the Clark and West (2007) MSFE-adjusted statistic. The sample period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Seed words:

- War: conflict, tension, terrorism, terrorist, war
- Pandemic: contagion, disease, epidemic, epidemiology, infection, outbreak, pandemic, public_health, quarantine, vaccination, vaccine, virus
- Natural Disaster: catastrophe, cyclone, destruction, drought, earthquake, flood, hurricane, landslide, mortality, natural_disaster, natural_hazard, storm, tornado, traumatic_exposure, tsunami, volcano, wildfire
- Panic: bank_failure, bank_panic, bank_run, crisis, depression, downturn, fear, financial_panic, hard_time, panic, recession
- Confidence: business_confidence, consumer_confidence
- Saving: compassion, family_morale, frugal, frugality, modesty, moral, poverty, saving
- Consumption: american_dream, conspicuous_consumption, consumption, equal_opportunity, equality, homeownership, luxury, patriotism, prosperity
- Money: bimetallism, devaluation, gold, gold_standard, inflation, monetary_standard, money, silver
- Tech: automate, computer, digital_divide, electronic_brain, invention, labor_save, labor_save_machine, machine, mechanize, network, technocracy, technological_unemployment, technology, unemployment
- Real Estate Boom: boom, bubble, flip, flipper, home_ownership, home_purchase, house_boom, house_bubble, land_boom, land_bubble, price_increase, real_estate_boom, real_estate_bubble, speculation
- Real Estate Crash: bust, crash, house_bust, house_crash, land_bust, land_crash, price_decrease, real_estate_bust, real_estate_crash
- Stock Bubble: advance_market, boom, bubble, bull, bull_market, bullish, earnings_per_share, inflate_market, margin, margin_requirement, market_boom, market_bubble, price_earn_ratio, price_increase, sell_short, short_sell, speculation, stock_market_boom, stock_market_bubble
- Stock Crash: bear, bear_market, bearish, bust, crash, fall_market, market_crash, stock_crash, stock_market_crash, stock_market_decline
- Boycott: anger, boycott, community, evil_business, excess_profit, fair_wage, moral, outrage, postpone_purchase, profiteer, protest, strike, wage_cut
- Wage: consumer_price, cost_of_live, cost_push, cost_push_inflation, high_wage, increase_wage, inflation, labor_union, rise_cost, wage, wage_demand, wage_lag, wage_price, wage_price_spiral

Table D.1
Predicting One-Month Market Returns:
More Seed Words for *Pandemic* and *Natural Disasters*

Panel A: In-Sample Results

	1871-2019	1871-1949	1950-2019	2000-2019
War	3.36 ***	2.56	4.37 ***	10.46 ***
t-stat	(3.03)	(1.59)	(2.68)	(3.57)
R^2	0.29	0.06	0.66	3.90
Pandemic	0.53	1.89	-2.03	6.60 **
t-stat	(0.43)	(1.15)	(-1.18)	(2.47)
R^2	-0.05	-0.02	0.05	1.29
Natural Disaster	1.36	1.66	1.66	-6.30 *
t-stat	(1.09)	(0.91)	(0.92)	(-1.88)
R^2	0.00	-0.04	-0.01	1.14
Panic	1.53	1.28	2.29	2.43
t-stat	(1.12)	(0.64)	(1.37)	(0.71)
R^2	0.02	-0.06	0.09	-0.19
Confidence	-0.19	-1.52	1.11	1.68
t-stat	(-0.15)	(-0.88)	(0.61)	(0.49)
R^2	-0.05	-0.05	-0.07	-0.31
Saving	-3.05 **	-2.60	-4.43 **	2.90
t-stat	(-2.39)	(-1.48)	(-2.35)	(1.26)
R^2	0.23	0.06	0.68	-0.09
Consumption	0.93	2.56	0.49	-4.44
t-stat	(0.67)	(1.31)	(0.31)	(-1.15)
R^2	-0.03	0.06	-0.11	0.35
Money	-1.78	-1.70	-1.29	1.52
t-stat	(-1.11)	(-0.76)	(-0.69)	(0.38)
R^2	0.04	-0.03	-0.05	-0.33
Tech	-1.70	-1.73	-2.29	-9.05 **
t-stat	(-1.17)	(-0.90)	(-1.19)	(-2.02)
R^2	0.03	-0.03	0.09	2.81
Real Estate Boom	0.31	-1.94	2.29	1.79
t-stat	(0.28)	(-1.32)	(1.44)	(0.60)
R^2	-0.05	-0.01	0.09	-0.30
Real Estate Crash	0.04	-0.69	1.07	-6.83 **
t-stat	(0.04)	(-0.45)	(0.57)	(-2.17)
R^2	-0.06	-0.09	-0.07	1.42
Stock Bubble	-0.56	-1.00	-0.17	-3.28
t-stat	(-0.43)	(-0.59)	(-0.09)	(-0.86)
R^2	-0.05	-0.08	-0.12	0.00
Stock Crash	-1.86	0.42	-3.53 **	-7.34 **
t-stat	(-1.55)	(0.27)	(-2.12)	(-2.57)
R^2	0.05	-0.10	0.39	1.70
Boycott	-1.63	-1.67	-2.26	3.54
t-stat	(-1.30)	(-1.02)	(-1.27)	(1.00)
R^2	0.03	-0.04	0.09	0.07
Wage	0.09	1.11	-1.37	0.38
t-stat	(0.07)	(0.58)	(-0.78)	(0.14)
R^2	-0.06	-0.07	-0.04	-0.42
PLS	4.71 ***	3.66 **	6.24 ***	10.17 ***
t-stat	(4.25)	(2.37)	(3.73)	(3.43)
R^2	0.62	0.23	1.47	3.65
Shiller PLS	4.14 ***	3.15 *	6.12 ***	2.22
t-stat	(3.32)	(1.81)	(3.65)	(0.65)
R^2	0.47	0.14	1.41	-0.23

Table D.1
Predicting One-Month Market Returns:
More Seed Words for *Pandemic* and *Natural Disasters*

Panel B: Out-of-Sample Results

	1881-2019	1881-1949	1950-2019	2000-2019
Panel A: OLS				
War	0.14***	-0.15	0.64***	1.39***
Pandemic	-0.11	-0.04	-0.22	0.07**
Natural Disaster	-0.07	-0.11	-0.01	-0.49
Panic	-0.03	-0.10	0.08	0.14
Confidence	-0.16	-0.14	-0.19	-0.12
Saving	0.17*	0.00	0.45**	-0.73
Consumption	-0.23	-0.19	-0.29	-0.20
Money	-0.05	-0.13	0.09	-0.39
Tech	-0.05	-0.10	0.04	0.51
Real Estate Boom	-0.09	0.00	-0.24	-0.03
Real Estate Crash	-0.15	-0.18	-0.10	-0.26
Stock Bubble	-0.13	-0.11	-0.16	-0.20
Stock Crash	0.01	-0.10	0.20**	0.37*
Boycott	-0.08	-0.15	0.05	-0.91
Wage	-0.13	-0.15	-0.11	-0.04
Panel B: PLS				
All Topics	-0.28**	-0.96	0.91***	1.21***
Shiller Topics	-0.67	-1.05	0.01	-1.13

Table D.2
Predicting One-Month Market Returns: More Seed Words for *War*

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month, x_t is one of the discourse topics or the PLS indexes, and β, the coefficient of interest, measures the strength of predictability. "Shiller PLS" uses only the 12 topics from Shiller (2019). Seed words for 15 topics are listed below. Panel A reports the in-sample results. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. Panel B reports the out-of-sample R^2 in percentages. All out-of-sample forecasts are estimated recursively using the data available in the expanding estimation window. Out-of-sample R^2 is tested using the Clark and West (2007) MSFE-adjusted statistic. The sample period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Seed words:

- War: army, battalion, battle, bomb, conflict, front_line, gun, military, munition, navy, officer, tension, terror, terrorism, terrorist, war, weapon
- Pandemic: contagion, disease, epidemic, epidemiology, infection, outbreak, pandemic, public_health, quarantine, vaccination, vaccine, virus
- Natural Disaster: catastrophe, cyclone, destruction, drought, earthquake, flood, hurricane, landslide, mortality, natural_disaster, natural_hazard, storm, tornado, traumatic_exposure, tsunami, volcano, wildfire
- Panic: bank_failure, bank_panic, bank_run, crisis, depression, downturn, fear, financial_panic, hard_time, panic, recession
- Confidence: business_confidence, consumer_confidence
- Saving: compassion, family_morale, frugal, frugality, modesty, moral, poverty, saving
- Consumption: american_dream, conspicuous_consumption, consumption, equal_opportunity, equality, homeownership, luxury, patriotism, prosperity
- Money: bimetallism, devaluation, gold, gold_standard, inflation, monetary_standard, money, silver
- Tech: automate, computer, digital_divide, electronic_brain, invention, labor_save, labor_save_machine, machine, mechanize, network, technocracy, technological_unemployment, technology, unemployment
- Real Estate Boom: boom, bubble, flip, flipper, home_ownership, home_purchase, house_boom, house_bubble, land_boom, land_bubble, price_increase, real_estate_boom, real_estate_bubble, speculation
- Real Estate Crash: bust, crash, house_bust, house_crash, land_bust, land_crash, price_decrease, real_estate_bust, real_estate_crash
- Stock Bubble: advance_market, boom, bubble, bull, bull_market, bullish, earnings_per_share, inflate_market, margin, margin_requirement, market_boom, market_bubble, price_earn_ratio, price_increase, sell_short, short_sell, speculation, stock_market_boom, stock_market_bubble
- Stock Crash: bear, bear_market, bearish, bust, crash, fall_market, market_crash, stock_crash, stock_market_crash, stock_market_decline
- Boycott: anger, boycott, community, evil_business, excess_profit, fair_wage, moral, outrage, postpone_purchase, profiteer, protest, strike, wage_cut
- Wage: consumer_price, cost_of_live, cost_push, cost_push_inflation, high_wage, increase_wage, inflation, labor_union, rise_cost, wage, wage_demand, wage_lag, wage_price, wage_price_spiral

