THREE ESSAYS ON THE TIME SERIES OF RETURNS

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

DAT Y MAI

Dr. Kuntara Pukthuanthong, Dissertation Supervisor

JULY 2023

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

THREE ESSAYS ON THE TIME SERIES OF RETURNS

presented by Dat Y Mai, a candidate for the degree of doctor of philosophy, and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Kuntara Pukthuanthong

Professor Michael O'Doherty

Professor Fred Bereskin

Professor Adam Yore

Professor Zack (J. Isaac) Miller

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor Dr. Kuntara Pukthuanthong for her valuable time, efforts, guidance, and support during my research process and my Ph.D. journey. I would also like to sincerely thank the other members of my dissertation committee: Dr. Michael O'Doherty, Dr. Fred Bereskin, Dr. Adam Yore, and Dr. Zack Miller for their endless teaching and continuous guidance on my research and career preparation. I am also grateful for the helpful comments and suggestions from the other graduate students and faculty in the department of Finance. Finally, I would like to thank the Trulaske College of Business for generously funding my conference travels and other research-related expenses during my Ph.D. study. Without all of you, I would not have been able to complete this dissertation and make it through my Ph.D. degree!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THREE ESSAYS ON THE TIME SERIES OF RETURNS

Dat Mai

Dr. Kuntara Pukthuanthong, Dissertation Supervisor

ABSTRACT

This dissertation consists of three essays on the time series of asset returns. The first essay in Chapter 1—**Time-Varying Drivers of Stock Prices**—provides novel evidence of the time-varying roles of subjective expectations in explaining stock price variations across the market and 30 industry portfolios monthly from 1976 to 2020. Cash flow expectations matter more under financial uncertainty and recessions, especially among the hardest-hit industries such as Telecommunications during the Dot-com Bubble, Financials during the Great Recession, and Healthcare during the Covid-19 pandemic. Conversely, discount rates explain more price variations during expansionary periods. Finally, inflation expectations, while accounting for 60% of price fluctuations in the high inflationary environment before 2000, play a negligible role thereafter.

In the second essay in Chapter 2—**Investor Sentiment and Asset Returns: Actions Speak Louder than Words**—I analyze daily predictability of investor sentiment across four major asset classes and compares sentiment measures based on news and social media with those based on trade information. For the majority of assets, trade-based sentiment measures outperform their text-based equivalents for both in-sample and out-of-sample predictions. This outperformance is particularly noticeable in long-term forecasts. However, real-time mean-variance investors can only achieve economic gains using Bitcoin trade sentiment, suggesting the challenge of transforming sentiment into daily profitable trading strategies.

In the last essay in Chapter 3—**War Discourse and Disaster Premia: 160 Years of Evidence from Stock and Bond Markets**—using a semi-supervised topic model on 7,000,000 *New York Times* articles spanning 160 years, I test whether topics of media discourse predict future stock and bond market returns to test rational and behavioral hypotheses about market valuation of disaster risk. Focusing on media discourse addresses the challenge of sample size even when major disasters are rare. Our methodology avoids look-ahead bias and addresses semantic shifts. War discourse positively predicts market returns, with an out-of-sample R^2 of 1.35%, and negatively predicts returns on short-term government and investment-grade corporate bonds. The predictive power of war discourse increases in more recent time periods.

Chapter 1

Time-Varying Drivers of Stock Prices

1.1 Introduction

"What explains stock price variations?" has been one of the central questions of financial research over at least the past four decades since Shiller (1981) and is of interests to both academic researchers and practitioners. Stock prices fluctuate in response to changes in either future cash flow expectations or future discount rates. Financial economists are thus interested in studying which of the two components is the main driver of stock variations (Campbell and Shiller 1988; Cochrane 2008). This paper contributes to the literature by providing novel evidence of the time-varying roles played by cash flows and discount rates in driving stock prices.

In this paper, I follow the recent literature in using *real-time subjective* expectations made by market participants to explain stock price fluctuations (Bordalo et al. 2020a; De La O and Myers 2021, 2022). More specifically, I use the earnings forecasts made by equity analysts to measure subjective cash flow expectations and extract the implied discount rates from subjective cash flow expectations and market prices. I then investigate if there are structural breaks in the comovements between these subjective expectations and stock prices. Unlike the traditional approach that relies on ex-post returns and cash flows, using subjective expectations made by market participants gives us a better understanding of the market dynamics in a real-time manner (e.g., if the stock market goes up today, is it because investors have a higher cash flow expectation or because they lower their discount rate?).

Ex-ante, it is unclear whether the roles of subjective expectations display any time-varying pattern. To tackle this empirical question, I employ a new Bayesian panel break method developed by Smith and Timmermann (2021a). The merits of this Bayesian break method lie in its ability to detect structural breaks in the data without the need to specify the number and locations of break points in advance. Instead of relying on one time series, this panel break method uses a panel of data to increase the power in identifying the structural breaks which might occur over time. This method has been used in Smith and Timmermann (2021a) to identify the breaks in the time series predictability of future returns using price ratios and in Smith and Timmermann (2021b) to uncover the structural breaks in the risk premia on common asset pricing factors.

In this paper, I apply the Bayesian panel break method to a panel of the aggregate market and 30 industry portfolios. I construct a panel of price-earnings ratios and subjective cash flow expectations at the monthly interval from 1976 to 2020. I then extract the implied subjective discount rates as the difference between subjective cash flow expectations and market prices.[1] Both of my cash flow and discount rate expectations are available monthly at the market and industry level. My sample is thus more granular than recent studies which use analyst forecasts to study price variations. For example, De La O and Myers (2021, 2022) use quarterly data on the aggregate market level, Bordalo et al. (2020a) only focus on the S&P 500 index, and Landier and Thesmar (2020) examine only three months after the Covid-19 outbreak. The use of monthly data enables the panel break model to identify the structural breaks in the time series with almost no delays (Smith and Timmermann 2021a).

Applying the Bayesian break method to subjective cash flow expectations, I identify six structural breaks related to changes in the real economy. The first break occurs in July 1986, one year before the 1987 Black Monday event. The second and third breaks are related to the Tech Bubble while the fourth and fifth ones center

[1] Recent studies such as Bordalo et al. (2020a) and De La O and Myers (2021) use as a measure of subjective discount rates the CFOs' expectations of returns on the S&P 500 over the next 12 months in the surveys conducted by Professors John Graham and Campbell Harvey. This measure is only available quarterly on the aggregate market level and is too stale to explain stock price variations.

on the Great Recession. The final break takes place right before the outbreak of the Covid-19 pandemic in the U.S.

The most striking result from the model is that the proportion of price variations explained by cash flow expectations rises sharply during the regimes associated with recessions. From 1976 to 2020, subjective cash flow (discount rate) expectations unconditionally explain 34% (66%) of price variations on the aggregate market. However, during the Tech Bubble, cash flow expectations explain 41% (75%) of price variations based on the Bayesian (OLS) estimate.[2] Such figures during the Great Recession are even higher at 76% (91%). The biggest jumps in the role of cash flow expectations are observed among the industries most closely associated with the recessions, such as Telecommunications and Business Equipment during the Tech Bubble, Financials during the Great Recession, and Healthcare during Covid-19. Outsides of these regimes, cash flow expectations explain just around 20% of price fluctuations. For the post 2000 sample as a whole, earnings growth expectations account for 80% of price-earnings variations because this 20-year period is strongly driven by the recession-related regimes.

During the 2000-2020 sample, subjective discount rates unconditionally explain 20% of price variations. However, outside of the three regimes related to the Tech Bubble, Great Recessions, and Covid-19 pandemic, subjective discount rates are the main driver of market movements. Hence, the role of discount rates should not be dismissed as implied in recent studies such as Bordalo et al. (2020a) and De La O and Myers (2021) who use the CFOs' expectations of returns on the S&P 500 as a measure of subjective discount rates. CFOs' expectations of returns on the S&P 500 are too stale to be correlated with actual market prices and their unconditional results mask substantial time-varying market dynamics.

[2] The OLS estimates are computed within each regimes identified by the Bayesian break method as a robustness check. As a shrinkage estimator, the Bayesian method produces smaller estimates than OLS, especially during short regimes.

As I empirically document that cash flow expectations become more important during regimes related to recessions, I hypothesize that cash flow expectations matter more under financial uncertainty. To formally test this hypothesis, I specify the roles of subjective expectations as a linear function of empirical uncertainty indexes including the financial uncertainty index from Jurado et al. (2015), macro uncertainty index from Ludvigson et al. (2021), economic policy uncertainty from Baker et al. (2016), consumption surplus from Cochrane (2017), geopolitical risk from Caldara and Iacoviello (2022), and implied volatility (VIX). I find that cash flow expectations become more important under uncertainty about financial markets, but not under macro uncertainty. Since Ludvigson et al. (2021) find that macro uncertainty is generally an endogenous response to output fluctuations while financial uncertainty is a cause of such output shocks, my findings indicate that market participants pay more attention to fundamentals expectations only when faced with exogenous uncertainty shocks. Furthermore, when using the level of price-earnings to capture economic states, I document that cash flow expectations become more important during bad economic states featured by high price-earnings. These conditional results confirm the time-varying roles of cash flow expectations as uncovered by the Bayesian break method.

In the last empirical analysis of the paper, I investigate whether subjective inflation expectations play an important role in explaining price variations. Consistent with De La O and Myers (2022), I find that unconditionally, inflation expectations explain almost half of price-earnings variations over the past 40 years. The impact of inflation expectations on stock prices is strongest among the industries offering consumer goods and services such as Food, Beer, Household, Retails, and Meals. However, when I apply the Bayesian break method to the panel of inflation expectations and price-earnings ratios, I document that the impact of inflation expectations on stock prices is concentrated before 1998. From 1998 to 2020, inflation expectations account for at most 6% of price variations across all industries. This finding implies that inflation expectations only matter under high inflation periods, especially during the 1970s and 1980s.

This paper makes important contributions to the literature on the dynamics of stock prices. First, I document that subjective cash flow expectations matter more under periods of heightened exogenous financial uncertainty. Recent recessions such as the Tech Bubble, Great Recession, and Covid-19 pandemic are such typical periods. Second, I break down the sources of variations to the industry level and find that the role of cash flow expectations is concentrated among industries hit the hardest by a recession such as Telecommunications during the Dot-com Bubble, Financials during the Great Recession, and Healthcare during the Covid-19 pandemic. Third, I find that unlike the recent view that downplays subjective discount rates (Bordalo et al. 2020a; De La O and Myers 2021), discount rates are indeed the main source of price variations during expansionary periods when uncertainty is low. Hence, the role of discount rates should not be ignored. Finally, inflation expectations only matter when actual inflation is high.

In terms of implications, the empirical evidence in this paper informs future macro-finance theory development that incorporates time-varying roles of subjective expectations in driving stock prices. From a practical standpoint, my findings suggest that investors and especially asset managers should closely track earnings expectations during periods of high financial uncertainty because during these periods, cash flow expectations are the main driver of stock prices.

Related literature. There are currently three main approaches to estimating the cash flow and discount rate components and examining their roles in explaining price variations.[3] The first one as employed in Campbell and Shiller (1988) is to use an econometric model to estimate the rational expectations of future cash flows and

[3] To say that this literature is vast is an understatement. Thus, I do not attempt to survey the literature here.

discount rates. This method relies on the state variables selected in the model and on the assumption that market participants have the same information set as econometricians. When this method is applied at the aggregate market level, the conclusion is that discount rates explain most of price variations (Campbell and Shiller 1988; Campbell 1991). However, when the analysis is conducted on stocks and anomaly portfolios, cash flows play a dominant role (Vuolteenaho 2002; Cohen et al. 2002; Lochstoer and Tetlock 2020).

The second approach is to use price ratios to predict ex-post returns and cash flows as in Cochrane (2005, 2008), Lettau and Van Nieuwerburgh (2008), Chen et al. (2012), Kelly and Pruitt (2013), Golez (2014), Golez and Koudijs (2018), Gao and Martin (2021), and Smith and Timmermann (2021a).[4] The general conclusion from this strand is that ex-post returns are predictable while realized cash flows become largely unpredictable in the post-war period (Golez and Koudijs 2018). While being able to predict future returns and cash flows is important, it is equally important to learn what drives stock prices in the real time (e.g. whether stock prices rise today because investors are more optimistic about future earnings growth). This quest gives rise to the third approach.

In this new approach, researchers study how real-time subjective cash flow and return expectations made by market participants explain price variations.[5] Subjective expectations offer several advantages as they are independent of model specifications, measurable and available in the real-time, and used by real-world market participants. Recent studies such as De La O and Myers (2021, 2022) and Bordalo et al. (2020a) use analyst earnings forecasts for cash flow expectations and CFO's surveys for return expectations on the S&P 500 index. They find that subjective cash flow

[4]For a survey on returns and cash flows predictability using valuation ratios, see Koijen and Van Nieuwerburgh (2011).

[5]This specific research question falls within an emerging literature in finance and macroeconomics that studies subjective expectations of investors and professional forecasters (see, e.g., Greenwood and Shleifer 2014; Bordalo et al. 2019, 2020b; Kuchler and Zafar 2019; Giglio et al. 2021).

expectations account for almost all of price ratio variations on the aggregate market. Landier and Thesmar (2020) also use analyst earnings forecasts and back out the implicit discount rates implied by market prices and document that revisions in analyst earnings forecasts account for all fluctuations of stock prices during the first three months of the Covid-19 outbreak. Chen et al. (2013) use analyst forecasts but employ a different price-ratio decomposition and find that subjective cash flow expectations play an important role at both the firm and market level.

Due to its several advantages, this paper uses the third approach and contributes by examining a largely unexplored angle: whether the roles of cash flows and discount rates change over time and whether there is any mechanism to explain such changes.

1.2 Method

In this section, I discuss first the present value framework, then the measures of subjective expectations, and finally the Bayesian break model used to estimate the dynamic relation between price ratios and subjective expectations.

1.2.1 Present-Value Relation

This paper employs the log-linearized present value relation introduced by Campbell and Shiller (1988). Consider the ex-post one-year return identity

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{\left(\frac{P_{t+1}}{D_{t+1}} + 1\right)\frac{D_{t+1}}{D_t}}{\frac{P_t}{D_t}}, \tag{1.1}$$

where P_t is current stock price and D_t is realized rolling 12-month dividend. Log-linearizing equation (1.1) around a long-term average of P/D gives

$$pd_t = \kappa + \Delta d_{t+1} - r_{t+1} + \rho \times pd_{t+1}, \tag{1.2}$$

where $pd_t \equiv \ln(P_t/D_t)$ is log price-dividend ratio, $\Delta d_{t+1} \equiv \ln(D_{t+1}/D_t)$ is log dividend growth, $r_{t+1} \equiv \ln(R_{t+1})$ is log return, and

$$\rho \equiv \frac{P/D}{1+P/D} \quad \text{and}$$
$$\kappa \equiv \ln(1+P/D) - \rho \times \ln(P/D)$$

are two constants. Following Cochrane (2005, 2008), I set the long-term average P/D to 25, which means $\rho = 0.96$.

Imposing the no-bubble condition, $\lim_{i\to\infty} \rho^j(pd_{t+j}) = 0$, equation (1.2) can be iterated forward to obtain

$$pd_t = \frac{1}{1-\rho}\kappa + \sum_{j=1}^{\infty} \rho^{j-1}\Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, \tag{1.3}$$

As equation (1.3) is based on the return identity (1.1), it should hold both ex-post and ex-ante. Taking conditional *subjective* expectations (e.g., by analysts) of both sides of equation (1.3) yields

$$pd_t = \frac{1}{1-\rho}\kappa + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1}\mathbb{E}_t^*\Delta d_{t+j}}_{\text{Cash flow } (CF_t)} - \underbrace{\sum_{j=1}^{\infty} \rho^{j-1}\mathbb{E}_t^* r_{t+j}}_{\text{Discount rate } (DR_t)}, \tag{1.4}$$

where $\mathbb{E}_t^*$ denotes the conditional *subjective* expectation. As noted by De La O and Myers (2021), given any set of dividend and return expectations, equation (1.4) holds as long as the no-bubble constraint is satisfied (i.e., $\lim_{j\to\infty} \rho^j \mathbb{E}_t^* pd_{t+j} = 0$ or investors do not expect the price-dividend ratio to go to infinity). Equation (1.4) says that the price-dividend ratio should go up if investors expect higher future dividend growths or lower future discount rates. Following the literature, I refer to the first term as the *cash flow (CF)* component and the second term as the *discount rate (DR)* component of price ratios.

Following recent papers in this field (Bordalo et al. 2020a; De La O and Myers 2021, 2022), I use analyst forecasts to compute subjective CF expectations. As the data on analyst earnings forecasts as recorded by IBES go back to 1976 while the data on dividend forecasts are only available from 2003 and the number of firms paying dividends is relatively small, similar to De La O and Myers (2022), I use analyst earnings, instead of dividend, forecasts in this paper.

As shown in De La O and Myers (2021), we can easily translate equation (1.4) from dividend growth into earnings growth. Specifically, using the log payout ratio de_t, I can substitute the identity $pe_t = pd_t + de_t$ into equation (1.2) to obtain

$$pe_t = \kappa + \Delta e_{t+1} - r_{t+1} + (1-\rho) \times de_{t+1} + \rho \times pe_{t+1}. \tag{1.5}$$

De La O and Myers (2021) find that, since $1-\rho = 0.04$ is very close to zero, changes in future dividend payout ratio explain less than 1% of the variations of the price-earnings ratio. Hence, variations in the dividend-payout ratio can be ignored and I can iterate equation (1.5) forward and take the conditional subjective expectations to obtain

$$pe_t = \frac{\kappa + (1-\rho) \times de}{1-\rho} + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t^* \Delta e_{t+j}}_{\text{Cash flow } (CF_t)} - \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t^* r_{t+j}}_{\text{Discount rate } (DR_t)}, \tag{1.6}$$

where de is the long-term average payout ratio.[6] Similar to equation (1.4), equation (1.6) requires the no-bubble condition that the price-earnings ratio do not explode in the future (i.e., $\lim_{j\to\infty} \rho^j \mathbb{E}_t^* pe_{t+j} = 0$). De La O and Myers (2021) relax this constraint and find that only 3% of the price-earnings ratio can be attributed to price bubbles.

[6]Since analysts make earnings forecasts in levels, not in log growths, an approximation is made here. Specifically, assume that log earnings growth $\Delta e_{t+1} = \ln(E_{t+1}/E_t)$ follows a normal distribution with subjective expectation $\mathbb{E}_t^* \Delta e_{t+1}$ and subjective variance σ^{*2}. Given the property of a normal random variable (i.e., $\mathbb{E}e^x = e^{\mu_x + 0.5\sigma_x^2}$), I have

$$\Delta e_{t+1}^* \equiv \underbrace{\ln(\mathbb{E}_t^* E_{t+1}) - \ln(E_t)}_{\text{what I have}} = \underbrace{\mathbb{E}_t^* \Delta e_{t+1}}_{\text{what I want}} + 0.5\sigma^{*2}$$

In other words, I approximate the subjective expectation of earnings growth by the log growth of earnings forecasts. As noted by De La O and Myers (2021), as long as the conditional subjective variance of earnings growth is countercyclical, i.e., $\text{cov}\left(pd_t, \sigma^{*2}\right) < 0$, accounting for the variance term only strengthens the covariance between subjective cash flows and price ratios reported in this paper.

Given equation (1.6), we can decompose the total variations of the price-earnings ratio into a CF component and a DR component by taking unconditional covariances with pe and divide both sides of equation (1.6) by the unconditional variance of pe as in Campbell and Shiller (1988) and Campbell (1991):

$$1 \approx \frac{\text{cov}\left(\sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t^* \Delta e_{t+j}, pe_t\right)}{\text{var}(\text{pe}_\text{t})} + \frac{\text{cov}\left(-\sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t^* r_{t+j}, pe_t\right)}{\text{var}(\text{pe}_\text{t})} \tag{1.7}$$

Equation (1.7) says that movements in the price-earnings ratio must come from changes in earnings growth expectations or changes in discount rates. This variance decomposition serves as the basis for a large literature on studying the behavior of stock prices dating back to Shiller (1981). Equation (1.7) thus serves as the backbone for the rest of the paper.

1.2.2 Measures of Subjective Expectations

In this paper, I use analyst earning forecasts to compute subjective earnings growth expectations. Since analysts do not make infinite earnings forecasts while the future cash flow component in equation (1.6) is an infinite sum, I follow De La O and Myers (2021, 2022) to assume that analyst earnings growth expectations follow a decay process:

$$\mathbb{E}_t^* \left[\Delta e_{t+1+j}\right] - \mu_e = \phi_e^j \left(\mathbb{E}_t^* \left[\Delta e_{t+1}\right] - \mu_e\right) + \epsilon_{t,j}^e, \tag{1.8}$$

Equation (1.6) can then be rewritten as follows

$$pe_t = k + \underbrace{\frac{1}{1-\rho\phi_e} \mathbb{E}_t^* \Delta e_{t+1}}_{\text{Cash flow } (CF_t)} - \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t^* r_{t+j}}_{\text{Discount rate } (DR_t)}, \tag{1.9}$$

where k is a constant consisting of the original constant term in equation (1.6) and the term involving μ_e in equation (1.8). As all of the analyses in this paper are about the covariance between price ratios and expectations, this constant term becomes irrelevant and can be disregarded. Under the assumption that earnings growth expectations follow a decay process, as shown in (1.9), the covariance between price-earning ratios and all future CF expectations becomes the covariance with just one-year earnings growth expectations scaled by the constant $\frac{1}{1-\rho\phi_e}$. This assumption is valid because De La O and Myers (2021) find that almost all price-earnings variations are driven by the short-term earnings growth expectations.

In this paper, I examine the time-varying comovements between price-earnings ratios and earnings growth expectations across the market and 30 industry portfolios. Thus, I separately estimate the earnings growth decay coefficient ϕ_e for each of the portfolios using the one- and two-year earnings growth expectations. I report these coefficients in Table C1 in Appendix 1.C. Accordingly, these coefficients range from -0.04 for the Games industry to 0.18 for the Construction industry. Combined with the assumption that ρ equals 0.96, these values imply the scale factor in equation (1.9) lies in the range of 0.96 for Games to 1.21 for Construction.

For the market portfolio, ϕ_e equals 0.09, close to the value of 0.06 reported in De La O and Myers (2021), and the scale factor is 1.09. In other words, the covariances between price-earnings ratios and one-year earnings growth expectations across all portfolios are scaled by a very small factor under the decay process assumption. In case the decay process assumption does not hold for any portfolio, the results can still be interpreted as the covariances between price-earnings ratios and one-year earnings growth expectations. Because one-year earnings growth expectations explain most of price variations De La O and Myers (2021), the main conclusions of the paper regarding the roles of CF expectations in driving stock prices do not alter in any material manner.

Because I take conditional subjective expectations by analysts of both sides of equation (1.3) and use analyst earnings forecasts to estimate the CF component, the remaining portion of market price ratios must be attributed to analyst subjective discount rates.[7] Since data on analyst discount rates are not available, I compute them indirectly as the difference between price-earnings and subjective CF expectations. In doing so, I ignore approximation constants and other unimportant components such as price bubbles and subjective payout ratios which have been shown to account for a negligible proportion of price variations in De La O and Myers (2021). Computing subjective DR as the difference between pe and subjective CF and calculating the covariance term between DR and pe is mathematically equivalent to subtract the CF covariance term from 1 in (1.7). Hence, in my empirical analyses, I only report the results for CF and the results for DR can be easily inferred.

My approach of directly estimating the CF component from analyst earnings forecasts and attributing the residual of price ratios to the DR component is similar in spirit to Campbell and Shiller (1988) and Campbell (1991) who use VAR to directly estimate rational DR expectations and assign the residual component to rational CF expectations. Landier and Thesmar (2020) also compute implied discount rates at the stock level using analyst forecasts. However, their method is different from my approach here. Landier and Thesmar (2020) rely on the Gordon growth present value model and compute the implied DR that gives market prices. In doing so, they require the DR to be constant across the discounting periods and use the historical average growth as a proxy for the long-term earnings growth. In contrast, I rely on the log-linearized relation (1.6) and the decay process assumption for the CF component to infer the DR component from market prices. My approach does not impose any restriction on the behavior of discount rates.

[7] In equation (1.4), I first start with the ex-post return relation and take conditional subjective expectations to decompose market price ratios into the subjective CF and DR components. In Appendix 1.A, I use an alternative approach to work directly with the subjective return relation. The intuition is the same under the two approaches: we can decompose the market price ratios into a subjective CF and a subjective DR component.

1.2.3 Panel Break Model

In this section, I describe how I estimate the covariances of price-earnings ratios with subjective cash flow and discount rate expectations defined above. To begin with, each component in the variance decomposition in equation (1.7) is equivalent to the coefficient β in the following univariate regression:

$$y_t = \alpha + \beta \times pe_t + e_t, \tag{1.10}$$

where y_t is either CF_t or DR_t.[8]

The previous literature on price variations assumes that β is constant over time. In this paper, I investigate whether the roles of cash flows and discount rates are time-varying. To achieve this goal, I employ the Bayesian break method in Smith and Timmermann (2021a,b), which uses a panel of data to increase the power of identifying the structural breaks in the data. The merit of this Bayesian break method is that we do not need to specify the number of break points or break locations in advance. Instead, we let the model determine the set of break points that best describes the data. Following their specification, I consider the following model:

$$y_{it} = \alpha_{ik} + \beta_{ik} \times pe_{it} + e_{it}, \quad e_{it} \sim \mathrm{N}(0, \sigma_{ik}^2), \quad t = \tau_{k-1} + 1, \ldots, \tau_k, \tag{1.11}$$

where $\tau_k \in (\tau_1, \tau_2, \ldots, \tau_K)$—a set of K unknown break points which gives rise to $K+1$ separate regimes identified in the panel of data, and α_{ik}, β_{ik}, and σ_{ik}^2 are the intercept, slope, and residual variance for each portfolio i in each regime k, respectively.

Following Smith and Timmermann (2021a), I place an inverse-gamma prior on the residual variance σ_{ik}^2 with the prior shape and rate parameters set so that the prior

[8] An alternative approach to estimating the variance decomposition (1.7) as used in recent studies (Bordalo et al. 2020a; De La O and Myers 2022) is to run the regression of pe_t on either CF_t or DR_t

$$pe_t = a + b \times x_t + e_t$$

and examine the regression R^2. As long as the coefficient b is constrained to equal 1 as in the structural relation (1.6), the R^2 of this approach is equal to β in regression (1.10). In this paper, I use the regression (1.10) as I can adapt it to the Bayesian break model and directly compute the estimation error.

mean residual variance equals the sample variance of y_{it} pooled across all portfolios. The intercept α_{ik} and the coefficient of interest β_{ik} are assumed to follow a prior multivariate normal distribution with the prior mean of α_{ik} set to the sample average of y_{it} pooled across all portfolios and the prior mean of β_{ik} set to zero. Hence, this setup creates a ridge-type regression model with the covariances between price-earnings ratios and subjective expectations shrunk towards zero, which helps guard against spurious relations showing up in short regimes. The prior variances of α_{ik} and β_{ik}, conditional on the residual variance σ^2_{ik}, are specified to equal one on average. This variance choice together with the zero prior mean gives the model freedom in uncovering either a positive or negative value of β_{ik}.

The intuition behind the panel break model is that given an initial random set of break points, the estimation procedure will iteratively move each of the break locations as well as randomly add or remove the breaks until its convergence to a set of break points that gives the highest probability of observing the data given the model. The full parameterization of the model is presented in Appendix 1.B.

1.3 Data

I use three sets of data in this paper. First, I obtain monthly analyst earnings forecasts for each company from IBES. To remain consistent with recent studies (Bordalo et al. 2020a; De La O and Myers 2021), I use the median analyst forecasts from the unadjusted summary file. Every month, analysts usually make forecasts for the next one to three financial years and provide a long-term earnings growth estimate. As the forecasts are based on the financial year basis, following De La O and Myers (2021), I interpolate the forecasts to be exactly twelve and twenty four months ahead. For example, in June 2018, analysts make earnings forecasts over financial year one (E^*_{FY1}) and two (E^*_{FY2}) for firm A with the fiscal year end in December. In this case, the interpolated 12-month earnings forecast for firm A is $\frac{E^*_{FY1}+E^*_{FY2}}{2}$.

Second, I merge analyst earnings forecasts with CRSP prices and outstanding shares. As IBES usually surveys analysts around the middle of every month, to ensure all of my analyses are done in a real-time manner, I use the CRSP daily database to find the market data for each stocks on the closest day before the day of earning forecasts. To remain consistent with the asset pricing literature, I include only stocks traded on NYSE/NASDAQ/Amex and having a share code of 10 or 11.

Third, I use the Compustat quarterly data to compute the realized rolling twelve-month (four-quarter) earnings for each company. I use the earnings release date to match analyst earnings forecasts with only earnings already available to the market. For firms which do not have the release dates in Compustat, similar to Hou et al. (2020), I assume that realized earnings are available to the market four months after the previous fiscal quarter end. To avoid any look-ahead bias, I keep the rolling four-quarter earnings constant between two release dates. To sort stocks into portfolios, I use Compustat SIC code in the previous quarter end and supplement with CRSP SIC code when missing.

As I construct 30 industry portfolios and all of the ratios used in this paper are in logs, it will be problematic if the industry-aggregated realized or forecast earnings are negative or close to zero. To combat this issue, I follow Vuolteenaho (2002), Lochstoer and Tetlock (2020), and others to convert each real company into a pseudo-company by buying 90% of that company's market value and investing the remaining 10% market value in Treasury bill. I then construct the price-earnings ratios, cash flow components, and discount rate components for each of the market and 30 industry portfolios using these pseudo-companies. More details on sample construction are presented in Appendix 1.C.

Table 1.1 reports the summary statistics of log price-earnings ratios and subjective one-year earnings growth expectations for the market and 30 industry portfolios from

1976 to 2020. The market has a mean *pe* of 2.94 with a standard deviation of 0.47. Its *pe* is highly correlated with the first-order autocorrelation of 98%. The average expected one-year earnings growth of the market is about 28% with a standard deviation of 19%. The top five industries having the highest *pe* are Healthcare, Mines, Services, Business Services, and Games while the top five with the highest expected earnings growth are Steel, Healthcare, Mines, Business Equipment, and Services. Clearly, Healthcare, Services, and Business Equipment have high price-earning ratios because they also have high earnings growth expectations. As reported in the last row of Table 1.1, the average company has a *pe* of 2.94 and its earnings is expected to grow by 28% annually.

Figure 1.1 plots the time series of subjective cash flow expectations and price-earnings ratios on the market portfolio. Accordingly, cash flow expectations move more in line with price-earnings ratios in the second half than in the first half of the sample. Especially, cash flow expectations move almost one-on-one with price ratios during the Great Recession, but not so much outside of the recessionary periods. It is reasonable to expect that the comovements of subjective expectations and price ratios are time-varying.

1.4 Empirical Results

In this section, I report the empirical comovements between price ratios and subjective expectations first under the Bayesian break model and then under a conditional model. In the last subsection, I discuss the comovements between price ratios and subjective inflation expectations.

1.4.1 Comovements of Earnings Growth Expectations and Price-Earnings Ratios

The first empirical analysis examines the time-varying relation between price-earnings ratios and subjective cash flow expectations. To uncover this time-varying relation, I apply the Bayesian panel break method described in Subsection 1.2.3 to the market and 30 industry portfolios from 1976 to 2020.

Figure 1.2 reports the posterior break locations with their posterior probabilities. Accordingly, the model identifies six structural breaks: the first one is in July 1986, one year before the event of Black Monday; the second and third break dates are in November 2000 and October 2004, which are related to the Tech Bubble in the early 2000s; the fourth and firth beaks takes place in November 2008 and July 2010, which includes the second half of the Great Recession; and the final one is in January 2020, right before the outbreak of the Covid-19 pandemic in the U.S. By construction, these break points mark the end of the preceding regime. As can be seen, the breaks identified by the model clearly link to major economic events which may substantially change many economic relations, including the one between price ratios and subjective expectations.

Table 1.2 reports the β_{ik} estimates from the Bayesian model for each portfolio in each regime. Specifically, I report the posterior mean and the posterior standard deviation of β_{ik} under the columns labeled *Bayesian*. For comparison, I also compute the normal OLS estimates of these β_{ik}'s for each portfolio in the respective regime and report them under the OLS columns. The leftmost header reports the results for both the Bayesian and OLS models using the whole sample (i.e., imposing no structural breaks in the data).

There are two main differences between the Bayesian model and the OLS approach. First, as discussed in Subsection 1.2.3, under the normal prior with zero mean, the posterior means of these β_{ik} estimates are typically smaller than the OLS counterparts. The differences are larger in regimes with few months as the prior plays a bigger role when the data is limited. Second, the Bayesian model only accounts for heteroscedasticity in the residuals of each portfolio across regimes. To mitigate this issue, I specify a large prior variance on β_{ik}. In contrast, the OLS estimates account for both heteroscedasticity and autocorrelation in the residuals of each portfolio within each regime. Thus, in long regimes with abundant data, the effect of the prior distribution is small so the Bayesian model produces posterior variances smaller than the OLS standard errors. However, in short regimes with scare data, the prior plays a bigger role and the Bayesian model thus produces larger posterior variances than the OLS counterparts.

I create two plots (Figure 1.3 and Figure 1.4) to visualize the information reported in Table 1.2. In Figure 1.3, I plot β_{ik}'s for the market and two portfolios most impacted by each of the three recent recessions: Tech Bubble, Great Recession, and Covid-19 pandemic. The last plot in Figure 1.3 is for the whole market.

Over the whole 1976-2020 sample with no breaks, both the Bayesian and OLS methods show that 34% of monthly price-earnings variations at the market level is attributed to changes in subjective cash flow expectations. This result is consistent with the number of 44% from 1976 to 2015 in De La O and Myers (2021).[9] As implied analyst DR is the difference between pe and CF, the DR component accounts for 66% of price-earnings variations from 1976 to 2020.

However, the unconditional whole sample estimate masks substantial structural breaks identified by the Bayesian model. Specifically, subjective CF expectations

[9] From their Table II, $CF_1 = 42\%$, and from their Table VII, $\phi_e = 0.06$, combining with $\rho = 0.96$, the cash flow component of price-earnings variation is $CF = \frac{CF_1}{1-\phi_e \times \rho} = 44\%$. There are three main reasons why my number is not the exactly same as theirs. First, I use monthly estimates from 1976 to 2020 while De La O and Myers (2021) use quarterly data from 1976 to 2015 because their data on return expectations are available quarterly. Second, I use all stocks that have earnings forecasts while they only focus on the S&P 500 companies. Third, I make sure that realized earnings used in my variables are available to the market in the real time by using earnings release dates; it is not clear in De La O and Myers (2021) whether they use fiscal dates or release dates when constructing their variables.

explain under 30% of price ratio variations during the two regimes before 2000. The role of CF expectations jumps sharply during the recent three recession-related regimes, reaching 41% (75%), 76% (91%), and 29% (66%) under these regimes with the Bayesian (OLS) approach. As noted before, under short regimes, the Bayesian estimate is smaller than the OLS one as the former is a shrinkage estimator. Outsides of the three recessions after 2000, earnings expectations fall back to explaining under 30% of price-earnings fluctuations. Interestingly, the unconditional OLS estimate using the 2000-2020 sample is around 80%, similar to the results obtained during the 2008-2010 regime as this event is so influential. It becomes clear that over 2000-2020, CF dominates unconditionally but its explanatory power is concentrated in recession-related regimes. Outside of these periods, DR is still the main driver of stock prices.

The first row of Figure 1.3 plots the results for Telecommunications and Business Equipment, two industries heavily associated with the Tech Bubble. Unconditionally, CF expectations explain 29% and 56% of price variations for Telecommunications and Business Equipment, respectively. The role of CF expectations rises markedly during the Tech Bubble, accounting for 80% (over 90%) of price fluctuations for Telecommunications (Business Equipment) during this regime. CF expectations also play a big role during the Great Recession for these two industries.

The second row of Figure 1.3 plots the results for Financials and Autos. The Great Recession has a clear impact on the relation between CF expectations and price ratios for Financials. Specifically, during the Great Recession, 100% of price variations are attributed to CF expectations and this strong result drives up the unconditional estimates. For Autos, as expected the role of earnings growth expectations surge during the Tech Bubble and Great Recession regimes.

The third row reports the results for Healthcare and Carry, two industries strongly associated with the Covid-19 pandemic. For Healthcare, the role of CF expectations rises to 77% (104%) under the Bayesian (OLS) method, the highest values across the regimes. Regarding Carry, the proportion of price variations explained by CF expectations also rise during 2020, but not so much compared the 2008-2010 regime.

Figure 1.3 clearly illustrates that CF expectations become more important for industries most impacted by specific recessions. In Figure 1.4, I plot the top five industries with the highest proportions of price variations driven by CF news over the whole sample and during the three recessions. Unconditionally, CF expectations dominate in Coal, Autos, Books, Steel, and Paper. During the Tech Bubble, five industries where CF news are important are Autos, Services, Textiles, Electrical Equipment, and Business Equipment. During the Great Recession, the five industries are unsurprisingly Financials, as well as Paper, Mines, Oil, and Steel. Finally, during the Covid-19 pandemic, CF expectations matter the most for Healthcare, Steel, Autos, Clothes, and Textiles.

To formally test whether the role of CF expectations significantly rises during recession-related regimes, I examine the posterior mean and standard deviation of the changes in β_{ik} over the two adjacent regimes and report the results in Table 1.3. The first row in Panel A shows that the proportion of pe variations explained by CF expectations for the market significantly goes up by 21% in the 2000-2004 regime and by 70% in the 2008-2010 regime. The proportion also significantly plunges by 34% and 59% over the following regime after the hikes. Row two reports the average change in β_{ik}'s across all 30 portfolios. Besides the two increases in the Tech Bubble and Great Recession, the role of CF expectations also rises during the 1986-2000 regimes. The third row reports the average changes across both the market and 30 portfolios and show similar results to those using the industries alone. Panel B reports the results for each individual industry.

Overall, the empirical results in this section show that CF expectations plays a bigger role in explaining pe variations during regimes related to recessions. Outside of these, DR is still the main driver of market prices.

1.4.2 Subjective Expectations under Uncertainty

The previous section reports that the proportion of price variations explained by cash flow expectations is countercyclical. Since uncertainty and risk aversion usually rise during recessions and fall during expansions, I hypothesize that the role of CF expectations in explaining price variations is an increasing function of time-varying uncertainty.

My hypothesis on the role of CF expectations is consistent with the earnings growth reversal model in De La O and Myers (2021). According to their model, the proportion of pe explained by subjective CF expectations is positively related to the coefficient of relative risk aversion (RRA) in a power utility setup. The only difference between their model and my hypothesis is that RRA in their model is constant because the role of CF is assumed to be time-invariant. In contrast, I find that the empirical relation between pe and CF expectations is time-varying.

To formally test this hypothesis, I specify β_{it} as a function of some state variable z_{t-1} capturing uncertainty and risk aversion at time $t-1$, i.e., $\beta_{it} = b_0 + b_1 \times z_{t-1}$.[10] In other words, instead of letting β_{ik} shift every regime, I allow β_{it} to change every time period as dictated by z_{t-1}. I expect b_1 to be significantly positive. Regression (1.11) then becomes

$$CF_{it} = \alpha + b_0 \times pe_{it} + b_1 \times z_{t-1} \times pe_{it} + e_{it}. \tag{1.12}$$

I use a number of proxies to capture uncertainty and risk aversion. First, I use the macro and financial uncertainty indexes from Jurado et al. (2015) and Ludvigson et al. (2021). These uncertainty indexes are constructed from the forecasting errors in predicting a large number of macroeconomic and financial variables using a stochastic

[10]Using contemporaneous state variable z_t yields consistent results.

process. I also use the CBOE Volatility Index (VIX), a commonly used measure of financial uncertainty. Next, I include the economic policy uncertainty (EPU) index from Baker et al. (2016). This measure is a weighted count of the number of news articles containing words related to a combination of economy, policy, and uncertainty. As a proxy for risk aversion, I use the consumption surplus which is the difference between current consumption and a moving habit as developed in Campbell and Cochrane (1999). I compute the empirical consumption surplus following Cochrane (2017). I also include the geopolitical risk index developed by Caldara and Iacoviello (2022). As seen in Figure 1.1, market price-earnings first fall but then rise sharply during recessions when corporate earnings are low. Hence, I also use $pe_{i,t-1}$ to capture economic states.

I estimate equation (1.12) using OLS separately for the market as well as for the panel of industry portfolios. When only the market is tested, I use the Newey and West (1987) standard errors. When I estimate a panel of portfolios, I use the standard errors clustered by both portfolios and months. To streamline the results, I standardize all uncertainty variables to zero mean and unit variance.

Panel A of Table 1.4 reports the results for CF expectations over the whole sample from 1976 to 2020. For the market in the first row, b_1 is positive and significant at the 5% level for financial uncertainty, indicating that the role of CF expectations increases in financial uncertainty. In the second and third row, the results for the industries show that β_{it} rises with financial uncertainty, VIX, macro uncertainty, and geopolitical risk. For both the market and industry portfolios, the role of CF also increases (all significant at the 1% level) with the current level of pe: when the whole economy or a specific industry i is in a bad state represented by a high pe value, market participants pay more attention to CF.

Panel B reports the results for CF expectations over the short sample from 2000 to 2020. The results are consistent with Panel A regarding financial uncertainty, VIX, and pe. Specifically, the results for the market are stronger and b_1 is significant for VIX. A notable observation is that b_1 on the remaining non-financial uncertainty measures is insignificant for both the market and the industry portfolios.

That the result is significant for financial uncertainty and VIX, and insignificant for macro uncertainty and EPU over the past 20 years is worth discussion. Ludvigson et al. (2021) find that macroeconomic and policy uncertainty in recessions is often an endogenous response to business cycle fluctuations while uncertainty about financial markets is a likely source of output fluctuations. These findings combined with what is reported in Panel B of Table 1.4 imply that the proportion of price variations driven by cash flow expectations only rises under financial uncertainty that causes exogenous shocks to the real economy.

In sum, the empirical results are consistent with the hypothesis that CF expectations matter more under uncertainty or in bad economic states. The results further show that CF expectations become more important under only uncertainty about financial markets which causes output fluctuations. These results are intuitive in that investors pay more attention to fundamentals expectations (i.e., earnings growth) when they face exogenous financial uncertainty.

1.4.3 Comovements of Inflation Expectations and Price-Earnings Ratios

The cash flow component in the present value relation (1.6) can be further decomposed into real earnings growth expectations and inflation expectations:

$$pe_t = \text{constant} + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t^* \Delta e_{t+j}^r}_{\text{Real cash flow } (CF_t^r)} + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t^* \pi_{t+j}}_{\text{Inflation } (\Pi_t)} - \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \mathbb{E}_t^* r_{t+j}}_{\text{Discount rate } (DR_t)}, \tag{1.13}$$

where $\mathbb{E}_t^* \pi_{t+j}$ is subjective inflation expectations. Hence, it is interesting to investigate whether inflation expectations play a large role in driving price ratio fluctuations.