Table D.2
Predicting One-Month Market Returns:
More Seed Words for *War*

Panel A: In-Sample Results

	1871-2019	1871-1949	1950-2019	2000-2019
War	3.24 ***	2.91 *	3.43 **	11.46 ***
t-stat	(2.98)	(1.79)	(2.17)	(3.92)
R^2	0.27	0.11	0.36	4.76
Pandemic	0.78	1.06	-0.52	6.64 **
t-stat	(0.65)	(0.66)	(-0.33)	(2.52)
R^2	-0.04	-0.08	-0.11	1.31
Natural Disaster	-0.41	-1.12	1.58	-4.94 *
t-stat	(-0.31)	(-0.55)	(0.88)	(-1.71)
R^2	-0.05	-0.07	-0.02	0.54
Panic	3.90 ***	4.90 **	2.59	0.06
t-stat	(2.79)	(2.34)	(1.57)	(0.02)
R^2	0.41	0.50	0.15	-0.42
Confidence	0.43	0.05	0.77	-1.74
t-stat	(0.39)	(0.03)	(0.49)	(-0.57)
R^2	-0.05	-0.11	-0.10	-0.30
Saving	-1.77	-1.50	-2.76	1.19
t-stat	(-1.35)	(-0.81)	(-1.63)	(0.42)
R^2	0.04	-0.05	0.19	-0.37
Consumption	-1.87	-1.32	-1.63	-3.54
t-stat	(-1.35)	(-0.69)	(-0.98)	(-0.98)
R^2	0.05	-0.06	-0.01	0.07
Money	-1.55	-1.56	-0.96	0.31
t-stat	(-0.99)	(-0.72)	(-0.50)	(0.07)
R^2	0.02	-0.04	-0.08	-0.42
Tech	-2.28	-2.49	-2.54	-7.54 **
t-stat	(-1.56)	(-1.22)	(-1.39)	(-2.12)
R^2	0.10	0.05	0.14	1.82
Real Estate Boom	2.02 *	0.20	3.75 **	3.02
t-stat	(1.70)	(0.13)	(2.24)	(0.93)
R^2	0.07	-0.10	0.45	-0.07
Real Estate Crash	-2.82 **	-3.40 **	-1.97	-4.62
t-stat	(-2.35)	(-1.97)	(-1.19)	(-1.46)
R^2	0.19	0.19	0.04	0.42
Stock Bubble	-2.46 *	-2.54	-2.37	-6.32 *
t-stat	(-1.88)	(-1.57)	(-1.16)	(-1.67)
R^2	0.13	0.06	0.11	1.15
Stock Crash	-1.18	-1.40	0.03	1.39
t-stat	(-0.97)	(-0.85)	(0.02)	(0.44)
R^2	-0.01	-0.06	-0.12	-0.35
Boycott	-1.13	-1.21	-1.68	4.66
t-stat	(-0.87)	(-0.73)	(-0.87)	(1.23)
R^2	-0.02	-0.07	-0.01	0.43
Wage	1.01	1.62	0.08	3.22
t-stat	(0.77)	(0.85)	(0.05)	(1.15)
R^2	-0.02	-0.04	-0.12	-0.01
PLS	5.61 ***	5.59 ***	5.63 ***	10.50 ***
t-stat	(4.97)	(3.32)	(3.38)	(3.43)
R^2	0.91	0.68	1.17	3.93
Shiller PLS	5.62 ***	5.93 ***	4.94 ***	5.46
t-stat	(4.22)	(2.97)	(2.95)	(1.51)
R^2	0.91	0.78	0.88	0.75

Table D.2
Predicting One-Month Market Returns:
More Seed Words for *War*

Panel B: Out-of-Sample Results

	1881-2019	1881-1949	1950-2019	2000-2019
Panel A: OLS				
War	0.08**	-0.15*	0.49***	1.35***
Pandemic	-0.11	-0.13	-0.06	0.13**
Natural Disaster	-0.12	-0.13	-0.11	0.03
Panic	0.19**	0.20*	0.16*	-0.34
Confidence	-0.09	-0.11	-0.05	-0.18
Saving	0.00	-0.06	0.12	-0.21
Consumption	0.01	-0.07	0.15	-0.05
Money	-0.08	-0.15	0.05	-0.19
Tech	0.03	-0.00	0.09	0.65
Real Estate Boom	0.05	-0.08	0.28**	0.17
Real Estate Crash	0.07*	0.06	0.10	0.59*
Stock Bubble	0.06	0.03	0.09	0.99*
Stock Crash	-0.08	-0.09	-0.06	-0.35
Boycott	-0.14	-0.21	-0.02	-0.67
Wage	-0.11	-0.13	-0.08	0.14
Panel B: PLS				
All Topics	0.03***	-0.52	0.98***	1.93***
Shiller Topics	-0.22**	-0.67	0.58**	0.30

Table D.3
Predicting One-Month Market Returns: Removing Duplicates in Seed Words

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month, x_t is one of the discourse topics or the PLS indexes, and β, the coefficient of interest, measures the strength of predictability. "Shiller PLS" uses only the 12 topics from Shiller (2019). Seed words for 15 topics are listed below. Panel A reports the in-sample results. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. Panel B reports the out-of-sample R^2 in percentages. All out-of-sample forecasts are estimated recursively using the data available in the expanding estimation window. Out-of-sample R^2 is tested using the Clark and West (2007) MSFE-adjusted statistic. The sample period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Seed words:

- War: army, battalion, battle, bomb, conflict, front_line, gun, military, munition, navy, officer, tension, terror, war, weapon
- Pandemic: contagion, disease, epidemic, epidemiology, infection, outbreak, pandemic, public_health, quarantine, vaccine, virus
- Natural Disaster: catastrophe, cyclone, destruction, drought, earthquake, flood, hurricane, landslide, mortality, natural_disaster, natural_hazard, storm, tornado, traumatic_exposure, tsunami, volcano, wildfire
- Panic: bank_failure, bank_panic, bank_run, crisis, depression, downturn, fear, financial_panic, hard_time, recession
- Confidence: business_confidence, consumer_confidence
- Saving: compassion, family_morale, frugal, modesty, moral, poverty, saving
- Consumption: american_dream, conspicuous_consumption, equal_opportunity, equality, home_ownership, luxury, patriotism, prosperity
- Money: bimetallism, devaluation, gold_standard, inflation, monetary_standard, money, silver
- Tech: automate, computer, digital_divide, electronic_brain, invention, labor_save_machine, mechanize, network, technocracy, technology, unemployment
- Real Estate Boom: flip, flipper, home_purchase, house_boom, house_bubble, land_boom, land_bubble, real_estate_boom, real_estate_bubble, speculation
- Real Estate Crash: house_bust, house_crash, land_bust, land_crash, real_estate_bust, real_estate_crash
- Stock Bubble: advance_market, bull_market, bullish, earnings_per_share, inflate_market, margin, market_bubble, price_earn_ratio, price_increase, short_sell, stock_bubble, stock_market_boom, stock_market_bubble
- Stock Crash: bear_market, bearish, fall_market, market_crash, stock_crash, stock_market_crash, stock_market_decline
- Boycott: anger, boycott, community, evil_business, excess_profit, fair_wage, moral, outrage, postpone_purchase, profiteer, protest, strike, wage_cut
- Wage: consumer_price, cost_of_live, cost_push, high_wage, increase_wage, inflation, labor_union, rise_cost, wage, wage_demand, wage_lag, wage_price_spiral