Indeed, De La O and Myers (2022) find that real cash flow expectations ($CF_t^r = CF_t - \Pi_t$) explains around 80% of price variations from 1976 to 2018. Real cash flow expectations play a larger role than (nominal) cash flow expectations because inflation expectations have a negative covariance with price ratios.

Since forecasters do not make infinite inflation forecasts, De La O and Myers (2022) specify a decay process for inflation expectations

$$\mathbb{E}_t^*[\pi_{t+1+j}] - \mu_\pi = \phi_\pi^j \left(\mathbb{E}_t^*[\pi_{t+1}] - \mu_\pi\right) + \epsilon_{t,j}^\pi. \tag{1.14}$$

To estimate ϕ_π, De La O and Myers (2022) use the next 12 months and average 10 years median inflation forecasts from the Survey of Professional Forecasters. The estimated value of ϕ_π is 0.96 with a standard error of 0.01. With the decay process (1.14), the inflation expectations component in equation (1.13) can be written as

$$\Pi_t = \frac{1}{1 - \rho\phi_\pi}\mathbb{E}_t^*\pi_{t+1}. \tag{1.15}$$

Since subjective inflation expectations explain a large portion of price ratio variations, in this section, I will investigate if there are any structural breaks in the relation between inflation expectations and price-earnings ratios. Because inflation expectations come at the quarterly frequency, I interpolate them to make a monthly series to utilize the monthly data on the market and industry portfolios.

Before applying the panel break method, I first plot the time series of inflation expectations Π_t against realized year-over-year inflation and price-earnings ratios on the market portfolio in Figure 1.5. As can be seen from Figure 1.5, Π_t is high throughout the 1970s and 1980s under the high inflationary environment as illustrated by realized year-over-year inflation. From 1976 until the mid of 1990s, inflation expectations have a strong negative correlation with price-earnings ratios. However, after that, Π_t seems very stale and does not have any comovement with the volatile pe. Thus, I expect there is at least one structural break around the middle of 1990s.

I then estimate the Bayesian break model with $-\Pi_t$ as the dependent variable.[11] Figure 1.6 plots the posterior probabilities of the structural break locations identified by the model. As expected, there is one break in December 1997. Additionally, the model uncovers four other preceding break points, namely October 1978, March 1982, August 1984, and May 1992.

Table 1.5 reports the estimates of β_{ik}'s for the market and 30 industries portfolios under each regime. To visualize the content of Table 1.5, I plot the estimates for the market and five industries having the largest whole sample estimates of β_{ik}, i.e., five industries on which inflation expectations have the largest effect. These industries include Food, Beer, Household, Retail, and Meals. It is unsurprising that inflation expectations have the strongest influence on the industries offering consumer goods and services.

The first plot in Figure 1.7 shows the results for the market. Unconditionally, the inflation expectation component explains 46% of price-earnings variations from 1976 to 2020. This result is consistent with that in De La O and Myers (2022). However, the unconditional estimate masks substantial changes among regimes. During the first regime from 1976-1978, inflation expectations have no impact on stock price variations. The second regime from 1978-1982 is the most interesting one as during this period price ratios increase in inflation expectations.[12] Inflation expectations explain around 20% of price-earnings fluctuations over the next two regimes 1982-1984 and 1984-1992. From 1998 until 2020, inflation expectations account for only 5% of price earnings variations. When I group all the regimes except the last one together (from 1976 to 1997), β_i is slightly above 60%. Combined with nominal cash flow expectations CF_t, real cash flow expectations always account for around 80% of price variations; yet inflation expectations dominate before 2000 while cash flow

[11] I use $-\Pi_t$ so that β_{ik}'s are positive.

[12] Note that as I use $-\Pi_t$, a negative value of β_{ik} means a positive covariance between Π_t and pe_{it}.

expectations take the lead thereafter. Thus, using a structural break model helps us uncover the dynamics masked by the whole sample estimates.

The plots for the five industries follow the same pattern of the market. Unconditionally, inflation expectations explain about 60% of price-earnings variations in these portfolios. However, these estimates are all driven by the period before 1998. Since then, inflation expectations account for less than 10% of price variations in these industries. Similar to the market, the price ratios of these industries also rise with higher inflation forecasts during the 1978-1982 regimes.

In Table 1.6, I test the significance of the changes in β_{ik}'s across the two adjacent regimes. As expected, there are large significant changes in β_{ik}'s in the 1978-1982 regime when inflation expectations go from having a negative effect on stock prices to having a positive one and in the 1982-1984 regime when the sign of β_{ik}'s reverts to its normality.

In sum, this section documents that inflation expectations have a strong negative impact on stock prices, especially among the industries offering consumer goods and services. However, the strong impact is mainly driven by the pre-1998 period under a high inflationary environment. Since then, inflation expectations play a minimal role in price ratio variations. Moreover, for a short period from 1978-1982, high inflation forecasts are indeed good news for stock prices.

1.5 Conclusion

This paper studies the empirical time-varying relation between price-earnings ratios and subjective expectations using the market and 30 industry portfolios at the monthly frequency from 1976 to 2020. I use analyst earnings forecasts to capture cash flow expectations and extract the implied discount rates from analyst earnings forecasts and market prices.

I apply new Bayesian panel break method to identifying the structural breaks in the relations between subjective expectations and price ratios. I find that breaks commonly occur around recessions and the proportion of price variations explained by cash flow expectations rises sharply during the recession-related regimes. The largest increases are observed in industries hit the hardest by recessions such as Telecommunications and Business Equipment during the Tech Bubble, Financials during the Great Recession, and Healthcare during the Covid-19 pandemic.

From 2000 to 2020, cash flow expectations unconditionally explains 80% of monthly price variations, yet the number falls to only 20% outside recessions. Over the same period, implied discount rates unconditionally account for 20% of price variations but their portion rises to over 50% during expansionary periods. Further tests show that cash flow expectations matter more for price variations under uncertainty about financial markets, but not under general endogenous macro uncertainty. The role of cash flow expectations also rises during bad economic states captured by the level of price-earnings.

Lastly, subjective inflation expectations unconditionally explain 46% of price variations from 1976 to 2020 with the strongest impacts on the consumer goods and services industries. However, the power of inflation expectations is concentrated before 1998 under the high inflationary environment. Since then, only about 5% of price variations is attributed to inflation expectations. Interestingly, for a short period from 1978 to 1982, high inflation expectations are good news for stock prices.

Figure 1.1. Time Series of Earnings Growth Expectations and Price-Earnings Ratios

This figure plots the time series of subjective earnings growth expectations (solid blue line) and price-earnings ratios (dotted red line) on the market portfolio. All variables have been demeaned for ease of visualization. Shades indicate NBER recessions. The sample is monthly from January 1976 to December 2020.

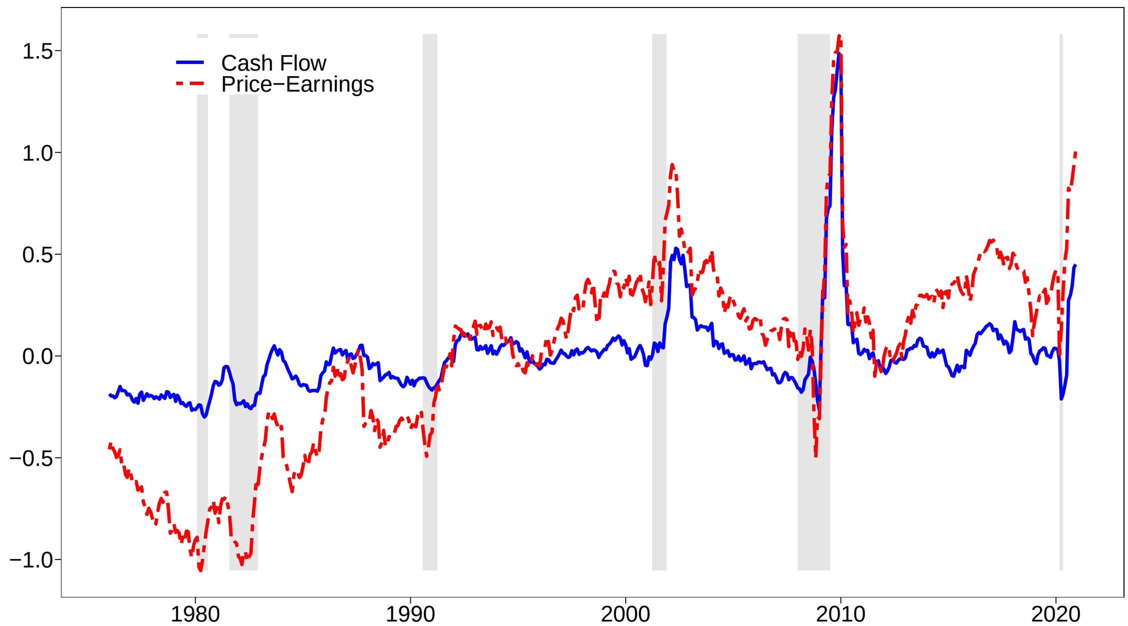


Figure 1.2. Posterior Break Locations in Comovements of Earnings Growth Expectations and Price-Earnings Ratios

This figure plots the posterior probability (y-axis) of the structural break locations (x-axis) identified by the following Bayesian panel break model:

$$CF_{it} = \alpha_{ik} + \beta_{ik} \times pe_{it} + e_{it}, \quad e_{it} \sim \mathrm{N}(0, \sigma_{ik}^2), \quad t = \tau_{k-1} + 1, \ldots, \tau_k,$$

where CF_{it} and pe_{it} are the cash flow component and log price-earnings ratio of portfolio i in month t. The K structural break locations are captured by the set $(\tau_1, \tau_2, \ldots, \tau_K)$ in the model. Shades indicate NBER recession dates. The sample is monthly from January 1976 to December 2020.

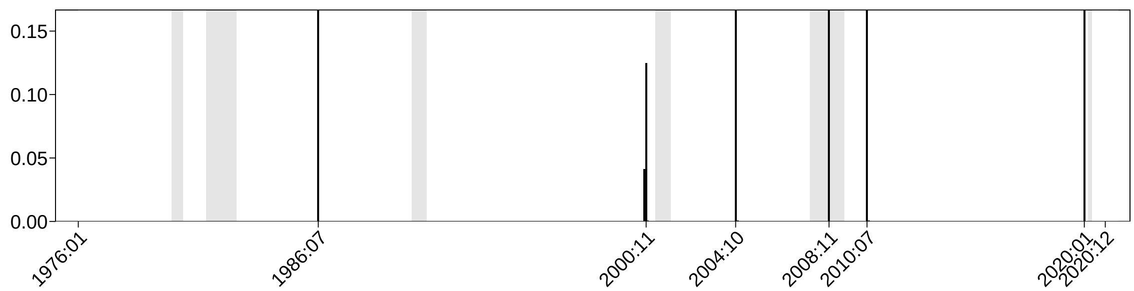


Figure 1.3. Comovements of Earnings Growth Expectations and Price-Earnings Ratios

This figure plots the proportion of price-earnings variations explained by subjective cash flow expectations (β_{ik} in regression (1.11)) under each regime for the market and six industry portfolios. The solid black line shows the Bayesian posterior mean of β_{ik} and the shade around it shows the 95% credible interval under each regime; and the dotted green line shows the corresponding OLS estimates. The dotted blue line shows the whole sample OLS estimate and the dotted red line shows the OLS estimate for the sample from 2000. The sample is monthly from January 1976 to December 2020.

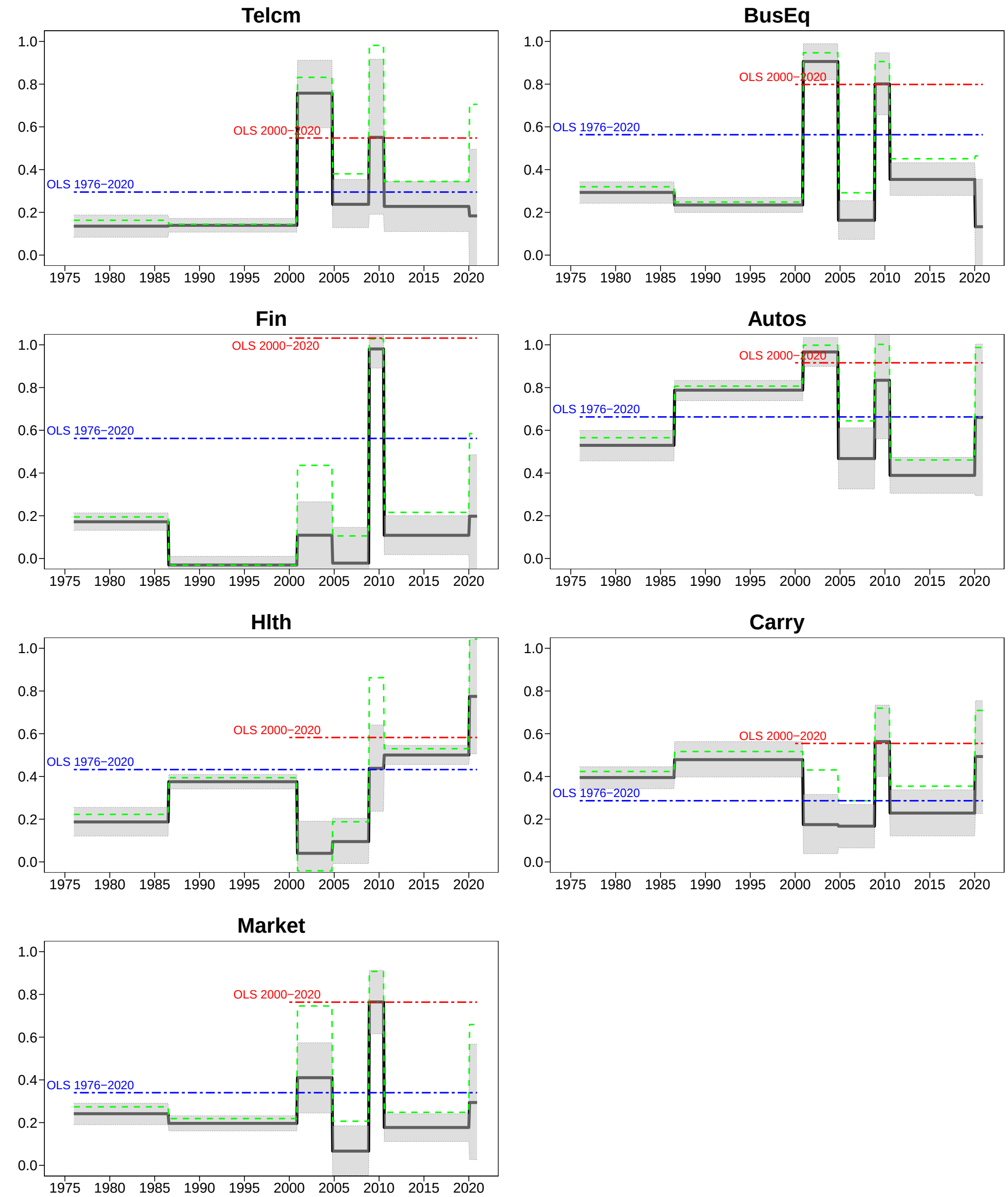


Figure 1.4. Top Five Industries with Highest Comovements of Earnings Growth Expectations and Price-Earnings Ratios during Crises

This figure plots the top five industries having the highest proportions of price-earnings variations explained by subjective cash flow expectations (β_{ik} in regression (1.11)) over the whole sample and the three regimes related to the three recessions (Tech Bubble, Great Recession, and Covid-19 pandemic). Black dots indicate posterior means and crossbars indicate 95% credible intervals. The sample is monthly from January 1976 to December 2020.

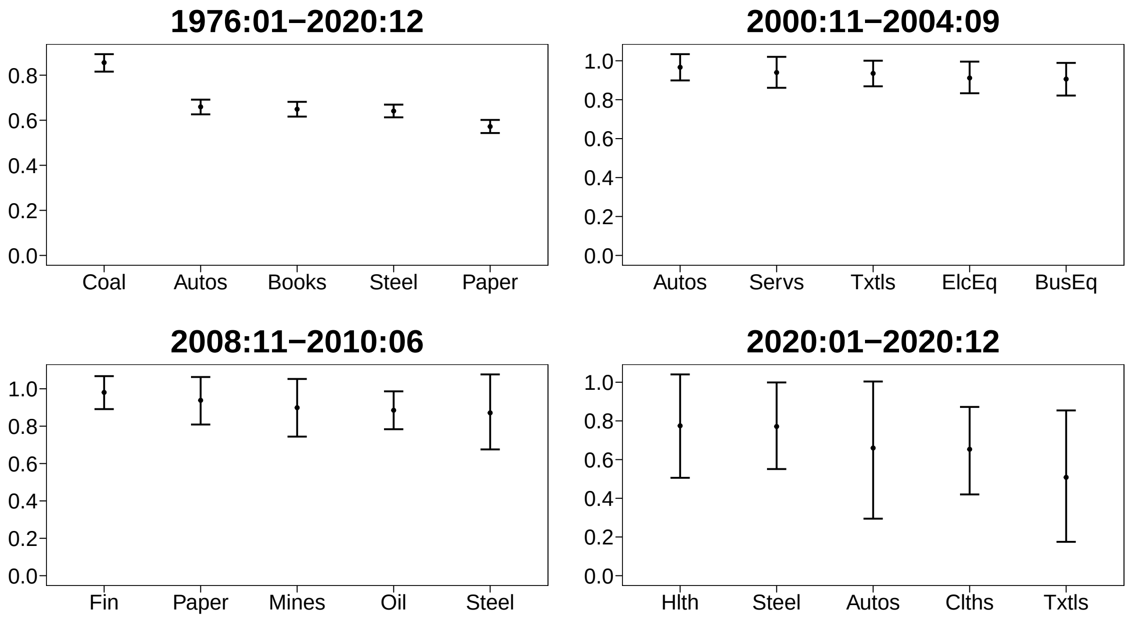


Figure 1.5. Time Series of Inflation Expectations, Realized Inflation, and Price-Earnings Ratio

This figure plots the time series of inflation expectations Π_t (solid blue line), realized year-over-year inflations (solid green line), and price-earnings ratios on the market portfolio (dotted red line). All variables have been demeaned for ease of visualization. Shades indicate NBER recessions. The sample is monthly from January 1976 to December 2020.

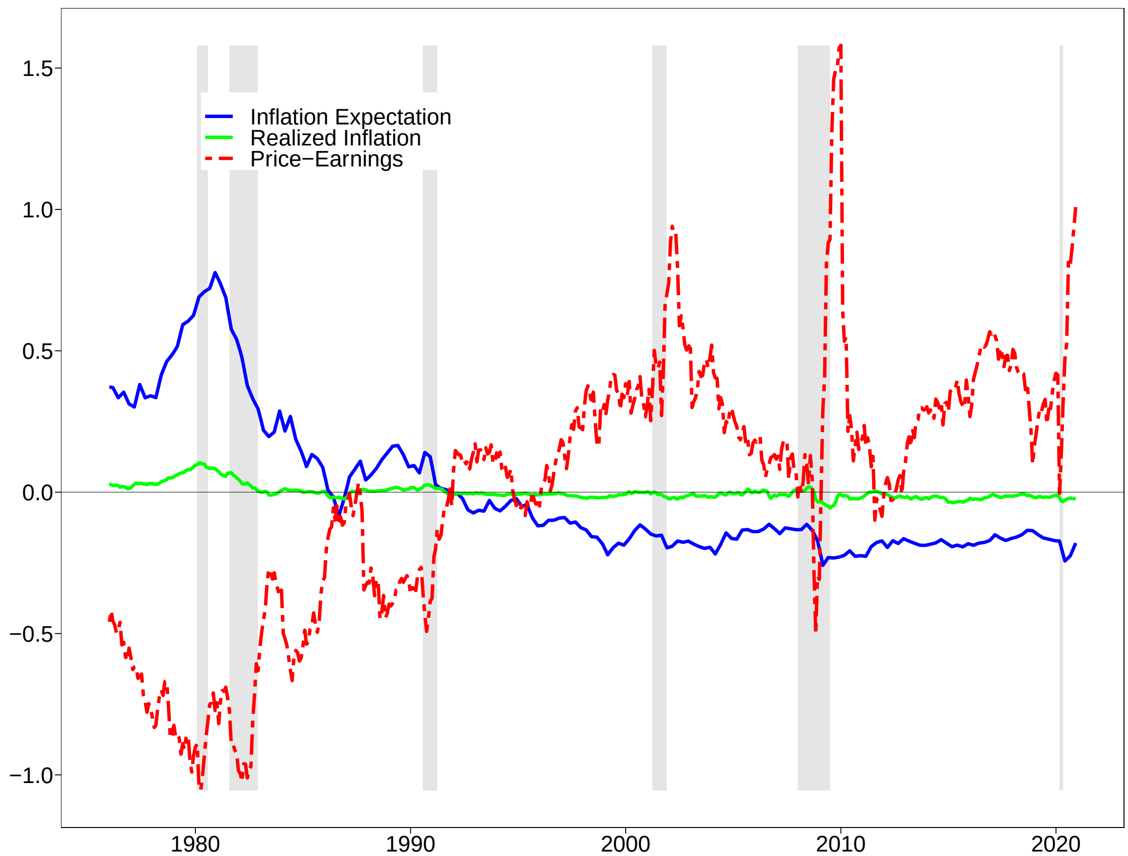


Figure 1.6. Posterior Break Locations in Comovements of Inflation Expectations and Price-Earnings Ratios

This figure plots the posterior probability (y-axis) of the structural break locations (x-axis) identified by the following Bayesian panel break model:

$$\Pi_t = \alpha_{ik} + \beta_{ik} \times pe_{it} + e_{it}, \quad e_{it} \sim \mathrm{N}(0, \sigma^2_{ik}), \quad t = \tau_{k-1} + 1, \ldots, \tau_k,$$

where Π_t and pe_{it} are the inflation expectation component and log price-earnings ratio of portfolio i in month t. The K structural break locations are captured by the set $(\tau_1, \tau_2, \ldots, \tau_K)$ in the model. Shades indicate NBER recession dates. The sample is monthly from January 1976 to December 2020.

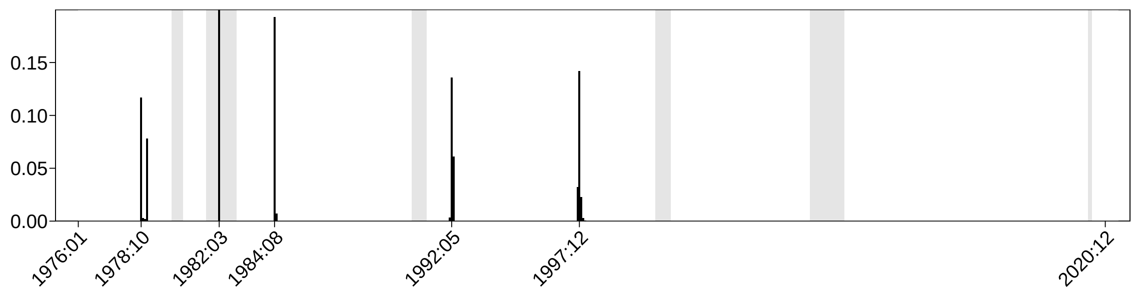


Figure 1.7. Comovements of Inflation Expectations and Price-Earnings Ratios

This figure plots the proportion of price-earnings variations explained by subjective inflation expectations (β_{ik} in regression (1.11)) under each regime for the market and five industry portfolios having the largest unconditional β_{ik}. The solid black line shows the Bayesian posterior mean of β_{ik} and the shade around it shows the 95% credible interval under each regime; and the dotted green line shows the corresponding OLS estimates. The dotted blue line shows the whole sample OLS estimate and the dotted red line shows the OLS estimate for the sample before 1998. The sample is monthly from January 1976 to December 2020.

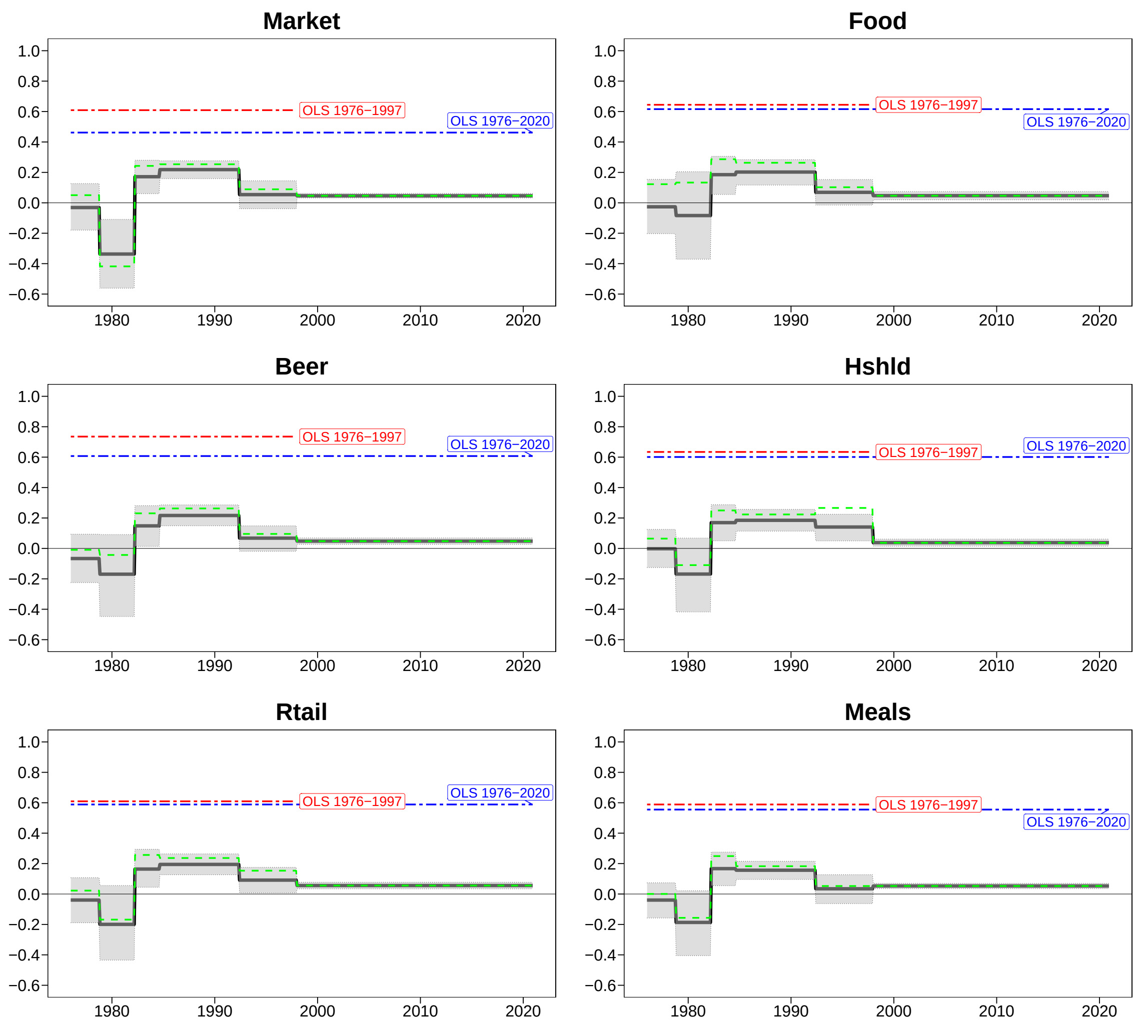


Table 1.1
Summary Statistic

This table reports the summary statistics of log price-earnings ratios (pe_t) in Panel A and subjective one-year earnings growth expectations ($\mathbb{E}_t^* \Delta e_{t+1}$) in Panel B for the market and 30 industry portfolios. All values for $\mathbb{E}_t^* \Delta e_{t+1}$ and the autocorrelation for pe_t are in percentages. The sample is monthly from January 1976 to December 2020.

	Panel A: Price-Earnings Ratio (pe_t)						Panel B: Expected One-Year Earnings Growth ($\mathbb{E}_t^*(\Delta e_{t+1})$)						
	Mean	SD	AC(1)	p25	p50	p75	Mean	SD	AC(1)	p25	p50	p75	Obs
0 Market	2.94	0.47	97.96	2.60	3.05	3.25	27.87	18.72	94.91	17.91	27.05	32.20	540
1 Food	2.86	0.37	98.67	2.66	2.96	3.11	19.87	8.67	92.42	14.19	19.14	23.79	540
2 Beer	2.78	0.35	97.53	2.55	2.85	3.00	12.72	12.47	91.28	7.16	11.42	17.05	540
3 Smoke	2.75	0.35	96.84	2.58	2.79	2.94	16.30	14.68	90.19	8.94	15.27	19.85	540
4 Games	3.15	0.60	94.34	2.87	3.31	3.52	31.91	26.94	87.59	18.66	33.24	45.38	526
5 Books	3.05	0.56	96.16	2.71	2.98	3.23	34.46	42.15	94.18	11.10	18.92	38.93	514
6 Hshld	2.88	0.37	98.07	2.66	2.99	3.12	19.90	9.63	92.90	13.18	20.02	24.78	540
7 Clths	2.74	0.50	96.61	2.48	2.84	3.07	20.59	16.75	87.87	11.98	18.45	24.20	540
8 Hlth	3.36	0.69	98.09	2.77	3.38	3.81	44.04	28.20	94.06	24.83	43.73	54.88	540
9 Chems	2.88	0.45	97.20	2.55	2.94	3.14	28.85	20.17	93.97	16.72	25.30	36.29	540
10 Txtls	2.79	0.62	92.65	2.49	2.82	3.10	33.01	40.08	86.71	15.22	28.74	47.22	519
11 Cnstr	2.81	0.42	96.68	2.55	2.82	3.05	26.19	19.37	93.06	14.58	23.98	33.62	527
12 Steel	2.94	0.77	94.73	2.32	2.85	3.47	47.00	52.52	90.62	10.78	34.93	79.65	501
13 FabPr	2.99	0.47	97.21	2.77	3.04	3.28	34.53	24.34	95.33	18.34	29.30	45.90	540
14 ElcEq	3.03	0.59	91.31	2.63	3.07	3.37	33.92	34.27	74.01	19.26	27.91	41.61	535
15 Autos	2.94	0.62	94.19	2.55	2.93	3.16	38.10	42.76	90.22	16.87	26.28	42.66	528
16 Carry	2.83	0.44	96.19	2.47	2.95	3.12	25.93	18.01	91.10	15.45	22.35	34.04	540
17 Mines	3.31	0.58	95.99	2.97	3.29	3.59	41.53	40.81	93.93	17.04	37.04	52.65	540
18 Coal	2.81	0.84	89.28	2.26	2.67	3.44	36.37	75.49	87.23	-3.92	24.06	88.79	462
19 Oil	3.05	0.62	93.66	2.68	3.06	3.32	34.96	38.20	89.10	15.28	28.94	44.70	515
20 Util	2.68	0.45	98.86	2.36	2.79	3.00	14.13	11.80	96.18	6.16	12.72	21.48	540
21 Telcm	3.04	0.59	95.11	2.65	3.09	3.29	21.64	24.81	73.40	10.87	18.81	27.42	514
22 Servs	3.25	0.65	92.73	2.81	3.29	3.66	40.54	37.61	80.84	22.67	36.71	52.40	528
23 BusEq	3.17	0.61	95.65	2.77	3.19	3.41	41.31	37.57	90.62	24.93	37.19	46.04	521
24 Paper	2.88	0.53	96.89	2.50	2.90	3.23	31.15	29.50	93.19	12.87	26.91	42.64	540
25 Trans	2.94	0.51	96.54	2.67	2.97	3.19	33.85	29.58	94.53	18.07	26.51	40.33	540
26 Whlsl	2.93	0.40	97.34	2.70	3.03	3.20	29.91	15.99	93.49	19.51	28.61	37.82	540
27 Rtail	2.88	0.37	97.69	2.68	3.00	3.13	22.73	11.78	94.04	15.26	21.75	28.18	540
28 Meals	3.02	0.37	95.95	2.78	3.11	3.27	23.99	13.45	90.31	17.22	21.49	30.48	540
29 Fin	2.73	0.56	95.98	2.40	2.83	2.98	23.48	36.03	92.41	14.20	18.04	22.80	540
30 Other	3.12	0.49	96.87	2.87	3.14	3.40	37.49	24.84	94.83	20.92	31.73	43.27	540
All Stocks	2.94	0.81	87.76	2.47	2.87	3.26	27.59	62.23	80.60	5.04	18.29	38.93	1132285

Table 1.2
Comovements of Earnings Growth Expectations and Price-Earnings Ratios

This table reports the proportions of price-earnings variations explained by cash flow expectations estimated in the following Bayesian panel break model:

$$CF_{it} = \alpha_{ik} + \beta_{ik} \times pe_{it} + e_{it}, \quad e_{it} \sim \mathrm{N}(0, \sigma_{ik}^2), \quad t = \tau_{k-1}+1, \ldots, \tau_k,$$

where CF_{it} and pe_{it} are the cash flow component and log price-earnings ratio of portfolio i in month t. The K structural break locations are captured by the set $(\tau_1, \tau_2, \ldots, \tau_K)$ in the model. Header (0) reports the results using the whole sample and each of the following headers reports the results for the corresponding regime identified by the model. Under each header, the posterior mean of β_{ik} together with its posterior standard deviation from the Bayesian break model is reported in the *Bayesian* column while the OLS estimate together with its standard error using data within the respective regime is reported in the *OLS* column. ***, **, * under the *Bayesian* columns indicate the posterior mean is at least 2.58, 1.96, and 1.65 times larger than the posterior standard deviation while ***, **, * under the *OLS* columns indicate significance at 1%, 5%, and 10% based on the Newey and West (1987) standard error with 6 lags. Blanks indicate industries having fewer than six observations during that regime. The sample is monthly from January 1976 to December 2020.

Regime	1976:1-2020:12 (0)		1976:1-1986:7 (1)		1986:8-2000:11 (2)		2000:12-2004:10 (3)		2004:11-2008:11 (4)		2008:12-2010:7 (5)		2010:8-2020:1 (6)		2020:2-2020:12 (7)	
	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS
0 Market	0.34 ***	0.34 ***	0.24 ***	0.27 ***	0.20 ***	0.22 ***	0.41 ***	0.75 ***	0.07	0.21 *	0.76 ***	0.91 ***	0.18 ***	0.25 ***	0.29 **	0.66 ***
	(0.01)	(0.06)	(0.03)	(0.04)	(0.02)	(0.03)	(0.09)	(0.10)	(0.06)	(0.11)	(0.08)	(0.04)	(0.03)	(0.04)	(0.14)	(0.12)
1 Food	0.13 ***	0.14 ***	0.09 ***	0.10 **	0.15 ***	0.18 ***	0.04	0.34 ***	0.05	0.23	0.36 ***	0.78 ***	0.20 ***	0.32 ***	0.08	0.55 ***
	(0.01)	(0.02)	(0.03)	(0.05)	(0.02)	(0.02)	(0.08)	(0.13)	(0.06)	(0.18)	(0.11)	(0.07)	(0.04)	(0.09)	(0.12)	(0.11)
2 Beer	0.13 ***	0.13 ***	0.06	0.08	0.21 ***	0.24 ***	0.21 ***	0.53 ***	0.02	0.24	0.42 ***	0.86 ***	0.23 ***	0.33 ***	0.16	0.61 ***
	(0.02)	(0.04)	(0.04)	(0.05)	(0.02)	(0.04)	(0.08)	(0.11)	(0.08)	(0.21)	(0.12)	(0.07)	(0.05)	(0.09)	(0.13)	(0.17)
3 Smoke	0.28 ***	0.28 ***	0.10 ***	0.13 ***	0.07 **	0.11 ***	0.41 ***	0.67 ***	0.17 *	0.63 **	0.32 ***	0.73 ***	0.04	0.07 *	0.29 **	0.75 ***
	(0.01)	(0.07)	(0.03)	(0.04)	(0.03)	(0.04)	(0.08)	(0.13)	(0.09)	(0.25)	(0.11)	(0.02)	(0.03)	(0.04)	(0.13)	(0.21)
4 Games	0.31 ***	0.32 ***	0.41 ***	0.43 ***	0.31 ***	0.35 ***	0.57 ***	0.77 ***	0.27 ***	0.42 ***	0.58 ***	0.68 ***	0.52 ***	0.59 ***	0.48 ***	0.60 ***
	(0.01)	(0.03)	(0.03)	(0.04)	(0.03)	(0.06)	(0.07)	(0.05)	(0.06)	(0.07)	(0.08)	(0.01)	(0.03)	(0.06)	(0.14)	(0.13)
5 Books	0.65 ***	0.65 ***	0.08 ***	0.10 ***	0.56 ***	0.64 ***	0.65 ***	0.78 ***	-0.08	-0.13	0.70 ***	1.01 ***	0.75 ***	0.76 ***		
	(0.02)	(0.05)	(0.02)	(0.03)	(0.03)	(0.08)	(0.05)	(0.03)	(0.09)	(0.14)	(0.18)	(0.08)	(0.03)	(0.05)		
6 Hshld	0.15 ***	0.15 ***	0.15 ***	0.17 ***	0.18 ***	0.21 ***	0.13 *	0.46 ***	0.10 *	0.32 ***	0.43 ***	0.65 ***	0.13 ***	0.24 ***	0.18	0.59 ***
	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.07)	(0.08)	(0.11)	(0.06)	(0.07)	(0.09)	(0.03)	(0.04)	(0.07)	(0.13)	(0.10)
7 Clths	0.22 ***	0.22 ***	0.22 ***	0.23 ***	0.28 ***	0.30 ***	0.07	0.34 ***	0.09 **	0.19 ***	0.57 ***	0.70 ***	0.34 ***	0.50 ***	0.65 ***	0.82 ***
	(0.01)	(0.05)	(0.02)	(0.03)	(0.02)	(0.05)	(0.09)	(0.07)	(0.05)	(0.02)	(0.08)	(0.05)	(0.05)	(0.15)	(0.11)	(0.07)
8 Hlth	0.43 ***	0.43 ***	0.19 ***	0.22 ***	0.38 ***	0.39 ***	0.04	-0.04	0.10 *	0.19 **	0.44 ***	0.86 ***	0.50 ***	0.53 ***	0.77 ***	1.04 ***
	(0.01)	(0.03)	(0.03)	(0.07)	(0.02)	(0.04)	(0.08)	(0.15)	(0.05)	(0.08)	(0.10)	(0.05)	(0.02)	(0.04)	(0.13)	(0.09)
9 Chems	0.42 ***	0.42 ***	0.33 ***	0.36 ***	0.36 ***	0.40 ***	0.74 ***	0.90 ***	0.17 ***	0.32 ***	0.58 ***	0.78 ***	0.31 ***	0.38 ***	0.38 ***	0.67 ***
	(0.01)	(0.04)	(0.03)	(0.04)	(0.03)	(0.06)	(0.06)	(0.04)	(0.05)	(0.05)	(0.09)	(0.04)	(0.03)	(0.04)	(0.12)	(0.11)
10 Txtls	0.55 ***	0.55 ***	0.54 ***	0.55 ***	0.49 ***	0.52 ***	0.94 ***	0.96 ***	0.73 ***	0.97 ***	0.08		0.58 ***	0.65 ***	0.51 ***	0.77 ***
	(0.02)	(0.10)	(0.03)	(0.07)	(0.03)	(0.05)	(0.03)	(0.02)	(0.08)	(0.07)	(0.18)		(0.03)	(0.05)	(0.18)	(0.15)

Table 1.2
Comovements of Earnings Growth Expectations and Price-Earnings Ratios (Continued)

Regime	1976:1-2020:12 (0)		1976:1-1986:7 (1)		1986:8-2000:11 (2)		2000:12-2004:10 (3)		2004:11-2008:11 (4)		2008:12-2010:7 (5)		2010:8-2020:1 (6)		2020:2-2020:12 (7)	
	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS
11 Cnstr	0.40 ***	0.41 ***	0.45 ***	0.48 ***	0.67 ***	0.72 ***	0.22 ***	0.45 ***	0.40 ***	0.46 ***	0.50 ***	0.93 ***	0.47 ***	0.57 ***	0.18	0.44 **
	(0.02)	(0.05)	(0.03)	(0.06)	(0.03)	(0.08)	(0.07)	(0.07)	(0.07)	(0.12)	(0.15)	(0.03)	(0.05)	(0.12)	(0.14)	(0.14)
12 Steel	0.64 ***	0.64 ***	0.62 ***	0.62 ***	0.57 ***	0.58 ***	0.86 ***	1.01 ***	0.15 ***	0.20 ***	0.87 ***	0.93 ***	0.77 ***	0.78 ***	0.77 ***	0.88 ***
	(0.01)	(0.03)	(0.02)	(0.05)	(0.02)	(0.03)	(0.14)	(0.11)	(0.05)	(0.07)	(0.10)	(0.09)	(0.02)	(0.05)	(0.11)	(0.08)
13 FabPr	0.47 ***	0.48 ***	0.62 ***	0.64 ***	0.60 ***	0.65 ***	0.67 ***	0.88 ***	0.25 ***	0.42 ***	0.81 ***	0.92 ***	0.36 ***	0.44 ***	0.22	0.59 ***
	(0.02)	(0.05)	(0.03)	(0.07)	(0.03)	(0.05)	(0.08)	(0.09)	(0.06)	(0.08)	(0.07)	(0.04)	(0.04)	(0.05)	(0.14)	(0.15)
14 ElcEq	0.48 ***	0.48 ***	0.28 ***	0.31 ***	0.46 ***	0.52 ***	0.91 ***	0.96 ***	0.18 ***	0.25 ***	0.74 ***	0.82 ***	0.19 ***	0.24 ***	0.09	0.25 ***
	(0.01)	(0.10)	(0.03)	(0.05)	(0.03)	(0.07)	(0.04)	(0.05)	(0.04)	(0.04)	(0.08)	(0.08)	(0.05)	(0.07)	(0.10)	(0.07)
15 Autos	0.66 ***	0.66 ***	0.53 ***	0.57 ***	0.79 ***	0.81 ***	0.97 ***	1.00 ***	0.47 ***	0.64 ***	0.83 ***	1.00 ***	0.39 ***	0.46 ***	0.66 ***	0.99 ***
	(0.02)	(0.07)	(0.04)	(0.06)	(0.02)	(0.06)	(0.03)	(0.02)	(0.07)	(0.10)	(0.13)	(0.04)	(0.04)	(0.06)	(0.18)	(0.17)
16 Carry	0.28 ***	0.29 ***	0.39 ***	0.42 ***	0.48 ***	0.52 ***	0.17 **	0.43 ***	0.17 ***	0.29 ***	0.56 ***	0.72 ***	0.23 ***	0.35 ***	0.49 ***	0.71 ***
	(0.01)	(0.04)	(0.03)	(0.05)	(0.04)	(0.12)	(0.07)	(0.07)	(0.05)	(0.05)	(0.09)	(0.05)	(0.05)	(0.11)	(0.13)	(0.10)
17 Mines	0.57 ***	0.57 ***	0.51 ***	0.52 ***	0.47 ***	0.50 ***	0.48 ***	0.89 ***	0.22 **	0.47 **	0.90 ***	1.03 ***	0.83 ***	0.85 ***	0.35 **	0.82 ***
	(0.02)	(0.09)	(0.03)	(0.05)	(0.03)	(0.04)	(0.09)	(0.05)	(0.09)	(0.21)	(0.08)	(0.04)	(0.02)	(0.04)	(0.14)	(0.11)
18 Coal	0.86 ***	0.86 ***	0.95 ***	0.97 ***	0.91 ***	0.92 ***	0.64 ***	0.72 ***	0.46 ***	0.51 **	0.55 ***	0.62 ***	0.99 ***	1.01 ***	0.03	
	(0.02)	(0.05)	(0.05)	(0.10)	(0.02)	(0.03)	(0.08)	(0.10)	(0.10)	(0.25)	(0.13)	(0.17)	(0.06)	(0.09)	(0.72)	
19 Oil	0.52 ***	0.52 ***	0.26 ***	0.30 ***	0.70 ***	0.73 ***	0.88 ***	1.05 ***	0.25 ***	0.41 ***	0.89 ***	0.94 ***	0.38 ***	0.41 ***	0.30 *	0.39 **
	(0.02)	(0.09)	(0.03)	(0.04)	(0.02)	(0.06)	(0.08)	(0.08)	(0.07)	(0.13)	(0.05)	(0.03)	(0.06)	(0.06)	(0.16)	(0.09)
20 Util	0.23 ***	0.23 ***	0.05	0.10 **	0.17 ***	0.21 ***	-0.01	0.03	0.17 **	0.40 ***	0.36 ***	0.71 ***	0.37 ***	0.49 ***	0.12	0.62 ***
	(0.01)	(0.01)	(0.03)	(0.05)	(0.02)	(0.03)	(0.07)	(0.17)	(0.07)	(0.13)	(0.11)	(0.02)	(0.03)	(0.04)	(0.13)	(0.11)
21 Telcm	0.29 ***	0.29 ***	0.14 ***	0.16 ***	0.14 ***	0.14 ***	0.76 ***	0.83 ***	0.24 ***	0.38 ***	0.55 ***	0.98 ***	0.23 ***	0.34 *	0.18	0.71 ***
	(0.01)	(0.09)	(0.03)	(0.04)	(0.02)	(0.05)	(0.08)	(0.12)	(0.06)	(0.13)	(0.19)	(0.07)	(0.06)	(0.19)	(0.15)	(0.12)
22 Servs	0.54 ***	0.55 ***	0.30 ***	0.34 ***	0.36 ***	0.37 ***	0.94 ***	0.99 ***	0.11 **	0.23 ***	0.48 ***	0.75 ***	0.35 ***	0.42 ***	0.23 **	0.50 ***
	(0.01)	(0.10)	(0.02)	(0.02)	(0.01)	(0.03)	(0.04)	(0.02)	(0.05)	(0.06)	(0.09)	(0.04)	(0.03)	(0.04)	(0.11)	(0.09)
23 BusEq	0.56 ***	0.56 ***	0.29 ***	0.32 ***	0.23 ***	0.25 ***	0.91 ***	0.95 ***	0.16 ***	0.29 ***	0.80 ***	0.91 ***	0.35 ***	0.45 ***	0.13	0.46 ***
	(0.01)	(0.09)	(0.03)	(0.05)	(0.02)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.07)	(0.01)	(0.04)	(0.07)	(0.11)	(0.10)
24 Paper	0.57 ***	0.58 ***	0.48 ***	0.51 ***	0.56 ***	0.58 ***	0.88 ***	1.08 ***	0.60 ***	0.74 ***	0.94 ***	1.03 ***	0.41 ***	0.55 ***	0.38 **	0.84 ***
	(0.02)	(0.06)	(0.03)	(0.07)	(0.03)	(0.05)	(0.08)	(0.11)	(0.06)	(0.11)	(0.06)	(0.02)	(0.04)	(0.06)	(0.15)	(0.13)
25 Trans	0.47 ***	0.47 ***	0.40 ***	0.42 ***	0.74 ***	0.78 ***	0.55 ***	0.73 ***	0.21 ***	0.48 **	0.72 ***	0.81 ***	0.24 ***	0.41 **	0.46 ***	0.64 ***
	(0.02)	(0.07)	(0.02)	(0.03)	(0.03)	(0.06)	(0.06)	(0.08)	(0.08)	(0.19)	(0.06)	(0.05)	(0.07)	(0.16)	(0.15)	(0.13)
26 Whlsl	0.30 ***	0.30 ***	0.34 ***	0.36 ***	0.32 ***	0.38 ***	0.32 ***	0.57 ***	0.23 ***	0.39 ***	0.62 ***	0.80 ***	0.40 ***	0.54 ***	0.27 **	0.58 ***
	(0.01)	(0.04)	(0.02)	(0.05)	(0.03)	(0.05)	(0.06)	(0.06)	(0.05)	(0.07)	(0.08)	(0.05)	(0.04)	(0.05)	(0.13)	(0.12)
27 Rtail	0.21 ***	0.22 ***	0.15 ***	0.17 ***	0.20 ***	0.24 ***	0.16 **	0.36 ***	0.20 ***	0.42 ***	0.58 ***	0.82 ***	0.04	0.12	0.31 **	0.63 ***
	(0.01)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.06)	(0.06)	(0.06)	(0.11)	(0.10)	(0.02)	(0.04)	(0.09)	(0.13)	(0.11)
28 Meals	0.17 ***	0.17 ***	0.06 **	0.07	0.31 ***	0.36 ***	0.08	0.29 ***	0.15 ***	0.31 **	0.47 ***	0.70 ***	0.29 ***	0.44 ***	0.44 ***	0.69 ***
	(0.01)	(0.04)	(0.03)	(0.05)	(0.03)	(0.04)	(0.07)	(0.08)	(0.06)	(0.12)	(0.09)	(0.03)	(0.05)	(0.11)	(0.14)	(0.11)
29 Fin	0.56 ***	0.56 ***	0.17 ***	0.19 ***	-0.03	-0.03	0.11	0.44 ***	-0.02	0.11	0.98 ***	1.03 ***	0.11 **	0.22 **	0.20	0.59 ***
	(0.02)	(0.17)	(0.02)	(0.02)	(0.02)	(0.04)	(0.08)	(0.12)	(0.08)	(0.31)	(0.04)	(0.03)	(0.05)	(0.10)	(0.15)	(0.13)
30 Other	0.40 ***	0.41 ***	0.21 ***	0.22 ***	0.75 ***	0.79 ***	0.25 ***	0.31 ***	0.35 ***	0.57 ***	0.26 ***	0.50 ***	0.72 ***	0.85 ***	0.33 ***	0.72 ***
	(0.01)	(0.04)	(0.02)	(0.03)	(0.02)	(0.03)	(0.06)	(0.10)	(0.06)	(0.09)	(0.09)	(0.06)	(0.04)	(0.07)	(0.12)	(0.10)

Table 1.3
Cross-Regime Changes in Comovements of Earnings Growth Expectations and Price-Earnings Ratios

This table reports the cross-regime changes in proportion of price-earnings variations explained by cash flow expectations. Each column reports the posterior mean and standard deviation of the difference between β_{ik} in that regime and $\beta_{i,k-1}$ in the preceding regime ($\beta_{ik} - \beta_{i,k-1}$ in regression (1.11)). Panel A reports the results for the market only (top row), the average across the 30 industries (second row), and the average across both the market and 30 industries (third row); Panel B reports the results for each industry individually. ***, **, * indicate the posterior mean is at least 2.58, 1.96, and 1.65 times larger than the posterior standard deviation. Blanks indicate industries having fewer than six observations during that regime. The sample is monthly from January 1976 to December 2020.

Regime	1986:8-2000:11	2000:12-2004:10	2004:11-2008:11	2008:12-2010:7	2010:8-2020:1	2020:2-2020:12
Panel A: Aggregate						
Market	-0.05	0.21 **	-0.34 ***	0.70 ***	-0.59 ***	0.12
	(0.03)	(0.09)	(0.11)	(0.09)	(0.08)	(0.14)
Industries	0.10 ***	0.07 ***	-0.27 ***	0.38 ***	-0.21 ***	-0.06
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)
Market + Industries	0.10 ***	0.08 ***	-0.27 ***	0.39 ***	-0.22 ***	-0.05
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)
Panel B: Individual Industry						
1 Food	0.06 *	-0.11	0.01	0.31 **	-0.15	-0.13
	(0.03)	(0.08)	(0.10)	(0.13)	(0.12)	(0.13)
2 Beer	0.15 ***	-0.00	-0.19 *	0.41 ***	-0.19	-0.08
	(0.04)	(0.08)	(0.11)	(0.14)	(0.13)	(0.14)
3 Smoke	-0.03	0.34 ***	-0.24 **	0.15	-0.28 **	0.25 *
	(0.04)	(0.09)	(0.12)	(0.14)	(0.11)	(0.13)
4 Games	-0.09 **	0.26 ***	-0.30 ***	0.31 ***	-0.06	-0.04
	(0.04)	(0.07)	(0.09)	(0.10)	(0.09)	(0.15)
5 Books	0.49 ***	0.09	-0.72 ***	0.78 ***	0.04	
	(0.04)	(0.06)	(0.10)	(0.21)	(0.18)	
6 Hshld	0.02	-0.05	-0.02	0.33 ***	-0.30 ***	0.05
	(0.03)	(0.08)	(0.10)	(0.10)	(0.09)	(0.14)
7 Clths	0.06 **	-0.20 **	0.02	0.48 ***	-0.22 **	0.31 **
	(0.03)	(0.09)	(0.10)	(0.09)	(0.09)	(0.12)
8 Hlth	0.19 ***	-0.34 ***	0.05	0.34 ***	0.06	0.27 **
	(0.04)	(0.08)	(0.10)	(0.12)	(0.11)	(0.14)
9 Chems	0.02	0.38 ***	-0.56 ***	0.41 ***	-0.28 ***	0.07
	(0.04)	(0.07)	(0.08)	(0.11)	(0.10)	(0.13)
10 Txtls	-0.05	0.45 ***	-0.20 **	-0.65 ***	0.50 ***	-0.07
	(0.05)	(0.05)	(0.09)	(0.20)	(0.18)	(0.18)
11 Cnstr	0.22 ***	-0.45 ***	0.18 *	0.09	-0.03	-0.28 *
	(0.04)	(0.08)	(0.10)	(0.17)	(0.16)	(0.15)
12 Steel	-0.05	0.30 **	-0.71 ***	0.72 ***	-0.11	0.01
	(0.03)	(0.14)	(0.15)	(0.12)	(0.10)	(0.12)
13 FabPr	-0.01	0.07	-0.43 ***	0.56 ***	-0.45 ***	-0.14
	(0.04)	(0.08)	(0.10)	(0.09)	(0.08)	(0.15)
14 ElcEq	0.18 ***	0.45 ***	-0.74 ***	0.57 ***	-0.55 ***	-0.10
	(0.04)	(0.05)	(0.06)	(0.09)	(0.09)	(0.11)
15 Autos	0.26 ***	0.18 ***	-0.50 ***	0.37 **	-0.45 ***	0.27
	(0.04)	(0.04)	(0.08)	(0.15)	(0.14)	(0.19)

Table 1.3
Cross-Regime Changes in Comovements of Earnings Growth Expectations and Price-Earnings Ratios (Continued)

Regime	1986:8-2000:11	2000:12-2004:10	2004:11-2008:11	2008:12-2010:7	2010:8-2020:1	2020:2-2020:12
Panel B: Individual Industry						
16 Carry	0.08 *	-0.30 ***	-0.01	0.40 ***	-0.33 ***	0.26 *
	(0.05)	(0.08)	(0.09)	(0.10)	(0.10)	(0.14)
17 Mines	-0.04	0.01	-0.25 *	0.68 ***	-0.07	-0.48 ***
	(0.04)	(0.10)	(0.13)	(0.12)	(0.08)	(0.15)
18 Coal	-0.04	-0.28 ***	-0.18	0.09	0.44 ***	-0.96
	(0.05)	(0.08)	(0.12)	(0.16)	(0.14)	(0.73)
19 Oil	0.44 ***	0.17 **	-0.63 ***	0.64 ***	-0.50 ***	-0.08
	(0.04)	(0.08)	(0.11)	(0.09)	(0.08)	(0.17)
20 Util	0.11 ***	-0.18 **	0.17 *	0.19	0.01	-0.25 *
	(0.04)	(0.08)	(0.10)	(0.13)	(0.11)	(0.14)
21 Telcm	0.00	0.62 ***	-0.52 ***	0.31	-0.32 *	-0.04
	(0.03)	(0.08)	(0.10)	(0.19)	(0.19)	(0.16)
22 Servs	0.05 *	0.58 ***	-0.83 ***	0.36 ***	-0.13	-0.12
	(0.03)	(0.04)	(0.06)	(0.10)	(0.10)	(0.11)
23 BusEq	-0.06 *	0.67 ***	-0.74 ***	0.64 ***	-0.45 ***	-0.22 *
	(0.03)	(0.05)	(0.06)	(0.09)	(0.08)	(0.12)
24 Paper	0.08 **	0.32 ***	-0.28 ***	0.34 ***	-0.52 ***	-0.03
	(0.04)	(0.09)	(0.10)	(0.09)	(0.08)	(0.16)
25 Trans	0.35 ***	-0.20 ***	-0.33 ***	0.51 ***	-0.48 ***	0.22
	(0.03)	(0.07)	(0.10)	(0.10)	(0.09)	(0.16)
26 Whlsl	-0.01	-0.00	-0.09	0.39 ***	-0.22 **	-0.14
	(0.04)	(0.07)	(0.08)	(0.10)	(0.09)	(0.14)
27 Rtail	0.05 *	-0.04	0.04	0.38 ***	-0.54 ***	0.27 *
	(0.03)	(0.07)	(0.09)	(0.12)	(0.11)	(0.14)
28 Meals	0.25 ***	-0.24 ***	0.07	0.32 ***	-0.17 *	0.15
	(0.04)	(0.08)	(0.09)	(0.10)	(0.10)	(0.15)
29 Fin	-0.20 ***	0.14 *	-0.13	1.00 ***	-0.87 ***	0.09
	(0.03)	(0.08)	(0.12)	(0.09)	(0.06)	(0.16)
30 Other	0.55 ***	-0.50 ***	0.10	-0.09	0.46 ***	-0.39 ***
	(0.03)	(0.06)	(0.08)	(0.11)	(0.10)	(0.13)

Table 1.4
Comovements of Subjective Expectations and Price-Earnings under Uncertainty

This table reports the results of the following regression

$$CF_{it} = \alpha + b_0 \times pe_{it} + b_1 \times z_{t-1} \times pe_{it} + e_{it},$$

Panel A reports the results over the whole sample 1976-2020 while Panel B reports the results over 2000-2020. Within each panel, the first row reports b_1's with their t-statistics computed with the Newey and West (1987) standard errors for only the market portfolio; the second row reports b_1's with their t-statistics computed with the two-way clustered standard errors for the 30 industry portfolios; and the third row reports b_1's with their t-statistics computed with the two-way clustered standard errors for both the market and 30 industry portfolios. Each z_{it} is reported in each column: FinUncer and MacroUncer are financial uncertainty from Ludvigson et al. (2021) and macro uncertainty from Jurado et al. (2015); VIX is CBOE Volatility Index; EPU is economic policy uncertainty from Baker et al. (2016); ConSurplus is consumption surplus from Cochrane (2017); GPR is geopolitical risk from Caldara and Iacoviello (2022); and pe is log price-earnings ratio. The data for EPU is only available from 1985 and the data for VIX is only available from 1990. ***, **, * indicates significance at 1%, 5%, and 10%, respectively.

	FinUncer	VIX	MacroUncer	EPU	ConSurplus	GPR	pe
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 1976:2020							
Market	0.01 **	0.01	0.01	0.00	0.03	0.00	0.20 ***
	(2.37)	(1.64)	(1.58)	(0.80)	(0.77)	(1.54)	(4.48)
Industries	0.01 **	0.01 *	0.01 **	-0.00	-0.01	0.01 ***	0.10 ***
	(2.34)	(1.82)	(2.30)	(-0.08)	(-0.61)	(2.72)	(5.00)
Market + Industries	0.01 **	0.01 *	0.01 **	-0.00	-0.01	0.01 ***	0.10 ***
	(2.39)	(1.84)	(2.34)	(-0.05)	(-0.55)	(2.72)	(5.08)
Panel B: 2000:2020							
Market	0.01 **	0.01 **	0.00	0.00	0.03	-0.00	0.14 ***
	(2.56)	(2.21)	(0.56)	(0.34)	(0.58)	(-0.91)	(3.28)
Industries	0.01 ***	0.01 ***	0.00	-0.00	-0.00	0.00	0.05 ***
	(3.99)	(2.71)	(0.45)	(-0.17)	(-0.08)	(0.18)	(3.28)
Market + Industries	0.01 ***	0.01 ***	0.00	-0.00	-0.00	0.00	0.05 ***
	(4.04)	(2.74)	(0.47)	(-0.15)	(-0.04)	(0.13)	(3.32)

Table 1.5

Comovements of Inflation Expectations and Price-Earnings Ratios

This table reports the proportions of price-earnings variations explained by inflation expectations estimated in the following Bayesian panel break model:

$$\Pi_t = \alpha_{ik} + \beta_{ik} \times pe_{it} + e_{it}, \quad e_{it} \sim \mathrm{N}(0, \sigma_{ik}^2), \quad t = \tau_{k-1} + 1, \ldots, \tau_k,$$

where Π_t and pe_{it} are the inflation expectation component and log price-earnings ratio of portfolio i in month t. The K structural break locations are captured by the set $(\tau_1, \tau_2, \ldots, \tau_K)$ in the model. Header (0) reports the results using the whole sample and each of the following headers reports the results for the corresponding regime identified by the model. Under each header, the posterior mean of β_{ik} together with its posterior standard deviation from the Bayesian break model is reported in the *Bayesian* column while the OLS estimate together with its standard error using data within the respective regime is reported in the *OLS* column. ***, **, * under the *Bayesian* columns indicate the posterior mean is at least 2.58, 1.96, and 1.65 times larger than the posterior standard deviation while ***, **, * under the *OLS* columns indicate significance at 1%, 5%, and 10% based on the Newey and West (1987) standard error with 6 lags. The sample is monthly from January 1976 to December 2020.

Regime	1976:1-2020:12 (0)		1976:1-1978:10 (1)		1978:11-1982:3 (2)		1982:4-1984:8 (3)		1984:9-1992:5 (4)		1992:6-1997:12 (5)		1998:1-2020:12 (6)	
	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS
0 Market	0.46 ***	0.46 ***	-0.03	0.05	-0.34 ***	-0.42 *	0.17 ***	0.24 ***	0.22 ***	0.25 ***	0.05	0.09	0.05 ***	0.05 ***
	(0.01)	(0.05)	(0.08)	(0.08)	(0.11)	(0.23)	(0.06)	(0.03)	(0.03)	(0.05)	(0.05)	(0.08)	(0.01)	(0.01)
1 Food	0.61 ***	0.62 ***	-0.03	0.12	-0.08	0.13	0.19 ***	0.29 ***	0.20 ***	0.26 ***	0.07 *	0.10 *	0.05 ***	0.05 *
	(0.01)	(0.03)	(0.09)	(0.14)	(0.14)	(0.40)	(0.06)	(0.05)	(0.04)	(0.06)	(0.04)	(0.06)	(0.01)	(0.03)
2 Beer	0.60 ***	0.61 ***	-0.07	-0.01	-0.17	-0.04	0.15 **	0.23 ***	0.22 ***	0.26 ***	0.07	0.10 *	0.05 ***	0.05 ***
	(0.02)	(0.05)	(0.08)	(0.05)	(0.14)	(0.31)	(0.07)	(0.08)	(0.03)	(0.06)	(0.04)	(0.05)	(0.01)	(0.02)
3 Smoke	0.55 ***	0.55 ***	-0.02	0.04	-0.21 *	-0.19	0.18 ***	0.27 ***	0.18 ***	0.21 ***	0.01	-0.02	0.03 ***	0.03 ***
	(0.02)	(0.08)	(0.07)	(0.06)	(0.11)	(0.26)	(0.06)	(0.04)	(0.03)	(0.03)	(0.05)	(0.08)	(0.01)	(0.01)
4 Games	0.35 ***	0.35 ***	-0.14 **	-0.14 **	0.09	0.18	0.14 ***	0.19 ***	0.08 ***	0.09 ***	-0.03	-0.08	0.03 ***	0.03 ***
	(0.01)	(0.04)	(0.07)	(0.07)	(0.09)	(0.11)	(0.05)	(0.04)	(0.02)	(0.03)	(0.04)	(0.06)	(0.01)	(0.01)
5 Books	0.29 ***	0.29 ***	-0.08	-0.01	-0.15	-0.08	0.17 **	0.27 ***	0.17 ***	0.21 ***	0.03	0.03	0.00	0.00
	(0.02)	(0.06)	(0.08)	(0.06)	(0.11)	(0.17)	(0.06)	(0.04)	(0.04)	(0.06)	(0.03)	(0.03)	(0.00)	(0.01)
6 Hshld	0.60 ***	0.60 ***	-0.00	0.06	-0.17	-0.11	0.17 ***	0.25 ***	0.18 ***	0.22 ***	0.14 ***	0.27 ***	0.04 ***	0.04
	(0.01)	(0.04)	(0.06)	(0.07)	(0.12)	(0.28)	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)	(0.02)	(0.01)	(0.03)
7 Clths	0.41 ***	0.41 ***	-0.04	-0.01	-0.41 ***	-0.52 **	0.16 ***	0.21 ***	0.14 ***	0.15 ***	0.10 ***	0.12 ***	0.03 ***	0.02 **
	(0.01)	(0.04)	(0.07)	(0.03)	(0.12)	(0.19)	(0.05)	(0.04)	(0.02)	(0.04)	(0.02)	(0.01)	(0.01)	(0.01)
8 Hlth	0.29 ***	0.29 ***	-0.04	0.03	-0.33 ***	-0.42 *	0.18 ***	0.29 ***	0.19 ***	0.22 ***	0.10 ***	0.14 ***	0.00	0.00
	(0.01)	(0.04)	(0.08)	(0.06)	(0.11)	(0.22)	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)	(0.00)	(0.01)
9 Chems	0.44 ***	0.44 ***	-0.04	0.02	-0.34 ***	-0.39 **	0.16 ***	0.21 ***	0.20 ***	0.23 ***	-0.06 **	-0.09 ***	0.03 ***	0.03 **
	(0.01)	(0.05)	(0.07)	(0.06)	(0.10)	(0.18)	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)	(0.02)	(0.01)	(0.01)
10 Txtls	0.29 ***	0.29 ***	-0.07	-0.06 *	-0.14	-0.10	0.11 ***	0.13 ***	0.14 ***	0.15 ***	0.06 **	0.08 ***	0.02 ***	0.02 **
	(0.01)	(0.06)	(0.05)	(0.03)	(0.10)	(0.17)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.00)	(0.01)

Table 1.5
Comovements of Inflation Expectations and Price-Earnings Ratios (Continued)

Regime	1976:1-2020:12 (0)		1976:1-1978:10 (1)		1978:11-1982:3 (2)		1982:4-1984:8 (3)		1984:9-1992:5 (4)		1992:6-1997:12 (5)		1998:1-2020:12 (6)	
	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS	Bayesian	OLS
11 Cnstr	0.41 ***	0.41 ***	-0.00	0.05	-0.30 ***	-0.32 **	0.14 ***	0.18 ***	0.14 ***	0.15 ***	-0.03	-0.04	0.03 ***	0.03 **
	(0.02)	(0.06)	(0.06)	(0.06)	(0.08)	(0.14)	(0.04)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.01)	(0.01)
12 Steel	0.17 ***	0.17 ***	-0.11	-0.09	-0.11	-0.08	0.07 ***	0.07 ***	0.07 ***	0.07 ***	-0.03 ***	-0.04 ***	0.02 ***	0.02 ***
	(0.01)	(0.03)	(0.08)	(0.08)	(0.09)	(0.16)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.00)	(0.00)
13 FabPr	0.38 ***	0.38 ***	-0.05	0.01	-0.35 ***	-0.39 ***	0.08 ***	0.08 ***	0.13 ***	0.14 ***	-0.03	-0.04 *	0.04 ***	0.04 ***
	(0.02)	(0.05)	(0.07)	(0.05)	(0.07)	(0.09)	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
14 ElcEq	0.31 ***	0.31 ***	0.01	0.08	-0.27 **	-0.31	0.14 ***	0.19 ***	0.18 ***	0.22 ***	-0.05	-0.15 **	0.01 ***	0.01
	(0.01)	(0.07)	(0.07)	(0.08)	(0.11)	(0.24)	(0.05)	(0.04)	(0.03)	(0.07)	(0.05)	(0.07)	(0.00)	(0.01)
15 Autos	0.22 ***	0.22 ***	-0.04	-0.01	-0.16 ***	-0.15 **	0.11 **	0.16 **	0.06 ***	0.06 ***	-0.02	-0.02 *	0.01 **	0.01 **
	(0.01)	(0.05)	(0.06)	(0.03)	(0.05)	(0.06)	(0.05)	(0.06)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
16 Carry	0.47 ***	0.47 ***	-0.05	-0.02	-0.30 ***	-0.33 **	0.13 ***	0.16 ***	0.15 ***	0.16 ***	-0.03	-0.04	0.03 ***	0.03
	(0.01)	(0.05)	(0.06)	(0.04)	(0.09)	(0.14)	(0.04)	(0.03)	(0.02)	(0.04)	(0.03)	(0.03)	(0.01)	(0.02)
17 Mines	0.27 ***	0.27 ***	-0.11	-0.12	0.16 **	0.23 ***	0.09 ***	0.11 ***	0.10 ***	0.11 ***	-0.03	-0.04 **	0.02 ***	0.02 **
	(0.01)	(0.06)	(0.08)	(0.10)	(0.07)	(0.06)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.00)	(0.01)
18 Coal	0.11 ***	0.11 ***	-0.04 **	-0.04 ***	0.08 ***	0.09 **	-0.02	-0.01	0.02	0.02	-0.02 ***	-0.02 ***	0.00	0.00
	(0.01)	(0.03)	(0.02)	(0.01)	(0.02)	(0.03)	(0.04)	(0.06)	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)	(0.00)
19 Oil	0.26 ***	0.26 ***	-0.08	-0.00	-0.37 ***	-0.44 ***	0.14 ***	0.17 ***	0.09 ***	0.10 ***	-0.05	-0.10 ***	0.02 ***	0.02 ***
	(0.01)	(0.05)	(0.09)	(0.08)	(0.08)	(0.07)	(0.05)	(0.03)	(0.02)	(0.03)	(0.04)	(0.03)	(0.00)	(0.00)
20 Util	0.48 ***	0.48 ***	-0.08	0.35	-0.24	-0.20	0.19 *	0.38 ***	0.16 ***	0.18 ***	0.02	0.01	0.00	0.00
	(0.01)	(0.04)	(0.12)	(0.26)	(0.15)	(0.49)	(0.10)	(0.11)	(0.03)	(0.04)	(0.06)	(0.11)	(0.01)	(0.02)
21 Telcm	0.31 ***	0.31 ***	-0.04	0.04	-0.20 *	-0.17	0.18 **	0.33 ***	0.18 ***	0.22 ***	0.09 ***	0.11 ***	-0.00	-0.00
	(0.01)	(0.06)	(0.07)	(0.07)	(0.12)	(0.19)	(0.07)	(0.04)	(0.03)	(0.04)	(0.03)	(0.01)	(0.00)	(0.01)
22 Servs	0.28 ***	0.28 ***	-0.11	-0.08	-0.33 ***	-0.39 **	0.15 ***	0.21 ***	0.18 ***	0.21 ***	0.09 ***	0.12 ***	0.00	0.00
	(0.01)	(0.07)	(0.08)	(0.07)	(0.09)	(0.18)	(0.05)	(0.03)	(0.03)	(0.05)	(0.03)	(0.02)	(0.00)	(0.01)
23 BusEq	0.27 ***	0.27 ***	-0.01	0.03	-0.32 ***	-0.38 *	0.14 ***	0.19 ***	0.16 ***	0.17 ***	0.05	0.08	0.01 ***	0.01 **
	(0.01)	(0.07)	(0.05)	(0.04)	(0.10)	(0.19)	(0.05)	(0.03)	(0.02)	(0.05)	(0.04)	(0.05)	(0.00)	(0.01)
24 Paper	0.36 ***	0.36 ***	-0.01	0.04	-0.11	-0.04	0.13 ***	0.16 ***	0.16 ***	0.17 ***	-0.05 **	-0.06 ***	0.02 ***	0.02 **
	(0.01)	(0.04)	(0.06)	(0.06)	(0.11)	(0.22)	(0.04)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)
25 Trans	0.35 ***	0.36 ***	0.01	0.07	-0.30 ***	-0.31 ***	0.14 ***	0.18 ***	0.11 ***	0.11 ***	-0.04 **	-0.04 **	0.04 ***	0.04 ***
	(0.01)	(0.05)	(0.06)	(0.07)	(0.06)	(0.09)	(0.05)	(0.04)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
26 Whlsl	0.53 ***	0.53 ***	-0.02	0.07	-0.33 ***	-0.39 *	0.15 ***	0.19 ***	0.24 ***	0.29 ***	0.01	0.00	0.04 ***	0.04 ***
	(0.01)	(0.04)	(0.08)	(0.08)	(0.11)	(0.22)	(0.05)	(0.03)	(0.03)	(0.04)	(0.05)	(0.11)	(0.01)	(0.01)
27 Rtail	0.58 ***	0.59 ***	-0.04	0.02	-0.20 *	-0.17	0.16 ***	0.26 ***	0.19 ***	0.24 ***	0.09 **	0.15 ***	0.06 ***	0.06 ***
	(0.01)	(0.04)	(0.07)	(0.06)	(0.12)	(0.23)	(0.06)	(0.05)	(0.03)	(0.06)	(0.04)	(0.05)	(0.01)	(0.02)
28 Meals	0.55 ***	0.56 ***	-0.04	0.00	-0.19 *	-0.16	0.17 ***	0.25 ***	0.16 ***	0.18 ***	0.03	0.05	0.05 ***	0.05 ***
	(0.02)	(0.05)	(0.06)	(0.04)	(0.11)	(0.17)	(0.06)	(0.03)	(0.03)	(0.04)	(0.05)	(0.06)	(0.01)	(0.01)
29 Fin	0.33 ***	0.33 ***	0.03	0.10	-0.18	-0.11	0.21 ***	0.32 ***	0.26 ***	0.32 ***	0.10 **	0.14 ***	0.03 ***	0.03 ***
	(0.01)	(0.11)	(0.07)	(0.08)	(0.14)	(0.24)	(0.07)	(0.03)	(0.04)	(0.05)	(0.04)	(0.03)	(0.00)	(0.00)
30 Other	0.38 ***	0.38 ***	-0.03	0.00	-0.21 ***	-0.21 ***	0.13 **	0.18 ***	0.19 ***	0.25 ***	0.04 **	0.04 ***	-0.00	-0.00
	(0.01)	(0.04)	(0.06)	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.04)	(0.07)	(0.02)	(0.01)	(0.01)	(0.01)

Table 1.6
Cross-Regime Changes in Comovements of Inflation Expectations and Price-Earnings Ratios

This table reports the cross-regime changes in proportion of price-earnings variations explained by inflation expectations. Each column reports the posterior mean and standard deviation of the difference between β_{ik} in that regime and $\beta_{i,k-1}$ in the preceding regime ($\beta_{ik} - \beta_{i,k-1}$ in regression (1.11)). The table reports the results for the market only (top row), the average across the 30 industries (second row), and the average across both the market and 20 industries (third row). ***, **, * indicate the posterior mean is at least 2.58, 1.96, and 1.65 times larger than the posterior standard deviation. The sample is monthly from January 1976 to December 2020.

Regime	1978:11-1982:3	1982:4-1984:8	1984:9-1992:5	1992:6-1997:12	1998:1-2020:12
Market	-0.31 **	0.51 ***	0.05	-0.16 ***	-0.01
	(0.14)	(0.13)	(0.06)	(0.06)	(0.05)
Industries	-0.16 ***	0.35 ***	0.01	-0.13 ***	0.00
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Market + Industries	-0.17 ***	0.35 ***	0.02	-0.13 ***	0.00
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)

Appendix

1.A Alternative Price Ratio Decomposition

In this appendix, I discuss a different approach to decomposing price ratios that works directly with subjection expectations. Start with the ex-post one-year return identity

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t}, \tag{1.A.1}$$

where P_t is current stock price and D_t is realized rolling 12-month dividend.

Next take conditional subjective expectations (by analysts) of both sides of equation (1.A.1)

$$\begin{aligned} \mathbb{E}_t^* R_{t+1} &= \mathbb{E}_t^* \left[\frac{P_{t+1} + D_{t+1}}{P_t} \right] \\ &= \frac{\mathbb{E}_t^* P_{t+1} + \mathbb{E}_t^* D_{t+1}}{P_t} \\ R_{t+1}^* &= \frac{P_{t+1}^* + D_{t+1}^*}{P_t} \end{aligned} \tag{1.A.2}$$

where the second equality follows from the (very weak) assumption that analysts know the current market price, i.e., $\mathbb{E}_t^* P_t = P_t$, and the third equality defines $R_{t+j}^* \equiv \mathbb{E}_t^* R_{t+j}$ as subjective return expectations or discount rates, $P_{t+j}^* \equiv \mathbb{E}_t^* P_{t+j}$ as subjective price expectations or target prices, and $D_{t+j}^* \equiv \mathbb{E}_t^* D_{t+j}$ as subjective dividend expectations or dividend forecasts with $\mathbb{E}_t^* D_t = D_t$. Equation (1.A.2) can be rewritten as

$$R_{t+1}^* = \frac{P_{t+1}^* + D_{t+1}^*}{P_t} = \frac{\left(\frac{P_{t+1}^*}{D_{t+1}^*} + 1 \right) \frac{D_{t+1}^*}{D_t}}{\frac{P_t}{D_t}}, \tag{1.A.3}$$

Further define $E_{t+j}^* \equiv \mathbb{E}_t^* E_{t+j}$ as subjective earnings expectations or earnings forecasts with $\mathbb{E}_t^* E_t = E_t$, $pe_{t+j}^* \equiv \ln(P_{t+j}^*/E_{t+j}^*)$ as log target price-earnings forecasts, $pd_{t+j}^* \equiv \ln(P_{t+j}^*/D_{t+j}^*)$ as log target price-dividend forecasts, $de_{t+j}^* \equiv \ln(D_{t+j}^*/E_{t+j}^*)$ as log dividend-earnings forecasts (i.e., expected payout ratio).

Log-linearizing equation (1.A.3) around a long-term average of P^*/D^* and substituting $pe_t^* = pd_t^* + de_t^*$ give

$$pe_t = k + \Delta e_{t+1}^* - r_{t+1}^* + \rho \times pe_{t+1}^*, \tag{1.A.4}$$

where $\Delta e_{t+1}^* \equiv \ln(E_{t+1}^*/E_t)$ is log growth of earnings forecasts, $r_{t+1}^* \equiv \ln(R_{t+1}^*)$ is log discount rate, and k and

$$\rho \equiv \frac{P^*/D^*}{1 + P^*/D^*}$$

are constants. Assuming again the no bubble condition, i.e., $\lim_{j\to\infty} \rho^j pe_{t+j}^* = 0$, I can iterate equation (1.A.4) forward to obtain

$$pe_t = k + \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} \Delta e_{t+j}^*}_{\text{Cash flow } (CF_t)} - \underbrace{\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}^*}_{\text{Discount rate } (DR_t)}, \tag{1.A.5}$$

I thus decompose the market price-earnings ratio into a cash flow (CF) component (an infinite sum of log growths of earnings forecasts) and a discount rate (DR) component (an infinite sum of log subjective discount rates). Note that these definitions are similar to, but not the same as, those in equation (1.6) in Subsection 1.2.1. In equation (1.6), I define CF (DR) as the conditional subjective expectation of an infinite sum of log earnings growths (log discount rates). However, as discussed in footnote 6, I approximate the CF component in (1.6) as the infinite sum of log growths of earnings forecasts as defined in equation (1.A.5). Therefore, both price decomposition approaches give similar empirical results.

1.B Bayesian Break Model

This appendix describes the full paramaterization of the Bayesian break model.

1.B.1 Likelihood

Consider the panel regression:

$$y_{it} = \alpha_{ik} + \beta_{ik} \times x_{it} + e_{it}, \quad e_{it} \sim \mathrm{N}(0, \sigma_{ik}^2), \quad t = \tau_{k-1}+1, \ldots, \tau_k, \tag{1.B.1}$$

where the intercepts α_{ik}, slopes β_{ik}, and residual variances σ_{ik}^2 are allowed to shift across the unknown $K+1$ regimes. The regime k^{th} has length $l_k = \tau_k - \tau_{k-1}$ and consists of the observations $\tau_{k-1}+1, \ldots, \tau_k$.

Let $\theta_{ik} = (\alpha_{ik}, \beta_{ik})$ be a column vector, $\theta_i = \begin{bmatrix} \theta'_{i1} \\ \vdots \\ \theta'_{i,K+1} \end{bmatrix}$, and $\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_N \end{bmatrix}$, where N is the number of portfolios. Denote $\sigma_i^2 = (\sigma_{i1}^2, \ldots, \sigma_{i,K+1}^2)$ and $\sigma^2 = (\sigma_1^2, \ldots, \sigma_N^2)$. Let $y_{ik} = (y_{i,\tau_{k-1}+1}, y_{i,\tau_{k-1}+2}, \ldots, y_{i,\tau_k})$, $y_i = (y_{i1}, y_{i2}, \ldots, y_{i,K+1})$, and $y = (y_1, y_2, \ldots, y_N)$. Finally, let $X_{ik} = \begin{bmatrix} 1 & x_{i,\tau_{k-1}+1} \\ \vdots & \vdots \\ 1 & x_{i,\tau_k} \end{bmatrix}$, $X_i = \begin{bmatrix} X_{i1} \\ \vdots \\ X_{i,K+1} \end{bmatrix}$, and $X = \begin{bmatrix} X_i \\ \vdots \\ X_N \end{bmatrix}$. Given a set of break locations $\tau = (\tau_1, \ldots, \tau_K)$, the likelihood of y is

$$p(y|X, \theta, \sigma^2, \tau) = \left(\prod_{i=1}^{N} \prod_{k=1}^{K+1} (2\pi\sigma_{ik}^2)^{-l_k/2} \right) \exp\left[\sum_{i=1}^{N} \sum_{k=1}^{K+1} -\frac{1}{2\sigma_{ik}^2} (y_{ik} - X_{ik}\theta_{ik})' (y_{ik} - X_{ik}\theta_{ik}) \right] \tag{1.B.2}$$

1.B.2 Priors

Following Smith and Timmermann (2021a,b), I place a Poisson prior distribution over the regime durations:

$$p(l_k|\gamma_k) = \frac{\gamma_k^{l_k} e^{-\gamma_k}}{l_k!}, \quad k = 1, \ldots, K+1, \tag{1.B.3}$$

where the Poisson intensity parameter γ_k has a conjugate Gamma prior distribution:

$$p(\gamma_k) = \frac{d^c}{\Gamma(c)} \gamma_k^{c-1} e^{-d\gamma_k}, \quad k = 1, \ldots, K+1. \tag{1.B.4}$$

Following Smith and Timmermann (2021a), I set $c = 240$ and $d = 2$ to reflect the prior belief that breaks occur every 10 years.

Next, for regimes $k = 1, \ldots, K+1$ and portfolios $i = 1, \ldots, N$, I place an inverse Gamma prior on the residual variance:

$$p(\sigma_{ik}^2) = \frac{b^a}{\Gamma(a)}(\sigma_{ik}^2)^{-(a+1)} \exp(-\frac{b}{\sigma_{ik}^2}) \tag{1.B.5}$$

where I set $a = 2$ and $b = \hat{\sigma}_y^s$ which is the sample variance of y. This specification sets the prior mean of the residual variance equal to the sample variance.

Finally, conditional on σ_{ik}^2, the regression coefficients θ_{ik} follow a multivariate normal prior distribution:

$$p(\theta_{ik}|\sigma_{ik}^2) = (2\pi\sigma_{ik}^2)^{-\kappa/2}|V_\theta|^{-1/2} \exp\left[-\frac{1}{2\sigma_{ik}^2}(\theta_{ik} - \mu_\theta)' V_\theta^{-1}(\theta_{ik} - \mu_\theta)\right], \tag{1.B.6}$$

where $\kappa = 2$ is the number of covariates (including the intercept), $\mu_\theta = (\bar{y}, 0)$, and $V_\theta = \begin{bmatrix} 1/b & 0 \\ 0 & 1/b \end{bmatrix}$. The value of μ_θ reflects the prior belief that subjective expectations do not comove with price-earnings ratios and thus the intercept in the model is expected to equal to the sample mean of y. Furthermore, a prior zero mean on β_{ik} creates a shrinkage estimator that helps guard against spurious covariances between subjective expectations and price-earnings ratios which may arise in short regimes. Lastly, this specification of V_θ sets the prior variance of β_{ik} to equal one on average. A large prior variance serves two purposes. Fist, a large prior variance combined with a zero prior mean gives the model the freedom to recover either a positive or negative relation between subjective expectations and price ratios identified in the data. This setup emphasizes the role of the data, not the prior belief. Second, a large prior variance combats the failure of the model to account for autocorrelations in the residuals. This implies that in short regimes with limited data, the model would produce a large posterior variance, once again guarding against any spurious comovements that may show up.