Table D.3
Predicting One-Month Market Returns: Removing Duplicates in Seed Words

Panel A: In-Sample Results

	1871-2019	1871-1949	1950-2019	2000-2019
War	3.36 ***	3.12 **	3.45 **	11.00 ***
t-stat	(3.13)	(2.04)	(2.13)	(3.71)
R^2	0.29	0.14	0.37	4.35
Pandemic	-0.54	-0.38	-1.56	5.04 *
t-stat	(-0.43)	(-0.21)	(-0.94)	(1.84)
R^2	-0.05	-0.10	-0.02	0.58
Natural Disaster	0.16	-0.36	1.25	-4.89 *
t-stat	(0.11)	(-0.18)	(0.71)	(-1.70)
R^2	-0.06	-0.10	-0.06	0.52
Panic	0.84	0.43	2.12	3.31
t-stat	(0.61)	(0.22)	(1.15)	(0.87)
R^2	-0.03	-0.10	0.06	0.01
Confidence	0.37	0.19	0.52	0.67
t-stat	(0.31)	(0.11)	(0.31)	(0.23)
R^2	-0.05	-0.10	-0.11	-0.41
Saving	-0.48	-0.22	-1.08	-3.29
t-stat	(-0.36)	(-0.12)	(-0.64)	(-1.05)
R^2	-0.05	-0.10	-0.07	0.00
Consumption	-0.82	-0.96	0.32	-1.44
t-stat	(-0.65)	(-0.51)	(0.19)	(-0.50)
R^2	-0.04	-0.08	-0.12	-0.34
Money	-1.53	-0.94	-1.78	0.05
t-stat	(-0.97)	(-0.42)	(-0.97)	(0.01)
R^2	0.02	-0.08	0.01	-0.42
Tech	-1.42	-0.67	-2.81	-9.90 **
t-stat	(-0.88)	(-0.29)	(-1.46)	(-2.31)
R^2	0.01	-0.09	0.20	3.44
Real Estate Boom	-1.70	-2.56	-0.80	-8.66 **
t-stat	(-1.32)	(-1.32)	(-0.46)	(-2.51)
R^2	0.03	0.06	-0.09	2.54
Real Estate Crash	-2.17	-3.20	-0.86	0.89
t-stat	(-1.56)	(-1.47)	(-0.53)	(0.27)
R^2	0.09	0.15	-0.09	-0.39
Stock Bubble	0.64	-1.32	2.77	2.81
t-stat	(0.53)	(-0.83)	(1.52)	(0.76)
R^2	-0.04	-0.06	0.19	-0.11
Stock Crash	1.07	2.41	-0.39	-1.78
t-stat	(0.86)	(1.36)	(-0.23)	(-0.53)
R^2	-0.02	0.04	-0.11	-0.30
Boycott	-2.09	-1.81	-3.11 *	3.13
t-stat	(-1.62)	(-1.08)	(-1.66)	(0.87)
R^2	0.08	-0.02	0.27	-0.04
Wage	1.61	2.21	1.43	6.31 **
t-stat	(1.01)	(0.96)	(0.85)	(2.54)
R^2	0.02	0.02	-0.04	1.15
PLS	4.65 ***	4.20 ***	5.36 ***	12.07 ***
t-stat	(4.22)	(2.61)	(3.29)	(4.22)
R^2	0.61	0.34	1.05	5.32
Shiller PLS	4.17 ***	4.06 **	5.03 ***	6.64 **
t-stat	(2.95)	(1.99)	(3.06)	(2.07)
R^2	0.48	0.31	0.91	1.32

Table D.3
Predicting One-Month Market Returns:
Removing Duplicates in Seed Words

Panel B: Out-of-Sample Results

	1881-2019	1881-1949	1950-2019	2000-2019
Panel A: OLS				
War	0.05**	-0.20*	0.50***	1.36***
Pandemic	-0.11	-0.13	-0.07	-0.14
Natural Disaster	-0.11	-0.14	-0.05	-0.12
Panic	-0.10	-0.16	-0.00	0.09
Confidence	-0.09	-0.10	-0.07	-0.05
Saving	-0.13	-0.19	-0.02	0.03
Consumption	-0.08	-0.10	-0.04	-0.04
Money	-0.10	-0.22	0.11	-0.14
Tech	-0.16	-0.28	0.05	0.41
Real Estate Boom	-0.03	0.04	-0.15	0.70***
Real Estate Bust	-0.04	0.02	-0.15	-0.61
Stock Bubble	-0.10	-0.08	-0.15	-0.07
Stock Crash	-0.15	-0.10	-0.25	-0.34
Boycott	-0.11	-0.27	0.18	-0.97
Wage	-0.08	-0.09	-0.05	0.39*
Panel B: PLS				
All Topics	-0.54*	-1.23	0.66***	1.57***
Shiller Topics	-1.07	-1.67	-0.03	-0.17

Table D.4
Predicting One-Month Market Returns: Removing Duplicates in Seed Words and Using 50 Unseeded Topics

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month, x_t is one of the discourse topics or the PLS indexes, and β, the coefficient of interest, measures the strength of predictability. "Shiller PLS" uses only the 12 topics from Shiller (2019). In this setup, we include 50 unseeded topics in addition to 15 seeded topics listed below. Panel A reports the in-sample results. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. Panel B reports the out-of-sample R^2 in percentages. All out-of-sample forecasts are estimated recursively using the data available in the expanding estimation window. Out-of-sample R^2 is tested using the Clark and West (2007) MSFE-adjusted statistic. The sample period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Seed words:

- War: army, battalion, battle, bomb, conflict, front_line, gun, military, munition, navy, officer, tension, terror, war, weapon
- Pandemic: contagion, disease, epidemic, epidemiology, infection, outbreak, pandemic, public_health, quarantine, vaccine, virus
- Natural Disaster: catastrophe, cyclone, destruction, drought, earthquake, flood, hurricane, landslide, mortality, natural_disaster, natural_hazard, storm, tornado, traumatic_exposure, tsunami, volcano, wildfire
- Panic: bank_failure, bank_panic, bank_run, crisis, depression, downturn, fear, financial_panic, hard_time, recession
- Confidence: business_confidence, consumer_confidence
- Saving: compassion, family_morale, frugal, modesty, moral, poverty, saving
- Consumption: american_dream, conspicuous_consumption, equal_opportunity, equality, home_ownership, luxury, patriotism, prosperity
- Money: bimetallism, devaluation, gold_standard, inflation, monetary_standard, money, silver
- Tech: automate, computer, digital_divide, electronic_brain, invention, labor_save_machine, mechanize, network, technocracy, technology, unemployment
- Real Estate Boom: flip, flipper, home_purchase, house_boom, house_bubble, land_boom, land_bubble, real_estate_boom, real_estate_bubble, speculation
- Real Estate Crash: house_bust, house_crash, land_bust, land_crash, real_estate_bust, real_estate_crash
- Stock Bubble: advance_market, bull_market, bullish, earnings_per_share, inflate_market, margin, market_bubble, price_earn_ratio, price_increase, short_sell, stock_bubble, stock_market_boom, stock_market_bubble
- Stock Crash: bear_market, bearish, fall_market, market_crash, stock_crash, stock_market_crash, stock_market_decline
- Boycott: anger, boycott, community, evil_business, excess_profit, fair_wage, moral, outrage, postpone_purchase, profiteer, protest, strike, wage_cut
- Wage: consumer_price, cost_of_live, cost_push, high_wage, increase_wage, inflation, labor_union, rise_cost, wage, wage_demand, wage_lag, wage_price_spiral

Table D.4
Predicting One-Month Market Returns:
Removing Duplicates in Seed Words and Using 50 Unseeded Topics

Panel A: In-Sample Results

	1871-2019	1871-1949	1950-2019	2000-2019
War	3.28 ***	3.26 **	2.89 *	9.50 ***
t-stat	(3.03)	(2.21)	(1.73)	(3.35)
R^2	0.27	0.16	0.22	3.14
Pandemic	2.98 **	3.45 **	1.73	4.86 *
t-stat	(2.33)	(2.18)	(1.01)	(1.79)
R^2	0.22	0.20	0.00	0.51
Natural Disaster	-0.28	0.37	-1.59	-7.51 **
t-stat	(-0.21)	(0.20)	(-0.85)	(-2.01)
R^2	-0.05	-0.10	-0.02	1.80
Panic	-0.34	-0.49	0.42	-0.99
t-stat	(-0.22)	(-0.23)	(0.18)	(-0.19)
R^2	-0.05	-0.10	-0.11	-0.39
Confidence	-0.98	-2.18	0.74	-2.51
t-stat	(-0.80)	(-1.28)	(0.43)	(-0.79)
R^2	-0.03	0.01	-0.10	-0.18
Saving	-0.41	-0.39	-0.84	1.14
t-stat	(-0.19)	(-0.12)	(-0.47)	(0.45)
R^2	-0.05	-0.10	-0.09	-0.37
Consumption	1.90	4.14	-1.63	-4.99
t-stat	(0.98)	(1.50)	(-0.88)	(-1.52)
R^2	0.05	0.33	-0.01	0.56
Money	-1.08	-2.28	0.96	3.62
t-stat	(-0.85)	(-1.33)	(0.56)	(1.09)
R^2	-0.02	0.03	-0.08	0.09
Tech	-0.33	2.13	-2.89	-5.57
t-stat	(-0.25)	(1.17)	(-1.52)	(-1.50)
R^2	-0.05	0.01	0.22	0.80
Real Estate Boom	-3.91 ***	-2.09	-5.44 ***	-4.26
t-stat	(-2.75)	(-1.03)	(-2.71)	(-1.34)
R^2	0.41	0.00	1.09	0.29
Real Estate Crash	-1.34	-2.12	-0.31	1.04
t-stat	(-1.16)	(-1.32)	(-0.18)	(0.33)
R^2	-0.00	0.01	-0.12	-0.38
Stock Bubble	0.39	-0.85	1.32	0.83
t-stat	(0.30)	(-0.41)	(0.74)	(0.25)
R^2	-0.05	-0.09	-0.05	-0.40
Stock Crash	1.14	0.12	2.96 *	2.27
t-stat	(0.81)	(0.06)	(1.91)	(0.79)
R^2	-0.02	-0.11	0.24	-0.22
Boycott	-0.09	1.28	-2.17	0.65
t-stat	(-0.07)	(0.79)	(-1.25)	(0.19)
R^2	-0.06	-0.06	0.07	-0.41
Wage	-2.03	-2.15	-2.93 *	-7.30 **
t-stat	(-1.41)	(-1.05)	(-1.67)	(-2.44)
R^2	0.07	0.01	0.23	1.68
PLS	4.55 ***	4.38 ***	4.73 ***	10.82 ***
t-stat	(4.22)	(2.97)	(2.74)	(3.79)
R^2	0.58	0.38	0.79	4.19
Shiller PLS	4.69 **	4.92 *	4.67 ***	3.26
t-stat	(2.55)	(1.69)	(2.61)	(1.17)
R^2	0.62	0.50	0.77	-0.01

Table D.4
Predicting One-Month Market Returns:
Removing Duplicates in Seed Words and Using 50 Unseeded Topics

Panel B: Out-of-Sample Results

	1881-2019	1881-1949	1950-2019	2000-2019
Panel A: OLS				
War	0.08**	-0.12**	0.41**	0.97**
Pandemic	0.19**	0.23**	0.14*	-0.23
Natural Disaster	-0.10	-0.14	-0.03	-0.07
Panic	-0.17	-0.19	-0.13	-0.24
Confidence	-0.11	-0.06	-0.20	0.07
Saving	-0.31	-0.46	-0.04	-0.06
Consumption	-0.06	0.20	-0.51	-0.61
Money	-0.06	0.01	-0.17	-0.33
Tech	-0.25	-0.15	-0.44	-0.06
Real Estate Boom	0.14**	-0.35	1.01***	0.05
Real Estate Bust	-0.01	0.05	-0.12	-0.22
Stock Bubble	-0.23	-0.31	-0.10	-0.18
Stock Crash	-0.08	-0.15	0.06	0.11
Boycott	-0.07	-0.05	-0.12	-0.08
Wage	0.02	-0.01	0.07	0.80*
Panel B: PLS				
All Topics	0.03**	-0.22*	0.46**	1.11**
Shiller Topics	-0.63	-0.49	-0.88	-1.56

Table D.5
Predicting One-Month Market Returns: Using One Seeded Topic (*War*) and 50 Unseeded Topics

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month, x_t is *War*, and β, the coefficient of interest, measures the strength of predictability. The only seed word for *War* is "war." Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. The last column reports the out-of-sample R^2 in percentages. All out-of-sample forecasts are estimated recursively using the data available in the expanding estimation window. Out-of-sample R^2 is tested using the Clark and West (2007) MSFE-adjusted statistic. The sample period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	β	t-stat	R^2	R^2_{OS}
1871-2019	3.80 ***	(3.46)	0.39	0.22 ***
1871-1949	4.05 ***	(2.62)	0.31	0.08 **
1950-2019	2.99 *	(1.78)	0.24	0.46 **
2000-2019	8.88 ***	(3.01)	2.69	1.05 **

Table D.6
Predicting One-Month Market Returns: Bootstrap

This table reports the empirical p-values generated by the bootstrap described in D. Column p_t reports the percentage of 10,000 iterations that yield a Newey-West t-statistic positively larger than the t-statistic generated by five War seed words in the corresponding sample period (2.6 for 1871-2019, 3.04 for 1950-2019, and 2.86 for 2000-2019). Column p_{IS-R^2} reports the percentage of 10,000 iterations that yield an in-sample R^2 larger than the in-sample R^2 generated by five War seed words in the corresponding sample period (0.26% for 1871-2019, 0.82% for 1950-2019, and 2.15% for 2000-2019). Column p_{OS-R^2} reports the percentage of 10,000 iterations that yield an out-of-sample R^2 larger than the out-of-sample R^2 generated by five War seed words in the corresponding sample period (0.08% for 1881-2019, 0.26% for 1950-2019, and 0.77% for 2000-2019). The last column of Table D.6 reports the percentage of 10,000 iterations that simultaneously beat all three statistic of five War seed words in the corresponding sample period. The full sample period is from January 1871 to October 2019.

	p_t	p_{IS-R^2}	p_{OS-R^2}	p_{all}
1871-2019	0.0185	0.0296	0.0320	0.0053
1950-2019	0.0021	0.0051	0.0398	0.0006
2000-2019	0.0148	0.0437	0.0099	0.0043

Table D.7
Predicting One-Month Market Returns: Using Topic Innovation

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month; x_t is innovation estimated from AR(1) for each of the discourse topics in the baseline model in the main text or the PLS indexes constructed from topic innovations; and β measures the strength of predictability. "Shiller PLS" uses only the 12 topics from Shiller (2019). Panel A reports the in-sample results. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. Panel B reports the out-of-sample R^2 in percentages. All out-of-sample forecasts are estimated recursively using the data available in the expanding estimation window. Out-of-sample R^2 is tested using the Clark and West (2007) MSFE-adjusted statistic. The sample period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table D.7
Predicting One-Month Market Returns: Using Topic Innovation

Panel A: In-Sample Results

	1871-2019	1871-1949	1950-2019	2000-2019
War	3.54 ***	4.25 **	2.42	6.11 *
t-stat	(2.93)	(2.30)	(1.63)	(1.90)
R^2	0.33	0.35	0.12	1.05
Pandemic	-2.51 **	-3.94 **	-1.38	-1.90
t-stat	(-1.98)	(-2.13)	(-0.81)	(-0.57)
R^2	0.14	0.29	-0.04	-0.28
Panic	1.05	1.87	0.64	1.88
t-stat	(0.89)	(1.14)	(0.41)	(0.59)
R^2	-0.02	-0.02	-0.10	-0.28
Confidence	0.71	-0.08	1.19	0.72
t-stat	(0.60)	(-0.05)	(0.71)	(0.23)
R^2	-0.04	-0.11	-0.06	-0.40
Saving	-0.62	-0.97	-0.40	0.10
t-stat	(-0.48)	(-0.54)	(-0.23)	(0.04)
R^2	-0.04	-0.08	-0.11	-0.42
Consumption	1.91	1.41	3.73 **	0.05
t-stat	(1.50)	(0.74)	(2.30)	(0.02)
R^2	0.06	-0.06	0.45	-0.42
Money	-2.64 **	-1.20	-3.99 **	-0.52
t-stat	(-1.97)	(-0.57)	(-2.41)	(-0.15)
R^2	0.16	-0.07	0.53	-0.41
Tech	-1.29	0.24	-3.63 *	-13.98 ***
t-stat	(-0.89)	(0.11)	(-1.93)	(-3.02)
R^2	-0.01	-0.10	0.42	7.28
Real Estate Boom	-2.86 **	-2.99 *	-3.10 *	-6.87 **
t-stat	(-2.37)	(-1.79)	(-1.85)	(-2.23)
R^2	0.20	0.12	0.27	1.44
Real Estate Crash	0.48	0.05	1.43	-1.50
t-stat	(0.40)	(0.03)	(0.86)	(-0.46)
R^2	-0.05	-0.11	-0.04	-0.33
Stock Bubble	0.02	-0.03	-0.07	-2.76
t-stat	(0.02)	(-0.02)	(-0.04)	(-0.93)
R^2	-0.06	-0.11	-0.12	-0.12
Stock Crash	0.83	2.06	-0.10	0.55
t-stat	(0.65)	(1.12)	(-0.06)	(0.18)
R^2	-0.04	0.00	-0.12	-0.41
Boycott	-1.79 *	-3.10 **	0.02	4.43
t-stat	(-1.67)	(-2.24)	(0.01)	(1.36)
R^2	0.04	0.14	-0.12	0.35
Wage	0.71	-0.26	2.15	9.27 ***
t-stat	(0.53)	(-0.14)	(1.26)	(3.15)
R^2	-0.04	-0.10	0.07	2.97
PLS	6.25 ***	7.11 ***	6.40 ***	9.49 ***
t-stat	(4.62)	(3.73)	(3.75)	(2.67)
R^2	1.14	1.17	1.55	3.13
Shiller PLS	4.60 ***	4.83 **	5.71 ***	7.31 **
t-stat	(3.39)	(2.51)	(3.39)	(2.13)
R^2	0.59	0.48	1.21	1.69

Table D.7
Predicting One-Month Market Returns: Using Topic Innovation

Panel B: Out-of-Sample Results

	1881-2019	1881-1949	1950-2019	2000-2019
Panel A: OLS				
War	0.26***	0.35**	0.11*	1.04**
Pandemic	0.09*	0.30**	-0.28	0.08
Panic	-0.03	0.00	-0.10	0.11
Confidence	-0.10	-0.11	-0.06	-0.10
Saving	-0.08	-0.09	-0.07	-0.02
Consumption	-0.01	-0.11	0.16*	0.00
Money	0.05*	-0.18	0.44**	-0.13
Tech	-0.39	-0.69	0.12*	0.52**
Real Estate Boom	0.15**	0.11	0.22*	1.12**
Real Estate Bust	-0.11	-0.17	-0.02	-0.09
Stock Bubble	-0.10	-0.06	-0.18	-0.03
Stock Crash	-0.06	0.07	-0.30	-0.02
Boycott	-0.06	0.02	-0.19	-0.48
Wage	-0.14	-0.17	-0.08	-0.43
Panel B: PLS				
All Topics	-0.47**	-0.42	-0.56**	1.56**
Shiller Topics	-0.70	-0.95	-0.26	0.65

E Topic Weights Constructed by Counts of Seed Words

In this appendix, we conduct a robustness check for the main empirical results in the paper. Specifically, we investigate whether the sLDA model adds economic insight beyond a simple count of seed words in the news.

While the majority of finance papers that employ textual analysis rely on simple counts of words from a predefined dictionary (for reviews, see Loughran and McDonald (2016) and Loughran and McDonald (2020)), recent studies have exploited statistical unsupervised topic modeling to extract thematic contents from textual data (e.g., Dyer, Lang, and Stice-Lawrence (2017), Choudhury et al. (2019), Brown, Crowley, and Elliott (2020), and Bybee et al. (2023)). This paper blends the two branches by employing a semisupervised model in which we inject seed words into the topic model to extract desired contents. Hence, a natural question is whether the sLDA model reveals any additional information beyond a simple count of those seed words in the news. To answer this question, we construct topic weights by simply counting the occurrences of seed words and scaling them by the total number of ngrams in the article.