1.B.3 Posterior Distribution

Multiplying the likelihood with the priors gives the following posterior distribution:

$$
\begin{aligned}
p(\theta, \sigma^2 | y, X, \tau) = & \left(\prod_{i=1}^{N} \prod_{k=1}^{K+1} (2\pi\sigma_{ik}^2)^{-l_k/2} \right) \exp\left[\sum_{i=1}^{N} \sum_{k=1}^{K+1} -\frac{1}{2\sigma_{ik}^2} \left(y_{ik} - X_{ik}\theta_{ik}\right)' \left(y_{ik} - X_{ik}\theta_{ik}\right) \right] \\
& \times \left(\prod_{i=1}^{N} \prod_{k=1}^{K+1} (2\pi\sigma_{ik}^2)^{-\kappa/2} |V_\theta|^{-1/2} \right) \exp\left[\sum_{i=1}^{N} \sum_{k=1}^{K+1} -\frac{1}{2\sigma_{ik}^2} (\theta_{ik} - \mu_\theta)' V_\theta^{-1} (\theta_{ik} - \mu_\theta) \right] \\
& \times \left(\prod_{i=1}^{N} \prod_{k=1}^{K+1} \frac{b^a}{\Gamma(a)} (\sigma_{ik}^2)^{-(a+1)} \right) \exp\left[\sum_{i=1}^{N} \sum_{k=1}^{K+1} -\frac{b}{\sigma_{ik}^2} \right] \qquad (1.B.7)
\end{aligned}
$$

From this posterior distribution, by rearranging the terms, I can derive the following posterior conditional distributions:

$$\sigma_{ik}^2 | y_{ik}, X_{ik} \sim \mathrm{IG}(a_k, b_{ik}) \quad \text{and} \qquad (1.B.8)$$

$$\beta_{ik} | \sigma_{ik}^2, y_{ik}, X_{ik} \sim N(\mu_{ik}, \sigma_{ik}^2 \Sigma_{ik}) \qquad (1.B.9)$$

for $k = 1, \ldots, K+1$ and $i = 1, \ldots, N$, where

$$\Sigma_{ik}^{-1} = V_\theta^{-1} + X_{ik}' X_{ik}, \qquad (1.B.10)$$

$$\mu_{ik} = \Sigma_{ik} \left(V_\theta^{-1} \mu_\theta + X_{ik}' y_{ik} \right), \qquad (1.B.11)$$

$$a_k = a + l_k/2, \quad \text{and} \qquad (1.B.12)$$

$$b_{ik} = \frac{1}{2} \left(2b + y_{ik}' y_{ik} - \mu_{ik}' \Sigma_{ik}^{-1} \mu_{ik} + \mu_\theta' \Sigma_\theta^{-1} \mu_\theta \right). \qquad (1.B.13)$$

It is then straightforward to use Gibbs sampler to draw from these conditionals. In implementing the Gibbs sampler, I sequentially make 5000 draws from the conditionals, discard the first 500 draws as burn-in, and keep every other draw as the final sample. Increasing the number of draws to 10000 or lowering it to 3000 does not change the results.

1.B.4 Break Locations

Smith and Timmermann (2021a) show that for a given set of break locations τ, the parameters θ and σ^2 can be integrated out from the posterior (1.B.7) to achieve the following marginal likelihood of the data:

$$p(y|X,\tau) = \prod_{i=1}^{N}\prod_{k=1}^{K+1}(2\pi)^{-l_k/2}\frac{b^a}{\Gamma(a)}\frac{\Gamma(a_k)}{b_{ik}^{a_k}}\frac{|\Sigma_{ik}|^{1/2}}{|V_\theta|^{1/2}} \tag{1.B.14}$$

Each set of break locations τ can be considered a Bayesian model. Smith and Timmermann (2021a) develop a reversible jump MCMC procedure that jumps around models with different locations of breaks and numbers of breaks and each time selects a new model using an acceptance ratio based on the marginal likelihood (1.B.14) of each model. This procedure is implemented until the set of break locations τ stabilizes. Please see Appendix D in Smith and Timmermann (2021a) for details.

1.C Sample Construction

This appendix provides more details on sample construction. As I construct 30 industry portfolios and all of the ratios used in this paper are in logs, it will be problematic if the industry-aggregated realized or forecast earnings are negative or close to zero. To combat this issue, I follow Vuolteenaho (2002), Lochstoer and Tetlock (2020), and others to convert each real company into a pseudo-company by buying 90% of that company's market value and investing the remaining 10% market value in Treasury bill. Specifically, the pseudo total realized earnings over the past 12 months of company j in period t is

$$TE_{jt}^{p} = 0.9 \times TE_{jt} + 0.1 \times P_{i,t-1} \times shrout_{i,t-1} \times tb_{t-1\rightarrow t}, \tag{1.C.1}$$

where TE_{jt}^{p} is the total pseudo-earnings, TE_{jt} is the total (*not* per share) realized rolling 12-month earnings before extraordinary items reported by Compustat, $P_{i,t-1}$ and $shrout_{i,t-1}$ are the price and shares outstanding 12 months (one year, t here indicates year) ago, and $tb_{t-1\rightarrow t}$ is the Treasury bill rate over the past 12 months. Similarly, the pseudo total earnings forecast for the next 12 months and pseudo target market value are

$$TE^{*p}_{j,t+1} = 0.9 \times E^{*}_{j,t+1} \times shrout_{jt} + 0.1 \times P_{jt} \times shrout_{jt} \times tb_{t\to t+1} \quad \text{and,} \tag{1.C.2}$$

where $E^{*}_{j,t+1} \equiv \mathbb{E}^{*}_{t} E_{j,t+1}$ is the analyst earnings-per-share forecast for the next 12 months. It is clear that the pseudo *realized* earnings include the income from the risk-free investment based on the firm's market value 12 months ago while the risk-free investment for the pseudo *forecast* earnings are based on the current market value. Hence, the pseudo portfolio is rebalanced every 12 months, similar to the construct in Vuolteenaho (2002).

For the pseudo total earnings forecast over months 13 to 24, I assume that the investor expects the current annual Treasury bill rate to be constant for the next 2 years, so I have

$$TE^{*p}_{j,t+2} = 0.9 \times E^{*}_{j,t+2} \times shrout_{jt} + 0.1 \times P_{jt} \times shrout_{jt} \times tb_{t\to t+1}, \tag{1.C.3}$$

where $E^{*}_{j,t+2} \equiv \mathbb{E}^{*}_{t} E_{j,t+2}$ is the analyst earnings-per-share forecast for months 13 to 24 ahead. Overall, the intuition behind the creation of the pseudo-company is that if the investor buys 100% of the market value of one company, he will receive 100% of the realized 12-month earnings and expect 100% earnings forecasts. In this case, because he buys only 90% of that firm's market value and invests the remaining 10% in Treasury bill, only 90% of realized earnings and forecast earnings come from that firm while the remaining 10% comes from the risk-free investment.

After creating the pseudo firms, to further minimize the impact of outliers and maximize the number of observations available for analyses, every month, I cross-sectionally winsorize each of these pseudo variables at their 5^{th} and 95^{th} percentiles before sorting them into industry portfolios. The results are robust to small changes in the 10% risk-free investment threshold or in the winsorization percentiles.

Unlike Bordalo et al. (2020a) and De La O and Myers (2021, 2022) who focus on the S&P 500 stocks, I retain all stocks with analyst forecasts to create the market portfolio. To implement the panel break method, I combine the market portfolio with the 30 industry portfolios based on the Fama-French industry classification. Smith and Timmermann (2021a,b) also use these 30 industry portfolios in their analyses. For portfolio i at time t, I construct the following variables

$$pe_{it} = \ln\left(\frac{\sum_{j=1}^{n_{it}} P_{jt} \times shrout_{jt}}{\sum_{j=1}^{n_{it}} TE_{jt}^{p}}\right), \tag{1.C.4}$$

$$E_t^* \Delta e_{i,t+1} \approx \Delta e_{i,t+1}^* = \ln\left(\frac{\sum_{j=1}^{n_{it}} TE_{j,t+1}^{*p}}{\sum_{j=1}^{n_{it}} TE_{jt}^{p}}\right), \tag{1.C.5}$$

$$E_t^* \Delta e_{i,t+2} \approx \Delta e_{i,t+2}^* = \ln\left(\frac{\sum_{j=1}^{n_{it}} TE_{j,t+2}^{*p}}{\sum_{j=1}^{n_{it}} TE_{j,t+1}^{*p}}\right), \tag{1.C.6}$$

$$CF_{it} = \frac{1}{1-\rho\phi_{ie}} E_t^* \Delta e_{i,t+1}, \quad \text{and} \tag{1.C.7}$$

$$DR_{it} = pe_{it} - CF_{it}, \tag{1.C.8}$$

which are the portfolio-level log price-earnings ratio, subjective expectation of earnings growth over the next 12 months, subjective expectation of earnings growth over months 13 to 24, subjective cash flow component, and subjective discount rate component, respectively. Note that although all variables are computed over a 12-month window, the frequency of the data is monthly.

Because analysts do not make long-term forecasts in the earlier periods of the sample and I use interpolation to compute earnings forecasts over the next 12 months and months 13 to 24, the number of companies with earnings forecasts for the second 12 months is substantially lower than that of the first 12 month forecasts. The data for the first 12 month earnings forecasts are available from January 1976 while the monthly number of firms with the second 12 months forecasts is not relatively large until January 1985. To resolve these data issues, I create two separate datasets.

The first one consists of companies having the earnings forecasts for the second 12 months and starts from January 1985. I use this sample to estimate the coefficients of the expected earnings growth decay process for each portfolio created from this dataset. Table C1 reports the coefficient estimates for these decay processes. The second dataset consists of all firms with the earnings forecasts for the next 12 months. This dataset begins from January 1976 and is the main sample of the paper. I apply the decay coefficients estimated for each portfolio using the first dataset to this one to construct the cash flow components. Both samples end at December 2020.

Figure C1 plots the monthly number of firms having the earnings forecasts for the next 12 months that go into the final empirical analyses. On average, the monthly number of firms having 12-month ahead earnings forecasts is about 2000 with a peak in the late 1990s.

The *pe* of the market portfolio in this paper has an 85% correlation with the *pe* on the S&P 500 computed from data on Professor Shiller's website. There are three reasons why the correlation is not higher. First, my *pe* is computed for all stocks having the analyst forecasts while Shiller's *pe* focuses only on the S&P 500 stocks. Seconds, my *pe* is computed from pseudo-companies as discussed above. Third, I use earnings release dates to match realized earnings with market data and keep the realized earnings constant between two release dates as my paper focuses on the relation between real-time price-earnings ratios and earnings forecasts. In contrast, Shiller computes quarterly earnings and interpolate between two quarters to have monthly values.

Figure C1. Time Series of Number of Firms having Analyst Forecasts

This figure plots the monthly number of firms having earnings forecasts for the next 12 months that go into the final empirical analyses. The sample is from January 1976 to December 2020.

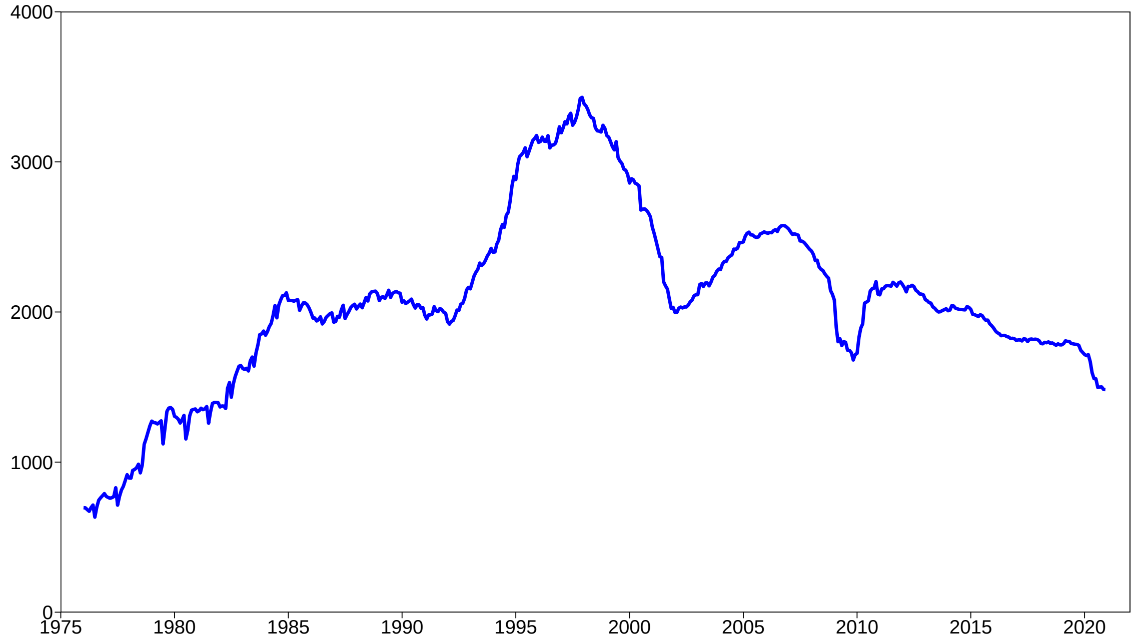


Table C1
Coefficient of Earnings Growth Expectation Decay Process

This table reports the decay coefficients for the market and 30 industry portfolios using the following regression:

$$\mathbb{E}_t^*\left[\Delta e_{i,t+1+j}\right] - \mu_{ie} = \phi_{ie}^j\left(\mathbb{E}_t^*\left[\Delta e_{i,t+1}\right] - \mu_{ie}\right) + \epsilon_{t,j}^e,$$

where $\mathbb{E}_t^*\left[\Delta e_{t+1+j}\right]$ is the subjective expectation of earnings growth over period $t+j$ for portfolio i. The sample is from January 1985 to December 2020. t-statistics are computed using the Newey and West (1987) standard errors with 6 lags. ***, **, * indicate significance at the 1%, 5%, and 10%, respectively.

		ϕ	t
0	Market	0.09 ***	(5.97)
1	Food	0.04 **	(2.21)
2	Beer	0.06 ***	(5.19)
3	Smoke	0.02	(1.40)
4	Games	-0.04	(-0.58)
5	Books	0.00	(0.03)
6	Hshld	-0.02	(-1.02)
7	Clths	0.05 ***	(3.37)
8	Hlth	0.14 ***	(7.38)
9	Chems	0.11 ***	(6.51)
10	Txtls	0.04 **	(2.40)
11	Cnstr	0.18 ***	(4.80)
12	Steel	0.06	(1.62)
13	FabPr	0.14 ***	(6.01)
14	ElcEq	0.02	(0.43)
15	Autos	0.11 ***	(3.82)
16	Carry	0.07 *	(1.85)
17	Mines	0.03 *	(1.91)
18	Coal	0.05	(1.32)
19	Oil	0.07 **	(2.27)
20	Util	0.04 *	(1.68)
21	Telcm	0.04 *	(1.76)
22	Servs	0.06 ***	(3.57)
23	BusEq	0.05 **	(1.98)
24	Paper	0.17 ***	(4.32)
25	Trans	0.06 ***	(2.64)
26	Whlsl	0.04 *	(1.78)
27	Rtail	0.04 **	(2.30)
28	Meals	0.05	(1.45)
29	Fin	0.10 ***	(11.68)
30	Other	0.02 *	(1.78)

Chapter 2

Investor Sentiment and Asset Returns: Actions Speak Louder than Words

2.1 Introduction

Sentiment has been shown to play an essential role in explaining asset returns (Baker and Wurgler 2006, 2007; Stambaugh et al. 2012) and in predicting the equity risk premium (Tetlock 2007; Huang et al. 2015; Jiang et al. 2019). The finance literature has introduced a plethora of investor sentiment measures ranging from those constructed from market-based information (Baker and Wurgler 2006; Zhu and Zhou 2009; Neely et al. 2014) to those from news and social media and surveys (Tetlock 2007; Garcia 2013; Calomiris and Mamaysky 2019; Obaid and Pukthuanthong 2022; Greenwood and Shleifer 2014). Numerous research has explored these sentiment sources individually. Still, no one has compared and contrasted them to my knowledge.

In this study, I compare the return predictability of trade-based sentiment (technical indicators) versus text-based sentiment (news and social media) across four major asset classes (Bitcoin, stocks, Treasury bonds (T-Bond), and gold). These two types of sentiment measures are readily available daily and are widely utilized by market participants, but they are fundamentally distinct. While news and social media are publicly available and reflect people's attention and beliefs (Shiller 2005), trading information such as prices and volumes indicate investors' trading activity and decisions. By comparing these two sources of information, I can identify which type of information is most useful in predicting market movements. Second, both sentiment types may be more relevant to different types of market participants. Trade sentiment may be more relevant to institutional investors with access to trade data and can execute trades based on that data. In contrast, text sentiment may be more relevant to individual investors who rely on news articles and social media for market information. By comparing these two types of sentiment, I shed light on which types of market participants are most likely to benefit from each type of sentiment data.

Third, trade and text sentiment may be more helpful in predicting market movements over different time horizons. Trade sentiment, which is based on actual trades, may be more beneficial for short-term predictions of market movements. In contrast, text sentiment may be more beneficial for longer-term predictions based on trends in market sentiment. By comparing these two types of sentiment across different time intervals, I may determine which sort of sentiment data is more valuable for specific time horizons. Different types of investors may have different needs. Real-time mean-variance investors may be more interested in short-term predictions, while longer-term investors may be more interested in longer-term forecasts. Therefore, comparing sentiment measures across different types of investors and investment horizons can be important.

Notably, I apply daily sentiment data to construct my trading strategy. Daily return is more challenging to predict than monthly returns. Fama and French (1988), Campbell and Thompson (2008), and Cochrane (2008) show that stocks are more likely to be predictable over long horizons. Cochrane (2009) shows the expected asset returns are a function of risk aversion, risk (consumption volatility), and the correlation between asset returns and the stochastic discount factor. He argues that these components are more likely to alter at a long horizon than on a short horizon like daily. In other words, asset returns are more likely to be predictable over long horizons. Since asset returns are more challenging to forecast daily, it is more complicated and crucial to developing profitable trading strategies daily. My asset allocation results support this conjecture.

For the text-based sentiment, I use the sentiment measures constructed by Refinitiv

MarketPsych Indices (RMI) based on thousands of news and social media sources. My measures are the most comprehensive text-based indexes to date. RMI provides four sentiment series for each asset class from different text sources: news articles, headlines, social media, and news and social media combined. I further consider each series's moving averages over various lag windows. In total, I examine 28 text-based sentiment indexes for each asset.

Regarding sentiment metrics based on trading activity, I create six measures utilizing prices and trading volumes. The first is price-based WRS (William's %R), created in the 1960s and widely used in technical analysis to measure overbought/oversold (high/low sentiment) conditions. The second is NHS (nearness to a recent high) introduced by Li and Yu (2012), who finds that closeness to recent high proxies for the degree of underreaction by stock traders and thus positively predicts future returns. The third component is the trading volume ratio (TVS), one of the six components of the well-known sentiment index developed by Baker and Wurgler (2006)'s, and the only one with daily data available. The following three measures are comparable to those presented by Neely et al. (2014): MAS (a moving average rule), MOM (a momentum-type measure), and OBV (a combination of both trading volumes and prices). I evaluate different specifications based on distinct lag windows for these six trade sentiment metrics. There is no theoretical guidance regarding which construction window produces the most remarkable prediction performance. Therefore, I evaluate 28 indices of trade sentiment for each asset.

Except for Bitcoin, I utilize ETFs to represent the remaining three asset classes since I need daily prices and trading volumes for each asset to create my trade sentiment measures.[1] Using ETFs ensures that real-time investors can utilize my forecasting practices. In addition, Bitcoin is chosen to represent all cryptocurrencies because it is the largest and most popular cryptocurrency and has the most extended available

[1] See Section 2.2 for details on the ETF list.

time series data. I aim to compare the return predictability of 28 trade sentiment metrics to 28 text variables for each asset.

Given these many predictors, I apply four prevalent dimension reduction techniques for parsimonious comparison. Utilizing information aggregation techniques decreases overfitting and improves out-of-sample return projections. The first and most straightforward technique is the simple average of all variables (AV) (Huang and Lee 2010; Dong et al. 2022). The second technique is the combination forecast (CF) (Rapach et al. 2010). Under this methodology, individual projections are averaged into a single forecast. The third is principal component analysis (PCA), a frequently employed technique for extracting common variations from a group of variables. My final technique is partial least squares (PLS), which extracts signals from a set of variables most closely associated with a prediction target. PLS has gained prominence in the literature on time series prediction (Kelly and Pruitt 2013, 2015; Huang et al. 2015). Except for CF, which is only used for out-of-sample return predictions, the other three approaches are applied, both in-sample and out-of-sample.

In my in-sample tests, trade sentiment outperforms text sentiment across prediction approaches and forecasting horizons for the four assets: Bitcoin, stocks, T-Bond, and gold. In addition, contrary to the recent literature demonstrating the strong return predictability of news sentiment (Tetlock 2007; Garcia 2013), I find that text sentiment from news, social media, and their combination have almost no predictive power over the next day's returns for all assets. Even with the use of information aggregation, the prediction of text sentiment remains dismal. I show that the inconsistent evidence is a result of the sample period. After the 2000s, for most of my sample period, text sentiment loses its significance.

I also demonstrate that sentiment has consistent predictions across asset classes. In particular, trade sentiment positively predicts future returns across asset classes. In other words, investors underreact to trade sentiment. Text sentiment presents another positive predictor than trade sentiment, but the power is much weaker.

My out-of-sample analysis considers whether the sentiment measures can forecast real-time asset returns. I use two commonly employed out-of-sample tests, the first of which is the out-of-sample R^2 (Campbell and Thompson 2008). As a benchmark, this metric compares the recursive return predictions generated by a predictor utilizing only information known up to each time point to the historical mean return. A positive out-of-sample R^2 indicates that the predictor variable has outperformed the historical mean benchmark. I also apply the forecast-encompassing test, a more direct comparison between trade and text sentiment (Harvey et al. 1998). This test examines the null hypothesis that a model i's forecast encompasses a competing model j's forecast. A failure to reject the null implies that the model j's forecast is not more informative than the model i's.

Under these two out-of-sample tests, trade sentiment generates more accurate return forecasts for Bitcoin and T-bond. Unsurprisingly, Bitcoin sentiment produces the most stable and robust out-of-sample results. Bitcoin is a highly speculative and difficult-to-value asset, and trading sentiment has been demonstrated to be highly predictive of such assets in the stock market (Baker and Wurgler 2006; Stambaugh et al. 2012). However, no sentiment metric surpasses the historical mean benchmark out-of-sample in forecasting T-bonds, stocks, and gold. Forecast encompassing tests present similar results. Trade sentiment encompasses text sentiment clearly in both short and long terms for Bitcoin, but only for long term for T-bond. There is no clear pattern for stocks and gold.

Finally, I investigate whether the daily return estimates generated by sentiment metrics can provide economic value to real-time mean-variance investors. Strikingly, Bitcoin is the only asset whose trade sentiment can achieve remarkable economic returns and a higher Sharpe ratio than a buy-and-hold investment strategy. Both trading and text sentiment cannot surpass the historical mean benchmark or the simple buy-and-hold strategy for the remaining assets. This finding demonstrates that converting technical indicators or media opinions into profitable trading strategies for stocks, bonds, and gold is intricate. Refinitiv MarketPsych Indices (RMI) is the largest and most influential sentiment provider together with Ravenspack. Despite their seemingly wide use by investors, I find little value in using RMI's sentiment measures to create profitable daily trading strategies. This is the implication of my finding.

My paper offers two significant contributions to the finance literature. First, I respond to Zhou (2018) by conducting the first exhaustive analysis on the comparative return predictability of trade and text sentiment. While the enormous literature on investor sentiment has studied trade and text sentiment individually, my study combines several sentiment sources and fills a gap in the literature. In contrast to the existing studies that primarily focus on equities, I analyze four distinct asset types. I find that sentiment influences on future returns vary across asset classes. However, regarding return predictability, technical indicators outperform news and social media tone. This shows that text-based sentiment appears too noisy, diminishing its ability to anticipate future returns. My findings extend the Cheap Talk model proposed by Stein (1989), Farrell (1995), and Farrell and Rabin (1996), which is mainly applied to the Fed announcements. I demonstrate that talk is cheap; trading sentiment outperforms text sentiment in both in-sample and out-of-sample.

I also contribute to the growing literature on cryptocurrencies and Bitcoin in particular. Financial economists are still attempting to ascertain the worth of digital assets. Theoretically, as demonstrated by Dwyer (2015), Athey et al. (2016), and Pagnotta and Buraschi (2018), the value of cryptocurrencies depends on a combination of usage, degree of acceptance, scarcity, and the importance of anonymity. However, empirically assessing digital assets is extremely difficult because fundamental metrics such as cash flows to stocks are missing. Since cryptocurrencies are more difficult to value than stocks, investor sentiment can explain returns on stocks that are difficult to value and costly to arbitrage (see, for example, Baker and Wurgler (2006), and Zhou (2018) and references therein); cryptocurrencies appear more susceptible to sentiment than stocks. Recent research by Detzel et al. (2021) examines the prediction of Bitcoin using technical indicators. I extend their study by comparing trade sentiment to text sentiment and demonstrate that trading sentiment is significantly more predictive of future Bitcoin returns than text sentiment. Moreover, real-time investors can use Bitcoin trade sentiment to achieve substantial economic benefits.

2.2 Data

2.2.1 Market Data

In this paper, I study the return predictability of sentiment across four asset classes. The first one is Bitcoin. I collect daily Bitcoin prices and trading volumes from bitcoinity.org.[2] Unlike stocks, Bitcoin is traded every day, and thus the prices and volumes are available seven days a week. As of October 2022, Bitcoin has a market capitalization of $391B.

Because I need both prices and trading volumes to construct my trade sentiment measures and investigate the trading strategies of all assets based on their sentiment signals, I use ETF trading data on the remaining three asset classes:

- Stocks: SPY, the world's first and largest ETF tracking the S&P 500 index since 01/22/1993 (having net asset value (NAV) of $326B as of October 2022);
- Long-term Treasury bonds: TLT, an iShares Treasury Bond ETF, tracks the investment results of an index composed of U.S. Treasury bonds with remaining maturities greater than twenty years since 07/22/2002 (having a NAV of $24B

[2]bitcoinity.org aggregates Bitcoin trading information from all major exchanges. The data can be downloaded here: https://data.bitcoinity.org/markets/price_volume/all/USD?t=lb&vu=curr.

as of October 2022); and

- Gold: GLD, an SPDR gold ETF tracking the price of gold since 11/18/2004 (having a NAV of $50B as of October 2022).

2.2.2 Text-Based Sentiment Data

My text-based sentiment measures are from Refinitiv MarketPsych Indices (RMI) managed under the umbrella of Refinitiv. The indices are constructed from news and social media content and identify specific sentiment, macroeconomic, and general buzz-related words relevant to the entity. Subsequently, the volume and tone of phrases and words are converted into measurable variables. In short, RMI sentiment measures are based on various news media and are computed from overall positive references net of negative references of each asset class. Michaelides et al. (2019) and Michaelides et al. (2015) apply RMI.[3] For each asset, I have four RMI sentiment measures constructed from different content sources: whole news articles (hereafter News), only news headlines (NewsHL), social media (Social), and news and social media combined (NewsSoc).

RMI indexes are updated either hourly or daily at 3:30 p.m. EST using data from the past 24 hours as their construction period. Because Bitcoin is traded continuously, the daily data for Bitcoin reports the price at midnight UTC as its closing price. Consistent with this convention, I use the RMI indexes for Bitcoin updated at midnight UTC. Specifically, my prediction analyses use Bitcoin sentiment measures computed from midnight UTC on the day t to midnight UTC on t+1 to forecast Bitcoin returns measured from midnight UTC on the day $t + 1$ onward.

I use the default RMI sentiment measures for all other assets, updated daily at 3:30 p.m. EST. According to RMI, this timestamp is selected so investors have enough time to update their investments based on the sentiment signals before the NYSE

[3]See Section 2.B for details on the construction of text sentiment.

market closes. During the weekends or national holidays when the stock market is closed, I average the sentiment scores over these days with the score on the day when the market is reopened (i.e., the forward average). For example, sentiment scores over Saturday and Sunday are combined with those on Monday to predict returns on Tuesday and beyond. I adopt this timing choice to ensure that my sentiment measures do not overlap with returns, which is especially important in out-of-sample analyses. Continuing with the above example, if sentiment scores on Saturday and Sunday are combined with sentiment scores on Friday (i.e., backward average) to predict Monday returns computed from closing prices on Monday and last Friday, sentiment and return are overlapped, resulting in information leakage in out-of-sample prediction.

It should be noted that only English-language texts were used for constructing RMI sentiment measures before February 2020, which might cause bias in countries where English is not used as an official language and articles are not primarily written with it. However, this bias is partly alleviated because even in non-English countries, the most critical business, finance, and economics-related news are often published in English. Moreover, informed traders will likely post significant news and events internationally in English. Since February 2020, RMI's real-time translation and analysis engine has included Arabic, Chinese, Japanese, and Portuguese-language news sources. In January 2021, Dutch, French, German, Indonesian, Italian, Korean, and Spanish language sources were added.

2.2.3 Trade-Based Sentiment Construction

Based on trading data, I construct sentiment measures analogous to the most famous investor sentiment index of Baker and Wurgler (2006, 2007). In constructing their index, Baker and Wurgler (2006) use six proxies that are unavailable daily for my assets. Consistent with their study, I define TVS, one of my six sentiment measures, based on trading volume. Mathematically, the log trading volume ratio over the past

L days is computed as

$$\mathrm{TVS}_t^i(L) = \log\left(\frac{\mathrm{TV}_t^i}{\mathrm{TV}_{t-L+1}^i}\right) \tag{2.1}$$

where TV_t^i is the trading volume of asset i on day t. A higher trading volume indicates greater sentiment.

The second trade sentiment measure is based on the most famous over-bought and over-sold technical analysis indicator, Williams' %R (WRS). The indicator was developed by Larry Williams in the 1960s and was popularized by his book (Williams 1976). Mathematically, I define

$$\mathrm{WRS}_t^i(L) = \frac{P_{max,t}^i(L) - P_t^i}{P_{max,t}^i(L) - P_{min,t}^i(L)} \tag{2.2}$$

where $P_{max,t}^i(L)$ and $P_{min,t}^i(L)$ are the highest and lowest daily prices of asset i over the window from day $t-L+1$ to day t. Intuitively, if $\mathrm{WRS}_t^i(L)$ is less than 20%, asset i is regarded as over-bought (high-sentiment) as the trading price, P_t^i is closest to its highest price. Likewise, when $\mathrm{WRS}_t^i(L)$ is greater than 80%, asset i is regarded as an over-sell (low-sentiment). The measure is very obvious when the asset price oscillates around a certain price level.

As a low value of $\mathrm{WRS}_t^i(L)$ indicates high sentiment, it has the opposite sign compared to all other sentiment measures. To remain consistent and streamline my empirical analyses, I define a new $\mathrm{WRS}_t^i(L)$ as the negative of $\mathrm{WRS}_t^i(L)$ in (2.2).

My third sentiment measure is based on only the current and max prices,

$$\mathrm{NHS}_t^i(L) = \frac{P_t^i}{P_{max,t}^i(L)} \tag{2.3}$$

Li and Yu (2012) interprets it as the nearness to the recent high and seems the first to use it in the stock market. It is clear that $\mathrm{NHS}_t^i(L)$ is closely related to $\mathrm{WRS}_t^i(L)$. However, they are different. $\mathrm{WRS}_t^i(L)$ depends on the minimum price and measures the incremental price move from the minimum. In contrast, $\mathrm{NHS}_t^i(L)$ is independent of the minimum price and measures the closeness of P_t^i to $P_{max,t}^i(L)$.

Note that no theories or economic intuitions guide the choice of L. As with studies in technical analysis, such as Brock et al. (1992), I consider several plausible choices, with $L = 5, 10, 20, 50, 100$, respectively. While it will then be difficult to select which L works the best in the future, the data can tell whether all the sentiment measures collectively can forecast returns out-of-sample. In my empirical results, I denote the $\mathrm{WRS}_t^i(L)$ and $\mathrm{NHS}_t^i(L)$ computed with a specific value of L as WRSL and NHSL, respectively.

My subsequent three sentiment measures are those used in Neely et al. (2014). The first is a moving average (MA) rule that generates a buy or sell signal ($\mathrm{MAS}_t^i = 1$ or $\mathrm{MAS}_t^i = 0$, respectively) for asset i on day t by comparing two moving averages:

$$\mathrm{MAS}_t^i(s,l) = \begin{cases} 1 & \text{if} \quad \mathrm{MA}_{s,t}^i \geq \mathrm{MA}_{l,t}^i, \\ 0 & \text{if} \quad \mathrm{MA}_{s,t}^i < \mathrm{MA}_{l,t}^i \end{cases} \tag{2.4}$$

where

$$\mathrm{MA}_{j,t}^i = \frac{1}{j}\sum_{k=0}^{j-1} P_{t-k}^i \quad \text{for} \quad j = s, l;$$

P_t^i is the price index of asset i on day t, and $s(l)$ is the length of the short (long) MA ($s < l$). Intuitively, the MA rule detects changes in price trends because the short MA will be more sensitive to recent price movements than the long MA. In my study, I consider daily MA rules with $s = 10, 20$ and $l = 50, 100$ so that I have four measures of MAS for each asset i. My results denote a MAS constructed with a specific value of s and l as MASs_l.

My fifth measure is a momentum-based strategy. A simple momentum rule generates the following signal:

$$\mathrm{MOM}_t^i(L) = \begin{cases} 1 & \text{if} \quad P_t \geq P_{t-L+1}, \\ 0 & \text{if} \quad P_t \geq P_{t-L+1}. \end{cases} \tag{2.5}$$

Intuitively, a current price higher than level L days ago indicates “positive” momentum, generating a buy signal. In this study, for each asset i, I compute MOM_t^i for $L = 5, 10, 20, 50, 100$ and denote them as MOML.

My final trade sentiment measure combines prices with trading volumes. More specifically, I first define

$$V_t^i = \sum_{k=1}^{t} \text{VOL}_k^i \text{D}_k^i, \tag{2.6}$$

where VOL_k^i is the trading volume of asset i on day k and D_k^i is an indicator equal to 1 if $P_k - P_{k-1} \geq 0$ and -1 otherwise. I then form a trading signal from V_t^i as follows

$$\text{OBV}_t^i(s,l) = \begin{cases} 1 & \text{if} \quad \text{MA}_{s,t}^{V,i} \geq \text{MA}_{l,t}^{V,i}, \\ 0 & \text{if} \quad \text{MA}_{s,t}^{V,i} < \text{MA}_{l,t}^{V,i} \end{cases} \tag{2.7}$$

where

$$\text{MA}_{j,t}^{V,i} = \frac{1}{j} \sum_{k=0}^{j-1} V_{t-k}^i \quad \text{for} \quad j = s, l;$$

Intuitively, relatively high volume and recent price increases indicate a strongly positive market trend and generate a buy signal. Similar to MAS, I consider $s = 10, 20$ and $l = 50, 100$ for OBV and denote them as OBVs_l.

2.2.4 Text-Based Sentiment Construction

Because I have 28 trade sentiment measures in total, I increase the number of RMI sentiment measures by computing various moving averages for each sentiment measure to provide a fair playground between trade and text sentiment. Specifically,

$$x_t^i(L) = \frac{1}{L} \sum_{k=0}^{L-1} x_{t-k}^i,$$

where x_t^i is one of News, NewsHL, Social, and NewsSoc sentiment measures for asset i on the day t, like trade sentiment measures, I consider $L = 5, 7, 10, 20, 50, 100$ and denote them as News$L$, NewsHLL, Social$L$, and NewsSoc$L$ for a specific choice of L in my empirical results. Thus, I have 28 text-based sentiment measures for each asset

i.

Considering the availability of both trade and text sentiment measures and using the first 100 days to construct various sentiment measures ($L = 100$ in my setups), my final sample period for stocks, bonds, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively, to September 2022.[4]

2.2.5 Sentiment Orthogonalization

Since my goal is to compare the true return predictability of sentiment measured from textual and trading information, I want to make sure other market forces do not impact the return predictability. To accomplish this goal, I follow the literature (Garcia 2013; Huang et al. 2015) to orthogonalize my sentiment measures before using them in the forecasting exercise. Specifically, following Garcia (2013), I use the residuals in the following regression as a measure of my orthogonalized sentiment:

$$\begin{aligned} x_t^i =& \big(1 - D_t\big)\big(\boldsymbol{\beta_1 L_{0-5}}(r_t^i) + \boldsymbol{\gamma_1 L_{0-5}}(r_t^{2,i}) + \boldsymbol{\delta_1 L_{1-5}}(x_t^i)\big) \\ & D_t\big(\boldsymbol{\beta_2 L_{0-5}}(r_t^i) + \boldsymbol{\gamma_2 L_{0-5}}(r_t^{2,i}) + \boldsymbol{\delta_2 L_{1-5}}(x_t^i)\big) + \boldsymbol{\eta} Z_t + \epsilon_t^i \end{aligned} \tag{2.8}$$

where x_t^i is a sentiment of asset i on day t, $\boldsymbol{L_{0-5}}(r_t^i)$ are returns of asset i on days t to $t-5$, $\boldsymbol{L_{0-5}}(r_t^{2,i})$ are squared returns of asset i on days t to $t-5$, $\boldsymbol{L_{1-5}}(x_t^i)$ are five lags of sentiment x_t^i, D_t is a dummy variable equal to 1 (0) if day t belongs to a recession (expansion) as defined by NBER, and Z_t is a list of control variables including a constant and weekday indicators.

Except in Table 2.1 where I report the contemporaneous correlations between the raw sentiment measures and asset returns,[5] I use the orthogonalized version in the remaining empirical analyses. In my out-of-sample tests, I recursively orthogonalize

[4] For example, although Bitcoin has been around since January 2009, its text sentiment measures from social media and news headlines are unavailable until mid-2013. After I use the first 100 days to construct my sentiment variables, the final sample for Bitcoin starts from January 2014.

[5] Because the contemporaneous return is included as an independent variable in regression (2.8), the contemporaneous correlations between the orthogonalized sentiment measures and asset returns are zero.

sentiment variables using data up to the estimation time to avoid the look-ahead bias. Since an estimation window may not contain any recessionary days, I exclude the recession dummy variable from (2.8) in the out-of-sample analyses.[6]

2.3 Method

2.3.1 Predictive Regression

Following most studies, I analyze the return predictability of sentiment via the standard predictive regression model,

$$r^i_{t+1\to t+h} = \alpha^i_j + \beta^i_j \times x^i_{j,t} + \epsilon^i_{t+1\to t+h}, \tag{2.1}$$

where $r^i_{t+1\to t+h}$ is the normal return of asset i over days from $t+1$ to $t+h$, $x^i_{j,t}$ is a sentiment j for asset i on day t, and $\epsilon^i_{t+1\to t+h}$ is the residual term. β^i_j measures the strength of predictability. In my empirical results, I report its significance based on the Newey and West (1987)'s standard error with h lags (if $h < 5$, I set lag to 5).

I generate out-of-sample forecasts of asset returns using an expanding estimation window. Given an initial estimation window length W, the out-of-sample forecasts of $r^i_{t+1\to t+h}$ is given by

$$\hat{r}^i_{t+1\to t+h} = \hat{\alpha}^i_{j,t} + \hat{\beta}^i_{j,t} \times x^i_{j,t}, \tag{2.2}$$

where $\hat{\alpha}^i_{j,t}$ and $\hat{\beta}^i_{j,t}$ are estimates of α^i_j and β^i_j in equation (2.1) respectively by using data up to time t for $t = W, W+1, \ldots, T-h$ with T being the sample size. I repeat this procedure until the end of the out-of-sample period. My empirical analyses use an initial estimation window of 750 days for all assets, i.e., $W = 750$.

[6]In unreported results, I find that using the raw sentiment measures yields robust predictability.

2.3.2 Information Aggregation

As discussed above, this paper compares the return predictability of 28 trade sentiment variables against 28 text sentiment measures for each asset. I employ several commonly used dimension reduction methods to provide a parsimonious comparison between these two groups of sentiment variables. Dimension reduction also helps reduce overfitting, resulting in better out-of-sample prediction performance.

The first and most simple dimension reduction method is the average variable (AV). This AV technique is used in Dong et al. (2022) to combine a large set of returns on long-short anomalies into a strong aggregate market predictor. Under this method, at each time point, I compute the cross-sectional average of all standardized variables within a group of predictors, i.e.

$$\mathrm{AV}^i_{g,t} = \frac{1}{n_g}\sum_{j=1}^{n_g} x^i_{j,t}, \tag{2.3}$$

where $n_g = 28$ and $x^i_{j,t}$ is one of the trade or text sentiment variables for asset i and sentiment group j.

In my out-of-sample forecasts, I also use the combination forecast (CF) method as in Rapach et al. (2010), which shows that this is a powerful filter to reduce noises in forecasts, leading to more accurate predictions. Under this method, at each point in the out-of-sample window, I compute the average of all return forecasts, i.e.

$$\bar{r}^i_{g,t+1\to t+h} = \frac{1}{n_g}\sum_{j=1}^{n_g} \hat{r}^i_{j,t+1\to t+h}, \tag{2.4}$$

where $\hat{r}^i_{j,t+1\to t+h}$ is a return forecast for asset i made by predictor j as in equation (2.2).

My third method is principal component analysis (PCA), a widely used tool for dimension reduction, reducing the set of factors to a smaller set of components that explains most of those factors' variance. Suppose I want to extract k principal components from p factors. I want to reduce the number of explanatory variables from p

to k, $k \leq p$. The k^{th} principal component PC_k is a normalized linear combination of the p factors: $PC_k \equiv Xv_k$ where X is the $T \times p$ centered matrix of factors and the $p \times 1$ vector v_k solves

$$\begin{aligned} \max_{v} \quad & \mathrm{Var}(Xv) \\ \text{s.t.} \quad & \|v\| = 1 \quad \text{and} \\ & (Xv)'(Xv_l) = 0, \quad l = 1, \ldots, k-1. \end{aligned} \tag{2.5}$$

From the objective function, it is easy to see that the first principal component PC_1 indicates the direction of the most significant sample variance in the column space of X; the second principal component PC_2 suggests the direction of the second largest sample variance among all normalized linear combination of the p factors; so on and so forth. The second constraint $(Xv)'(Xv_l) = 0$ ensures the k^{th} principal component is uncorrelated to all previous $k - 1$ principal components. A great advantage when applying PCA is that it does not require any prior information about the rate of returns to be explained. Following the standard practice, I standardize all predictors to zero mean and unit variance before implementing PCA.

My final method is partial least squares (PLS), which has gained popularity in the return prediction literature (Kelly and Pruitt 2013, 2015; Huang et al. 2015, 2020). Like PCA, PLS is also a dimension reduction method by constructing a (smaller) set of linear combinations of original factors. But unlike PCA, which focuses only on the components' high variance, PLS transforms the factors to capture both the high variance of the extracted components and the high correlation with asset returns to be predicted. Mathematically, the k^{th} PLS component is $PLS_k \equiv X\theta_k$ where θ_k solves

$$\begin{aligned} \max_{\theta} \quad & \mathrm{Corr}^2(y, X\theta) Var(X\theta) \\ \text{s.t.} \quad & \|\theta\| = 1 \quad \text{and} \\ & (X\theta)'(X\theta_l) = 0, \quad l = 1, \ldots, k-1, \end{aligned} \tag{2.6}$$

where y is the $T \times 1$ vector of returns to be predicted. When the variance term in the objective function dominates, PLS will behave much like PCA. However, when the correlation between X and y is strong, I expect PLS to perform better because it incorporates such information in dimension reduction. At the same time, PCA ignores the correlation and might return components with a high variance but little covariance with y.

2.3.3 Out-of-Sample R^2

I use two measures to evaluate the performance of out-of-sample forecasts: the out-of-sample R^2 (Campbell and Thompson 2008) and the realized utility gains for a risk-averse investor in the mean-variance world (Campbell and Thompson 2008; Rapach et al. 2010).