Table E.1 reports the summary statistics for these topic weights. *War* is still the most frequently mentioned and most volatile topic with a monthly mean of 0.12% and standard deviation of 0.09%. It implies, on average, 0.12% of monthly *NYT* words are related to five *War* seed words (war, tension, conflict, terrorism, and terrorist).[10] This might seem low, but it shows the limitation of the frequency count approach. It relies on a list of comprehensive words, and their sources can be subjective. In contrast, sLDA lets the machine captures the words co-occurring with the seed words; thus, it has less subjectivity than the words count approach. *War* has the first-order autocorrelation of 96%, much higher than the percentage (78%) obtained via the sLDA one. To remain consistent with the sLDA model, we also construct the PLS index from all topics.[11] Once again, the PLS index heavily loads on *War* and strongly correlates with this topic with a correlation coefficient of 99%.

[10] On average, every month, we have about 2 million ngrams in the *NYT* (the product of 4000 articles and 500 ngrams per article). For War, 0.12% of this number means 2400 mentions of the five War seed words.

[11] Comparing to the PLS weight from sLDA, the PLS weight of the seed word count is much lower due to its low topics weight. Recall that the PLS weight is the slope from regressing the topic weight on market returns; thus, the different scales of the dependent variable result in the different scale of the slope.

To investigate whether manually constructed topics have the same market implications as the sLDA topics, we first use them to predict the monthly market returns in the sample. Table E.2 shows that, in general, both *War* and the PLS index can powerfully and positively predict market excess returns one month ahead, consistent with the sLDA results. The other manually constructed topics, similar to the sLDA topics, do not display any consistent predictability pattern.

In Panel A of Table E.3, we find that the manually constructed PLS index is not significant after controlling for specific economic predictors (book-market, long-term yield, and term spread), and, in Panel B, the manually counted topic index loses its significance when controlling for other uncertainty variables.

The in-sample predictability results can be biased if the predictors are highly consistent, which is the case for the manually counted *War* and PLS index. Hence, in Table E.4, we report the out-of-sample R^2 computed with the frequency-based topics. Unsurprisingly, over the whole evaluation period of 1881–2019, the frequency count *War* index produces a much lower R^2_{OS} than the sLDA one: 0.08% versus 0.17%. The sLDA one continues to outperform in each subperiod. Similarly, the manually constructed topics via PLS greatly underperform their sLDA counterparts across all samples.

In sum, topic weights constructed with simple seed word counts yield monthly in-sample prediction results in line with the sLDA ones but substantially underperform in out-of-sample predictability. Moreover, the frequency-based topic index does not contain additional economic insights beyond the well-known economic and uncertainty predictors. These results indicate that the limited set of seed words fails to capture the whole universe of terms belonging to the same topic, and, hence, we need a statistical way to uncover and cluster them.

Table E.1
Summary Statistics:
Topic Weights Constructed by Counts of Seed Words

This table presents the summary statistic of the time series of 14 monthly topic weights from January 1871 to October 2019 constructed by frequency counts of seed words. Panel A reports the first and second moments; Panel B reports the autocorrelations from first- to fourth-order; Panel C reports the loading on each topic in constructing a partial least square (PLS) topic index, and Panel D report the correlations among topics. All numbers (except sample size) are in percentages.

	N	Mean	SD	Q1	Median	Q3	AC(1)	PLS Weights	Corr PLS
War	1784	0.12	0.09	0.07	0.09	0.12	95.78	0.10	99.29
Pandemic	1784	0.00	0.00	0.00	0.00	0.00	58.34	0.00	0.39
Panic	1784	0.05	0.02	0.03	0.04	0.05	92.92	0.02	-2.40
Confidence	1784	0.00	0.00	0.00	0.00	0.00	79.75	0.00	-1.10
Saving	1784	0.02	0.01	0.02	0.02	0.03	69.60	0.00	-0.86
Consumption	1784	0.02	0.00	0.01	0.02	0.02	73.99	0.00	14.21
Money	1784	0.12	0.05	0.09	0.12	0.14	87.86	-0.01	-51.43
Tech	1784	0.06	0.04	0.02	0.04	0.09	97.68	0.02	4.96
Real Estate Boom	1784	0.01	0.00	0.01	0.01	0.02	79.31	-0.00	-16.47
Real Estate Crash	1784	0.01	0.00	0.00	0.01	0.01	69.89	0.00	-3.77
Stock Bubble	1784	0.02	0.01	0.02	0.02	0.03	68.50	-0.00	-20.57
Stock Crash	1784	0.04	0.01	0.03	0.04	0.04	71.40	-0.00	-22.43
Boycott	1784	0.10	0.03	0.08	0.10	0.12	72.28	0.01	12.96
Wage	1784	0.03	0.02	0.02	0.02	0.03	84.77	0.01	26.48
PLS	1784	76.83	86.68	29.03	57.68	84.71	95.97		

Table E.2
Predicting One-Month Market Returns: Topic Weights Constructed by Counts of Seed Words

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month, x_t is one of the topics or the PLS index constructed by frequency counts of seed words, and β is the coefficient of interest that measures the strength of predictability. Returns are annualized percentages, and the independent variable is standardized to zero mean and unit variance presented in percentage. Adjusted R^2 is in percentage and t-stat is computed with the Newey and West (1987) standard errors. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	1871-2019	1871-1949	1950-2019	2000-2019
War	3.18 ***	3.64 **	4.80 ***	8.07 ***
t-stat	(2.60)	(2.17)	(3.04)	(2.86)
R^2	0.26	0.23	0.82	2.15
Pandemic	0.05	-0.61	0.73	6.96 ***
t-stat	(0.05)	(-0.39)	(0.54)	(2.96)
R^2	-0.06	-0.10	-0.10	1.49
Panic	2.48	2.29	2.20	1.53
t-stat	(1.38)	(0.48)	(0.88)	(0.29)
R^2	0.13	0.03	0.08	-0.33
Confidence	1.23	0.57	0.88	-0.45
t-stat	(1.00)	(0.22)	(0.51)	(-0.14)
R^2	-0.01	-0.10	-0.09	-0.42
Saving	1.67	3.23	-0.17	5.06
t-stat	(0.93)	(0.99)	(-0.09)	(1.47)
R^2	0.03	0.16	-0.12	0.59
Consumption	1.82	3.98 *	1.03	2.29
t-stat	(1.13)	(1.74)	(0.59)	(0.63)
R^2	0.05	0.29	-0.08	-0.22
Money	-0.58	0.18	-1.00	-0.55
t-stat	(-0.37)	(0.08)	(-0.52)	(-0.12)
R^2	-0.05	-0.10	-0.08	-0.41
Tech	1.21	-0.32	0.69	0.17
t-stat	(1.03)	(-0.12)	(0.42)	(0.05)
R^2	-0.01	-0.10	-0.10	-0.42
Real Estate Boom	-0.68	-0.47	-3.67 *	-2.98
t-stat	(-0.47)	(-0.27)	(-1.78)	(-0.60)
R^2	-0.04	-0.10	0.43	-0.07
Real Estate Crash	2.43 **	0.50	3.09 **	2.47
t-stat	(2.17)	(0.22)	(2.05)	(0.93)
R^2	0.12	-0.10	0.27	-0.18
Stock Bubble	-0.98	-1.88	-1.19	-3.38
t-stat	(-0.86)	(-1.20)	(-0.77)	(-1.07)
R^2	-0.03	-0.02	-0.06	0.03
Stock Crash	-0.06	-1.54	2.54	-1.23
t-stat	(-0.04)	(-0.85)	(1.22)	(-0.25)
R^2	-0.06	-0.05	0.14	-0.36
Boycott	1.15	1.01	0.22	7.29 ***
t-stat	(0.86)	(0.56)	(0.13)	(2.96)
R^2	-0.02	-0.08	-0.12	1.67
Wage	2.13	3.34	-0.21	2.73
t-stat	(1.50)	(1.64)	(-0.10)	(0.93)
R^2	0.08	0.18	-0.12	-0.13
PLS	3.13 **	3.40 **	4.78 ***	7.92 ***
t-stat	(2.54)	(2.03)	(2.93)	(2.72)
R^2	0.24	0.19	0.81	2.05
Shiller PLS	1.83	0.77	2.60	4.12
t-stat	(1.43)	(0.39)	(1.49)	(1.64)
R^2	0.05	-0.09	0.16	0.25

Table E.3
Predicting One-Month Market Returns:
Discourse Topics versus Economic, Sentiment, and Uncertainty Predictors
Topic Weights Constructed by Counts of Seed Words

This table presents the results of the following univariate predictive regression:

$$R^e_{t+1} = \alpha + \gamma z_t + \epsilon_{t+1},$$

and the following bivariate predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \gamma z_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month and x_t is the discourse topic PLS index. In Panel A, z_t is one of the 14 economic predictors from Goyal and Welch (2008), output gap from Cooper and Priestley (2009), and short interest from Rapach, Ringgenberg, and Zhou (2016); in Panel B, z_t is one of the uncertainty variables (financial and macro uncertainty indexes from Jurado, Ludvigson, and Ng (2015), economic policy uncertainty index from Baker, Bloom, and Davis (2016), disagreement index from Huang, Li, and Wang (2020), implied volatility (VIX), and news implied volatility (NVIX) from Manela and Moreira (2017)), sentiment variables (news sentiment, investor sentiment from Baker and Wurgler (2006), aligned sentiment from Huang et al. (2015), and manager sentiment from Jiang et al. (2019)), or Shiller's confidence indexes: one-year confidence index and crash confidence index. The last row of each panel reports the results using the PLS index constructed from all predictors in that panel. Returns and correlations are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table E.3
Predicting One-Month Market Returns:
Discourse Topics versus Economic, Sentiment, and Uncertainty Predictors
Topic Weights Constructed by Counts of Seed Words (Cont.)