The out-of-sample R^2, R^2_{OS}, evaluates the performance of out-of-sample forecasts of asset return $\hat{r}^i_{t+1\to t+h}$ against its historical average $\bar{r}^i_{t+1\to t+h}$ within the estimation window, with $\bar{r}^i_{t+1\to t+h} = \frac{1}{t-h}\sum_{s=1}^{s=t-h} r^i_{s+1\to s+h}$. More specifically, the R^2_{OS} statistics is given by

$$R^{i,2}_{OS}(h) = 1 - \frac{\sum_{t=W}^{t=T-h}(r^i_{t+1\to t+h} - \hat{r}^i_{t+1\to t+h})^2}{\sum_{t=W}^{t=T-h}(r^i_{t+1\to t+h} - \bar{r}^i_{t+1\to t+h})^2}, \tag{2.7}$$

If $R^2_{OS} > 0$, the constructed forecast $\hat{r}^i_{t+1\to t+h}$ beats the historical average forecast since $\hat{r}^i_{t+1\to t+h}$ has smaller mean squared prediction error (MSPE) than the latter does. I examine the statistical significance of the R^2_{OS} with the Clark and West (2007) adjusted MSPE test.

2.3.4 Utility Gains

Even though R^2_{OS} is widely used in forecast evaluation, it has its limitation: it does not consider investors' risk preferences. It, therefore, does not directly speak to the economic significance of the forecasts. To address this, I also calculate the average utility gain of a risk-averse investor in a mean-variance world (Campbell and Thompson 2008). After forming a forecast of the next period's return on risky asset i, at the end of period t, a mean-variance investor can decide how to allocate the total wealth between a risk-free asset and a risky asset. The weight on asset i, $\hat{w}^i_{t+1}$, is decided according to the following formula

$$\hat{w}^i_{t+1} = \frac{1}{\gamma}\left(\frac{\hat{r}^i_{t+1} - r^f_{t+1}}{\hat{\sigma}^{2,i}_{t+1}}\right), \tag{2.8}$$

where r^f_{t+1} is the risk-free rate on the day $t+1$, $\hat{r}^i_{t+1}$ is the forecast of the return on asset i on the day $t+1$, $\hat{\sigma}^{2,i}_{t+1}$ is a 20-day rolling window estimate of the variance of the excess return, $r^i_{t+1} - r^f_{t+1}$, and γ is the coefficient of risk aversion. My empirical results consider γ equal to 3 and 5. Following Campbell and Thompson (2008), I constrain $\hat{w}^i_{t+1}$ between 0 and 1.5 to exclude short-selling and limit the portfolio leverage to only 50%.

The realized rate of return of this portfolio on the day $t+1$ is

$$r^{i,p}_{t+1} = \hat{w}_{t+1} r^i_{t+1} + (1 - \hat{w}^i_{t+1}) r^f_{t+1}. \tag{2.9}$$

The realized average utility of this risk-averse investor exploiting such a portfolio strategy is calculated as follows.

$$\hat{u}^i = \hat{\mu}^i - \frac{1}{2}\gamma\hat{\sigma}^{2,i}, \tag{2.10}$$

where $\hat{\mu}^i$ and $\hat{\sigma}^{2,i}$ are the sample mean and sample variance of the realized return $r^{i,p}_{t+1}$, respectively, of the constructed portfolio over the out-of-sample period.

The evaluation benchmark is still the historical average forecast. If using the historical average forecast, this same investor will allocate $\bar{w}^i_{t+1}$ portion of total wealth to asset i according to the following formula

$$\bar{w}^i_{t+1} = \frac{1}{\gamma}\left(\frac{\bar{r}^i_{t+1} - r^f_{t+1}}{\hat{\sigma}^{2,i}_{t+1}}\right), \tag{2.11}$$

where $\bar{r}^i_{t+1}$ is the historical mean forecast of return. The realized rate of return of this portfolio is

$$r_{t+1}^{i,p} = \bar{w}_{t+1} r_{t+1}^{i} + (1 - \bar{w}_{t+1}^{i}) r_{t+1}^{f}. \tag{2.12}$$

The realized average utility of this risk-averse investor using a historical average forecast is given by

$$\bar{u}^{i} = \bar{\mu}^{i} - \frac{1}{2}\gamma \bar{\sigma}^{2,i}. \tag{2.13}$$

The utility gain is measured as the difference between $\hat{u}^i$ and $\bar{u}^i$

$$\text{Utility Gain} = 25000 \times (\hat{u}^i - \bar{u}^i) \tag{2.14}$$

I multiply this difference by 25000 to report it in the average annualized percentage return, assuming 250 trading days in a year. In addition to the utility gain, I present the annualized Sharpe ratio of the realized portfolio returns $r_{t+1}^{i,p}$. I examine the statistical significance of the utility gains and investigate whether the Sharpe ratio of the portfolio using the predictive model is greater than that of the portfolio using the historical mean forecasts with the tests in DeMiguel et al. (2009).

2.4 Empirical Results

2.4.1 Contemporaneous Correlations

Before presenting the empirical prediction results, I evaluate the contemporaneous correlations between my trade sentiment and text sentiment variables and asset returns. As sentiment measurements, I anticipate a favorable correlation between them and returns.

[Insert Table 2.1 here.]

Panel A of Table 2.1 displays the results for measures of trade sentiment. Notable is that while aberrant trading volume (TVS) is connected with higher returns for

Bitcoin, it is inversely correlated with returns on the remaining assets, with stocks exhibiting the most substantial negative associations. Only Bitcoin and T-bond returns are favorably connected with sentiment indicators based on moving averages of prices (MAS) and a combination of prices and trading volumes (OBV). Three price-based sentiment groups (WRS, NHS, and MOM) exhibit a substantial positive association with the returns of all assets.

Panel B presents the concurrent connections between text sentiment and returns. The link between sentiment from news headlines and asset returns is the greatest among the four sources of text sentiment. Gold has the strongest positive correlation with returns, at 18%, and is the asset most affected by text sentiment. On the other hand, Treasury bonds exhibit only positive relationships between news headline sentiment and their returns.

As expected, Table 2.1 shows that trade and text sentiment measures positively correlate with asset returns.

2.4.2 In-Sample Predictions with Individual Sentiment Measures

Next, I examine the in-sample return predictability of sentiment variables across asset classes. I analyze the returns forecast for the following day ($h = 1$) and 20 days ($h = 20$).

[Insert Table 2.2 here.]

Panel A of Table 2.2 displays the results for measures of trade sentiment. As Bitcoin is a highly speculative asset, it is unsurprising that measures of trade sentiment, including WRS and MOM, can favorably predict Bitcoin returns over the next 1 to 20 days, while OBV and MAS yield good predictions for the long term, with results weakest among the NHS and TVS variables. My findings indicate a momentum pattern in Bitcoin return forecasts: high trading sentiment today portends high returns over the following twenty days.

T-bonds exhibit a momentum-like relationship between trade sentiment and long-term returns. However, the prediction coefficient for next-day returns is negative for TVS, WRS, and NHS, but positive for MAS and OBV.

The relationship between trade sentiment and future returns in the stock market is similar: high trade sentiment is now associated with higher returns over the following 1 to 20 days. While WRS provides strong stock prediction results over the next 1 to 20 days, MOM provides more accurate short-term results than OBV for the long term. Some exceptions for TVS5, NHS10, and OBV20_50 predict negative returns. Lastly, gold is almost unpredictable based on indicators of trade sentiment, except for WRS and MO, which can modestly positively forecast gold returns over a 20-day horizon.

Panel B displays the results of sentiment analysis on text predictions. Surprisingly, news and social media tone measurements fail to predict Bitcoin and gold returns. I observe the prediction results for Treasury bonds and stocks, but they are not strong.

At first glance, it appears odd that news and social media tone cannot forecast market returns, given the literature's overwhelming evidence of predictability (Tetlock 2007; Garcia 2013). I use the news sentiment measure developed by Garcia (2013) and published on the author's website to resolve this discrepancy.[7] Then, I apply the predictive regression specification from Tetlock (2007) and Garcia (2013) to three sample periods: 1905 to 2005, 1984 to 1999, and 2000 to 2005. Following Tetlock (2007) and Garcia (2013), I utilize five lags of sentiment and control for five lags of returns, five lags of squared returns, and weekday indicators for this regression and show the results in Table 2.A.1. Accordingly, I can replicate Tetlock (2007) and

[7]I thank Diego Garcia for making this data available.

Garcia (2013) findings that negative (positive) sentiment is a substantial negative (positive) predictor of the next day's return. Nonetheless, over the sample 2000-2005, the predictability of news sentiment nearly vanishes, corresponding with the findings shown in Table 2.2. This finding suggests that news tone has been fully integrated into stock prices on the same day during the past two decades, rendering its future predictability null and void.

Overall, individual trade sentiment measures are significantly more accurate in predicting future returns on all assets than individual text sentiment measures. Indicators of trade sentiment exhibit momentum for all assets. Short-term text sentiment is favorable only for T-bonds and stocks.

2.4.3 In-Sample Predictions with Aggregate Sentiment Measures

Examining univariate regressions with individual sentiment variables, the preceding section demonstrates that trade sentiment measures can predict returns better than text sentiment measures. In this section, I formally compare trade and text sentiment measurements. I begin by developing indexes that summarize the differences between these two categories of sentiment indicators. As stated in Section 2.3, my first and most straightforward technique is to compute the cross-sectional means of trade and text sentiment variables separately (AV). My more advanced methods involve creating the PCA and PLS indexes.

[Insert Figure 2.1 here.]

Figure 2.1 illustrates each asset's PCA and PLS weights. For Bitcoin, WRS and MOM measures dominate the PLS weight. This is expected given the results in Table 2.2 because PLS weights are scaled correlations between sentiment measures and next one-day returns. PCA weights are spread throughout all trading sentiments. For text sentiment, PCA weights are distributed through all sentiments, while NewsHL dominates the PLS weights.

Except for TVS, PCA weights are distributed equally across all trade sentiment techniques for T-bonds. Only MAS, OBV, and MOM have positive PLS weights. For text, PLS weights are concentrated in News, NewsHL, and NewsSoc. Only TVS measures have positive weights for stocks, yet they are low. While PCA weights are distributed equally across measures for text sentiment, News and NewsHL dominate the PLS weights.

For gold, PLS weights are dominated by TVS and MOM. Interestingly, all trade sentiments have negative PCA weight. For text, PCA weights are spread across text sentiments, whereas PLS weights are negative for all except NewsHL.

After creating my AV, PCA, and PLS indexes from trade and text sentiment measurements for each asset, I employ them in univariate predictive regressions (2.1) and bivariate predictive regressions using one trade and one text sentiment index. As there is no restriction on the sign of each component in PCA and PLS, I reverse the sign of the aggregate PCA and PLS indexes when appropriate to ensure that the trade (text) PCA and PLS indexes are always positively associated with the trade (text) average-variable (AV) index. Since a high value of each sentiment measure indicates a high level of sentiment, their averages (i.e., AV indexes) are still sentiment. Thus, the PCA and PLS indices can still be regarded as a measure of sentiment when inverted.

[Insert Table 2.3 here.]

For Bitcoin in Panel A of Table 2.3, all three measures of aggregate trade sentiment predict Bitcoin returns over the following 1 to 20 days, whether examined alone or in conjunction with the text sentiment variables. In contrast, the text sentiment indexes fail to forecast future returns when used alone or with trade sentiment, as seen in Table 2.2.

The results for Treasury bonds in Panel B are comparable to those for Bitcoin in that trade sentiment indicators are positive return predictors, especially for the next five days. Text sentiment measurements predict future returns only for the next day under AV and PCA. Under PLS, trade sentiment can predict returns better than text.

Panel C presents the stock forecasting results. Trade sentiment indicators can positively predict future returns, with shorter periods producing more substantial impacts. When employed alone, text sentiment indexes such as AV and PCA cannot forecast future returns. Still, they become positive predictors (significant at 10%) and increase the statistical significance of trade indices when the two groups are compared. When predicting future stock returns, trading sentiment outperforms text sentiment, exhibiting a reversal prediction pattern for AV and PCA but the positive pattern for PLS.

Under PLS, trade sentiment is a positive predictor of future returns for gold in Panel D. When compared, trade sentiment predicts the following one to twenty days with more accuracy than text sentiment.

In conclusion, trading sentiment surpasses text sentiment in predicting future returns for all four assets. Trade sentiment exhibits strong and persistent predictive power for Bitcoin. Text sentiment works only for stocks, yet the result is weak.

2.4.4 Out-of-Sample R^2

In the previous section, I discover that trade sentiment measures predict returns more accurately than text sentiment measures across all four asset classes. I expand my empirical analysis to an OOS setting by examining the common metric R^2_{OS}. Out-of-sample prediction is more complicated than in-sample predictability, and most economic forecasters have failed to predict out-of-sample returns (Goyal and Welch 2008). In addition to the three dimension reduction strategies employed in the in-sample tests, I consider the combination forecast (CF), wherein I cross-sectionally average the forecasts generated by individual trade and text sentiment variables.

[Insert Table 2.4 here.]

Panel A of Table 2.4 demonstrates that trade sentiment measurements under all four reduction procedures produce strong R^2_{OS} with a one-day Bitcoin return. For instance, the one-day R^2_{OS} ranges between 0.18% and 1.57% for trade PCA and AV, respectively. Except for trade PCA, the significant outcomes continue to the 20-day horizon. In contrast, none of the text-reduction techniques can accurately predict Bitcoin returns out-of-sample. These results confirm the in-sample findings that only trade sentiment metrics apply to Bitcoin.

For Treasury bonds in Panel B, trade sentiment gives better out-of-sample forecasts than text sentiment, consistent with the in-sample results. Trade sentiment generates significantly favorable OOS R^2 values for the subsequent five-day prediction under CF and PCA. Only the text sentiment PCA indexes generate a positive OOS R^2.

For stocks in Panel C and gold in Panel D, neither trading nor text sentiment generates a substantial OOS R^2 over any horizon, regardless of the time horizon. As reported in Goyal and Welch (2008) and Goyal et al. (2022), this exemplifies the difficulty of accurately predicting asset returns out-of-sample.

Overall, out-of-sample results still favor Bitcoin and Treasury bond trade sentiment measures. No sentiment metric exceeds the historical mean benchmark at the 5% significant level or greater for out-of-sample stock and gold predictions.

2.4.5 Forecast Encompassing Tests

While the previous subsection provides empirical support for trade sentiment measures over text sentiment measures in out-of-sample forecasts for Bitcoin and Treasury bonds, such evidence is indirect because when using the R^2_{OS} metric, I compare the forecasts made by trade sentiment and text sentiment against the standard historical mean benchmark. In this part, I directly compare the informational content of trade forecasts with text sentiment forecasts.

Following Rapach et al. (2010), I employ the modified version of the test statistic created by Harvey et al. (1998) (MHLN). This MHLN test tests the null hypothesis that the model i forecast encompasses the model j forecast to the alternative hypothesis that the model i forecast does not encompass the model j prediction. I have a total of eight return projections for each asset (four dimension reduction methods applied to trade sentiment and four methods applied to text sentiment). Thus, I examine if a prediction method's forecast incorporates those of the remaining seven prediction methods. I conduct this test for the 1-day and 20-day forecasts.

[Insert Table 2.5 here.]

Table 2.5 presents the p-values for testing the null hypothesis that the model's forecast in a column encompasses the model's forecast in the corresponding row. The forecast-encompassing test (Chong and Hendry 1986) implies that the predictor that encompasses another offer's information regarding potential future movements of the testing asset that is absent from another predictor. I emphasize the entries where null is accepted. In Panel A, except trade PLS, all of Bitcoin trade sentiment measures encompass text sentiment measures over a one-day horizon ($h = 1$). With the 20-day return projection, except trade PCA and PLS, the other two trade sentiment forecasts encompass text sentiment measure estimates again. This result is consistent with the R^2_{OS} mentioned in the preceding section. Text sentiment encompasses trade only for the PLS index.

The results for Treasury bonds in Panel B indicate that text sentiment encompasses trade sentiment over the short term, but the relation reverts for the twenty-day horizon.

For equities in Panel C, trade and text sentiment predictions do not encompass each other. Regarding gold in Panel D, trade and text sentiment estimates overlap under all approaches except PLS for the following one-day horizon. In contrast, trade and text sentiment estimates do not overlap over the long term. The results indicate that trade and text sentiment forecasts include valuable information regarding future returns.

In short, I discover that, for Bitcoin, the projections from trade sentiment measures encompass those from text sentiment measures for both short and long horizons. For Treasury bonds, the same pattern appears only for the long term. These results are consistent with prior findings in that trade sentiment measurements produce more prominent (and significant) R^2_{OS}'s for Bitcoin and government bonds. For future stock returns, combination forecasts based on trade sentiment provide the most valuable information, whereas, for gold, the combination helps for a long-term forecast.

2.4.6 Utility Gains

This final empirical investigation examines whether sentiment measures may produce effective trading strategies. I estimate the utility gains and Sharpe ratio for a mean-variance investor that allocates a portfolio between a risky and a risk-free asset based on the prediction method's *one-day return forecast* for each asset. The portfolio based on the historical mean forecast is the benchmark for utility gains. For comparison, I also include the Sharpe ratio for a buy-and-hold strategy.

[Insert Table 2.6 here.]

The most striking result from Table 2.6 is that, except for Bitcoin (Panel A), it is not easy to construct daily profitable trading strategies using sentiment measures, both trade, and text. Trade sentiment measures under all four prediction models deliver impressive economic gains for Bitcoin. For example, with the risk coefficient of three ($\gamma = 3$), trade PLS delivers a superb annualized utility gain of 50% with an annualized Sharpe ratio of 1.73, much higher than the buy-and-hold Sharpe ratio of 0.76. On the other hand, text sentiment measures under all prediction models do not generate any economic gains for a mean-variance investor.

No sentiment measure can generate economic values over the historical mean for Treasury bonds, equities, and gold in Panels B, C, and D and for bonds and gold, none of them produces a Sharpe ratio exceeding the buy-and-hold strategy. This corresponds to the OOS R^2 results provided in Table 2.4. Nevertheless, for stocks, even though the generated portfolios fail to surpass the historical mean benchmark, they have more excellent Sharpe ratios than the buy-and-hold strategy.

In short, Bitcoin is the sole asset whose trade sentiment metrics can be leveraged to provide exceptional economic values for real-time investors. Only trade sentiment under PLS delivers real-time economic gains for Treasury bonds; however, trading based on sentiment signals does not outperform the historical mean benchmark for stocks and gold.

2.5 Conclusion

This article examines the return predictability of investor sentiment across four asset classes. I contrast two types of investor sentiment: those based on trading data and those based on news and social media. For each asset, I derive 28 trade sentiment measures from its time series of prices and trading volumes and 28 text sentiment measures from news articles, headlines, social media, and news and social media combined.

Then, I employ four dimension reduction techniques, namely average variable (AV), mean combination forecast (CF), principal component analysis (PCA), and partial least squares (PLS), to facilitate a parsimonious setting in which trading and text sentiment can be easily compared. These approaches of information aggregation also improve out-of-sample return forecasts when several predictors are available.

In my example configuration, trade sentiment performs better than text sentiment across all prediction methodologies and forecasting horizons for all four assets. I also demonstrate that all asset returns consistently underreact to trade sentiment.

As for out-of-sample tests, I find that trade sentiment produces superior return projections based on the out-of-sample R^2 measure and forecast-encompassing test for Bitcoin and government bonds. Trade and text sentiment fail to beat the historical mean benchmark in out-of-sample predictions for stocks and gold.

Finally, I investigate whether sentiment-based return projections may provide real-time mean-variance investors with economic value. According to my findings, Bitcoin is the only asset whose trade sentiment can create exceptional economic gains.

Figure 2.1. PCA and PLS Weights

This figure plots the PCA (blue) and PLS (screen) weights of each orthogonalized individual sentiment measure. See Section 1.2 for the construction of trade sentiment measures and PCA and PLS indexes. The number(s) trailing the trade sentiment variables indicate(s) the rolling window used to compute that variable. Text sentiment measures are constructed from news articles (*News*), news headlines (*NewsHL*), social media (*Social*), and both news and social media (*NewsSoc*). The number L trailing the text sentiment measures means moving average over the past L days. Returns are in percentages and independent variables are standardized to zero mean and unit standard deviation. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Figure 2.1. PCA and PLS Weights (Cont.)

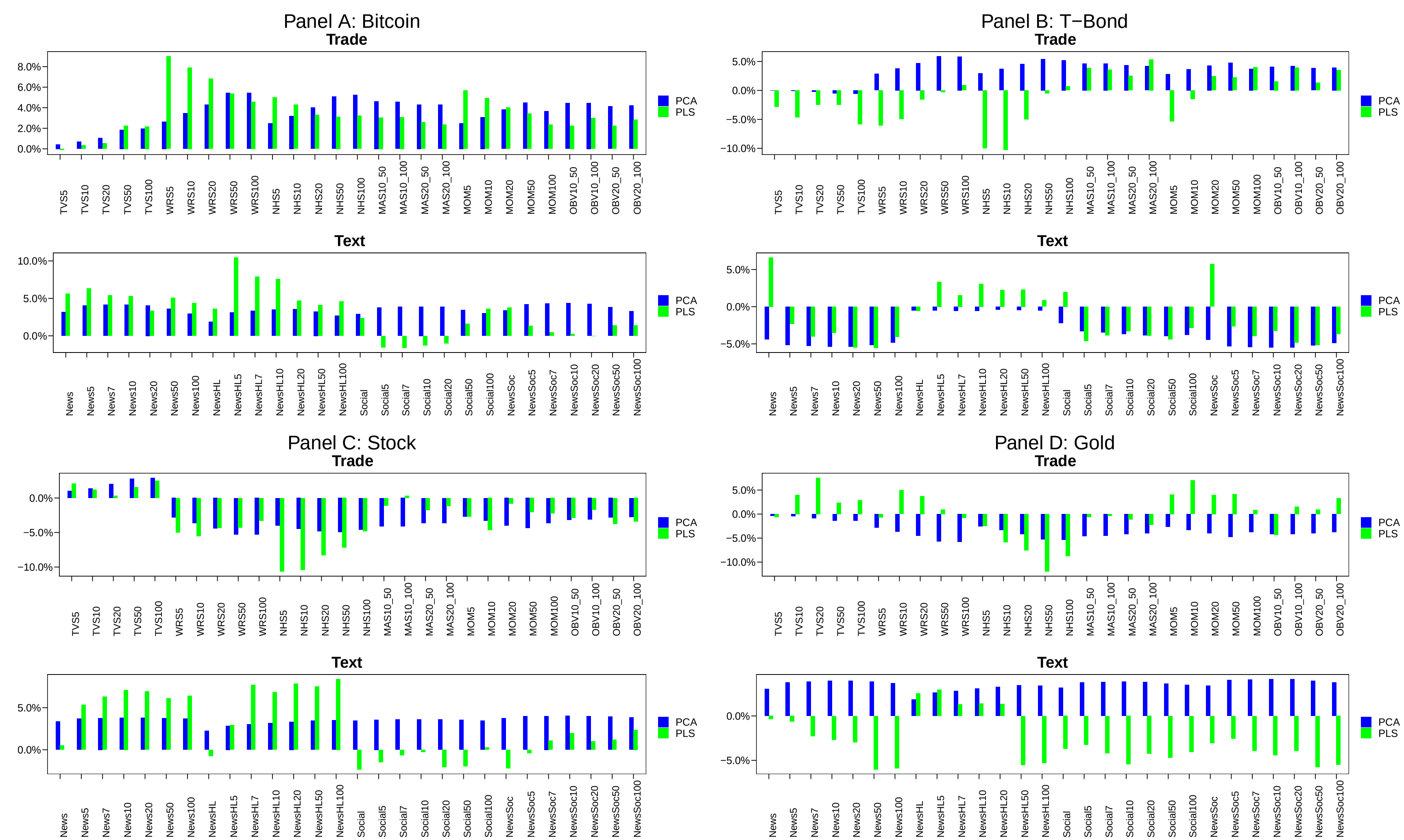


Table 2.1
Contemporaneous Correlations Between Sentiment Measures and Returns

This table reports contemporaneous correlations between trade (text) sentiment measures and returns in Panel A (B). See Section 1.2 for the construction of trade sentiment measures. The number(s) trailing the trade sentiment variables indicate(s) the rolling window used to compute that variable. Text sentiment measures are constructed from news articles (*News*), news headlines (*NewsHL*), social media (*Social*), and both news and social media (*NewsSoc*). The number L trailing the text sentiment measures means moving average over the past L days. ***, **, and * indicate significance at 1%, 5%, and 10% respectively. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Panel A: Trade Sentiment Measures

	Bitcoin	T-Bond	Stock	Gold
TVS5	0.00	-0.02	-0.15 ***	-0.03 **
TVS10	0.02	-0.02	-0.13 ***	-0.03 *
TVS20	0.02	-0.02	-0.13 ***	-0.02
TVS50	0.05 ***	-0.02	-0.11 ***	-0.02
TVS100	0.05 **	-0.02	-0.09 ***	-0.03 *
WRS5	0.57 ***	0.62 ***	0.56 ***	0.60 ***
WRS10	0.46 ***	0.49 ***	0.44 ***	0.47 ***
WRS20	0.36 ***	0.37 ***	0.34 ***	0.36 ***
WRS50	0.26 ***	0.26 ***	0.24 ***	0.25 ***
WRS100	0.20 ***	0.20 ***	0.18 ***	0.19 ***
NHS5	0.54 ***	0.57 ***	0.55 ***	0.55 ***
NHS10	0.41 ***	0.42 ***	0.39 ***	0.41 ***
NHS20	0.31 ***	0.33 ***	0.27 ***	0.31 ***
NHS50	0.23 ***	0.23 ***	0.18 ***	0.22 ***
NHS100	0.19 ***	0.17 ***	0.14 ***	0.18 ***
MAS10_50	0.07 ***	0.03 **	0.01	0.02
MAS10_100	0.08 ***	0.03 **	0.01	0.01
MAS20_50	0.06 ***	0.03 **	-0.00	0.01
MAS20_100	0.05 ***	0.03 **	0.00	-0.00
MOM5	0.37 ***	0.37 ***	0.34 ***	0.35 ***
MOM10	0.25 ***	0.23 ***	0.21 ***	0.24 ***
MOM20	0.18 ***	0.17 ***	0.14 ***	0.16 ***
MOM50	0.12 ***	0.10 ***	0.10 ***	0.09 ***
MOM100	0.08 ***	0.06 ***	0.07 ***	0.07 ***
OBV10_50	0.06 ***	0.02	-0.01	0.00
OBV10_100	0.07 ***	0.03 **	-0.00	0.01
OBV20_50	0.05 ***	0.02	-0.01	0.01
OBV20_100	0.07 ***	0.03 *	-0.02	0.01

Panel B: Text Sentiment Measures

	Bitcoin	T-Bond	Stock	Gold
News	0.11 ***	0.02	0.08 ***	0.15 ***
News5	0.05 ***	0.01	0.03 **	0.05 ***
News7	0.04 **	0.00	0.03 **	0.04 **
News10	0.03 *	-0.01	0.02 *	0.02
News20	0.02	-0.01	0.02	0.00
News50	0.02	-0.01	0.01	-0.02
News100	0.02	-0.01	0.01	-0.02
NewsHL	0.11 ***	0.17 ***	0.08 ***	0.18 ***
NewsHL5	0.06 ***	0.08 ***	0.02 *	0.08 ***
NewsHL7	0.06 ***	0.06 ***	0.02 *	0.06 ***
NewsHL10	0.05 ***	0.05 ***	0.02 *	0.05 ***
NewsHL20	0.03 *	0.04 ***	0.02 *	0.03 **
NewsHL50	0.02	0.02 *	0.02	-0.01
NewsHL100	0.02	0.01	0.02	-0.01
Social	0.11 ***	0.01	0.09 ***	0.13 ***
Social5	0.02	-0.00	0.02 *	0.03 **
Social7	0.02	-0.01	0.02	0.02
Social10	0.01	-0.01	0.01	0.00
Social20	0.00	-0.01	0.00	-0.00
Social50	0.01	-0.01	-0.00	-0.01
Social100	0.01	-0.01	0.00	-0.01
NewsSoc	0.13 ***	0.02	0.09 ***	0.15 ***
NewsSoc5	0.04 **	0.01	0.03 **	0.04 ***
NewsSoc7	0.03	-0.00	0.02 *	0.03 **
NewsSoc10	0.02	-0.00	0.02	0.01
NewsSoc20	0.01	-0.01	0.01	0.00
NewsSoc50	0.01	-0.01	0.00	-0.02
NewsSoc100	0.01	-0.01	0.01	-0.02

Table 2.2
Predicting Returns with Individual Sentiment Measures

This table reports β from the following regression

$$r_{t+1\to t+h} = \alpha + \beta \times x_t + \epsilon_{t+1\to t+h},$$

where $r_{t+1\to t+h}$ is the cumulative asset returns over the next h days and x_t is an orthogonalized trade (text) sentiment measure in Panel A (B). See Section 1.2 for the construction of trade sentiment measures. The number(s) trailing the trade sentiment variables indicate(s) the rolling window used to compute that variable. Text sentiment measures are constructed from news articles (*News*), news headlines (*NewsHL*), social media (*Social*), and both news and social media (*NewsSoc*). The number L trailing the text sentiment measures means moving average over the past L days. Returns are in percentages and independent variables are standardized to zero mean and unit standard deviation. ***, **, and * indicate significance at 1%, 5%, and 10% respectively based on Newey and West (1987)'s standard errors with h lags (if $h < 5$, lag is set to 5). The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Panel A: Trade Sentiment Measures

	Bitcoin		T-Bond		Stock		Gold	
	h=1	h=20	h=1	h=20	h=1	h=20	h=1	h=20
TVS5	-0.00	0.27	0.00	-0.08 *	0.01	-0.11 **	0.01	0.06
TVS10	-0.05	0.12	-0.01	-0.07	-0.01	-0.10 *	0.02	0.07
TVS20	-0.05	0.12	-0.00	-0.09 *	-0.01	-0.06	0.02	0.12
TVS50	0.02	0.69 **	0.00	-0.10 *	-0.01	-0.04	-0.00	0.03
TVS100	0.01	0.42	-0.01	-0.13 **	0.01	-0.02	0.01	0.02
WRS5	0.32 ***	1.25 ***	0.02	-0.00	0.04 **	0.04	0.01	0.14 *
WRS10	0.29 ***	1.14 ***	0.02	-0.02	0.06 ***	0.12 *	0.02	0.16 **
WRS20	0.18 ***	1.09 ***	0.02	0.06	0.06 **	0.13 **	0.02	0.10
WRS50	0.10	0.86 **	0.01	0.05	0.04	0.14 **	0.03 *	0.15 **
WRS100	0.04	0.68 **	0.00	0.10 **	0.03	0.15 **	0.02	0.05
NHS5	-0.04	-0.31	0.00	0.02	0.00	-0.06	-0.00	-0.03
NHS10	-0.06	-0.33	-0.01	-0.02	0.00	-0.14 **	-0.01	-0.07
NHS20	-0.04	-0.19	-0.02	0.07	0.05 *	-0.01	-0.02	-0.05
NHS50	-0.05	-0.20	-0.01	0.08	0.01	-0.01	-0.00	-0.08
NHS100	-0.12 *	-0.08	-0.02	0.08	-0.01	0.05	-0.01	-0.09
MAS10_50	0.09 *	0.74 **	0.01	0.06	0.00	0.03	0.01	-0.02
MAS10_100	0.05	0.45 *	-0.00	0.07	0.01	0.01	0.00	-0.04
MAS20_50	0.06	0.56 *	0.01	0.02	-0.01	-0.04	-0.01	-0.02
MAS20_100	0.06	0.31	0.03 **	0.15 ***	0.04 *	-0.05	0.02	-0.05
MOM5	0.19 ***	0.78 **	0.01	0.01	0.05 ***	0.08	0.02	0.15 **
MOM10	0.13 ***	0.85 ***	0.02	-0.00	0.01	0.08	0.03 *	0.05
MOM20	0.11 **	0.70 **	0.00	0.03	0.03 *	0.09	0.02	0.05
MOM50	0.03	0.39	0.00	0.08 *	0.02	0.01	0.02	0.02
MOM100	0.04	0.23	0.02 **	0.09 *	-0.03 *	-0.01	0.03	0.01
OBV10_50	0.04	0.35	0.01	0.02	0.00	-0.05	0.01	-0.08
OBV10_100	-0.02	0.64 **	0.02	0.07	0.04 *	-0.04	-0.00	-0.03
OBV20_50	0.02	0.69 **	0.03 **	0.02	-0.00	-0.14 **	-0.01	-0.02
OBV20_100	0.10 *	0.75 **	0.01	0.05	-0.00	-0.12 *	-0.00	-0.07

Panel B: Text Sentiment Measures

	Bitcoin		T-Bond		Stock		Gold	
	h=1	h=20	h=1	h=20	h=1	h=20	h=1	h=20
News	0.01	-0.11	0.03 **	0.03	-0.01	0.07	0.00	-0.02
News5	0.01	-0.16	0.02	0.01	-0.03	0.09	0.01	0.01
News7	-0.00	-0.20	0.01	0.01	-0.01	0.08	0.00	-0.02
News10	-0.02	-0.09	0.01	0.01	-0.01	0.09	0.02	-0.00
News20	0.02	-0.29	0.02	0.01	-0.03 *	0.10	0.00	0.06
News50	-0.00	-0.26	0.03 **	0.00	-0.03 *	0.05	0.00	-0.00
News100	0.02	-0.25	0.03 **	-0.03	-0.04 ***	0.11	0.02	-0.01
NewsHL	-0.01	-0.10	0.01	0.03	-0.00	0.06	0.02	0.02
NewsHL5	0.00	-0.08	0.02	0.04	-0.01	0.10	0.01	0.04
NewsHL7	0.02	-0.12	0.01	0.03	0.00	0.09	0.01	0.00
NewsHL10	-0.04	-0.04	0.02	0.03	-0.00	0.09	0.02	0.03
NewsHL20	0.04	-0.27	0.02 *	0.01	-0.01	0.07	0.01	0.09
NewsHL50	0.01	-0.16	0.01	0.05	-0.02	0.04	0.01	0.01
NewsHL100	-0.04	-0.19	0.01	0.04	-0.01	0.12 *	0.03 *	-0.01
Social	0.01	-0.40	0.01	-0.05	0.01	0.03	0.00	-0.04
Social5	0.01	-0.51	0.01	-0.03	-0.01	0.05	-0.01	-0.01
Social7	-0.06	-0.52	-0.00	-0.04	-0.01	0.05	0.01	-0.01
Social10	0.00	-0.40	0.02	-0.03	-0.01	0.05	-0.02	-0.03
Social20	0.03	-0.37	0.01	-0.02	-0.02	0.04	-0.01	-0.03
Social50	0.04	-0.32	0.01	-0.05	-0.04 **	-0.06	-0.00	-0.06
Social100	-0.01	-0.91 **	0.01	-0.07	-0.05 ***	-0.06	-0.00	-0.06
NewsSoc	0.01	-0.35	0.03 **	0.03	-0.00	0.04	-0.00	-0.02
NewsSoc5	0.01	-0.42	0.01	0.01	-0.03 **	0.05	-0.00	0.02
NewsSoc7	-0.02	-0.49	0.01	0.00	-0.02	0.05	0.00	0.00
NewsSoc10	-0.02	-0.34	0.01	0.02	-0.02	0.05	-0.01	-0.00
NewsSoc20	0.03	-0.35	0.02	0.01	-0.03 *	0.05	-0.02	0.03
NewsSoc50	0.02	-0.30	0.03 **	0.01	-0.04 ***	-0.02	-0.01	-0.03
NewsSoc100	-0.01	-0.54	0.03 **	-0.03	-0.06 ***	0.03	0.01	-0.02

Table 2.3
Predicting Returns with Aggregate Sentiment Measures

This table reports the following regression

$$r_{t+1\to t+h} = \alpha + \beta \times x_t + \epsilon_{t+1\to t+h},$$

where $r_{t+1\to t+h}$ is the cumulative asset returns over the next h days and x_t is an orthogonalized aggregate sentiment measure. Average Variable (AV) means average across all trade sentiment variables (Trade AV) and separately average across all text sentiment variables (Text AV). Principal Component Analysis (PCA) uses the PCA index constructed with all trade sentiment variables (Trade PCA) and separately with all text sentiment variables (Text PCA). Partial Least Squares (PLS) is constructed using the next one day returns to supervise the constructions. PCA and PLS indexes for trade (text) sentiment measures are adjusted to have the same sign as Trade (Text) AV to remain the sentiment interpretation. Returns are in percentages and independent variables are standardized to zero mean and unit standard deviation. Reported in parenthesis are t-statistics corrected for Newey and West (1987)'s standard errors with h lags (if $h < 5$, lag is set to 5). ***, **, and * indicate significance at 1%, 5%, and 10% respectively. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Table 2.3
Predicting Returns with Aggregate Sentiment Measures (Cont.)

	Panel A: Bitcoin				Panel B: T-Bond				Panel C: Stock				Panel D: Gold			
	h=1	h=5	h=10	h=20	h=1	h=5	h=10	h=20	h=1	h=5	h=10	h=20	h=1	h=5	h=10	h=20
Average Variable																
Trade AV	0.16 **	0.66 ***	1.03 ***	1.44 ***	0.02	0.09 ***	0.06	0.07	0.05 **	0.05	0.02	-0.01	0.03	0.08 *	0.08	0.06
	(2.53)	(4.04)	(3.72)	(3.37)	(0.93)	(2.92)	(1.34)	(1.03)	(2.15)	(1.19)	(0.34)	(-0.11)	(1.63)	(1.94)	(1.28)	(0.66)
IS R2 (%)	0.23	0.61	0.66	0.52	0.02	0.24	0.03	0.01	0.14	0.02	-0.01	-0.02	0.04	0.09	0.03	-0.01
Text AV	0.00	-0.10	-0.24	-0.52	0.03 **	0.01	0.00	0.00	-0.03 *	0.05	0.05	0.09	0.01	0.01	0.01	-0.00
	(0.08)	(-0.63)	(-0.81)	(-1.18)	(2.27)	(0.34)	(0.10)	(0.06)	(-1.79)	(1.43)	(1.18)	(1.15)	(0.36)	(0.33)	(0.17)	(-0.03)
IS R2 (%)	-0.03	-0.02	0.01	0.04	0.08	-0.02	-0.02	-0.02	0.04	0.02	0.01	0.02	-0.02	-0.02	-0.02	-0.02
Trade AV	0.16 **	0.67 ***	1.05 ***	1.47 ***	0.02	0.09 ***	0.06	0.07	0.05 **	0.05	0.02	-0.01	0.03	0.08 *	0.08	0.06
	(2.54)	(4.13)	(3.83)	(3.51)	(0.87)	(2.93)	(1.34)	(1.03)	(2.16)	(1.18)	(0.33)	(-0.12)	(1.62)	(1.94)	(1.28)	(0.66)
Text AV	-0.00	-0.14	-0.30	-0.61	0.03 **	0.01	0.00	0.00	-0.03 *	0.05	0.05	0.09	0.01	0.01	0.01	-0.00
	(-0.09)	(-0.88)	(-1.03)	(-1.39)	(2.21)	(0.24)	(0.06)	(0.03)	(-1.83)	(1.41)	(1.18)	(1.15)	(0.32)	(0.29)	(0.15)	(-0.04)
IS R2 (%)	0.20	0.61	0.69	0.59	0.09	0.22	0.01	-0.01	0.18	0.04	-0.01	0.00	0.02	0.07	0.01	-0.03
PCA																
Trade PCA	0.14 **	0.47 ***	0.71 ***	0.81 **	0.01	0.11 ***	0.07 *	0.09	0.05 **	0.09 **	0.07	0.13 **	0.01	0.09 **	0.07	0.05
	(2.00)	(2.77)	(2.72)	(2.07)	(0.44)	(3.08)	(1.66)	(1.49)	(2.14)	(2.22)	(1.40)	(2.09)	(0.83)	(2.40)	(1.22)	(0.77)
IS R2 (%)	0.18	0.30	0.30	0.14	-0.01	0.30	0.06	0.04	0.14	0.11	0.02	0.06	-0.01	0.12	0.01	-0.01
Text PCA	0.00	-0.10	-0.24	-0.49	0.03 **	0.01	0.00	0.01	-0.03 *	0.05	0.05	0.09	0.01	0.01	0.01	-0.00
	(0.09)	(-0.64)	(-0.81)	(-1.14)	(2.15)	(0.26)	(0.06)	(0.11)	(-1.82)	(1.35)	(1.15)	(1.14)	(0.32)	(0.33)	(0.16)	(-0.01)
IS R2 (%)	-0.03	-0.02	0.01	0.03	0.07	-0.02	-0.02	-0.02	0.04	0.02	0.01	0.02	-0.02	-0.02	-0.02	-0.02
Trade PCA	0.14 **	0.47 ***	0.72 ***	0.81 **	0.01	0.11 ***	0.07 *	0.09	0.05 **	0.09 **	0.07	0.13 **	0.01	0.09 **	0.07	0.06
	(2.00)	(2.78)	(2.75)	(2.09)	(0.38)	(3.09)	(1.67)	(1.49)	(2.15)	(2.22)	(1.40)	(2.09)	(0.81)	(2.40)	(1.23)	(0.78)
Text PCA	0.00	-0.11	-0.25	-0.50	0.03 **	0.00	-0.00	0.00	-0.03 *	0.05	0.05	0.08	0.01	0.01	0.01	-0.00
	(0.06)	(-0.68)	(-0.85)	(-1.16)	(2.09)	(0.10)	(-0.02)	(0.04)	(-1.85)	(1.34)	(1.14)	(1.14)	(0.27)	(0.20)	(0.10)	(-0.05)
IS R2 (%)	0.15	0.29	0.31	0.18	0.06	0.28	0.04	0.02	0.19	0.12	0.03	0.08	-0.03	0.10	-0.01	-0.03
PLS																
Trade PLS	0.28 ***	0.59 **	1.10 ***	2.44 ***	0.04 ***	0.08 *	0.12 *	0.03	0.05 ***	0.07	0.04	0.21 *	0.03 *	0.03	0.17	0.35 **
	(4.82)	(2.57)	(2.74)	(3.38)	(3.10)	(1.77)	(1.74)	(0.25)	(2.97)	(1.31)	(0.51)	(1.79)	(1.74)	(0.53)	(1.55)	(2.21)
IS R2 (%)	0.79	0.47	0.76	1.56	0.16	0.16	0.20	-0.01	0.17	0.07	-0.00	0.18	0.05	-0.01	0.20	0.50
Text PLS	0.08	0.19	-0.15	-0.10	0.01	-0.03	-0.02	-0.00	0.03 **	0.02	0.03	0.05	-0.01	-0.04	0.03	0.04
	(1.34)	(0.81)	(-0.38)	(-0.15)	(0.67)	(-0.85)	(-0.34)	(-0.02)	(2.09)	(0.35)	(0.30)	(0.32)	(-0.74)	(-0.59)	(0.28)	(0.23)
IS R2 (%)	0.04	0.02	-0.02	-0.03	-0.01	0.01	-0.01	-0.02	0.05	-0.01	-0.01	-0.00	-0.01	0.00	-0.02	-0.02
Trade PLS	0.28 ***	0.59 ***	1.10 ***	2.44 ***	0.04 ***	0.08 *	0.12 *	0.03	0.05 ***	0.07	0.04	0.21 *	0.03 *	0.03	0.17	0.35 **
	(4.84)	(2.58)	(2.74)	(3.38)	(3.10)	(1.77)	(1.74)	(0.25)	(2.96)	(1.31)	(0.51)	(1.79)	(1.74)	(0.53)	(1.56)	(2.21)
Text PLS	0.09	0.20	-0.14	-0.09	0.01	-0.03	-0.02	-0.00	0.03 **	0.02	0.03	0.05	-0.01	-0.04	0.03	0.04
	(1.39)	(0.83)	(-0.36)	(-0.14)	(0.67)	(-0.86)	(-0.35)	(-0.02)	(2.09)	(0.34)	(0.30)	(0.32)	(-0.73)	(-0.59)	(0.29)	(0.24)
IS R2 (%)	0.83	0.50	0.75	1.53	0.15	0.17	0.19	-0.03	0.22	0.06	-0.01	0.18	0.04	-0.00	0.19	0.48

Table 2.4
Out-Of-Sample R^2

This table reports the out-of-sample R^2_{OS} in percentages from recursively predicting the out-of-sample returns over the next h days using the following methods: average of forecasting variables (AV), average of individual forecasts (CF), principal component analysis (PCA), and partial least squares (PLS). The estimation employs an expanding window with an initial training window of 750 days. ***, **, and * indicate significance at 1%, 5%, and 10% respectively, based on the Clark and West (2007)'s statistic. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Panel A: Bitcoin					Panel B: T-Bond				
	h=1	h=5	h=10	h=20		h=1	h=5	h=10	h=20
Trade AV	1.57 ***	2.09 ***	2.71 ***	3.65 ***	Trade AV	-0.05	-0.19	-0.11	-0.09
Trade CF	0.22 ***	0.19 ***	0.15 **	0.13 ***	Trade CF	-0.02	0.06 **	0.03	0.04 *
Trade PCA	0.18 **	0.09 **	-0.11 *	-0.25	Trade PCA	-0.13	0.16 **	-0.03	-0.12
Trade PLS	0.48 ***	-0.35 ***	-0.54 *	-0.53 **	Trade PLS	-0.73	-0.06	-0.12	-0.08 **
Text AV	-0.03	-0.21	-0.33	-0.74	Text AV	-0.06	-0.16	-0.39	-0.69
Text CF	-0.03	-0.04	-0.02	-0.00	Text CF	0.02	-0.03	-0.03	-0.03
Text PCA	-0.04	-0.16	-0.08	-0.02	Text PCA	0.06 **	-0.06	-0.06	-0.05
Text PLS	-0.58	-0.08	0.04	-0.04	Text PLS	-0.27	-0.23	-0.15	-0.29
Panel C: Stock					**Panel D: Gold**				
	h=1	h=5	h=10	h=20		h=1	h=5	h=10	h=20
Trade AV	-0.01 *	-0.28	-0.68	-1.67	Trade AV	-0.11	-0.58	-0.95	-1.27
Trade CF	0.05	0.01	-0.02	0.00	Trade CF	-0.03	0.03	0.05 *	0.03
Trade PCA	0.12 *	-0.08	-0.09	-0.14	Trade PCA	-0.06	0.07 *	-0.04	-0.06
Trade PLS	-0.55	-0.61	-0.80	-0.57 *	Trade PLS	-1.07	-0.79	-0.66 *	-0.53 *
Text AV	-0.09	-0.38	-0.85	-1.59	Text AV	-0.17	-0.77	-1.17	-0.77
Text CF	0.00	0.01	0.01	0.04 *	Text CF	-0.04	-0.07	-0.07	-0.13
Text PCA	-0.04	-0.01	-0.02	-0.07	Text PCA	-0.09	-0.15	-0.23	-0.42
Text PLS	-0.03 **	-0.28	-0.31	-0.24	Text PLS	-0.27	-0.38	-0.29	-0.35

Table 2.5
Forecast Encompassing Tests

This table reports the p-value of testing the null hypothesis that the out-of-sample forecasts over the next h-day returns made by the models in the columns encompass the forecasts made by the models in the rows. p-value greater than 10% are in bold, i.e. those entries indicates the failure to reject the null that the forecast made by the model in the column encompasses the forecast made by the model in the row. For trade and text sentiment measures separately, the four prediction models include average of forecasting variables (AV), average of individual forecasts (CF), principal component analysis (PCA), and partial least squares (PLS). Grey areas indicate where the trade and text sentiment measures are tested against each other. The estimation employs an expanding window with an initial training window of 750 days. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Panel A: Bitcoin

	Trade AV	Trade CF	Trade PCA	Trade PLS	Text AV	Text CF	Text PCA	Text PLS
				h=1				
Trade AV		0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trade CF	**0.34**		**0.14**	0.00	0.00	0.00	0.00	0.00
Trade PCA	**0.28**	**0.23**		0.00	0.03	0.03	0.03	0.00
Trade PLS	0.00	0.00	0.00		0.00	0.00	0.00	0.00
Text AV	**0.53**	**1.00**	**0.38**	0.00		**0.46**	**0.31**	0.00
Text CF	**0.46**	**1.00**	**0.37**	0.00	**0.33**		**0.22**	0.00
Text PCA	**0.46**	**1.00**	**0.37**	0.00	**0.36**	**0.60**		0.00
Text PLS	**0.64**	**1.00**	**0.80**	0.01	**0.95**	**0.98**	**0.96**	
				h=20				
Trade AV		0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trade CF	**0.45**		0.01	0.00	0.00	0.01	0.02	0.02
Trade PCA	**0.45**	**0.56**		0.00	0.00	**0.24**	**0.22**	**0.17**
Trade PLS	**0.20**	0.10	0.03		0.00	0.03	0.02	0.02
Text AV	**0.56**	**1.00**	**0.41**	0.00		**1.00**	**1.00**	**1.00**
Text CF	**0.47**	**0.96**	0.02	0.00	0.00		**0.30**	**0.18**
Text PCA	**0.43**	**0.88**	0.02	0.00	0.00	**0.51**		**0.18**
Text PLS	**0.53**	**0.82**	0.02	0.00	0.00	**0.46**	**0.32**	

Panel B: T-Bond

	Trade AV	Trade CF	Trade PCA	Trade PLS	Text AV	Text CF	Text PCA	Text PLS
				h=1				
Trade AV		**0.61**	0.07	0.00	**0.21**	**0.90**	**0.89**	0.00
Trade CF	**0.14**		0.03	0.00	**0.14**	**0.89**	**0.84**	0.00
Trade PCA	**0.62**	**0.88**		0.00	**0.45**	**0.88**	**0.89**	0.01
Trade PLS	**0.82**	**0.91**	**0.87**		**0.76**	**0.88**	**0.89**	**0.30**
Text AV	**0.33**	**0.50**	**0.10**	0.00		**0.68**	**0.71**	0.00
Text CF	0.02	0.06	0.02	0.00	0.05		**0.82**	0.00
Text PCA	0.01	0.03	0.01	0.00	0.02	0.06		0.00
Text PLS	**0.53**	**0.65**	**0.28**	0.00	**0.45**	**0.82**	**0.86**	
				h=20				
Trade AV		**0.60**	0.10	0.01	0.00	**0.39**	**0.29**	0.00
Trade CF	0.03		0.06	0.01	0.00	0.02	0.02	0.00
Trade PCA	**0.14**	**0.52**		0.05	0.00	**0.30**	**0.24**	0.02
Trade PLS	0.01	0.08	0.03		0.00	0.03	0.02	0.00
Text AV	**0.36**	**0.67**	**0.23**	0.06		**0.59**	**0.58**	**0.11**
Text CF	0.09	**0.93**	0.10	0.01	0.00		**0.12**	0.00
Text PCA	**0.12**	**0.91**	**0.11**	0.01	0.00	**0.75**		0.00
Text PLS	**0.41**	**0.94**	**0.25**	0.04	0.00	**0.89**	**0.79**	

Panel C: Stock

	Trade AV	Trade CF	Trade PCA	Trade PLS	Text AV	Text CF	Text PCA	Text PLS
				h=1				
Trade AV		**0.13**	0.06	0.00	0.04	**0.12**	0.08	0.01
Trade CF	0.04		**0.33**	0.00	0.01	**0.11**	0.05	0.01
Trade PCA	0.01	**0.11**		0.00	0.03	0.06	0.04	0.01
Trade PLS	0.09	**0.44**	**0.90**		**0.18**	**0.26**	**0.20**	0.08
Text AV	**0.15**	**0.94**	**0.45**	0.00		**0.88**	**0.42**	0.07
Text CF	0.10	**0.72**	**0.31**	0.00	0.01		**0.11**	0.03
Text PCA	**0.13**	**0.66**	**0.33**	0.00	0.06	**0.65**		0.04
Text PLS	0.01	0.08	0.08	0.00	0.01	0.07	0.03	
				h=20				
Trade AV		**0.47**	**0.23**	0.00	0.00	**0.45**	**0.20**	**0.13**
Trade CF	0.00		0.00	0.00	0.00	**0.47**	0.01	0.00
Trade PCA	0.00	**0.88**		0.00	0.00	**0.73**	0.06	0.01
Trade PLS	0.00	**0.14**	**0.15**		0.00	0.09	0.02	0.01
Text AV	0.00	**1.00**	**0.98**	0.02		**1.00**	**0.98**	**0.88**
Text CF	0.00	0.08	0.00	0.00	0.00		0.01	0.00
Text PCA	0.00	0.10	0.01	0.00	0.00	**0.60**		0.00
Text PLS	0.00	**0.51**	**0.12**	0.00	0.00	**0.54**	0.04	

Panel D: Gold

	Trade AV	Trade CF	Trade PCA	Trade PLS	Text AV	Text CF	Text PCA	Text PLS
				h=1				
Trade AV		**0.83**	**0.66**	0.00	**0.20**	**0.78**	**0.42**	0.02
Trade CF	0.04		**0.11**	0.00	**0.12**	**0.35**	**0.12**	0.01
Trade PCA	**0.13**	**0.80**		0.00	**0.16**	**0.68**	**0.25**	0.01
Trade PLS	**0.70**	**0.86**	**0.79**		**0.57**	**0.78**	**0.72**	**0.32**
Text AV	**0.39**	**0.61**	**0.54**	0.00		**0.60**	**0.45**	0.07
Text CF	0.06	**0.50**	**0.18**	0.00	**0.12**		0.07	0.01
Text PCA	**0.28**	**0.75**	**0.57**	0.00	**0.17**	**0.87**		0.02
Text PLS	**0.25**	**0.43**	**0.36**	0.00	**0.21**	**0.53**	**0.34**	
				h=20				
Trade AV		**0.84**	**0.76**	0.01	0.00	**0.68**	**0.32**	**0.30**
Trade CF	0.00		0.01	0.00	0.00	0.02	0.00	0.00
Trade PCA	0.00	**0.98**		0.00	0.00	**0.11**	0.01	0.01
Trade PLS	0.00	0.08	0.03		0.00	0.02	0.00	0.00
Text AV	0.00	0.03	0.02	0.00		0.01	0.00	0.00
Text CF	0.00	**0.93**	**0.71**	0.00	0.00		0.01	0.02
Text PCA	0.00	**0.90**	**0.82**	0.00	0.00	**0.94**		**0.18**
Text PLS	0.00	**0.71**	**0.57**	0.00	0.00	**0.57**	0.07	

Table 2.6
Asset Allocation

This table reports utility gains for a mean-variance investor who allocates between a risky asset and a risk-free asset using the one-day return forecast by a predictive model against the forecast using the historical mean benchmark. **UG** means annualized utility gains in percentages while **SR** means annualized Sharpe ratio of the resulting portfolio. ***, **, and * indicate significance at 1%, 5%, and 10% respectively, based on the tests in DeMiguel et al. (2009). Stars under the UG columns test the significance of the utility gains and stars under the SR columns test whether the Sharpe ratio of the portfolio using the predictive model is different from that using the historical mean return. The estimation employs an expanding window with an initial training window of 750 days. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Panel A: Bitcoin

	$\gamma = 3$		$\gamma = 5$	
	UG	SR	UG	SR
Trade AV	61.72 ***	2.00 ***	48.42 ***	2.08 ***
Trade CF	16.17 ***	0.92 ***	13.33 ***	0.90 ***
Trade PCA	35.76 ***	1.49 ***	33.36 ***	1.72 ***
Trade PLS	50.29 ***	1.73 ***	37.06 ***	1.77 ***
Text AV	2.53	0.50	2.27 **	0.41 *
Text CF	-1.49	0.38	-0.87	0.29
Text PCA	1.06	0.47	2.01	0.44 *
Text PLS	-11.60	0.10	-9.05	-0.00
Buy-Hold		0.76		0.76

Panel B: T-Bond

	$\gamma = 3$		$\gamma = 5$	
	UG	SR	UG	SR
Trade AV	-0.37	0.22	-0.39	0.22
Trade CF	-0.21	0.23	-0.02	0.25
Trade PCA	-2.20	0.10	-2.44	0.05
Trade PLS	-0.40	0.21	-0.67	0.24
Text AV	-1.20	0.19	-1.23	0.20
Text CF	0.03	0.25	0.11	0.26
Text PCA	1.22	0.32	1.16	0.35
Text PLS	-1.07	0.16	-1.52	0.15
Buy-Hold		0.32		0.32

Panel C: Stock

	$\gamma = 3$		$\gamma = 5$	
	UG	SR	UG	SR
Trade AV	-0.04	0.49	0.21	0.52
Trade CF	-0.37	0.49	-0.89	0.40
Trade PCA	-1.66	0.38	-2.21	0.29
Trade PLS	-1.49	0.39	-1.37	0.40
Text AV	-1.23	0.41	-0.78	0.43
Text CF	-0.82	0.44	-0.95	0.39
Text PCA	-1.59	0.38	-1.70	0.32
Text PLS	-0.88	0.43	-0.35	0.48
Buy-Hold		0.36		0.36

Panel D: Gold

	$\gamma = 3$		$\gamma = 5$	
	UG	SR	UG	SR
Trade AV	-0.77	0.26	-1.23	0.25
Trade CF	-0.40	0.27	-0.49	0.28
Trade PCA	-0.44	0.27	-0.42	0.28
Trade PLS	-2.15	0.15	-3.22	0.12
Text AV	-1.95	0.13	-0.46	0.19
Text CF	-0.93	0.24	-0.37	0.28
Text PCA	-1.85	0.19	-1.11	0.23
Text PLS	-0.17	0.25	-1.30	0.21
Buy-Hold		0.45		0.45

Appendix

2.A Replication of Garcia (2013)

Table 2.A.1
Predicting Returns with News Sentiment from Garcia (2013)

This table reports the results of the following predictive regression:

$$r_t = \alpha + \beta' L_{1-5}(X_t) + \gamma' Z_t + \epsilon_t,$$

where r_t is the return of the DJIA index on the day t, $L_{1-5}(X_t)$ is five lags of one of news sentiment variables—positive (POS), negative (NEG), or POS-NEG (SENT)—from Garcia (2013), and Z_t is a set of control variables including five lags of daily returns, five lags of squared daily returns, and weekday indicators. To conserve space, only β's are reported. Returns and in-sample R^2's are in percentage points. Reported in parenthesis are t-statistics corrected for Newey and West (1987) standard errors with five lags. ***, **, and * indicate significance at 1%, 5%, and 10% respectively. Panel A reports the sample from 1905 to 2005 as in Garcia (2013), Panel B reports the sample from 1984 to 1999 as in Tetlock (2007), and Panel C reports the sample from 2000 to 2005.

	β_1	β_2	β_3	β_4	β_5	R^2	Obs
			Panel A: 1905-2005				
POS	0.04 ***	0.00	-0.01	-0.01 *	-0.00	1.83	27427
	(5.43)	(0.55)	(-0.90)	(-1.85)	(-0.54)		
NEG	-0.04 ***	0.00	0.01	0.01	0.01	1.82	27427
	(-5.24)	(0.36)	(0.70)	(1.29)	(1.56)		
SENT	0.05 ***	-0.00	-0.01	-0.02 *	-0.01	1.91	27427
	(6.53)	(-0.12)	(-0.98)	(-1.89)	(-1.51)		
			Panel B: 1984-1999				
POS	0.02	0.01	-0.00	0.00	-0.00	4.41	4044
	(1.52)	(0.45)	(-0.21)	(0.08)	(-0.22)		
NEG	-0.05 ***	-0.03 **	-0.01	0.02	0.01	4.72	4044
	(-3.04)	(-1.99)	(-0.39)	(1.51)	(0.48)		
SENT	0.05 ***	0.03 *	0.00	-0.02	-0.01	4.69	4044
	(3.03)	(1.78)	(0.22)	(-1.08)	(-0.57)		
			Panel C: 2000-2005				
POS	-0.03	-0.02	-0.00	-0.03	0.07 *	0.38	1501
	(-0.83)	(-0.61)	(-0.12)	(-0.81)	(1.89)		
NEG	0.01	0.03	0.04	-0.06 **	0.01	0.46	1501
	(0.31)	(0.94)	(1.44)	(-2.21)	(0.41)		
SENT	-0.02	-0.03	-0.04	0.04	0.02	0.32	1501
	(-0.54)	(-1.00)	(-1.21)	(1.38)	(0.62)		

2.B Refinitiv MarketPsych Indices (RMI) *Text Sentiment* Construction

Text sentiment is a news content index gathered by Refinitiv MarketPsych Indices (RMI). It is based on word recognition techniques designed to extract relevant economics, finance, business, and psychology data from financial news, social media, conference call transcripts, and executive interviews. Since 1998, statistics have been accessible for 45 currencies, over 12,000 firms from over 75 countries, and index-level data for the 15 largest markets. Also accessible are data for 187 nations and regions. Only English-language text is used for analysis.

2.B.1 Data Analysis

Most previous research uses a lexical analysis technique in which an article's text is compared to a context-relevant vocabulary that has been pre-specified, and their frequencies are then tallied. This method's major shortcoming is focusing solely on one component of *Text sentiment* (positive versus negative) while ignoring other partially linked variables. Also, certain open-source *Text sentiment* dictionaries may misclassify topics linked to business and money. RMI word recognition technique uses extensive customization and curation of lexicons. It aims to identify and score hundreds of *Text sentiment* dimensions using different grammatical frameworks to various text sources based on their unique characteristics and takes sentence and article structure into account. The convergence of the text processing techniques described below, when applied to text, yields approximately 400 psychological variables ($PsychVars$)), each of which has the potential to be applied to a different entity. It should be noted that the method seeks to find word interrelationships and calculates the score based on these instead of employing frequency analysis of individual words. Following is a discussion of the word recognition and content quantification techniques used in data generation.

2.B.2 Source Type Differences

The fact that RMI data are taken from traditional news and social media sources creates a problematic environment for word recognition techniques. Due to the differences in communication patterns between social and news media, each source's data analysis is subsequently tailored.

In contrast to mainstream media, which employ standard terminology and acceptable tone in their material, social media are rife with sarcasm, irony, colloquial meanings, incomplete thoughts, incorrect punctuation, misspellings, non-standard syntax, case insensitivity, and harsh language. The expressions of emotions differ between the news and social media. In social media, authors frequently express themselves through the use of emoticons (e.g., ":)") and acronyms (e.g., "LOL").

Moreover, unlike journalists of traditional news media, who are educated to provide many perspectives on the underlying issue and undergo an editorial process, authors in social media tend to convey their thoughts and emotional states more directly.

The duty of journalists in the news is to describe the emotional states of individuals being reported on. Consequently, social media content is often less inclusive of opposing perspectives and more emotionally expressive from the first-person perspective than news information.

Due to these variances, text analytical models are utilized to calibrate text analytics by source type, and separate models are employed for news, social media forums, tweets, SEC filings, and earnings conference call transcripts.

2.B.3 Entity Identification

A list including more than 60,000 entity names and their aliases, such as language-specific writing forms of locations and businesses, is used to identify entities.

RMI utilizes anti-correlate filters to eliminate irrelevant elements. For instance, gold and silver are typically referenced every two years at the Olympic Games, although they have no relation to gold and silver commodities. Similarly, allusions to the South Korean Won may allude to the successes of South Korean athletes or the country's currency. Additionally, the entities must have accurate co-references. For instance, "breakfast oats" is not a valid reference to the commodity Oats unless it is accompanied by crucial identification correlates such as "prices" or "futures."

2.B.4 Timing

RMI has various expectations-related variables. The text-analytics program is tuned to identify verb tenses in each phrase and recognize whether references are future-oriented and related to anticipations. For instance, "Optimism" distinguishes between references to future-oriented positive and negative statements, but "Uncertainty" excludes references to historical uncertainties from its analysis.

2.B.5 Modifiers and Negations

By changing the impact of a phrase or sentence, modifiers alter its meaning or tone. For instance, words or phrases that raise the importance of an adjective, such as "big" (e.g. "great loss"), are multiplicative on the weighting of the modified word, but minimizers such as "a little" and "a handful" multiply the meaning score by 0.5. Modifiers may enhance (maximizers) or decrease (minimizers) the score of a critical term, influencing the scoring of textual meaning. In the case of frequent interpretations based on fundamental word relationships, such as "new," a multiplier is employed to reduce the significance of the meaning. In the Innovation index, the multiplier for the word "new" is 0.1 when used with this meaning. The data also considers the distinctions between news headlines and article bodies. In the data creation, headlines have a multiplicative weight of 3, subtitles a weight of 2, and bodies a weight of 1. In addition to maximization and minimization, the data construction considers four negations. For instance, "I'm not worried about the earnings release" implies that the author is not fearful; hence, the extracted fear score is nullified (-1).

2.B.6 Source Texts

The RMI measures are derived from two categories of sources, news, and social media, and the data consists of three feeds: a news feed, a social media feed, and a combined news and social media feed. RMI analyses about two million articles every day, and the data are updated on a minute-by-minute basis. Each minute number represents the average of the previous 24 hours' observations regarding the target entity.

RMI uses data from the largest and most influential news organizations, global online news coverage, and a wide variety of trustworthy social media sites as its sources. The RMI news indexes are derived from content provided by Thomson Reuters News Feed Direct and two Thomson Reuters news archives: a Reuters-only archive from 1998 to 2002 and one containing Reuters and chosen third-party wires from 2003 onwards. In 2005, Moreover Technologies' aggregate news feed was added to the data, increasing the number of sources by 50,000 online news sites. In addition, the data includes hundreds of financial news websites and finance-specific sources that professional investors extensively read. Since 1998, MarketPsych has downloaded social media content from public social media sites. In 2009, RMI added data from Moreover Technologies' aggregate social media feed, collected from more than four million social media sites. v

2.B.7 Quantifying the Articles

Each RMI variable constitutes a combination of minor components called *PsychVars*. To form an RMI variable, the absolute values of *PsychVars* are first determined using 24 hours of observations. These absolute values are then summed for all constituents to get the Buzz variable. More specifically, Buzz is calculated as follows:

$$Buzz_t(a) = \sum_{c \in C(a), p \in P(s)} |PsychVar_{p,t}(c)|,$$

where t denotes time, a is the entity, $C(a)$ is the set of all constituents of a (for example, for indices, $C(a)$ consists of all individual assets comprising the index and for respective companies $C(a) = a$) and $P(s)$ is the set of all *PsychVars* relevant to a particular RMI variable s.

The value of an RMI variable s of an entity an at time t is calculated as follows:

$$RMI_{s,t}(a) = \frac{\sum_{c \in C(a), p \in P(s)} (I_t(s,p) * PsychVar_{p,t}(c))}{Buzz_t(a)}$$

where $I_t(s,p)$ defines whether a *PsychVar* $p \in P(s)$ is additive or subtractive to an RMI variable s at time t:

$$I_t(s,p) = \begin{cases} +1, \text{if additive at time t} \\ -0, \ \text{if subtractive at time t} \end{cases}$$

A single *PsychVar* can contribute to multiple RMI variables. For example, EarningsUp_f *PsychVar* (see below for an example) is a constituent of all, EarningsForecast, Text Sentiment, Optimism, and FundamentalStrength RMI variables.

2.B.8 Sentence-level Example

The following shows an example of the RMI data-generating process using the above-mentioned principles. Consider the following sentence: “Analysts expect Mattel to report much higher earnings next quarter.” The language analyzer performs the following sequence:

[1] Associates ticker symbol MAT with entity reference “Mattel.”

[2] Identifies “earnings” as an Earnings word in the lexicon.

[3] Identifies “expect” as a future-oriented word and assigns future tense to the

phrase.

[4] Identifies “higher” as an Up-Word (positive reference).

[5] Multiplies “higher” by 2 due to presence of the modifier word “much.”

[6] Associates “higher” (Up-Word) with “earnings” (Earnings) due to proximity.

The analysis algorithm will report the following:

Date	Time	Ticker	PsychVar	Score
20110804	15:00.123	MAT	EarningsUp_f	2

In the example above, 2 is the raw score produced for EarningsUp_f.

Chapter 3

War Discourse and Disaster Premia: 160 Years of Evidence from Stock and Bond Markets

3.1 Introduction

I study here the asset pricing implications of the prospect of disasters, such as wars, pandemics, and political crises. Rare disaster risks are one of the leading possible rational explanations for major asset pricing puzzles (Barro 2006, 2009). A basic implication of rational disaster models is that high disaster risk will receive a risk premium, and therefore will predict high future stock market returns. Two behavioral hypotheses offer a similar implication. The first is attentional and belief-based: that investors overestimate the probability of rare disasters owing to the high salience of extreme outcomes. This possibility is supported by evidence that people overestimate the probabilities of rare events (Fischhoff et al. 1977; Snowberg and Wolfers 2010). The second is preference-based: that investors overweight low probabilities, as in the expected value function of cumulative prospect theory (Tversky and Kahneman 1992).

The rarity of major disasters is a well-known obstacle to empirically testing the relationship between their occurrence and future stock returns.[1] This limits the power to identify effects. In this paper, I circumvent the obstacle of small sample size by using data on investors' attention to rare disaster risks derived from news. This provides a much larger sample of changing perceptions of disaster probabilities over 160 years.

[1]An international political crisis occurs on average once every 15 years, a full-scale war once every 74 years, fighting on home territory once every 119 years, and a pandemic once every 100 years.

Shifts in media topic coverage over time can potentially capture investor assessments of future prospects, including the risk of rare disasters. For example, Shiller (2019) argues that economic narratives are subject to occasional outbreaks when they spread rapidly and widely through the population, influencing behavior and decisions. He argues that such shifts can be captured by media discussions.

The novelty of my approach to capturing perceptions of disaster risk is twofold. First, I am the first study to compare the effects of disaster- and non-disaster-related topics of discourse systematically for the pricing of the aggregate stock market. Second, I apply a novel approach, *seeded Latent Dirichlet Allocation* (Lu et al. 2011, henceforth sLDA), to extract topics of popular discourse over time. This method has several key advantages over existing empirical application of language models to asset pricing.

On the first point, for non-disaster-related topics, I consider the 12 "narratives" (topics) discussed in Shiller (2017, 2019). To obtain interpretable findings, rather than using a purely statistical procedure to extract topics, I draw upon the topics of Shiller (2019). I make necessary adjustments to these topics to effectively implement sLDA. The disaster-related topics I consider are war and pandemics.[2]

I find that non-disaster-focused topics predict stock market returns in-sample and have limited out-of-sample predictive power.[3] The *Pandemics* topic has limited and inconsistent predictive power even in-sample. *War* has the most predictive power both in- and out-of-sample. My study is the first to show that war as a discourse topic is more powerful than non-disaster-focused topics in predicting returns.

[2]Both can have massive human and economic costs and highly uncertain outcomes, as exemplified by the Covid-19 pandemic. Oleg Itskhoki, the winner of the 2022 John Bates Clark Medal, suggests in an interview with Bloomberg on August 2, 2022, that existing wars at that time presented an even greater economic risk than Covid (https://www.bloomberg.com/news/articles/2022-08-02/clark-medal-winner-oleg-itskhoki-says-war-is-a-bigger-economic-risk-than-covid?leadSource=uverify%20wall).

[3]The *Real Estate Booms* topic also yields significant out-of-sample predictive results, but since 1950, its out-of-sample predictive power is weaker that that of the *War* topic. *War* also outperforms the *Real Estate Booms* topic in terms of creating value for real-time investors. For details, see Section 3.6.

On the second point, topic modeling is a prevalent dimension-reduction strategy in the machine learning and natural language processing literature that compresses large amounts of text into a limited set of topics. The topic model I use, sLDA, is a recent extension of the canonical unsupervised LDA model (Blei et al. 2003; Griffiths and Steyvers 2004).[4] Since I use seven million articles published in the *New York Times (NYT)* over 160 years, it is crucial to apply a method that can process a large body of materials with low cost, reasonable speed, and limited error and subjectivity.

Under traditional unsupervised LDA, the model arbitrarily gathers common phrases and themes based on word frequencies. In contrast, under the semisupervised model or sLDA, the creation of themes allows control over the content of themes to be extracted. sLDA fits my research goal of testing the consequences of disaster-focused and non-disaster-focused themes in media discussions. In this approach, I feed the model with the seed words associated with each topic and let the algorithm choose the phrases that often appear with these seed words.[5,6]

My semisupervised topic model performs two key tasks: (1) it classifies market attention throughout the 160 years history of *The New York Times* into several disaster-focused and non-disaster-focused themes, and (2) it traces the evolution of media attention to these themes. These provide new quantitative measures of the market attention to topics of public discourse. Specifically, the model estimates the fraction of an article's text devoted to each topic. Aggregating over articles, these proportions measure the amount of news coverage each topic receives. This makes it

[4]LDA is burgeoning in popularity in computer science and other social science fields. For surveys, see Steyvers and Griffiths (2007), Blei (2012), and Boyd-Graber et al. (2017).

[5]Recent papers in natural language processing, such as Lu et al. (2011), Jagarlamudi et al. (2012), Eshima et al. (2020), and Watanabe and Zhou (2020), have documented the advantages of a (semi)supervised LDA model over the unsupervised one. Among other preferable features, a guided LDA model ensures the interpretability of topics and avoids the need to label extracted topics ex-post to interpret them.

[6]A third group of topic models is fully supervised methods (see for example Mcauliffe and Blei (2007) and Ramage et al. (2009)). These supervised models extract topics predictive of document tags or labels so they need labeled documents to train with. These models are not suitable for us because I want to use topics from news to predict market returns not article titles.

feasible to assess, for instance, the relationship between news and asset returns.

Central to sLDA is the identification words that co-occur with the seed words. My topic weights quantify the market's interest in each topic based on the frequency of terms that co-occur with it. I gather seed words from well-known media and publications, such as *Nature* for disaster-focused topics and from Shiller's book *Narrative Economics* for other topics. As I study 2 disaster-focused topics *War* and *Pandemic* and 12 non-disaster-focused topics discussed in Shiller (2019), I have 14 seeded topics in total. I add one unseeded topic to gather everything not captured by the seeded ones.[7]

There are two key challenges for estimating the predictive power of topics of discourse. First is the need to avoid look-ahead bias; second is the need to address the effects of semantic changes over time. To avoid look-ahead bias, parameter estimates at date t must be based only on the data available before date t.

This problem is intertwined with the second problem, that the meanings of words and phrases evolve over time.[8] Such semantic shifts are extensive since I use 160 years of data. An empirical approach that pools the entire sample to identify word lists and estimate the model parameters would not address such semantic shifts.

To address this issue, my analysis regularly updates the word list that constitutes

[7] Gentzkow et al. (2019) discuss the potential arbitrariness of the number of topics selected for study. In principle, one can employ optimization to determine the number of topics. However, doing so using data for the entire sample period would introduce look-ahead bias, which I seek to avoid. Furthermore, extracting the number of topics each month is unsuitable, as the common topic weights could fluctuate depending on each month's total topic count. Consequently, the weight of topics would not accurately represent market attention; it would instead vary due to the total number of topics for a given month. See also Subsection 3.5.8.

[8] The meanings of words evolve over time spans of a century or more. For instance, "inflation" once referred to an increase in the money supply, but since the early 20$^{\text{th}}$ century it has referred to a general increase in the prices of goods and services in an economy (Homer and Sylla 1996). Another example is "amortization," which once referred only to the reduction of debt over time (Dictionary 1993), but in the 20$^{\text{th}}$ century has come to also refer to the gradual reduction in the value of an asset (Financial Accounting Standards Board (FASB) started this definition in 1973). The definition of "budget" has evolved substantially over time. In the 18$^{\text{th}}$ century, it referred to a financial statement outlining a government's anticipated expenses and revenues for the coming year. By the 1850s, the term expanded to include nongovernmental contexts, eventually encompassing the financial accounts of families or individuals (https://www.merriam-webster.com/words-at-play/financial-word-origins).

topic weights. Although the list of seed words remains unchanged, the model is re-estimated monthly to address these two issues using data from the past ten years of news articles (including the current month). Therefore, the words clustered inside the topics reflect those popularly used during that time and vary monthly depending on semantic shifts. This ability to reflect semantic shifts over time is a key advantage of sLDA. In contrast, monthly rolling estimation is impossible under unsupervised LDA because it is not designed to yield consistent thematic content across estimations. This makes it impossible to address these two challenges.

My estimation process addresses semantic changes by recalculating the topic weight on a monthly rolling forward basis. The output is an article-level weight vector with elements representing the proportion of content (or attention) devoted to the corresponding topic. I then compute the economy-wide monthly time series of topic weights for each topic from the article-level topic weights and use this aggregate time series in my tests of return predictability.

I use news media text to quantify perceived rare disaster risk. In principle, macroeconomic variables can be good proxies for the state variables of conditional asset pricing (Cochrane 1996). Empirically, however, macroeconomic variables perform poorly, with variation that is too low to match asset returns. News presents an alternative proxy of state variables that can potentially address this issue.

My tests of conditional asset pricing are premised on state variables being captured by what investors read in the news. News is available at a high frequency comparable to that of asset returns. While the current literature on rare disaster risks uses contemporaneous price data (Ferguson 2006; Le Bris 2012; Oosterlinck and Landon-Lane 2006) to understand stock returns and bond returns during wartime, we use news, which allows us to continuously and quantitatively track the market attention to, or probability of, disaster risks.[9]

[9] Le Bris (2012) uses an event study approach to infer the effects of disasters. Berkman et al. (2011) use a count of crisis events to proxy for rare disaster risk.

My main results are based on data from the *NYT*; I also verify with robustness tests based on the *Wall Street Journal (WSJ)*. The *NYT* and the *WSJ* are the two largest national media outlets, and thereby can reflect the attention and perception of a large general audience.

Using all articles from *NYT* over 160 years provides several benefits.[10] First, although there is evidence that news about rare disasters affects investors' expectations and the equity return premium (see, for instance, Rietz (1988), Barro (2006), and Julliard and Ghosh (2012)), the rarity of extreme disasters makes it valuable to have a large enough sample to draw clear inferences. Second, long time-series data provides a testing ground for verifying the robustness of effects and tests whether the results are consistent over time (Schwert 1990). The *NYT* has been published since the 1860s, offering a comprehensive historical sample from a continuously available source. Therefore, I use data from the inception of the media to ensure a more accurate representation of the entire spectrum of news topics and their potential impact on stock market returns.

Also, using a broader dataset may be more representative of the topics of actual concern to investors. To my knowledge, this is the first paper that analyzes news articles from *all* newspaper sections of *NYT* since the beginning of its inception. My sample includes nearly seven million articles from the *NYT* and six hundred thousand from the *WSJ*, making it a relatively comprehensive and extensive dataset.

Among the topics extracted from the *NYT*, I find that *War* is the strongest market predictor. The predictive power of *War* for the equity risk premium increases over time. Over 160 years, a one-standard-deviation increase in *War* predicts a 3.80% increase in annualized excess returns in the next month, and the monthly in-sample R^2 is 0.39%. In comparison, over the past 20 years, the respective numbers are 9.83%

[10] I use the first ten years of news data to train my sLDA model to obtain the first month's topic weights and continue rolling the window forward. My market index is unavailable from the Global Financial Data until 1871; thus, my training period starts in 1861, ten years after the *NYT*'s inception in 1851.

and 3.39%. *War* is significant for both subperiods (1871–1949 and 1950–2019). This is economically substantial in comparison with the average annualized monthly excess stock market return over the same period, 3.96%. As another benchmark, the average R^2 of the 40 well-known predictors is only 0.73% in-sample and -1.01% out-of-sample (see Goyal et al. (2021), who discuss the weak out-of-sample performance of most economic return predictors). The R^2 of *War* indicates that its predictive power is economically significant. According to the time-varying disaster model, expected market excess returns should increase with the probability of rare disasters. My results thus provide support for this theory.

In addition to *War*, I construct a discourse topic index from all 14 topics via the two-step Partial Least Squares (PLS) approach of Kelly and Pruitt (2013, 2015). The PLS index heavily loads on *War* with a correlation of 82%, so I interpret the PLS index as a more robust version of *War* that inherits its predictive power. Indeed, the monthly predictive regression of market returns with the PLS index yields a slope of 6.07% and an R^2 of 1.07% over the whole sample and a slope of 12.17% and an R^2 of 5.42% over the past 20 years. For the subperiods (1871–1949, 1950–2019), the PLS index remains a significant predictor at least the 1% level. The predictive power of the PLS index loaded mostly on *War* over market returns is not subsumed by common macroeconomic, sentiment, and uncertainty variables introduced in the literature.

I conduct standard out-of-sample tests as in the return predictability literature to investigate whether discourse topics create value for real-time investors. With expanding window estimation, *War* outperforms all individual economic predictors studied in this paper in terms of out-of-sample R^2 (R^2_{OS}), which compares the forecasting power of a predictor against the historical mean return used as a forecast. The out-of-sample R^2 of *War* is strong, with strongest return predictability in the last twenty years.

Turning to asset allocation implications, I consider a mean-variance investor who forms a portfolio allocating between stocks and a riskfree asset using either the return predictive model or the historical mean return to choose portfolio weights. Using *War* alone or combining topics to guide my portfolio decisions allows us to achieve a higher Sharpe ratio than a simple buy-and-hold strategy. Using a risk aversion coefficient of three following Huang et al. (2015) and Huang et al. (2020), I find the economic gains for the investor utilizing discourse topics in forming portfolios increase over time, consistent with the R^2_{OS} results.

I also find that the perception of rare disaster risks predicts bonds' excess returns. In the model of Gabaix (2012), higher disaster risk implies a higher premium on long-term bonds. My evidence is consistent with this prediction. I find that *War* positively predicts excess returns on mid- to long-term high-yield corporate bonds. Due to flight to quality, investors demand higher premiums to hold these high-yield assets. In contrast, *War* negatively predicts excess returns on safer investment instruments such as short-term government bonds and investment-grade corporate bonds. My bond prediction results hold both in- and out-of-sample.

Contributions. This paper contributes to several lines of research. First, it contributes to social finance and narrative economics in testing how topics of media discourse are related to stock market pricing.

More specifically, my paper contributes to a line of research on the relationship of news content to economic and financial outcomes. In a related study that also uses the sLDA approach, Hirshleifer et al. (2023) find that a factor that is based on my *War* variable explains the cross section of stock returns across a wide range of testing assets, and that leading benchmark factors as well as other media-based uncertainty measures do not subsume its explanatory power. My paper differs in focusing on the time series predictability of the aggregate market return.

My focus on predicting the aggregate market return distinguishes my paper from most existing research on news content and financial outcomes. Bybee et al. (2021) use

LDA on news content to fit *contemporaneous* financial and macroeconomic variables. My paper differs in testing whether disaster-related media discourse predicts *future* stock market returns.[11]

Bybee et al. (2023) aggregate disaster- and non-disaster-focused topics into a set of factors, where *Recession* is the most important topic for explaining the cross section of expected stock returns. Although not the main focus of their paper, they also report that their topics have little power to predict the aggregate market return. In contrast, my focus is on aggregate market return predictability. I consider disaster- and non-disaster-focused topics, and identify that *War* is a powerful predictor of aggregate market returns.

Applying data from many US local newspapers over a century, van Binsbergen et al. (2022) construct a measure of economic sentiment and find that it predicts future economic fundamentals, such as GDP, consumption, and employment growth. My paper differs in comparing the effects of disaster- and non-disaster-related topics of discourse using media discourse to predict future stock and bond returns rather than economic fundamentals.

Perhaps the most closely related papers to this paper are Adämmer and Schüssler (2020) and Manela and Moreira (2017). Both examine how news media can be used to predict aggregate stock market returns. Adämmer and Schüssler (2020) employ a variation of the unsupervised LDA model to extract topics from news articles and use them to forecast market returns. My approach differs from theirs in three key respects. First, they focus on economic news in the *NYT* and *Washington Post* from 1980 to 2018, whereas my study utilizes all *NYT* sections starting from more

[11] There are several other differences between my paper and theirs. First, Bybee et al. (2021) employ the traditional unsupervised LDA model and use cross-validation to select 180 topics in the model. Their machine-selected multiple topics enable them to match contemporaneous economic variables closely. In contrast, I employ a semisupervised LDA model to inject initial seed words to extract the desired themes. My model uses only 15 topics (14 seeded plus one unseeded topic) as I need to extract only the specific topics of interest. Second, while I use all sections of the *NYT* from 1861 to 2019, Bybee et al. (2021) focus on economic news in the *WSJ* from 1984 to 2014.

than a century earlier. As emphasized by Lundblad (2007), since stock returns are highly volatile, it is crucial to consider long time series data to reliably test for return predictability.

Second, the method of Adämmer and Schüssler (2020), as with LDA, generates outputs that are challenging to interpret. Using statistical criteria to optimize over number of topics, the authors choose 100 topics for their model and identify one of them (topic 20) as the most important. The authors conclude that this topic represents geopolitical risk, but it is not obvious on prior conceptual grounds that these words would be the top choices for identifying such risk. (The most important words for this topic are "east," "west" and "German.") In contrast, my approach identifies investor concern with rare disasters. and in particular war, as crucial for return prediction.

Third, since they use an unsupervised topic model, their approach does not address semantic changes over time.[12] In contrast, my use of a semisupervised LDA model allows us to address semantic changes via monthly estimation of the topic model.

The top-line predictability reported in their study is a remarkable out-of-sample R^2 of 6.52%. However, their initial training window for return prediction of three years is too short for a reliable out-of-sample R^2 estimate.[13] In addition to finding a strong new predictor of market returns, my study identifies investor concern with disaster risk, and in particular war, as a determinant of asset pricing.

Manela and Moreira (2017) study how news events affect foreseen volatility and equity risk premia. They construct news implied volatility (NVIX) from the front

[12] They train their topic model using news data from 1980 to 1995 and apply the trained model to extract topic weights from news articles from 1996 to 2018. They argue that because their sample is short, language change is not a concern (their footnote 7).