Panel A: Economic Predictors

	Univariate		Bivariate			
Economic Predictor	γ	R^2	β	γ	R^2	Period
Dividend-price ratio (DP)	1.39	0.00	2.99 **	0.98	0.22	187101-201910
Dividend yield (DY)	2.03	0.07	2.90 **	1.61	0.27	187102-201910
Earnings-price ratio (EP)	2.46	0.13	2.70 **	1.83	0.29	187101-201910
Dividend payout ratio (DE)	-1.00	-0.03	3.07 **	-0.74	0.21	187101-201910
Stock variance (SVAR)	-0.08	-0.06	3.13 **	-0.06	0.19	187101-201910
Book-to-market Ratio (BM)	5.19	0.58	2.40	4.80	0.63	192103-201910
Net equity expansion (NTIS)	-4.11	0.31	3.11 *	-3.91	0.45	192612-201910
Treasury bill rate (TBL)	-3.68 **	0.36	2.23 *	-3.02 *	0.44	187101-201910
Long term bond yield (LTY)	-2.82	0.11	2.50	-1.90	0.16	191901-201910
Long term bond return (LTR)	3.36 *	0.18	3.50 **	3.51 *	0.38	192601-201910
Term spread (TMS)	-2.70	0.10	2.55	-1.77	0.15	191901-201910
Default yield spread (DFY)	2.87	0.12	3.29 **	2.97	0.30	191901-201910
Default return spread (DFR)	2.30	0.04	3.34 **	2.29	0.22	192601-201910
Inflation (INFL)	-3.32	0.20	3.00 **	-3.98 *	0.34	191302-201910
Output Gap (OG)	-3.39	0.20	3.40 **	-3.65	0.40	191902-201910
Short Interest (SI)	-5.70 **	0.94	5.18 **	-6.63 ***	1.65	197301-201412
Economic PLS	4.40 *	0.48	5.85 ***	6.16 **	1.37	197301-201412

Table E.3
Predicting One-Month Market Returns:
Discourse Topics versus Economic, Sentiment, and Uncertainty Predictors
Topic Weights Constructed by Counts of Seed Words (Cont.)

Panel B: Uncertainty and Sentiment Predictors

	Correlations	Univariate		Bivariate			
Economic Predictor	Corr. with PLS	γ	R^2	β	γ	R^2	Period
Financial uncertainty	-7.56 **	-5.75 **	1.15	3.29 *	-5.50 *	1.43	196007-201910
Macro uncertainty	-15.02 ***	-4.30	0.58	3.13	-3.83	0.81	196007-201910
Economic policy uncertainty	17.59 ***	4.03	0.38	3.84	3.36	0.69	198501-201910
Implied volatility (VIX)	3.18	0.40	-0.27	4.93 **	0.25	0.46	199001-201910
News implied volatility (NVIX)	5.70 **	0.03	-0.07	2.90 **	-0.13	0.10	188907-201603
Disagreement	5.21	-8.43 ***	2.43	4.02 *	-8.63 ***	2.85	196912-201812
News sentiment	8.37 ***	-0.52	-0.05	3.20 **	-0.78	0.21	186612-201910
Investor sentiment (BW)	7.18 *	-2.50	0.08	3.70 *	-2.77	0.44	196507-201812
Investor sentiment (PLS)	-12.05 ***	-7.32 ***	1.86	2.66	-7.00 ***	1.97	196507-201812
Manager sentiment	-42.60 ***	-9.06 ***	3.32	2.37	-8.05 **	2.99	200301-201712
Shiller's one-year confidence index	12.50 *	-4.77	0.48	7.85 **	-5.75 **	2.54	200107-201910
Shiller's crash confidence index	16.65 **	-2.07	-0.28	7.69 ***	-3.35	1.64	200107-201910
Uncertainty PLS	-17.38 **	2.38	-0.39	6.24 *	3.46	0.59	200301-201603

Table E.4
Out-Of-Sample R^2:
Topic Weights Constructed by Counts of Seed Words

This table reports the out-of-sample R^2_{OS} statistic (Campbell and Thompson 2008) in predicting the monthly excess market return using the economic topics constructed by frequency counts of seed words. Panels A and B report results using OLS and PLS, respectively. All out-of-sample forecasts are estimated recursively using data available in the expanding estimation window. All numbers are in percentages. ***, **, and * indicate 1%, 5%, and 10% significance of the Clark and West (2007) MSFE-adj statistic. The evaluation period begins in January 1881, and the whole sample is from January 1871 to October 2019.

	1881-2019	1881-1949	1950-2019	2000-2019
Panel A: OLS				
Dividend-price ratio (DP)	-0.60	-0.81	-0.25	0.05
Dividend yield (DY)	-0.48	-0.39	-0.64	0.04
Earnings-price ratio (EP)	-0.14	-0.07	-0.26	-0.35
Dividend payout ratio (DE)	-0.83	-1.12	-0.33	-1.06
Stock variance (SVAR)	-1.68	-2.18	-0.79	-0.86
Treasury bill rate (TBL)	0.07 **	-0.05	0.26 **	0.45
War	0.08 ***	-0.02 **	0.26 *	0.77 **
Pandemic	-0.11	-0.13	-0.06	-0.41
Panic	-0.48	-0.71	-0.08	-2.37
Confidence	-0.45	-0.48	-0.39	-1.52
Saving	-0.13	-0.01	-0.34	0.36 *
Consumption	-0.20	-0.02 *	-0.52	0.11
Money	-0.17	-0.26	-0.01	-0.02
Tech	-0.34	-0.42	-0.22	-0.64
Real Estate Boom	-0.14	-0.13	-0.16	-0.14
Real Estate Crash	-0.09	-0.31	0.31 **	-1.05
Stock Bubble	-0.08	-0.04	-0.14	0.12
Stock Crash	-0.18	-0.17	-0.21	-0.10
Boycott	-0.26	-0.42	0.02	0.30 ***
Wage	-0.32	-0.39	-0.20	0.19
Panel B: PLS				
Economic	-0.84	-1.11	-0.38	-0.55
All Topics	-0.26 *	-0.53 *	0.22 *	0.63 *
Shiller Topics	-0.69	-1.14	0.10	-0.59

F Discourse Topics from the *WSJ*

In this appendix, we check whether discourse topics extracted from 660 thousand *WSJ* articles predict stock market returns over the period 2000–2019.[12] WSJ is another mainstream media in the US, and readers are stock market participants. We apply sLDA to *WSJ* as an out-of-sample check. We focus on the past 20 years because this is the period where the *NYT* topics show the most robust predictability. We apply the same estimation method described in Section 3 to obtain the 14 time series of topic weights from the *WSJ* data.

Before extracting the 14 topics from the *WSJ* articles, we also conduct text-processing steps. Similar to the procedure applied to the *NYT* articles, we remove articles with limited content indicated by the pattern of the section they belong to if the section label is available and then by the pattern of their title. These section and title patterns are constructed by manually examining the articles and are available upon request. See Appendix B for the description of text processing, cleaning, and converting into ngrams.

We plot the word clouds and time series of each topic in Figures F.1 and F.2. We report the summary statistics for these topics in Table F.1.

Table F.2 reports results in predicting the excess market returns one month ahead using all *WSJ* topics. Consistent with the *NYT* results, *War* constructed from *WSJ* is a strongly positive market predictor over 2000-2019, significant at the 1% level. Specifically, a one standard deviation increase in *War* attention is associated with an 8.3% annualized increase in market returns next month. Its R^2_{OS}, constructed in an expanding window fashion with an initial 60-month training period, is 1.53% (also significant at the 1% level). Besides *War*, *Stock Bubble* also shows significant prediction results (at the 5% level), although it is a negative predictor. Its R^2_{OS} is 0.89%, significant at the 10% level.

We also aggregate the topics with the PLS technique using all 14 topics (the "PLS" row) and 12 topics from Shiller (2019) (the "Shiller PLS" row). Both indexes display in-sample solid predictability. However, as the sample is small and the PLS method has many parameters to

[12] Following the method in Section 3, we use the first 120 months from 1990 to 1999 to construct the first monthly topic weights.

estimate, it yields poor OOS results.

Table F.3 shows prediction results over the long horizons. In line with the *NYT* results, *WSJ War* can predict the stock market returns up to 36 months ahead. *Stock Bubble* has predictability for up to 12 months, with the strongest result obtained within three months. The PLS index can strongly predict the market for up to 36 months, significant at the 1% level across all horizons. This result is expected because the PLS index is constructed to optimize its in-sample predictability over this 20-year sample.

Overall, we have found consistent results between the *NYT* and *WSJ* topics over the past 20 years. Across the two national newspapers, attention paid to the *War* topic has been a strong market predictor since 2000. In addition, we also document that *Stock Bubble* is a negative predictor for the *WSJ* articles. We conjecture that Stock Bubble captures the stock market state: news talks more about Stock Bubble when the market is overvalued, foreshadowing future corrections, resulting in pessimistic predictions.

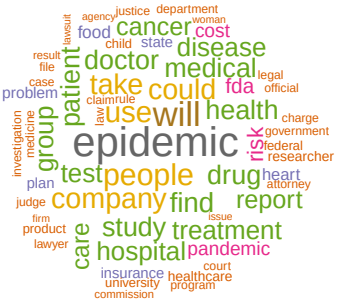
Figure F.1. Discourse Topic Contents from the *WSJ*

This figure plots the over-time frequencies of n-grams per each topic constructed according to the sLDA model. The size of each n-gram indicates its frequency. The sample period is from January 2000 to October 2019.

War

Panic

Pandemic



Monetary Standard

Real Estate Booms

Boycotts and Evil Business

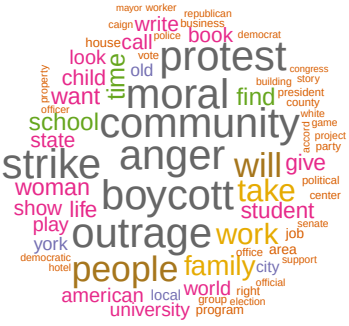
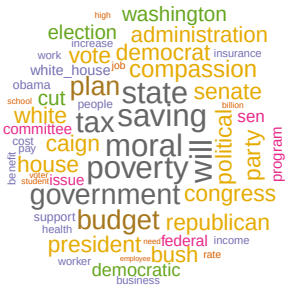


Figure F.1. Discourse Topic Contents from the *WSJ* (Cont.)

This figure plots the over-time frequencies of n-grams per each topic constructed according to the sLDA model. The size of each n-gram indicates its frequency. The sample period is from January 2000 to October 2019.

Confidence

Savings



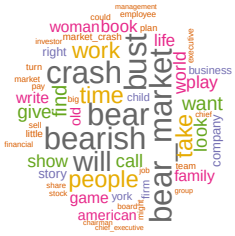
Conspicuous Consumption

Technology Replacing Jobs

Real Estate Crashes

Stock Market Bubbles

Stock Market Crashes



Wage and Labor Unions

Figure F.2. Time Series of Topic Weights from the *WSJ*

This figure plots the time series of monthly topic weights. The solid line is topic weight while the dashed line is excess market return; both have been demeaned for ease of visualization. The shades indicate NBER-dated recessions. The sample period is from January 2000 to October 2019.

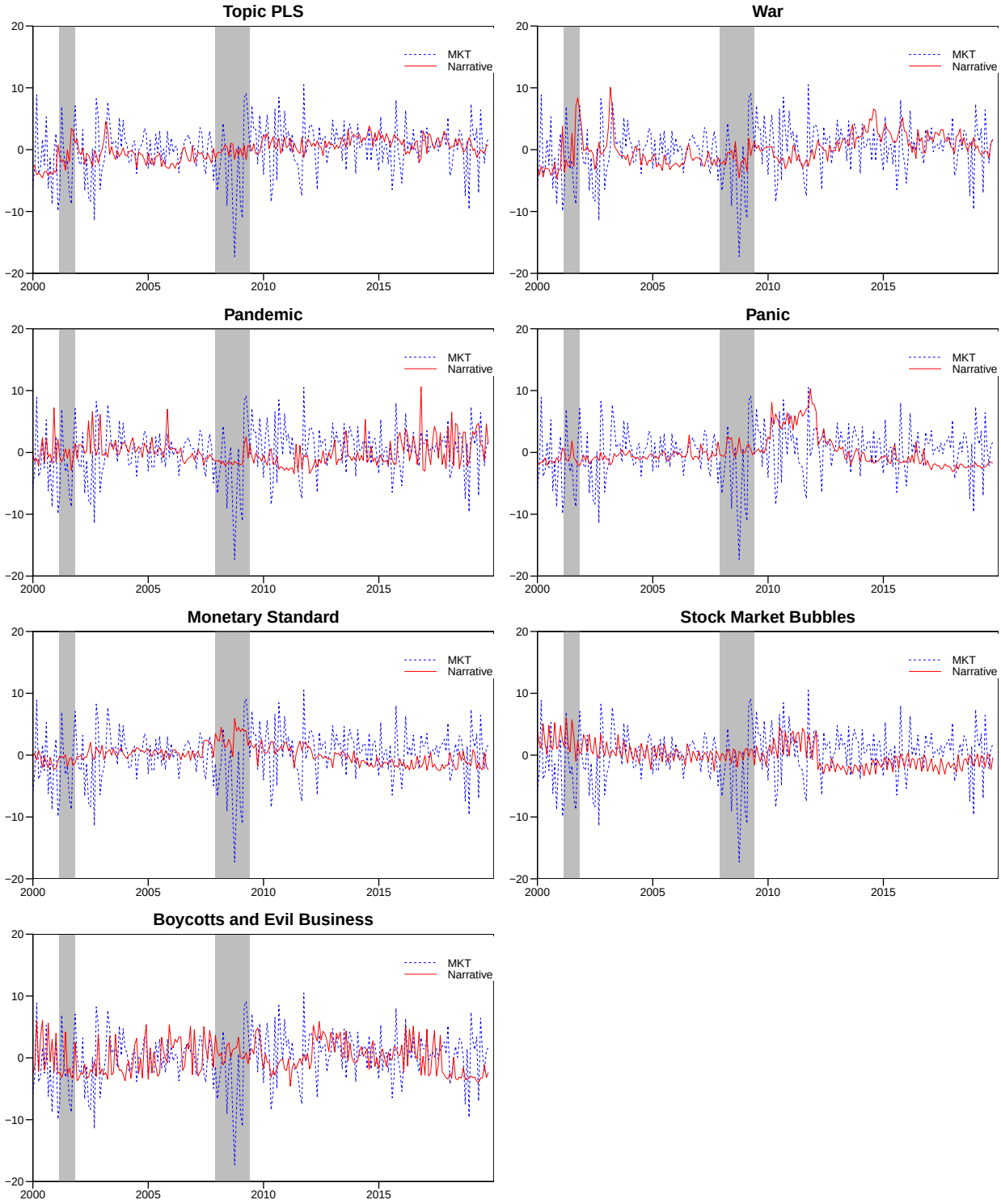


Figure F.2. Time Series of Topic Weights from the *WSJ* (Cont.)

This figure plots the time series of monthly topic weights. The solid line is topic weight while the dashed line is excess market return; both have been demeaned for ease of visualization. The shades indicate NBER-dated recessions. The sample period is from January 2000 to October 2019.

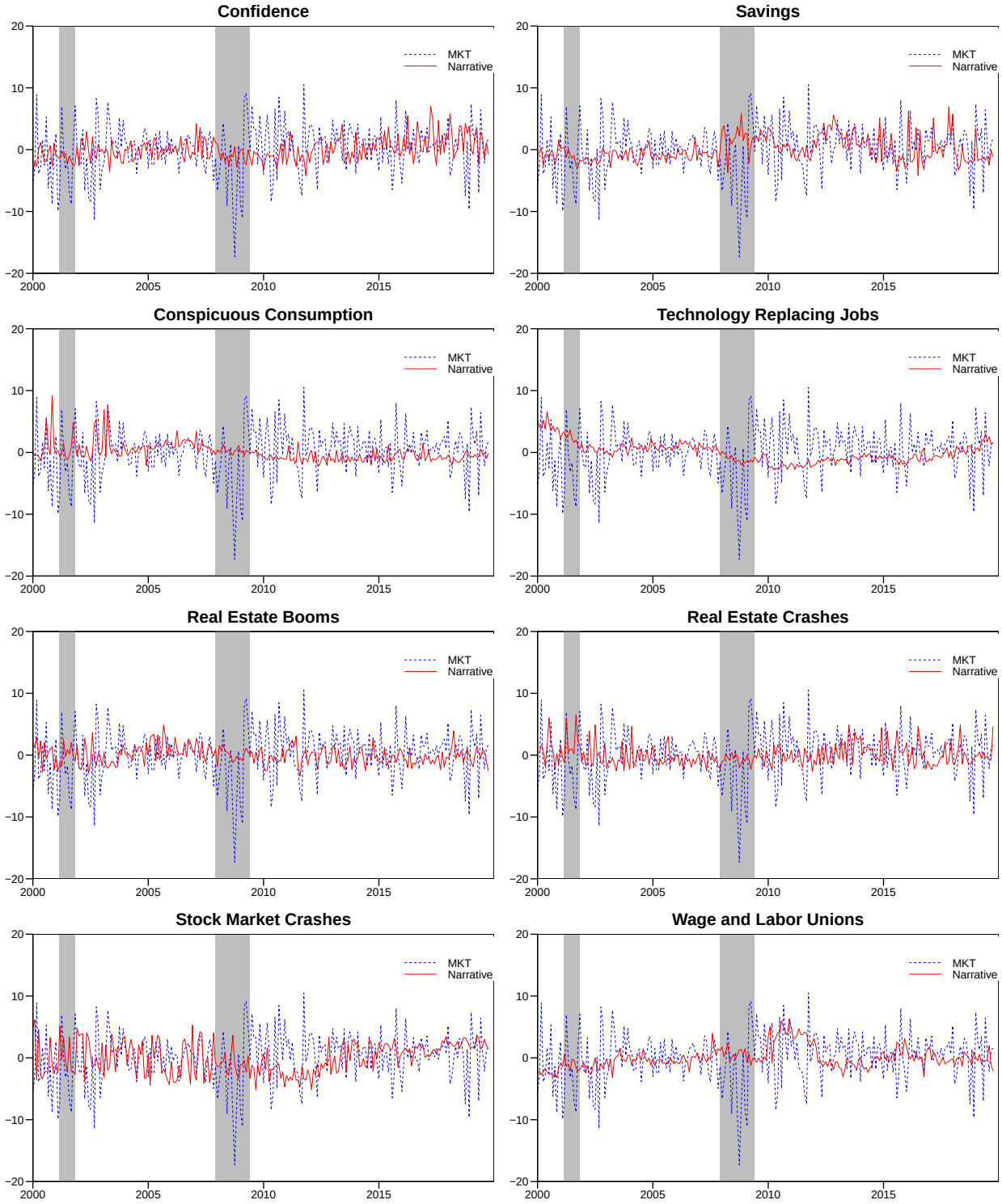


Table F.1
Summary Statistics of Topics from the *WSJ*

This table presents the summary statistics for the time series of 14 monthly topic weights constructed from the *WSJ* articles according to the sLDA model described in Section 3. All numbers (except sample size) are expressed as percentages. The sample period is from January 2000 to October 2019.