[13] The initial window used in my study is 10 years. An out-of-sample R^2 is computed by comparing the return forecast of a given model against the forecast using the historical mean return, which cannot be reliably estimated with only three years of data. Consistent with this, their out of sample R^2 is highly sensitive to different start dates of the evaluation sample. As reported in their Table AV, moving the start date of evaluation period from 1999 to 1998 or to 2001 reduces the out-of-sample R^2 to 2.3%, which is comparable to my results over the 2000-2019 sample period.

page of *WSJ* starting from 1890. Using the data provided on the authors' website, I find that over 1900-2016 a standard deviation increase in NVIX^2 is associated with an increase in the annualized market excess return by 0.22% over the next month. A one standard deviation increase in my *War* variable is associated with an annualized market excess return of 2.7%.[14]

My paper differs in several ways. First, my key return predictor is perceived disaster risk, as proxied by textual media, rather than general uncertainty. Perceived disaster risk could potentially be left tail risk rather than variance as studied by Manela and Moreira (2017).

Second, based on their support vector regression model training procedure, their risk measure captures only terms that appeared in the last 20 years of their sample period. This potentially induces look-ahead bias. In contrast, I estimate my model on a monthly rolling basis using data from the preceding 10 years. This allows us to address semantic changes over the 160-year sample period while avoiding look-ahead bias.

Third, I obtain stronger and more robust return predictability. I find that *War* predicts stock market return from one month to three years while their predictor, NVIX^2, does not predict returns until six months ahead. *War* is a powerful predictor in both in- and out-of-sample; Manela and Moreira (2017) do not report out-of-sample R^2.[15]

My paper also contributes to the literature on rare disaster risks, which incorporates disaster probabilities and loss into the standard consumption-based model to explain the high equity premium (Barro 2006, 2009; Gabaix 2012; Wachter 2013). Compared to other studies, I present the most comprehensive test of the predictions

[14] I report these results in Table 3.6. In their paper, Manela and Moreira (2017) report prediction results over 1945-2019 using NVIX^2.

[15] NVIX's in-sample predictive power almost disappears during the last twenty years of their sample period. Also, when I control for NVIX or NVIX^2, *War* is still a substantial and significant return predictor over the entire sample.

made by the time-varying rare disaster model or of the behavioral hypothesis that investors overweight rare disasters.

This paper is also related to the recent literature on extracting measures of political risk from textual data in relation to firm-level hiring and investment (see, e.g., Baker et al. (2016); Hassan et al. (2019); Caldara and Iacoviello (2022)). Several studies, such as Pástor and Veronesi (2013) and Brogaard and Detzel (2015), document that the economic policy uncertainty (EPU) index in Baker et al. (2016) positively predicts the aggregate market return over long horizons. In contrast, my *War* and PLS topic index are stronger predictors of one-month returns. So my index captures a different aspect of textual discourse and risk. My time-series prediction results remain strong after controlling for the geopolitical risk measure (GPR) developed in Caldara and Iacoviello (2022), and yields stronger predictability for stock and bond returns than the dictionary approach of Caldara and Iacoviello (2022) (see Internet Appendix 3.D).

Finally, my paper contributes to the burgeoning literature on applications of modern natural language processing tools to business and economic research. A growing body of research utilizes advanced topic modeling tools to extract thematic content from texts.[16] Unlike most finance papers that use the traditional unsupervised LDA model, my semisupervised LDA model allows us to extract a predefined set of topics in the news, which enhances interpretability.

3.2 Method

In this section, I briefly discuss the setup of the sLDA model (Lu et al. 2011) and my implementation of it to extract news topics.

[16] See, for example, Dyer et al. (2017), Hansen et al. (2018), Larsen and Thorsrud (2019), Choudhury et al. (2019), Brown et al. (2020), and Bybee et al. (2021).

3.2.1 The Stochastic Topic Model

Stochastic topic models are based on the core idea that documents can be described as mixtures of topics, where each topic is associated with a probability distribution over words (Steyvers and Griffiths 2007; Blei 2012). In this approach, latent topic weights are extracted from news articles. To do so, I assume that each text document is generated by a simple stochastic process that 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.

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. I 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 et al. (2003) and further developed by Griffiths and Steyvers (2004). Under LDA, a document is generated under the hierarchical process described above. Each word in the document is selected by first randomly selecting a topic from the document-topic distribution, and then for that topic, the word is selected from the topic-word distribution. Under LDA, the document-topic distribution (a vector of probabilities over the topics) for each document is selected from a prior Dirichlet distribution (see Appendix Section 3.A for details of the LDA and sLDA methodologies). The topic-word distribution is global; it does not depend on the document. It is also assumed to be drawn from a prior Dirichlet distribution. Since the topic-word distribution is a set of probabilities for drawing each possible word, the distribution for the number of instances of each word in an entire document is multinomial with these probabilities, with N being the number of words in the document.

The unknown parameters of the multinomial distributions are estimated using the frequencies of different words in the documents in the sample.[17] Specifically, I use Gibbs sampling to simulate the posterior distribution of words and documents and estimate the two hidden model parameters, namely the document-topic distribution (τ_d) and the topic-word distribution (ω_k).[18]

In traditional unsupervised LDA, only the number of topics K is prespecified; the model clusters words into these topics based on word frequencies in a completely unsupervised manner. The model automatically extracts underlying topics. 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).

Since I am interested in uncovering the effects of specific and interpretable topics relating to rare disaster risk, I instead 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. 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.

3.2.2 Seed Words

A key component of an sLDA model is the set of seed words representing the prior knowledge of each topic. As emphasized by Watanabe and Zhou (2020), a dictionary of seed words needs to be carefully chosen based on field-specific knowledge independent of word frequencies in the collection of texts used. Table 3.1 lists the lemmatized seed words for each topic. (Lemmatization is the removal of word endings such as

[17] An exception is that the two hyperparameters of the two prior Dirichlet distributions are taken from LDA topic modeling literature.

[18] 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.

s, es, ing, ed.) My seed words for *War* include *conflict, tension, terrorism, war* and seed words for *Pandemic* include *epidemic, pandemic.*[19] The seed words need to be general fundamental concepts that have reasonably stable meanings over very long periods. My methodology allows for the fact that the meanings of other words (such as "nuclear") may evolve over time or may even be neologisms that do not exist early in the sample.[20]

The seed words for the non-disaster-focused topics are manually collected from Shiller (2019). These words are discussed extensively in Shiller (2019). I also add certain words that help define the themes of the topics. Importantly, to avoid any look-ahead bias, in selecting the seed words, I exclude any words that were only introduced recently, such as *bitcoin*, *machine learning*, or *great recession.* As shown in Table 3.1, I have reclassified the 9 topics from Shiller (2019) into 12 topics to facilitate my estimation. Specifically, as *Panic* and *Confidence* are opposing notions, I split them into two topics. Similarly, *Frugality versus Conspicuous Consumption* is split into *Frugality* and *Conspicuous Consumption.* I further divide *Real Estate Booms and Bursts* into two separate topics, namely *Real Estates Booms* and *Real Estates Crashes.*[21] In addition to *Stock Market Bubbles*, I add *Stock Market Crashes.* In contrast, because of their similarities, I combine *Labor Saving Machines* and *Automation*

[19] In the setup of LDA, "tension(s)" tends to not be assigned to *War* in documents that talk little about war (such as articles about tension headaches), and to be assigned to *War* in documents that talk a lot about war (such as articles about international tensions).

My results remain robust, and *War* retains similar predictive power, when I include additional seed words such as *battle, front line, army, navy, weapons, military, officer, munitions, bombs, guns* and *battalion.*

[20] During the early 20$^{\text{th}}$ century, the term "nuclear" was primarily employed within the realm of atomic structure and nuclear physics (see Rutherford (2012)). As the mid-20$^{\text{th}}$ century approached, the development and utilization of nuclear weapons during World War II led to an association between "nuclear" and the immense destructive force of such armaments (Rhodes (2012)). In the aftermath of World War II and throughout the Cold War era, "nuclear" was increasingly linked to the application of nuclear technology for energy production (Walker (2004)). Advancing into the late 20$^{\text{th}}$ and early 21$^{\text{st}}$ centuries, the scope of "nuclear" broadened to encompass the concept of nuclear families (Cherlin (2010)).

[21] I replace the term "bursts" with "crashes," as the phrase "real estate burst" is not common in popular usage, and the word "burst" might be taken to mean a burst of positive activity, which is not the intended meaning.

and Artificial Intelligence into one topic.

In addition to the 14 topics discussed above, I include one additional "garbage collector" to absorb everything else in the news unrelated to these topics.

3.2.3 Estimation

Figure 3.1 illustrates the rolling estimation scheme used in the paper. At the end of each month t, I run the sLDA model using all news data over the preceding 120 months (months $t-119$ to t). I use ten years of news data in the monthly estimation to balance the amount of news data required to estimate the model and computational costs. On average, every ten years of historical data consists of around 460,000 articles, which should be sufficient to extract the topic weights at the time of estimation reliably. Notably, within topic models, rolling estimation is viable only under the sLDA model because the seed words that guide this approach provide consistency of thematic content over time.

I use Gibbs sampling to estimate the parameters of the model. I draw 200 drawings from the posterior distribution of z_{dv}, the realized topic for word location v in document d in the sLDA model, where I am conditioning on observed word frequencies.[22] In each drawing, I condition on the estimated values of the parameters of the model derived from previous drawing (where in the first draw, the initial estimate comes from a random number generator). In the last draw, I estimate my final value of the document-topic weights τ_d; that is, I estimate one 14×1 vector $\tau_d = [\tau_d^1, \tau_d^2, \ldots, \tau_d^{14}]$ for each news article, d, in the estimation window.

I then provide estimates of model parameters that condition on month t within the dataset. I compute the global monthly weights of each topic k ($k = 1, 2, \ldots, 14$) as the average weight of each topic across all articles in month t, weighted by the

[22] In addition to the number of topics and articles, the number of samples drawn from the posterior distribution is a computational cost consideration in any topic model.

length $L(d)$ of each article:

$$\tau_t^k = \frac{\sum_{d=1}^{n_t} \tau_d^k L(d)}{\sum_{d=1}^{n_t} L(d)}, \tag{3.1}$$

where τ_t^k is the weight of topic k in month t, n_t is the total number of news articles in month t, and $L(d)$ is the total number of unigrams (one-word terms), bigrams (two-word terms), and trigrams (three-word terms) in article d.[23] (Equal weighting of topic weights across articles yields similar results.)

Although ten years of news articles are used to estimate the model each month, the final topic weights in month t are computed from the news articles of that month only. The final output of the estimation process is a time series of monthly weights for each of the 14 topics. This time series is used for my economic and financial forecasting applications.

3.3 Data

I leverage the richness of full newspaper texts using articles since the beginning of the *NYT* inception. I still remove articles with limited content, such as those that contain mostly numbers, names, or lists. I then conduct standard text processing steps, as reported in detail in the Internet Appendix 3.B and Table 3.C.1. My analysis is based on all articles from the *NYT* since 1861 and all *WSJ* articles since 1990.[24,25] (This

[23] An n-gram is a sequence of n words. For instance, "San Diego" is a bigram, and "A study of topics is needed" is a 6-gram.

[24] The *NYT* has been used in other finance research (see, e.g., Garcia (2013) and Hillert and Ungeheuer (2019)). The *NYT* has received 130 Pulitzer Prizes, almost double that of its nearest competitor. See `https://www.nytco.com/company/prizes-awards/`. Articles from the first ten years of the *NYT* since its inception are used to estimate the first monthly topic weights. I start my *NYT* sample in 1871 as the S&P 500 data is available from that year.

[25] Bybee et al. (2021) use all articles in the *WSJ* with a sample that starts in 1987. My sample starts one hundred years earlier. Manela and Moreira (2017) include the articles from the *WSJ* since 1890 but focus only on the headline and title of articles on the front page, whereas I cover all articles. Baker et al. (2016) and Caldara and Iacoviello (2022) study an extensive collection of newspapers but apply a dictionary approach. Caldara and Iacoviello (2022) count, each month, the number of articles discussing rising geopolitical risks. They do not consider the number of words in each article whereas my methodology does.

data was provided to us by a private company.)

Then, for each month t, I create a document term matrix containing all articles over the preceding ten years through the current month. Each row of the matrix is an article, each column is a n-gram, and each entry is the count of that term in the article. The document-term matrix and topic-based seed words are input into the sLDA model to estimate monthly topic weights as described in the previous section. To streamline the presentation, I report the results for the *WSJ* in Internet Appendix 3.E.

Panel A of Figure 3.C.1 plots the time series of monthly article counts in my sample. After removing articles with limited content, since 1871, my *NYT* data has more than 6.8 million news articles with a monthly average of 3,800.[26] Before 1900, the *NYT* published about 2,000 articles a month. The number of monthly articles increased gradually after 1900, hovering between 4,000 and 6,000 until the end of the twentieth century. Amidst industry-wide struggles related to declining ad revenues and subscriber bases beginning in the 2000s, the *NYT* started scaling down its publishing capacity to around 2,000 articles a month during the 2010s.[27] However, the number of monthly articles surges back to just under 4,000 toward the end of the sample. A newspaper strike occurred from 1902 to 1903, and news articles spiked at the start of World War I.

Panel B of Figure 3.C.1 reports the average monthly article length, which is defined as the total count of unigrams (one-word terms), bigrams (two-word terms), and trigrams (three-word terms). (Internet Appendix 3.B provides more details on the construction of n-grams.) While Bybee et al. (2021) consider only unigrams and bigrams, I extend the analysis to trigrams as a majority of the seed words have three words. Examples include *real estate boom, stock market bubble*, and *cost push*

[26]Data are missing for September and October 1978 (due to strikes) and thus are excluded from Figure 3.C.1.

[27]For more details, see https://www.pewresearch.org/journalism/fact-sheet/newspapers/.

inflation. Over 1871–2019, articles have an average length of 493 n-grams. Articles tended to have around 500 n-grams until the 1920s. After that, the number hovered just above 400 n-grams until the 1960s. Since then, article length has increased, reaching about 600 n-grams during the 2010s.

The second set of data concerns stock market outcomes. I obtain the total S&P 500 index from Global Financial Data (GFD) with monthly data from January 1871.[28] Monthly riskfree rates are downloaded from Professor Kenneth French's website. For monthly riskfree rates before 1927, I use the series from Goyal and Welch (2008).

3.4 Discourse Topics

I examine the contents of news topics in Subsection 3.4.1. In Subsection 3.4.2, I discuss the summary statistics of my topics and in Subsection 3.4.3, I discuss their time series.

3.4.1 Contents of News Topics

In a semisupervised topic model such as sLDA, the favored approach in the literature to evaluate the choice of seed words is to investigate the most common terms within a topic post-estimation to determine whether the topics feature the desired contents (Lu et al. 2011; Watanabe and Zhou 2020). Hence, to investigate the contents of the 14 extracted topics, during every monthly estimation of the sLDA model, I retain the 30 most common n-grams per topic, that is, those having the highest probabilities in that topic. Then the most important words for each topic are identified as those that have the highest frequency over time.

[28] The GFD description is as follows: "The S&P 500 Total Return Index is based upon GFD calculations of total returns before 1971 [. . .] Beginning in 1871, data are available for stock dividends for the S&P Composite Index from the Cowles Commission and from S&P itself. I used this data to calculate total returns for the S&P Composite using the S&P Composite Price Index and dividend yields through 1970, official monthly numbers from 1971 to 1987, and official daily data from 1988."

To visualize each topic, I create word clouds using the top words from each topic; the higher the frequency of a word in the topic, the larger the word size. I report the word clouds of six main topics (based on their weights in the PLS index discussed next) in Figure 3.2, and the remaining topics in Figure 3.C.2 in the Internet Appendix.

As indicated by Figure 3.2, the sLDA model seems to perform well at extracting these topics from the *NYT* articles. For example, the most common terms for *War* extracted by the model are *conflict, war, government, tension* and for *Panic* are *panic, fear, crisis, depression, recession, hard_time*, all of which strongly overlap with the seed words. Although both *War* and *Panic* feature stress and anxiety, they capture distinct themes, and their correlation is -17% as reported in Table 3.C.2. The top words for *Monetary* are *money, gold, silver, inflation, bank*; for *Real Estate Booms* are *bubble, boom, speculation, price increase*; and for *Boycott* are *boycott, outrage, strike, moral, anger, community, protest.* Except for *Pandemic*, the thematic contents of these extracted topics are consistent with the predefined list of seed words.

3.4.2 Summary Statistics

I report the summary statistics for the 14 topic weights in Table 3.2. *War* receives the most attention on average, with a mean time-series weight of 9.7%. About 10% of the monthly *NYT* articles use one of the *War* words at least once. Table 3.2 shows that *War* is also the most volatile topic with a standard deviation of 3.7%, followed by *Stock Market Crash* at 3.1%.

For predicting stock market returns, I create a composite topic index by extracting and combining the signals most relevant to return prediction from all topics via the two-step PLS method, which has recently gained wide popularity in the literature (Kelly and Pruitt 2013, 2015; Huang et al. 2015, 2020). As a first step, the time series of each topic weight is regressed on the time series of next-month market returns using the whole sample. Second, in each period t, the vector of topic weights is regressed on the vector of slopes obtained in the first step. The slope in the second step regression is a value of the PLS index in period t. The construction of the PLS index for in-sample analyses uses the total sample from 1871 to 2019 as in the approach of Huang et al. (2015) and Huang et al. (2020). For the out-of-sample analysis, I recursively reconstruct the PLS index every month using only the information available up to that month.

The second to last column of Table 3.2 reports the PLS loadings (the slope in the time-series regressions) for all topics. In this methodology, only the relative PLS weights of the components are meaningful. *War* receives the highest weight, and its positive loading indicates that *War* is a positive predictor of market returns. Other essential topics in the PLS index include *Real Estate Booms*, *Pandemic*, and *Panic*. Surprisingly, the topics receiving the smallest weights are *Stock Market Bubbles* and *Stock Market Crashes*. These facts are potentially useful for future theorizing about economic narratives and the stock market.

The last column of Table 3.2 reports the correlations between the 14 topics and the PLS index. As expected, the PLS index is highly correlated with *War* with a correlation coefficient of 82%.

3.4.3 Time Series of Discourse Topics

Next, I examine fluctuations in topic weights over time. I plot the time series of each demeaned topic weight against excess market returns from January 1871 to October 2019. The results for the six main topics and the PLS index are displayed in Figure 3.3, and the remaining topics are shown in Figure 3.C.3 in the Internet Appendix. As can be seen from the graphs, except for *War*, the topics do not display any clear patterns. Thus, I focus my discussion on the time series of *War*.

Figure 3.3 describes the time series of *War*. *War* spiked in the 1870s, the Reconstruction period after the American Civil War. It also surged during the 1890s, a period that featured the Spanish-American War in 1898 and the Philippine-American War of 1899–1902. *War* rose to its highest since the start of the sample during World War I from 1917 to 1918. It remained low during the 1920s and 1930s before surging again during World War II. *War* reached its all-time high in 1963 due to major developments of the Vietnam War.

As shown later in this paper, the predictive power of *War* for the stock market has been stronger in recent decades. In Figure 3.4, I focus on the time series of *War* over the last 30 years of the sample. I track the ten articles with the most significant contributions to the ten highest monthly scores of *War* hikes since 1990.[29] Over the last 30 years, *War* spiked in the early 1990s during the Gulf War, and surged again at the end of 2001 after the 9/11 terrorist attack. In recent years *War* has remained high, especially from 2014 to 2018. During this time, the important articles reflect the climate of the period: stories are full of international tensions, notably including the Russian annexation of Crimea and the nuclear weapons threat from North Korea.

3.5 Discourse Topics as Stock Market Predictors

I next address the primary research question of the paper: do rare disasters and non-disaster-focused discourse topics predict the U.S. stock market return? In Subsection 3.5.1,we consider one-month return prediction. In Subsection 3.5.2 I consider long-horizon prediction. In the later subsections, I control for standard return predictors from past literature and examine whether discourse topics have incremental predictive power.

[29]Each month, the most influential article is the article with the highest product of article-level topic weight and article length, i.e., the numerator in Equation (3.1). Equal weighting, ignoring the article length, can help one identify slightly different influential articles. Still, these other articles are generally thematically similar to the most influential articles reported here.

3.5.1 Predicting One-Month-Ahead Returns

To investigate the return predictive power of discourse topics, I run the following standard predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1}, \tag{3.1}$$

where R^e_{t+1} is the annualized excess market return over the next month, x_t is one of the topics or the topic PLS indexes standardized to zero mean and unit variance, and β, the coefficient of interest, measures return predictability. The reported t-statistics are computed with Newey and West (1987) standard errors.

Table 3.3 reports the results. Over the whole 1871-2019 sample, among the 14 topics, *War* is the strongest positive predictor, with the coefficient statistically significant at the 1% level. Economically, a one-standard-deviation increase in *War* is associated with a 3.8% increase in the annualized excess return in the next month.

In addition to the full sample analysis, I run predictive regressions over three subperiods: 1871-1949, 1950-2019 and 2000-2019. This addresses possible concerns about text quality in the earlier part of the sample. Furthermore, it is interesting to examine whether financial market behavior is different during the latest two decades, with the rise of internet usage and new communication technologies. The results during this period may be the most relevant for the future, as emphasized in Goyal and Welch (2008).

The positive association between *War* and future market returns remains in both subperiods with significance at the 5% level. Furthermore, *War* yields an impressive forecasting power over the past two decades with a coefficient of 9.8%, significant at the 1% level, and an in-sample R^2 of 3.4%.

Among the remaining economic discourse topics, *Pandemic* and *Real Estate Boom* are negative return predictors over the whole sample, both significant at the 5% level. In contrast, *Panic* is a positive predictor of market returns, significant at the

10% level. Among these non-war topics, only *Real Estate Boom* yields meaningful predictions across all subsamples.

The last portions of Table 3.3 report return prediction results using the PLS method. The PLS index constructed from all 14 topics predicts returns more strongly than *War* alone. Over the total sample, a one-standard-deviation increase in the PLS index is associated with a 6.1% increase in the annualized return in the next month, with an in-sample R^2 above 1%. Moving from earlier to later subsamples, the PLS index displays increasingly strong predictive results, significant at the 1% level even in the early subsamples. This suggests that the combined information in all topics has predictive power for long time-series data.

I also examine the predictive power of only the topics discussed by Shiller (2019). To do so, I construct the "Shiller PLS" index by excluding *War* and *Pandemic* and report the prediction results using this index in the last row of Table 3.3. Accordingly, the Shiller PLS index displays similar prediction patterns as the composite PLS index, albeit with smaller magnitudes. For example, over the whole sample, the Shiller PLS has a prediction coefficient of 4.8% and an R^2 of 0.6% compared to 6.1% and 1.1%, respectively, of the composite PLS index.

Following Golez and Koudijs (2018), I compute the cumulative in-sample R^2 in predicting the next month return, reported in Panel A of Figure 3.5. An upward trend indicates a predictor performs well during the sample period. Both *War* and the composite PLS index experience poor performances during 1910–1930 but strongly recover after that. Again, both cumulative R^2 suffer from a slight decline for a short period before 2000.

Overall, Table 3.3 indicates that *War* and the PLS index are strong market predictors, and their forecasting power increases in more recent periods. The predictive power of *War* and the PLS index is most pronounced from 2000–2019. I conjecture that the digitization of news and the technology that accelerates the diffusion of information drive this result. This result is consistent with that of **?**, who find strong market predictive power in the sentiment of photos and text starting in 2010s.

3.5.2 Predicting Long-Horizon Returns

I have found that *War* and the PLS index predict market returns at a one-month horizon. I now examine the long-horizon predictive power of *War* and the PLS index by running the predictive regression:

$$R^e_{t+1\to t+h} = \alpha + \beta x_t + \epsilon_{t+1\to t+h}, \tag{3.2}$$

where $R^e_{t+1\to t+h}$ is the annualized excess market return over the following h months, x_t is either *War* or the PLS indexes, and β, the coefficient of interest, measures the strength of predictability. In addition to the composite PLS index constructed using the whole sample reported in Table 3.4, to avoid look-ahead bias I also use an "expanding PLS" index computed every month using only the data available up to that month. To account for potential autocorrelations of the residuals in the long-horizon predictive regressions, I compute the Newey and West (1987) standard errors with corresponding h lags.

The first row of each panel in Table 3.4 repeats the results for $h = 1$ for comparison. Panel A of Table 3.4 reports the results for the full sample from 1871 to 2019. Over these 159 years, *War* and the two PLS indexes significantly predict market returns up to 36 months ahead. The PLS using the whole sample is built to optimize in-sample predictive power. As expected, the expanding PLS index yields smaller coefficient estimates and in-sample R^2.

In the subsample analysis, the predictive power of *War* is relatively weak during the first half of the sample period (significant at the 5% level for the one-month horizon). Still, it is significant at the 5% level from one- to six-month horizons during the second subperiod (1950 to 2019). The predictive power of the whole-sample PLS index is significant at the 1% level one month ahead and at the 5% level for other horizons (except 24 months) during the first half of my sample period. The results become stronger during the second half.

The strongest effects are obtained starting from the year 2000. *War* yields impressive predictive power over the last 20 years of the sample; its in-sample adjusted R^2 ranges from 3.4% (1 month) to 18% (36 months). During this period, the whole-sample PLS index yields strong results (significant at the 1% level) across all forecasting periods. As for economic magnitudes, a one-standard deviation increase in *War* is associated with an annualized increase of 9.8% in next month return over the 2000-2019 period. The corresponding number for the PLS index is 12.2%. The mean S&P500 annualized return during the same period is 3.96% suggesting the predictive power of *War* and the PLS index is economically substantial.

3.5.3 Predicting One-Month-Ahead Returns: *War* versus Economic and Topic Predictors

The previous two subsections show that *War* is a strong predictor of stock market returns. I next investigate whether *War* has predictive power beyond standard economic predictors and the remaining 13 topics studied in this paper. For economic predictors, I include the dividend-price ratio (DP), earnings-price ratio (EP), dividend payout ratio (DE), stock variance (SVAR), and T-bill rate (TBL) from Goyal and Welch (2008). I include these variables since they are available for my full sample period of 1871 to 2019. I run the following bivariate regression:

$$R^e_{t+1} = \alpha + \beta War_t + \gamma z_t + \epsilon_{t+1}, \tag{3.3}$$

where z_t is either each of the economic or remaining topic predictors. All independent variables are standardized to zero mean and unit variance and t-statistics are computed with the Newey and West (1987) standard error.

Panel A of Table 3.5 reports the results for the economic predictors. In all bivariate regressions between *War* and each economic predictor, *War* remains significant at the 1% level with economic magnitude larger than those of the economic predictors. In the final column of Panel A, when I run a kitchen sink regression that includes *War* and all economic predictors,[30] *War* is still significant at the 5% level.[31]

In Panel B of Table 3.5, when *War* is tested against the remaining topic predictors, its statistical and economic significances remain intact in either bivariate or kitchen sink regressions.

Overall, I find that investors' perception of war risks as captured by my *War* index is a robust predictor of stock market returns and outperforms other common economic variables and non-disaster-related discourse topics in predicting the next one-month market returns.

3.5.4 Predicting One-Month-Ahead Returns: *War* versus Other Media-Based Uncertainty Indexes

In the previous subsection, I show that *War* as a measure of disaster risks has stronger predictive power than common economic predictors and non-disaster-focused discourse topics. However, the literature has introduced other news-based proxies for disaster risks, notably including the news implied volatility (NVIX) from Manela and Moreira (2017) and the geopolitical risks (GPR) from Caldara and Iacoviello (2022).[32] I now investigate whether my *War* contains incremental predictive power over these two measures.

In Panel A of Table 3.6, I use *War*, NVIX^2, and GPR to predict the excess market return one month ahead as in Equation (3.3). I first run univariate regressions and

[30] I exclude DE to avoid perfect collinearity because it is a linear combination of DP and EP.

[31] In Table 3.C.3, I document that the predictive power of *War* remains intact when I control for market returns, conditional skewness, and conditional volatility.

[32] I thank the authors of these papers for making their data available.

then compare *War* and GPR against NVIX^2. I do not directly compare *War* with GPR because the two variables are highly correlated (correlation of 60%) over the sample period January 1900 to March 2016.[33] In contrast, *War* and NVIX^2 have a -5% correlation.

Over this period, both *War* and GPR are positive return predictors, significant at the 5% and 10% levels, respectively, where *War* yields a larger economic magnitude. In contrast, NVIX^2 does not predict the market at the one-month horizon, consistent with the results in Manela and Moreira (2017). I observe the same results over the sub-sample 1950-2016. Over the most recent sample, 2000-2016, *War* dominates in terms of both economic and statistical significance.[34]

Overall, *War* produces stronger short-term predictive power for market returns than the other two media-based disaster risk measures, especially after 2000.

3.5.5 Predicting One-Month-Ahead Returns: *War* versus Crisis Event Counts

In a previous study, Berkman et al. (2011) measureinvestor perceptions of disaster risks by counting the number of crisis events each month. I now run a horse-race test of the effectiveness of this measure, which is based on actual crisis events, versus my media-based measure, which is based on textual discourse, in predicting returns.[35] I include the aggregate crisis index as this is the main variable studied in Berkman et al. (2011) and the war count index as this is mostly related to my *War*. In Panel B of Table 3.6, I show that over the 100-year period 1918-2018, both news-based (*War*) and event-based (*CWar*) war indexes predict next-month market returns, both significant at the 5% level. However, over the two subsamples 1950-2018 and 2000-2018, *War*

[33]NVIX is only available until March 2016. Following their paper, I use NVIX^2 in my analyses. Using NVIX yields almost the same results.

[34]In unreported results, when I put three predictors together in the same regression, *War* drives out the significance of GPR during the period 2000-2016.

[35]The data is updated to 2018 and available at https://sites.duke.edu/icbdata/.

dominates the event count indexes.[36]

I conclude that investors' perception of rare disaster risks as extracted from news media is an important predictor of stock market return, even after controlling for event-count variables.

3.5.6 Predicting One-Month-Ahead Returns: Controlling for Economic Variables

In the previous subsections, I show that *War* outperforms 5 economic variables, 13 topics, and numerous crisis event count as indexes in predicting next month stock returns. In this subsection, I extend the list of economic variables and consider the following bivariate predictive regression:

$$R^e_{t+1} = \alpha + \beta x_t + \gamma z_t + \epsilon_{t+1}, \tag{3.4}$$

where z_t is one of the economic predictors. Following Huang et al. (2020), I 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 et al. (2016).

The "Univariate" column in Panel A of Table 3.7 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).

[36] In Appendix Section 3.C, I explore a larger set of real crisis events obtained from Global Financial Data. I create dummy variables to capture the occurrences of the following events: recessions, bank failures, wars, natural disasters, epidemics, and any of them. As reported in Table 3.C.5, *War* retains its predictive power after controlling for these events.

War is 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.

3.5.7 Predicting One-Month-Ahead Returns: Controlling for Uncertainty and Sentiment Variables

The previous section documented that the topic PLS index contains valuable insights into market returns. I now test whether the topic PLS index has incremental ability to predict returns in comparison with other well-known uncertainty or sentiment variables. Recently, ample measures have been introduced into the literature, notably the financial and macro uncertainty indexes from Jurado et al. (2015), the economic policy uncertainty index from Baker et al. (2016), and the disagreement index from Huang et al. (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*, I 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, I also include the two U.S. stock market confidence indexes introduced by Shiller: the one-year confidence index and the crash confidence index.[37]

The "Correlations" column of Panel B of Table 3.7 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 et al. (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. I 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 et al. (2015) is a negative predictor, significant at the 5% level. The economic policy uncertainty index of Baker et al. (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, I 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 3.7 reports the results with the PLS index constructed from all the uncertainty

[37] 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.

and sentiment variables, against which the PLS index remains significant at the 5% level.

The results in Table 3.7 indicate that the discourse topic index contains valuable information about market returns after controlling for the strong market predictors recently proposed in the literature.

3.5.8 Robustness Checks

I next consider the robustness of my results with respect to several features of the empirical design. Lu et al. (2011) suggest using a number of topics equal to the number of seeded topics plus one.[38] It would not be appropriate to optimize the number of topics 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 that are central to my approach.

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

[38]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.

To evaluate the robustness of my findings with respect to the number of topics and the number of seed words, I focus on the 2000-2019 sample since this period shows the strongest return predictive power. For the number of topics, I experiment by increasing the number of unseeded topics to 50 while keeping the seeded topics unchanged and find that the predictive power for *War* is robust.

For the number of seed words, Lu et al. (2011) do not provide a guideline but mention in their footnote 6 that the number of seed words can be expanded depending on the size of the topic. I try varying numbers of seed words, duplicates in seed words within and across the topics, and the types of seed words such as unigrams, bigrams, or trigrams. I experiment with each of these choices subsequently while keeping other specifications of the model unchanged and find that the predictive power of *War* remains robust.[39]

As detailed in Internet Appendix 3.B, I create n-grams within punctuation boundaries before removing stop words. I then remove bigrams containing stop words because these bigrams add no additional value to topic estimation beyond the unigrams contained therein. I also remove trigrams containing stop words unless the stop word is a preposition in the middle position. My n-gram creation method ensures I only consider meaningful terms in the text data. However, I also experiment with removing stop words *before* creating n-grams and retaining all the resulting n-grams and find consistent results for *War* over 2000-2019.

When cleaning the texts, I remove articles containing mostly numbers, names, and lists (i.e., articles having limited content) by manually examining the patterns of their titles. As a robustness check, I keep all news articles in my data set and find qualitatively similar results for *War*.

I also examine the strategy of using a very large number of topics. In such a case,

[39]Specifically, I first vary the number of seed words while keeping other specifications unchanged. I then vary duplicates of seed words while keeping other specifications unchanged. Finally, I vary the types of seed words and while keeping everything else unchanged.

the weights of seeded topics can be approximated by the frequency of the seed words in the corpus. So, I investigate this case by constructing topic weights as the counts of seed words scaled by the article length and present the results in Internet Appendix 3.D. Frequency-based topic weights still yield results consistent with the sLDA ones, but their out-of-sample performance is weaker.

3.6 Out-of-Sample Analysis

The predictability results in Section 3.5 are obtained by pooling within the 150-year sample. Such tests are subject to look-ahead bias, as with past studies that perform in-sample predictability tests or that use in-sample information to construct return predictors. To address this concern, I now perform an out-of-sample analysis, as is required to offer real-time economic value to investors (Goyal and Welch 2008). I conduct two standard out-of-sample tests to investigate whether discourse topics can help investors make better investment decisions: out-of-sample R^2 and certainty equivalent return (CER) gains.

3.6.1 Out-of-Sample R-Squared

Following Campbell and Thompson (2008), I compute the following well-known out-of-sample R^2 statistic:

$$R^2_{OS} = 1 - \frac{\sum_{t=p}^{T-1}\left(R^e_{t+1} - \hat{R}^e_{t+1}\right)^2}{\sum_{t=p}^{T-1}\left(R^e_{t+1} - \bar{R}^e_{t+1}\right)^2}, \tag{3.1}$$

where R^e_{t+1} is the realized excess market return, $\hat{R}^e_{t+1} = \hat{f}_t(x_t)$ is the predicted excess return with $\hat{f}_t(x_t)$ being a function of the predictors recursively estimated using only the training window, $\bar{R}^e_{t+1}$ is the historical mean excess return computed over the training window, and p is the size of the initial training window. I employ an expanding estimation window to incorporate all available information into formulating future forecasts and begin the evaluation period in January 1881 (10 years from the sample's start).

I benchmark the out-of-sample results of the 14 topics against the six predictors from Goyal and Welch (2008), including dividend-price ratio, dividend yield, earnings-price ratio, dividend payout ratio, stock variance, and Treasury-bill rate, all of which are available from 1871.

I use two approaches to estimate the function $f_t(x_t)$ recursively. First, I specify $f_t(x_t)$ to be a linear function of the 14 topics and 6 other predictors. Second, I specify $f_t(x_t)$ to be a function of all 14 topics or all 6 other predictors estimated via PLS as described in Section 3.4. Also, recall that my topic weights are extracted monthly using data over the past ten years, so there is no look-ahead bias in the out-of-sample analysis.

When a predictor outperforms the historical mean benchmark in forecasting future returns, it produces a lower mean squared forecast error (MSFE) than the historical mean. Thus, the R^2_{OS} will be greater than zero. To test the significance of R^2_{OS}, I report the Clark and West (2007) MSFE-adjusted statistic.

Panel A of Table 3.8 reports the results from OLS regressions using individual predictors. Among the six economic predictors, only Treasury Bill produces a positive and significant R^2_{OS} over the whole evaluation period, yet the magnitude is tiny at 0.07%. Meanwhile, among the 14 topics, during 1881–2019, *War*, *Pandemic*, and *Real Estate Boom* yield a significant R^2_{OS} (0.17%, 0.08%, and 0.19%, respectively). Except for *Pandemic*, *War* and *Real Estate Boom* continue to deliver out-of-sample (henceforth, OOS) predictive power over the past 20 years with magnitudes much larger than the whole-sample results, at 1.35% and 1.13%, respectively. Consistent with the in-sample results in Section 3.5, *War* displays strong out-of-sample predictive power in recent periods.

Panel B of Table 3.8 combines the signals of individual predictors via PLS. Combining all six other predictors produces negative R^2's across all sample periods. Over the whole sample, combining all topics via PLS yields a negative R^2_{OS}. However, in the two most recent subsamples, the topic PLS method delivers strong predictive power, producing R^2_{OS}'s of 0.95% over 1950–2019 and 2.23% over 2000–2019, both significant at the 1% level. The predictive power of the topic PLS provides an economically substantial superior performance compared to the R^2_{OS} of 1.7% from Gómez-Cram (2022) using macroeconomic indicators over the last twenty years. In the last row of Panel B, I use only the 12 topics from Shiller (2019) in the PLS estimation, which yields negative R^2_{OS}'s in all samples.

Panel B of Figure 3.5 plots the cumulative out-of-sample R^2 for *War* and the PLS method using all 14 topics. An upward trend indicates good performance during that period. Consistent with Table 3.8, *War* and the PLS method do not perform well during the first half of the sample, especially from 1910 to 1930, in which both display a steep downward slope in the cumulative R^2_{OS}. From 1930 to 1990, both *War* and the PLS method feature steadily upwards trends, with the PLS method having a much steeper slope. However, both encounter a decline during the 1990s before having a turnaround during the last two decades of the sample.

Overall, I find that discourse topics outperform standard return predictors in out-of-sample prediction, especially during recent decades. These findings corroborate the in-sample results of earlier sections that the predictive power of discourse topics is stronger in recent periods.

Recent research has documented that sentiment variables have stronger predictive power during recessions and high sentiment periods when mispricing is likely to be prevalent owing to short-sale constraints (see, e.g., Garcia (2013), Huang et al. (2015), and Jiang et al. (2019)). I investigate whether my news topics follow the same pattern by decomposing the whole sample into expansions and recessions as well as high and low sentiment periods and compute the in- and out-of-sample R^2's during each subperiod. As reported in Table 3.C.6, discourse topics better predict the market return during low sentiment periods. However, I find no evidence of different predictive powers across the business cycles. These results highlight that the predictive power of discourse topics operates via a different channel from that of sentiment.

3.6.2 Asset Allocation Implications

I next examine the economic value of news topics from an asset allocation perspective. Following Campbell and Thompson (2008), I 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

$$w_{t+1} = \frac{1}{\gamma}\frac{\hat{R}^e_{t+1}}{\hat{\sigma}^2_{t+1}} \tag{3.2}$$

of the portfolio to equities during period $t+1$, where risk aversion coefficient γ is set to three following Huang et al. (2020),[40] $\hat{R}^e_{t+1}$ 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

$$R^p_{t+1} = w_{t+1}R^e_{t+1} + R^f_{t+1}, \tag{3.3}$$

where R^f_{t+1} is the riskfree return. Following Campbell and Thompson (2008), I 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.

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

The CER of the portfolio is

$$CER_p = \hat{\mu}_p - 0.5\gamma\hat{\sigma}_p^2, \tag{3.4}$$

where $\hat{\mu}_p$ and $\hat{\sigma}_p^2$ 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. I 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, I compute the annualized monthly Sharpe ratios of the portfolio's realized returns. I test the statistical significance of the CER gain and the Sharpe ratios (against the historical mean benchmark) using the test statistics in DeMiguel et al. (2009).

Panel A of Table 3.9 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 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 3.9 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 3.9 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.

3.7 *War* and Predictability of Bond Returns

I 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. I now test these predictions.

Specifically, I run the following predictive regression:

$$R^e_{t+1\to t+h} = \alpha + \beta War_t + \epsilon_{t+1\to t+h}, \tag{3.1}$$

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. I 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, I report the Newey and West (1987) standard error with

h lags.

Panel A of Table 3.10 reports the results for Treasury bonds. I 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. I 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.

I also perform the out-of-sample predictions of bond excess returns using *War* and report the results in Table 3.11. To facilitate comparison among different types of bond indexes, I limit the sample to the range of January 1993 to October 2019. I employ an expanding estimation window using the first 10 years of the sample as the initial estimation window. I 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, my 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.

3.8 Conclusion

I test the hypothesis that rare disaster risk is priced (or mispriced) by extracting market attention to rare disasters from news media. This helps overcome the challenge of scant data on realized disasters. It also has the advantage (from a behavioral perspective) of focusing on investor attention to and perceptions of disaster, which may differ from objective risks. I provide the most comprehensive analysis for empirically testing for pricing effects of disaster risk. In addition to two topics covering rare disaster risks (*War* and *Pandemic*), I also examine 12 non-disaster-focused narratives from Shiller (2019).

I employ an advanced natural language processing tool called sLDA to extract discourse topics from nearly seven million *New York Times* articles over the past 160 years. I create a list of topic-based seed words to input into the sLDA model to guide the topic extraction process. I employ a rolling estimation scheme to include only historical news data at every estimation time; thus, my measure avoids look-ahead bias and addresses changes in semantic usage over time.

Among the discourse topics considered, the most important is *War*, which encompasses various themes related to the danger of armed conflict. I find that *War* and an index constructed from all topics are strong positive predictors of the stock market return. I find the predictive power of *War* increases through the sample period and the predictive power of discourse topics holds at both the market and portfolio levels.[41] The predictive power of *War* remains even when extracted from a different media outlet, the *WSJ*.

That *War* positively predicts excess market returns is consistent with the literature on rare disaster risks, or with the behavioral hypothesis that investors overweight the prospect of rare disasters. Barro (2009) finds that the probability of rare disasters can explain the high equity premium. During times when the probability of a rare disaster is higher, the equity premium should be higher, which is consistent with my finding that *War* is associated with higher subsequent stock returns. Alternatively, if investors overweight rare risks (either owing to overestimation of probability owing to salience or the overweighting of low probability events in the cumulative prospect theory utility function), I again expect higher war media discourse to be associated with high future returns.

My results also confirm the prediction of Gabaix (2012) (or of a behavioral setting where agents overweight rare disasters). I find that *War* increases the excess returns on mid- to long-term high-yield corporate bonds. In contrast, *War* negatively predicts excess returns on safer investment instruments such as short-term government bonds and investment-grade corporate bonds.

[41] See Table 3.C.7 for portfolio prediction results.