	N	Mean	SD	Q1	Median	Q3	AC(1)	PLS Weights	Corr PLS
War	238	8.87	2.42	7.13	8.61	10.21	70.54	9.48	77.05
Pandemic	238	5.77	2.16	4.35	5.31	6.74	14.11	-4.32	-23.65
Panic	238	5.95	2.40	4.47	5.30	6.34	87.48	2.54	11.90
Confidence	238	6.39	1.88	5.04	6.08	7.58	15.07	-1.19	12.53
Saving	238	7.60	2.04	6.23	7.11	8.68	39.02	2.90	24.23
Consumption	238	5.14	1.52	4.14	4.87	5.71	30.22	-0.11	-27.67
Money	238	6.59	1.52	5.46	6.48	7.34	67.08	-1.14	-17.30
Tech	238	6.41	1.66	5.21	6.19	7.32	87.43	-3.28	-63.22
Real estate boom	238	6.12	1.47	5.08	6.03	7.01	16.16	-2.35	-31.04
Real estate crash	238	5.40	1.74	4.22	4.96	6.23	2.45	0.56	4.64
Stock bubble	238	5.72	1.91	4.33	5.59	6.67	43.85	-6.47	-43.81
Stock crash	238	8.01	2.54	5.85	8.05	10.07	31.80	1.41	23.18
Boycott	238	8.65	2.65	6.32	8.27	10.48	24.26	-4.90	-30.75
Wage	238	6.91	1.75	5.90	6.63	7.44	72.83	3.32	28.17
PLS	238	12.17	17.43	2.07	13.76	24.30	66.35		

Table F.2
Predicting One-Month Market Returns with *WSJ* Topics

This table presents the results of the following predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1},$$

where R^e_{t+1} is the excess market return over the next month, x_t is one of the topics or the PLS indexes constructed from the *WSJ* articles, and β, the coefficient of interest, measures the strength of predictability. "Shiller PLS" uses only the topics from Shiller (2019), excluding *War* and Pandemic. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. The out-of-sample R^2 (R^2_{OS}) is computed using an expanding window with the initial estimation window of 60 months and is evaluated based on the Clark and West (2007) MSFE-adjusted statistic. The sample is from January 2000 to October 2019. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

	β (%)	t-stat	R^2 (%)	R^2_{OS} (%)
War	8.31 ***	(2.75)	2.30	1.53 ***
Pandemic	-4.24	(-1.24)	0.29	-2.63
Panic	2.24	(0.68)	-0.23	-2.64
Confidence	-1.34	(-0.48)	-0.35	-2.42
Saving	3.02	(0.86)	-0.06	-1.98
Consumption	-0.15	(-0.05)	-0.42	-3.31
Money	-1.60	(-0.38)	-0.32	-4.42
Tech	-4.19	(-1.23)	0.27	-1.02
Real estate boom	-3.40	(-1.07)	0.03	-1.36
Real estate crash	0.68	(0.27)	-0.41	-0.66
Stock bubble	-7.17 **	(-2.14)	1.61	0.89 *
Stock crash	1.17	(0.35)	-0.37	-0.67
Boycott	-3.93	(-1.45)	0.18	-0.85
Wage	4.03	(1.15)	0.22	-1.46
PLS	12.17 ***	(4.52)	5.42	-3.58
Shiller PLS	9.86 ***	(3.10)	3.41	-7.27

Table F.3
Predicting Long-Horizon Market Returns with *WSJ* Topics

This table presents the results of the following predictive regression:

$$R^e_{t+1\to t+h} = \alpha + \beta x_t + \epsilon_{t+1\to t+h},$$

where $R^e_{t+1\to t+h}$ is the excess market return over the next h months, x_t is either *War*, Stock Bubble or the PLS index constructed from the *WSJ* articles, and β, the coefficient of interest, measures the strength of predictability. Returns are expressed as annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors using the corresponding h lags. The sample is from January 2000 to October 2019. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

	War (%)	t-stat	R^2 (%)	Stock Bubble (%)	t-stat	R^2 (%)	PLS (%)	t-stat	R^2 (%)	N
$h = 1$	8.31 ***	(2.75)	2.30	-7.17 **	(-2.14)	1.61	12.17 ***	(4.52)	5.42	238
$h = 3$	6.94 ***	(3.26)	4.93	-5.63 ***	(-2.77)	3.10	8.49 ***	(4.37)	7.59	238
$h = 6$	4.45 **	(2.32)	3.43	-4.71 **	(-2.22)	3.89	6.20 ***	(3.63)	7.03	238
$h = 12$	3.68	(1.49)	4.25	-4.53 **	(-2.19)	6.64	5.68 ***	(2.76)	10.70	238
$h = 24$	3.70 **	(2.01)	7.49	-3.02	(-1.35)	4.83	6.01 ***	(3.27)	20.50	230
$h = 36$	3.77 **	(2.24)	11.80	-1.85	(-0.89)	2.51	6.45 ***	(4.82)	35.47	218

References

Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *Journal of Finance*, *61*(4), 1645–1680.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, *131*(4), 1593–1636.

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

Barro, R. J. (2009). Rare Disasters, Asset Prices, and Welfare Costs. *American Economic Review*, *99*(1), 243–64.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Boudoukh, J., Israel, R., & Richardson, M. (2022). Biases in long-horizon predictive regressions. *Journal of Financial Economics*, *145*(3), 937–969.

Brogaard, J., & Detzel, A. (2015). The asset-pricing implications of government economic policy uncertainty. *Management Science*, *61*(1), 3–18.

Brown, N. C., Crowley, R. M., & Elliott, W. B. (2020). What are You Saying? Using Topic to Detect Financial Misreporting. *Journal of Accounting Research*, *58*(1), 237–291.

Bybee, L., Kelly, B. T., Manela, A., & Xiu, D. (2023). Business News and Business Cycles. *Working Paper*.

Campbell, J. Y., & Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, *107*(2), 205–251.

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

Choudhury, P., Wang, D., Carlson, N. A., & Khanna, T. (2019). Machine Learning Approaches to Facial and Text Analysis: Discovering CEO Oral Communication Styles. *Strategic Management Journal*, *40*(11), 1705–1732.

Clark, T. E., & West, K. D. (2007). Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. *Journal of Econometrics*, *138*(1), 291–311.

Cooper, I., & Priestley, R. (2009). Time-Varying Risk Premiums and the Output Gap. *Review of Financial Studies*, *22*(7), 2801–2833.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies*, *22*(5), 1915–1953.

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The Evolution of 10-K Textual Disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, *64*(2-3), 221–245.

Engle, R. F., & Bollerslev, T. (1986). Modelling the Persistence of Conditional Variances. *Econometric Reviews*, *5*(1), 1–50.

Ferson, W. E., Sarkissian, S., & Simin, T. T. (2003). Spurious regressions in financial economics? *The Journal of Finance*, *58*(4), 1393–1413.

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

Garcia, D. (2013). Sentiment During Recessions. *Journal of Finance*, *68*(3), 1267–1300.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as Data. *Journal of Economic Literature*, *57*(3), 535–74.

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

Griffiths, T. L., & Steyvers, M. (2004). Finding Scientific Topics. *Proceedings of the National academy of Sciences*, *101*(suppl 1), 5228–5235.

Huang, D., Li, J., & Wang, L. (2020). Are Disagreements Agreeable? Evidence from Information Aggregation. *Journal of Financial Economics*, *141*(1), 83–101.

Huang, D., Jiang, F., Tu, J., & Zhou, G. (2015). Investor Sentiment Aligned: A powerful Predictor of Stock Returns. *Review of Financial Studies*, *28*(3), 791–837.

Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager Sentiment and Stock Returns. *Journal of Financial Economics*, *132*(1), 126–149.

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

Loughran, T., & McDonald, B. (2011). When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, *66*(1), 35–65.

Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, *54*(4), 1187–1230.

Loughran, T., & McDonald, B. (2020). Textual Analysis in Finance. *Annual Review of Financial Economics*, *12*(1), 357–375.

Lu, B., Ott, M., Cardie, C., & Tsou, B. K. (2011). Multi-Aspect Sentiment Analysis with Topic Models. *2011 IEEE 11th International Conference on Data Mining Workshops*, 81–88.

Lundblad, C. (2007). The Risk Return Tradeoff in the Long Run: 1836–2003. *Journal of Financial Economics*, *85*(1), 123–150.

Manela, A., & Moreira, A. (2017). News Implied Volatility and Disaster Concerns. *Journal of Financial Economics*, *123*(1), 137–162.

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, *41*(5), 867–887.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, *59*(2), 347–370.

Newey, W., & West, K. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, *55*(3), 703–708.

Pástor, L., & Veronesi, P. (2013). Political Uncertainty and Risk Premia. *Journal of Financial Economics*, *110*(3), 520–545.

Rapach, D. E., Ringgenberg, M. C., & Zhou, G. (2016). Short Interest and Aggregate Stock Returns. *Journal of Financial Economics*, *121*(1), 46–65.

Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. *The Review of Financial Studies*, *23*(2), 821–862.

Shiller, R. J. (2019). *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*. Princeton University Press.

Stambaugh, R. F. (1999). Predictive Regressions. *Journal of Financial Economics*, *54*(3), 375–421.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The Short of It: Investor Sentiment and Anomalies. *Journal of Financial Economics*, *104*(2), 288–302.

Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In *Handbook of latent semantic analysis* (pp. 439–460). Psychology Press.

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

Zakoian, J.-M. (1994). Threshold Heteroskedastic Models. *Journal of Economic Dynamics and control*, *18*(5), 931–955.