Figure 3.1. Estimation Scheme

This figure plots the rolling estimation scheme for the sLDA model. Every month t, news articles in the previous 120 months (including month t) are used to estimate the sLDA model, and then articles in month t are used to compute topic weights in that month.

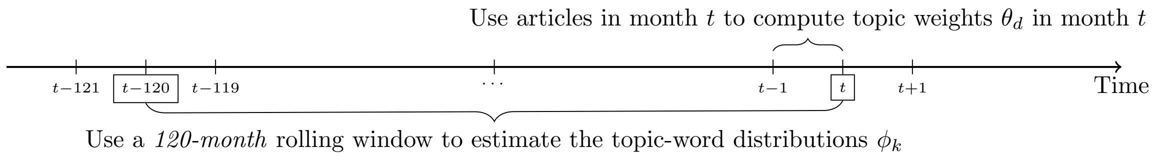
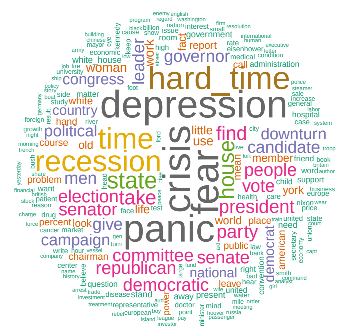


Figure 3.2. Discourse Topic Contents

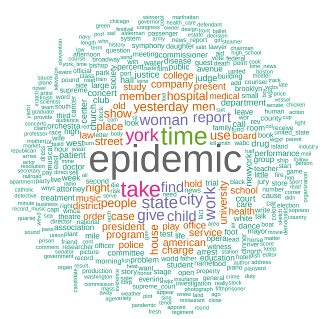
This figure plots the frequencies of n-grams per topic over time. Frequencies are constructed according to the sLDA model described in Section 3.2, and the size of each n-gram indicates its frequency. The sample period is from January 1871 to October 2019.

War

Panic

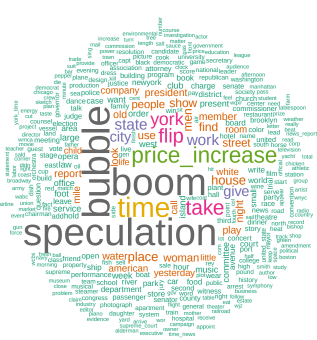


Pandemic



Monetary Standard

Real Estate Booms



Boycotts and Evil Business

Figure 3.3. Time Series of Discourse Topic Weights

This figure plots the time series of monthly topic weights constructed according to the sLDA model described in Section 3.2. The solid line represents the topic weight, and the dashed line represents the excess market return; both have been demeaned to improve visualization. The gray-shaded areas represent NBER-defined recessions. The sample period is from January 1871 to October 2019.

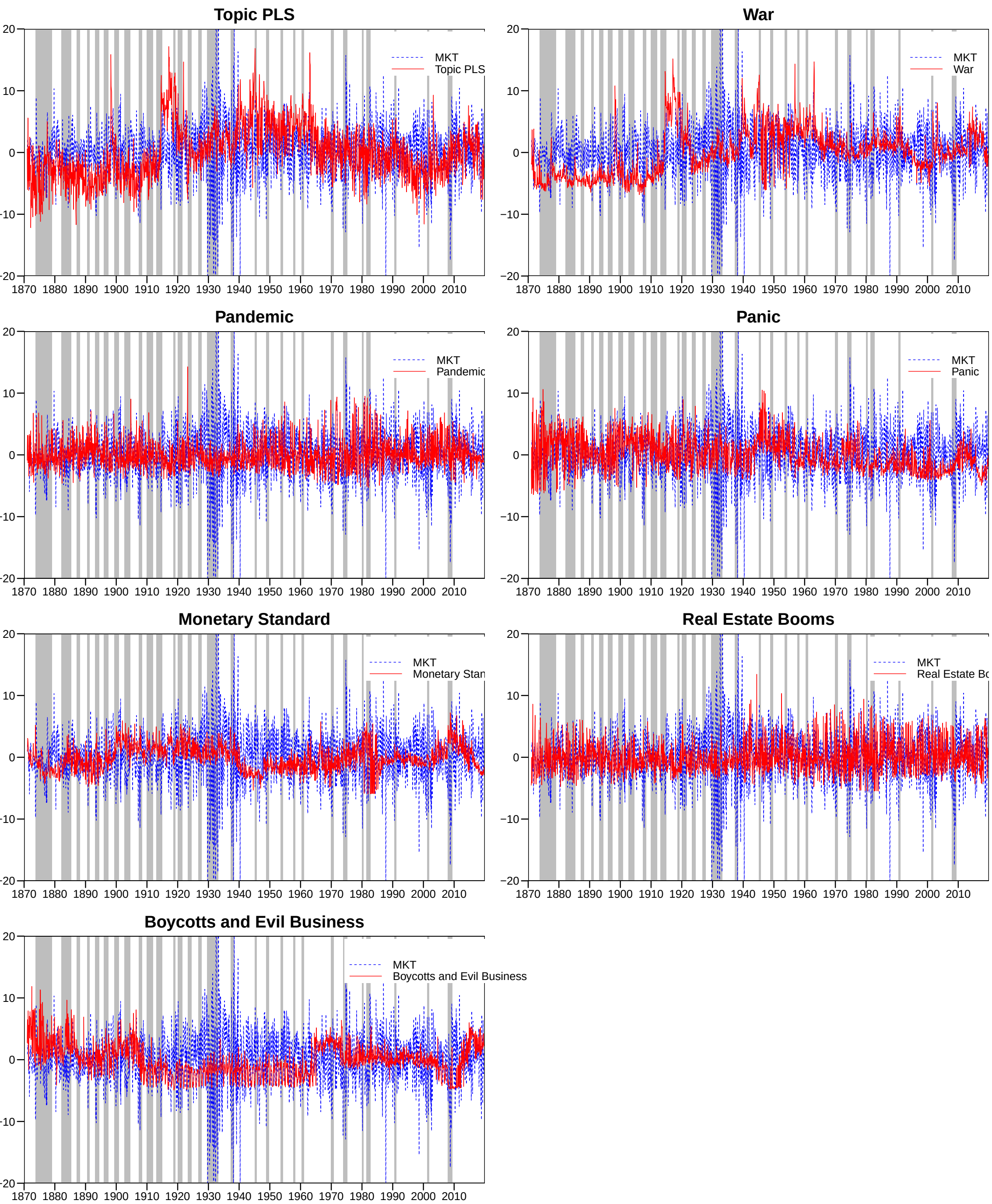


Figure 3.4. Articles Making the Biggest Contribution to *War* Spikes since 1990

This figure plots the ten articles that have contributed significantly to ten monthly heights of *War* since 1990. Topic weights are demeaned. The gray-shaded areas represent NBER-defined recessions. The sample period is from January 1990 to October 2019.

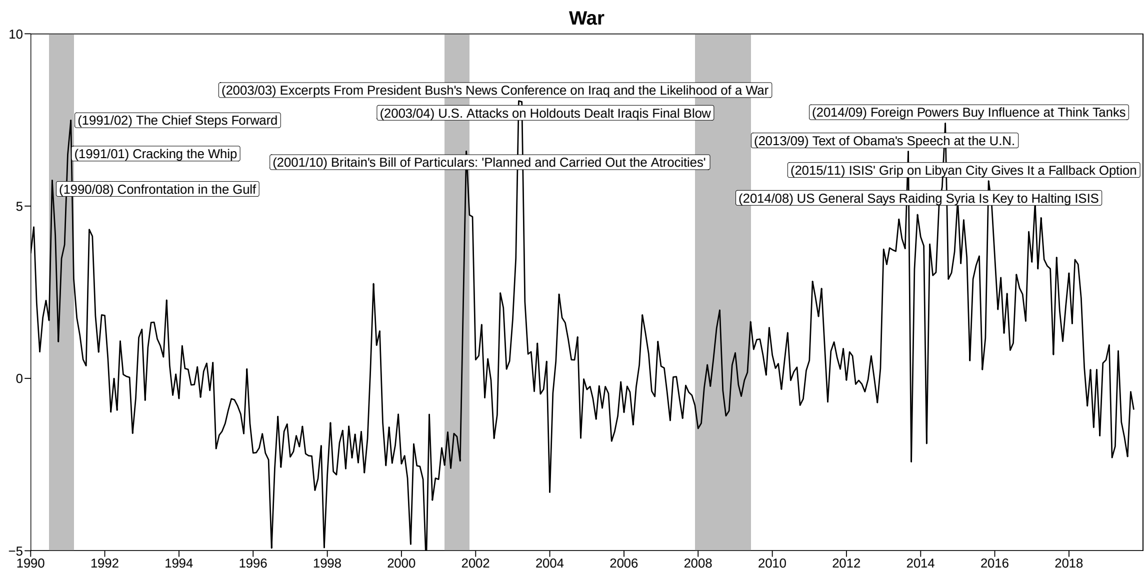


Figure 3.5. Cumulative R-Squared in One-Month Return Prediction

Panel A plots the cumulative in-sample R^2 computed as

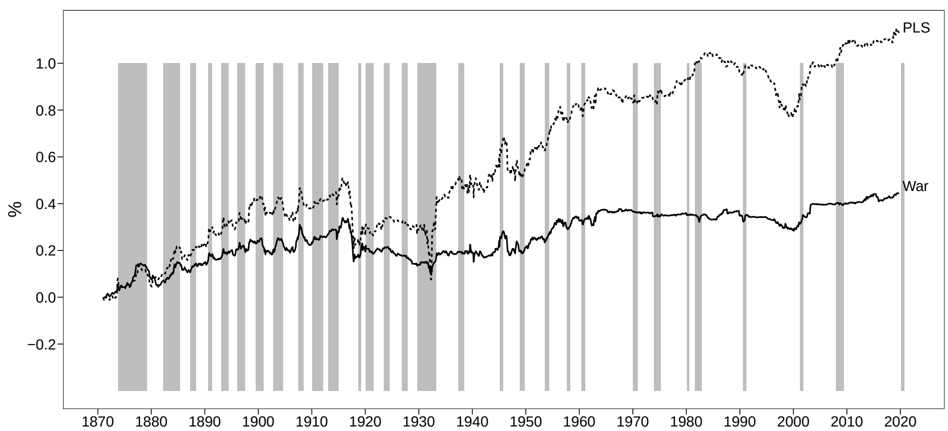
$$\left(\sum_{s=1}^{t} \left(R_s^e - \bar{R}^e \right)^2 - \sum_{s=1}^{t} \left(R_s^e - \hat{R}_s^e \right)^2 \right) / \sum_{s=1}^{T} \left(R_s^e - \bar{R}^e \right)^2 ,$$

where $\bar{R}^e$ is the sample mean of excess return and $\hat{R}_s^e$ is the fitted value from regression (3.1). The sample period is from January 1871 to October 2019. Panel B plots the cumulative out-of-sample R^2 computed as

$$\left(\sum_{s=1}^{t} \left(R_s^e - \bar{R}_s^e \right)^2 - \sum_{s=1}^{t} \left(R_s^e - \hat{R}_s^e \right)^2 \right) / \sum_{s=1}^{T} \left(R_s^e - \bar{R}_s^e \right)^2 ,$$

where $\bar{R}_s^e$ and $\hat{R}_s^e$ are the historical mean and predicted value, estimated based on the preceding estimation window. The evaluation period is from January 1881 to October 2019.

Panel A: In-Sample R^2



Panel B: Out-of-Sample R^2

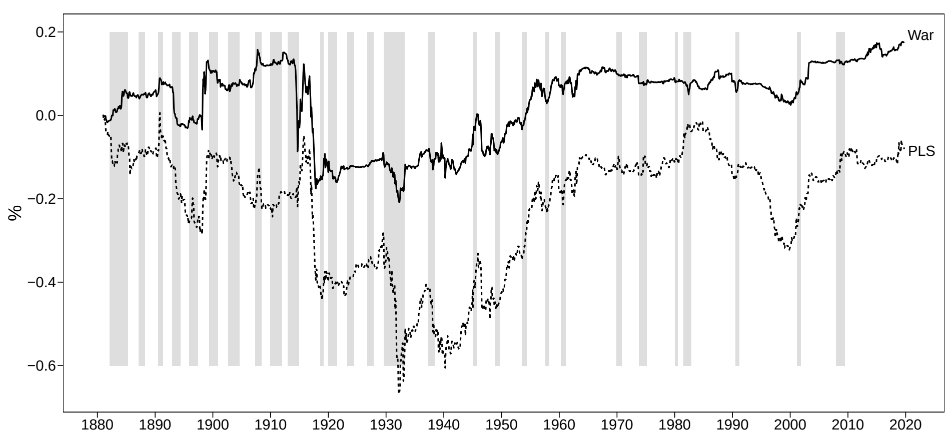


Table 3.1
Seed Words

This table lists the *lemmatized* seed words for each of the 14 discourse topics. The first column presents the full name of the topic, and the second column reports the short name used in the paper.

Narrative	Short Name	Seed Words
War	War	conflict, tension, terrorism, war
Pandemic	Pandemic	epidemic, pandemic
Panic	Panic	bank failure, bank panic, bank run, crisis, depression, downturn, fear, financial panic, hard time, panic, recession
Confidence	Confidence	business confidence, consumer confidence
Frugality	Saving	compassion, family morale, frugal, frugality, modesty, moral, poverty, saving
Conspicuous Consumption	Consumption	american dream, conspicuous consumption, consumption, equal opportunity, equality, homeownership, luxury, patriotism, prosperity
Monetary Standard	Money	bimetallism, devaluation, gold, gold standard, inflation, monetary standard, money, silver
Techmology Replacing Jobs	Tech	automate, computer, digital divide, electronic brain, invention, labor save, labor save machine, machine, mechanize, network, technocracy, technological unemployment, technology, unemployment
Real Estate Booms	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 Crashes	Real_estate_crash	bust, crash, house bust, house crash, land bust, land crash, price decrease, real estate bust, real estate crash
Stock Market Bubbles	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 Market Crashes	Stock_crash	bear, bear market, bearish, bust, crash, fall market, market crash, stock crash, stock market crash, stock market decline
Boycotts and Evil Business	Boycott	anger, boycott, community, evil business, excess profit, fair wage, moral, outrage, postpone purchase, profiteer, protest, strike, wage cut
Wage and Labor Unions	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 3.2
Summary Statistics

This table presents the summary statistics for the time series of 14 monthly topic weights constructed according to the sLDA model described in Section 3.2. All numbers (except sample size) are expressed as percentages. The sample period is from January 1871 to October 2019.

	N	Mean	SD	Q1	Median	Q3	AC(1)	PLS Weights	Corr PLS
War	1784	9.71	3.73	6.56	9.64	11.92	84.84	5.22	82.09
Pandemic	1784	5.72	2.34	4.25	5.38	6.72	7.54	-2.26	-28.53
Panic	1784	8.30	2.70	6.21	7.94	10.13	37.03	2.19	15.74
Confidence	1784	5.72	2.45	3.97	5.39	6.99	7.61	0.60	-0.47
Saving	1784	5.84	2.17	4.39	5.51	6.84	29.78	-1.09	-34.12
Consumption	1784	7.36	2.85	5.46	6.72	9.23	27.62	0.88	-4.58
Money	1784	6.58	2.06	5.21	6.46	7.87	60.55	-1.77	-15.26
Tech	1784	6.61	2.52	4.99	6.57	8.07	54.58	-0.79	-8.15
Real estate boom	1784	5.95	2.52	4.20	5.60	7.22	9.31	-2.82	-32.73
Real estate crash	1784	5.57	2.16	4.23	5.41	6.49	23.16	0.47	-1.12
Stock bubble	1784	5.79	2.30	4.28	5.74	7.23	48.98	-0.40	-14.67
Stock crash	1784	7.40	3.13	5.06	6.86	9.60	26.56	0.86	-2.31
Boycott	1784	5.79	2.62	4.14	5.51	7.37	67.00	-1.47	-41.69
Wage	1784	7.90	2.59	6.05	7.60	9.50	38.53	1.07	37.33
PLS	1784	49.99	44.73	18.75	46.68	77.28	70.35		

Table 3.3
Predicting One-Month Market Returns

This table presents the results of the following predictive regression:

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

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. Adjusted R^2 is expressed as a percentage, and t-statistics are computed with Newey and West (1987) standard errors. The sample period is from January 1871 to October 2019. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	1871-2019	1871-1949	1950-2019	2000-2019
War (%)	3.80 ***	3.49 **	4.06 **	9.83 ***
t-stat	(3.35)	(2.02)	(2.56)	(3.43)
R^2 (%)	0.39	0.20	0.55	3.39
Pandemic (%)	-2.61 **	-3.98 **	-1.55	-2.31
t-stat	(-2.06)	(-2.14)	(-0.91)	(-0.70)
R^2 (%)	0.15	0.29	-0.02	-0.21
Panic (%)	2.21 *	3.06 *	2.57 *	3.36
t-stat	(1.76)	(1.71)	(1.69)	(1.20)
R^2 (%)	0.09	0.13	0.15	0.02
Confidence (%)	0.66	-0.14	1.13	0.77
t-stat	(0.55)	(-0.08)	(0.67)	(0.24)
R^2 (%)	-0.04	-0.11	-0.07	-0.40
Saving (%)	-1.37	-1.62	-1.44	-0.98
t-stat	(-1.05)	(-0.92)	(-0.81)	(-0.37)
R^2 (%)	0.00	-0.04	-0.04	-0.39
Consumption (%)	0.84	0.63	2.58	0.07
t-stat	(0.66)	(0.32)	(1.61)	(0.02)
R^2 (%)	-0.03	-0.10	0.15	-0.42
Money (%)	-2.33	-1.23	-3.41 *	-0.70
t-stat	(-1.64)	(-0.61)	(-1.76)	(-0.17)
R^2 (%)	0.11	-0.07	0.36	-0.40
Tech (%)	-0.85	0.54	-3.10	-14.25 ***
t-stat	(-0.58)	(0.26)	(-1.62)	(-3.24)
R^2 (%)	-0.03	-0.10	0.27	7.59
Real estate boom (%)	-3.03 **	-3.51 **	-3.02 *	-6.57 **
t-stat	(-2.50)	(-2.05)	(-1.82)	(-2.16)
R^2 (%)	0.23	0.20	0.25	1.28
Real estate crash (%)	0.59	0.59	0.93	-2.58
t-stat	(0.48)	(0.34)	(0.55)	(-0.78)
R^2 (%)	-0.05	-0.10	-0.08	-0.16
Stock bubble (%)	-0.47	-1.67	0.39	-4.59
t-stat	(-0.37)	(-0.96)	(0.21)	(-1.31)
R^2 (%)	-0.05	-0.04	-0.11	0.41
Stock crash (%)	0.75	2.08	-0.23	-0.83
t-stat	(0.54)	(1.02)	(-0.14)	(-0.26)
R^2 (%)	-0.04	0.00	-0.12	-0.40
Boycott (%)	-1.52	-2.36	-0.43	5.33
t-stat	(-1.23)	(-1.41)	(-0.23)	(1.38)
R^2 (%)	0.01	0.04	-0.11	0.70
Wage (%)	1.12	0.70	1.87	8.62 ***
t-stat	(0.74)	(0.32)	(1.04)	(2.81)
R^2 (%)	-0.02	-0.09	0.02	2.51
PLS (%)	6.07 ***	5.86 ***	6.43 ***	12.17 ***
t-stat	(4.59)	(2.98)	(3.93)	(4.33)
R^2 (%)	1.07	0.76	1.57	5.42
Shiller PLS (%)	4.76 ***	5.98 ***	5.37 ***	7.46 **
t-stat	(3.15)	(2.71)	(3.23)	(2.30)
R^2 (%)	0.64	0.80	1.06	1.77

Table 3.4
Predicting Long-Horizon Market Returns

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. "Expanding PLS" is the PLS index recursively estimated every month using only the data available up to that month. 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 period is from January 1871 to October 2019. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	War (%)	t-stat	R^2 (%)	PLS (%)	t-stat	R^2 (%)	Expanding PLS (%)	t-stat	R^2 (%)	N
				Panel A: 1871-2019						
$h = 1$	3.80 ***	(3.35)	0.39	6.07 ***	(4.59)	1.07	4.65 ***	(3.52)	0.58	1664
$h = 3$	2.87 ***	(2.81)	0.57	4.20 ***	(3.88)	1.29	3.72 ***	(3.55)	0.96	1664
$h = 6$	3.03 ***	(2.83)	1.36	3.69 ***	(3.58)	2.03	3.28 ***	(3.26)	1.56	1664
$h = 12$	2.79 **	(2.48)	1.93	3.11 ***	(2.92)	2.42	2.76 ***	(2.66)	1.86	1664
$h = 24$	2.09 **	(2.08)	1.91	2.25 **	(2.30)	2.22	2.39 **	(2.39)	2.55	1656
$h = 36$	2.28 **	(2.22)	3.07	2.63 ***	(2.59)	4.12	3.06 ***	(3.10)	5.72	1644
				Panel B: 1871-1949						
$h = 1$	3.49 **	(2.02)	0.20	5.86 ***	(2.98)	0.76	3.92 *	(1.82)	0.24	828
$h = 3$	2.51	(1.60)	0.26	4.01 **	(2.44)	0.83	3.00 *	(1.75)	0.37	828
$h = 6$	2.88 *	(1.77)	0.94	3.71 **	(2.39)	1.63	3.00 *	(1.85)	0.95	828
$h = 12$	2.87 *	(1.73)	1.59	3.47 **	(2.23)	2.36	2.76 *	(1.74)	1.36	828
$h = 24$	1.90	(1.38)	1.25	2.55 *	(1.94)	2.32	2.54 *	(1.73)	2.27	828
$h = 36$	2.11	(1.45)	2.17	2.90 **	(2.09)	4.20	3.32 **	(2.31)	5.63	828
				Panel C: 1950-2019						
$h = 1$	4.06 **	(2.56)	0.55	6.43 ***	(3.93)	1.57	5.33 ***	(3.35)	1.04	836
$h = 3$	3.18 **	(2.41)	1.06	4.50 ***	(3.78)	2.24	4.40 ***	(3.64)	2.14	836
$h = 6$	2.93 **	(2.24)	1.64	3.61 ***	(3.13)	2.55	3.41 ***	(2.83)	2.26	836
$h = 12$	2.21	(1.50)	1.64	2.49 *	(1.92)	2.11	2.56 *	(1.94)	2.23	836
$h = 24$	1.85	(1.24)	1.93	1.66	(1.21)	1.53	1.99	(1.54)	2.25	828
$h = 36$	1.96	(1.36)	2.85	2.09	(1.55)	3.27	2.49 *	(1.90)	4.70	816
				Panel D: 2000-2019						
$h = 1$	9.83 ***	(3.43)	3.39	12.17 ***	(4.33)	5.42	9.61 ***	(3.58)	3.22	238
$h = 3$	7.64 ***	(4.07)	6.07	7.25 ***	(4.25)	5.42	5.24 ***	(3.12)	2.63	238
$h = 6$	6.08 ***	(3.31)	6.76	6.35 ***	(4.33)	7.40	4.05 ***	(3.13)	2.76	238
$h = 12$	5.42 **	(2.47)	9.68	5.21 ***	(3.28)	8.91	3.10 **	(2.51)	2.89	238
$h = 24$	4.78 **	(2.28)	12.78	4.35 ***	(2.66)	10.51	2.47 *	(1.96)	3.11	230
$h = 36$	4.59 ***	(2.86)	17.72	4.83 ***	(3.88)	19.64	3.24 ***	(3.43)	8.61	218

Table 3.5
Predicting One-Month Market Returns: *War* versus Other Predictors

This table presents the results of the following predictive regressions:

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

where R^e_{t+1} is the excess market return over the next month and z_t is one of the economic predictors from Goyal and Welch (2008) (Panel A) or the remaining topics (Panel B). The last column reports the results when *War* is tested against all predictors in each panel. 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 from January 1871 to October 2019. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Panel A: Economic Predictors

	(1)	(2)	(3)	(4)	(5)	(6)
War	3.85 ***	3.55 ***	3.75 ***	3.80 ***	3.23 ***	2.52 **
	(3.36)	(3.08)	(2.97)	(3.36)	(2.81)	(1.98)
DP	1.54					-0.90
	(0.80)					(-0.34)
EP		2.03				3.88
		(1.29)				(1.44)
DE			-0.25			
			(-0.11)			
SVAR				-0.17		-0.28
				(-0.04)		(-0.07)
TBL					-3.09 *	-4.21 ***
					(-1.92)	(-2.76)
R^2	0.40	0.46	0.33	0.33	0.61	0.75

Table 3.5
Predicting One-Month Market Returns: *War* versus Other Predictors (Cont.)

Panel B: Other Topic Predictors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
War	3.64 ***	4.30 ***	3.87 ***	3.67 ***	4.31 ***	3.60 ***	3.78 ***	3.74 ***	3.89 ***	3.79 ***	4.11 ***	3.69 ***	3.73 ***	5.38 **
	(3.21)	(3.74)	(3.41)	(3.13)	(3.51)	(3.16)	(3.35)	(3.30)	(3.39)	(3.34)	(3.48)	(3.08)	(3.20)	(2.29)
Pandemic	-2.38 *													-1.48
	(-1.88)													(-0.89)
Panic		2.94 **												3.33 *
		(2.31)												(1.67)
Confidence			0.97											1.63
			(0.81)											(0.89)
Saving				-0.55										0.48
				(-0.41)										(0.28)
Consumption					1.97									2.12
					(1.44)									(1.13)
Money						-1.97								-1.82
						(-1.37)								(-1.04)
Tech							-0.79							-0.29
							(-0.55)							(-0.15)
Real_estate_boom								-2.97 **						-2.05
								(-2.45)						(-1.24)
Real_estate_bust									0.98					1.33
									(0.79)					(0.80)
Stock_bubble										-0.09				0.72
										(-0.07)				(0.41)
Stock_crash											1.56			2.18
											(1.10)			(0.96)
Boycott												-0.33		0.11
												(-0.25)		(0.06)
Wage													0.27	0.55
													(0.18)	(0.26)
R^2	0.50	0.59	0.36	0.34	0.44	0.45	0.35	0.60	0.36	0.33	0.40	0.33	0.33	0.67

Table 3.6
Predicting One-Month Market Returns:
***War* versus Other Media-Based Uncertainty and Crisis Event Count Indexes**

This table presents the results of the following bivariate predictive regressions:

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

where R^e_{t+1} is the excess market return over the next month and x_t is *War*. In Panel A, z_t is either NVIX^2 from Manela and Moreira (2017), or geopolitical risk (GPR) from Caldara and Iacoviello (2022); in Panel B, z_t is either Crisis (monthly count of real-word crisis events), or CWar (monthly count of real-word war events) from Berkman et al. (2011). Returns are expressed as annualized percentages, and the independent variables are 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 in Panel A is from January 1900 to March 2016 and in Panel B is from January 1918 to December 2018.

Panel A: Other Media-Based Uncertainty Indexes

			1900-2016		
War	2.69 **			2.70 **	
	(1.99)			(2.00)	
NVIX^2		0.07		0.22	0.07
		(0.03)		(0.10)	(0.03)
GPR			2.48 *		2.48 *
			(1.72)		(1.72)
R^2	0.12	-0.07	0.09	0.05	0.02
			1950-2016		
War	3.96 **			3.90 **	
	(2.42)			(2.29)	
NVIX^2		1.01		0.67	0.96
		(0.32)		(0.21)	(0.31)
GPR			3.51 *		3.50 *
			(1.91)		(1.91)
R^2	0.50	-0.09	0.37	0.39	0.28
			2000-2016		
War	10.08 ***			10.19 ***	
	(3.11)			(3.05)	
NVIX^2		-0.63		-1.42	-0.07
		(-0.11)		(-0.25)	(-0.01)
GPR			7.32 **		7.32 **
			(2.27)		(2.32)
R^2	3.19	-0.50	1.44	2.75	0.92

Table 3.6
Predicting One-Month Market Returns:
***War* versus Other Media-Based Uncertainty and Crisis Event Count Indexes**

Panel B: Crisis Event Count Indexes

	1918-2018			
War	3.09 **			2.66 *
	(2.08)			(1.75)
Crisis		1.60		-0.36
		(0.99)		(-0.20)
CWar			3.75 **	3.52 *
			(1.99)	(1.69)
R^2	0.15	-0.02	0.26	0.27
	1950-2018			
War	4.12 **			4.28 ***
	(2.51)			(2.58)
Crisis		1.01		0.11
		(0.66)		(0.07)
CWar			1.58	1.93
			(0.97)	(1.13)
R^2	0.57	-0.08	-0.02	0.48
	2000:2018			
War	10.23 ***			10.05 ***
	(3.30)			(2.97)
Crisis		-2.59		-2.00
		(-0.92)		(-0.64)
CWar			-3.42	-0.37
			(-1.17)	(-0.11)
R^2	3.64	-0.18	0.01	2.96

Table 3.7

Predicting One-Month Market Returns after Controlling for 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 et al. (2016); in Panel B, z_t is one of the uncertainty variables (financial and macro uncertainty indexes from Jurado et al. (2015), economic policy uncertainty index from Baker et al. (2016), disagreement index from Huang et al. (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 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 3.7
Predicting One-Month Market Returns after
Controlling for Economic, Sentiment, and Uncertainty Predictors (Cont.)

Panel A: Economic Predictors

	Univariate		Bivariate			
Economic Predictor	$\gamma(\%)$	$R^2(\%)$	$\beta(\%)$	$\gamma(\%)$	$R^2(\%)$	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 3.7
Predicting One-Month Market Returns after
Controlling for 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 3.8
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 monthly excess market return using economic predictors or discourse topics. "Shiller Topics" uses only the topics from Shiller (2019), excluding *War* and Pandemic. Panels A and B report individual OLS regressions and the PLS method results, respectively. 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 1881, and the whole sample is from January 1871 to October 2019. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$ (based on the Clark and West (2007) MSFE-adjusted statistic).

	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	0.17 ***	-0.10 **	0.65 ***	1.35 ***
Pandemic	0.08 *	0.27 **	-0.23	0.18
Panic	0.04	0.11	-0.06	0.28
Confidence	-0.09	-0.11	-0.07	-0.09
Saving	-0.03	-0.05	-0.00	0.02
Consumption	-0.10	-0.14	-0.03	-0.01
Money	0.01	-0.18	0.33 *	-0.19
Tech	-0.45	-0.69	-0.01	0.12
Real estate boom	0.19 **	0.21 *	0.14 *	1.13 **
Real estate crash	-0.11	-0.15	-0.03	-0.15
Stock bubble	-0.12	-0.05	-0.23	-0.27
Stock crash	-0.10	-0.03	-0.23	-0.07
Boycott	-0.03	0.02	-0.11	-1.24
Wage	-0.16	-0.27	0.03	0.32 **
Panel B: PLS				
Economic	-0.84	-1.11	-0.38	-0.55
Topics	-0.08 ***	-0.67	0.95 ***	2.23 ***
Shiller Topics	-0.88	-1.03	-0.62	-0.17

Table 3.9
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 et al. (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 crash	-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 **
Topics	0.69	0.66	0.70	4.11 **	0.43	0.31 *	0.55	0.54 **
Shiller Topics	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 3.10
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 3.10
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 3.10
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 3.11
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

Appendix

3.A Seeded Latent Dirichlet Distribution

In this appendix, I 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 (Steyvers and Griffiths 2007; Blei 2012). 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 et al. (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: $\omega_k \sim \text{Dirichlet}(\beta)$.[42]

- 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.[43]
- 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

[42] 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.

[43] 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.

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.[44]

Users of the traditional unsupervised LDA developed by Blei et al. (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.

Since I am interested in uncovering some specific topics, I 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, I specify the topic-word distribution as follows:

$$\omega_k \sim \text{Dirichlet}\left(\beta + C_w\right)_{w \in V}, \tag{3.A.1}$$

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

[44] 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.

process.

Estimation is implemented by the seededla package in R and run on a high-performance computing (HPC) cluster. Full estimation of the model parallelized on 80 computational nodes requires about one week to complete. Following standard practice, I 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.

3.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.

I 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 3.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. I 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/`. I 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 my 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. I 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. I 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, I 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. I put all ngrams into one document-frequency matrix. To mitigate the impact of outliers on document-topic distribution, I remove tokens appearing in fewer than 0.2% and tokens appearing in more than 90% of all documents during each estimation window.

3.C Additional Figures and Tables for Discourse Topics from the *NYT*

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

Table 3.C.1 reports the number of *NYT* articles left after each screening step. Table 3.C.2 reports the correlation matrix of the topics and the PLS index.

3.C.1 *War* versus Conditional Volatility and Skewness

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

$$R^e_{t+1} = \alpha + \beta War_t + \gamma z_t + \epsilon_{t+1}, \tag{3.C.1}$$

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

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

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, I construct the monthly conditional skewness as follows:

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

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

I report the results in Table 3.C.3. Panel A reports the results for the whole sample from 1928 to 2019. I find that the predictability of *War* is not affected by any of the return moments. When I control for all three moments in the last column, the predictive power of *War* is still intact. I 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.

3.C.2 *War* as a Proxy for Time-Varying Risk Aversion

In the main text, I 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, I further show that *War* proxies for time-varying risk aversion, lending empirical support to the ICAPM. I first briefly discuss the ICAPM framework and then present the empirical results.

Before hypothesizing that *War* captures time-varying risk aversion, I 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, $\sigma_{MF,t}$:

$$\mathbb{E}_t[R_{M,t+1} - R_{f,t+1}] = \left[\frac{-J_{WW}W}{J_W}\right]\sigma^2_{M,t} + \left[\frac{-J_{WF}}{J_W}\right]\sigma_{MF,t}, \tag{3.C.4}$$

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 (3.C.4) 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 (3.C.4) 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, I consider a univariate version of Equation (3.C.4):

$$\mathbb{E}_t[R_{M,t+1} - R_{f,t+1}] = \lambda_0 + \lambda_1 \times \sigma^2_{M,t}, \tag{3.C.5}$$

where I 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 (3.C.5) with the popular GARCH-in-mean framework developed Bollerslev (1986) and Engle and Bollerslev (1986). Specifically, I consider first the following mean equation for the return-volatility trade-off:

$$R_{M,t+1} - R_{f,t+1} = \lambda_0 + \lambda_1 \times \sigma^2_{M,t} + \epsilon_{t+1}, \tag{3.C.6}$$

where ϵ_{t+1} has a mean of zero with conditional variance $\sigma^2_{M,t}$. Empirical tests of Equation (3.C.6) 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 (3.C.5) and Equation (3.C.6) assumes that the coefficient of relative risk aversion, λ_1, is time-invariant. However, I 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 I assume time-varying relative risk aversion, then I can specify the risk-return trade-off as a linear function of some state variable, x_t:

$$R_{M,t+1} - R_{f,t+1} = \lambda_0 + (\lambda_1 + \lambda_2 \times x_t) \times \sigma^2_{M,t} + \epsilon_{t+1}. \tag{3.C.7}$$

I hypothesize that *War* proxies for time-varying relative risk aversion, and, thus, I replace the state variable, x_t, with *War* in Equation (3.C.7). Hence, $\lambda_t = \lambda_1 + \lambda_2 \times War_t$. If this hypothesis holds with real-world data, then I expect (1) the adjusted R^2 of Equation (3.C.7) to be higher than that of (3.C.6), as the former is a more proper representation of the risk-return trade-off, and (2) the coefficient λ_2 in Equation (3.C.7) to be significantly positive as risk aversion is expected to rise when *War* is high.

To complete the GARCH-M framework, I need a specification for the conditional volatility equation. Following Lundblad (2007), I consider four different volatility specifications, namely, GARCH (Bollerslev, 1986), IGARCH (Engle and Bollerslev, 1986), TGARCH (Zakoian, 1994), and EGARCH (Nelson, 1991):

$$\begin{aligned}
\text{GARCH}(1,1): \sigma^2_{M,t} &= \delta_0 + \delta_1\epsilon^2_t + \delta_2\sigma^2_{M,t-1} \\
\text{IGARCH}(1,1): \quad \sigma^2_{M,t} &= \delta_0 + \delta_1\epsilon^2_t + (1-\delta_1)\sigma^2_{M,t-1} \\
\text{TGARCH}(1,1): \quad \sigma^2_{M,t} &= \delta_0 + \delta_1\epsilon^2_t + \delta_3 D_t\epsilon^2_t + \delta_2\sigma^2_{M,t-1} \\
\text{EGARCH}(1,1): \quad ln(\sigma^2_{M,t}) &= \delta_0 + \delta_1\left(\frac{|\epsilon_t|}{\sigma_{M,t}}\right) - \delta_3\left(\frac{\epsilon_t}{\sigma_{M,t}}\right) + \delta_2 ln(\sigma^2_{M,t-1}),
\end{aligned} \tag{3.C.8}$$

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

Panel A of Table 3.C.4 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, I 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, I 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.

I 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 3.C.4 report the results with different specifications for the volatility equation. I 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.

3.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, I 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 3.C.4 plots these events over the past 150 years.

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

$$R^e_{t+1} = \alpha + \beta \times War_t + \gamma^j \times D^j_t + \epsilon_{t+1}, \tag{3.C.9}$$

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 3.C.5 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, I 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.

3.C.4 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 (Stambaugh et al., 2012; Huang et al., 2015; Jiang et al., 2019). Huang et al. (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 et al., 2020). While these observations lean toward the behavioral channel, the predictability of my topics is more risk-based, so whether I can observe similar subsample concentrations in predictability remains unclear.

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

$$R_c^2 = 1 - \frac{\sum_{t=1}^{T} I_t^c \left(\hat{\epsilon}_t\right)^2}{\sum_{t=1}^{T} I_t^c \left(R_t^e - \bar{R}^e\right)^2}, \quad c = exp, \;\; rec, \;\; high, \;\; low, \tag{3.C.10}$$

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 (3.1); $\bar{R}^e$ is the full sample mean of the excess market return; and T is the number of observations for the full sample of 1871–2019. I classify months into expansions and recessions based on the National Bureau of Economic Research (NBER) business cycles. For sentiment periods, I follow Stambaugh et al. (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 (3.C.10), I compute the out-of-sample R^2_{OS} for each period. Similar to the previous out-of-sample analyses, I use the expanding estimation window, and the evaluation period began in January 1891.

Panel A of Table 3.C.6 reports the results with the in-sample R^2. 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.

I 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 I do not find evidence of different predicting powers of topics across the business cycles as commonly documented in the literature, I 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.

3.C.5 Predicting Returns on Characteristic Portfolios

In the main text, I document that *War* and the discourse topic index predict market returns. In this appendix, I investigate whether the return predictability of topics holds at the individual portfolio level. Following Huang et al. (2015), I 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, I run the following predictive regression:

$$R^e_{i,t+1} = \alpha_i + \beta_i x_t + \epsilon_{i,t+1}, \quad i = 1, \ldots, 40 \tag{3.C.11}$$

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 3.C.7 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 3.C.7 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 3.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.

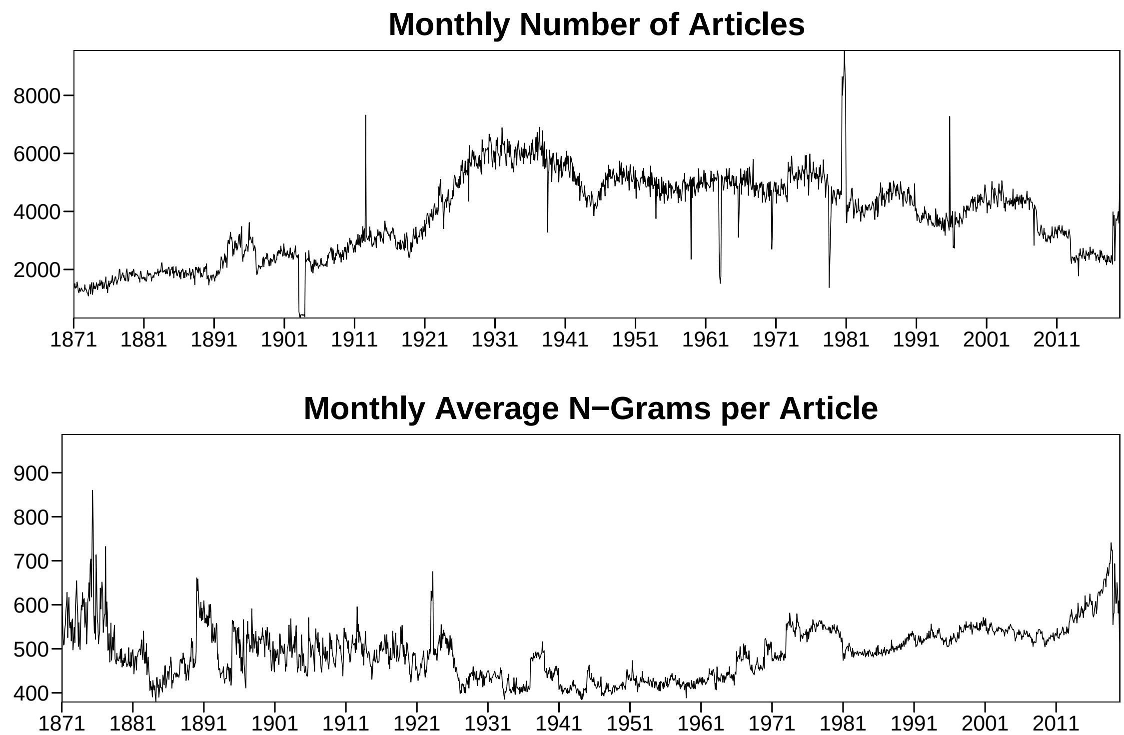
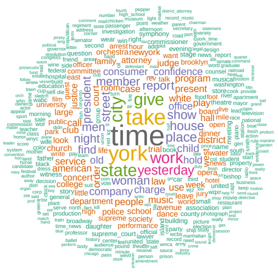


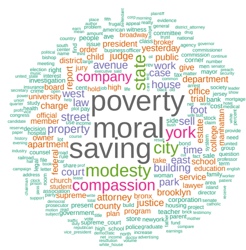
Figure 3.C.2. Discourse Topic Contents

This figure plots the frequencies of n-grams per topic over time. Frequencies are constructed according to the sLDA model described in Section 3.2, and the size of each n-gram indicates its frequency. The sample period is from January 1871 to October 2019.

Confidence

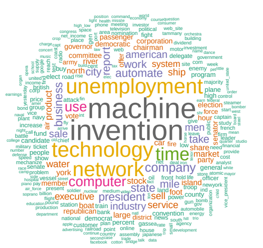


Savings

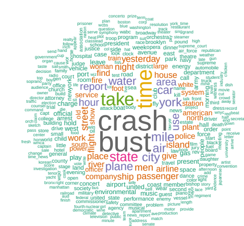


Conspicuous Consumption

Technology Replacing Jobs



Real Estate Crashes



Stock Market Bubbles

Stock Market Crashes

Wage and Labor Unions

Figure 3.C.3. Time Series of Topics Weights from the *NYT*

This figure plots the time series of monthly topic weights constructed according to the sLDA model as described in Section 3.2. 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 1871 to October 2019.

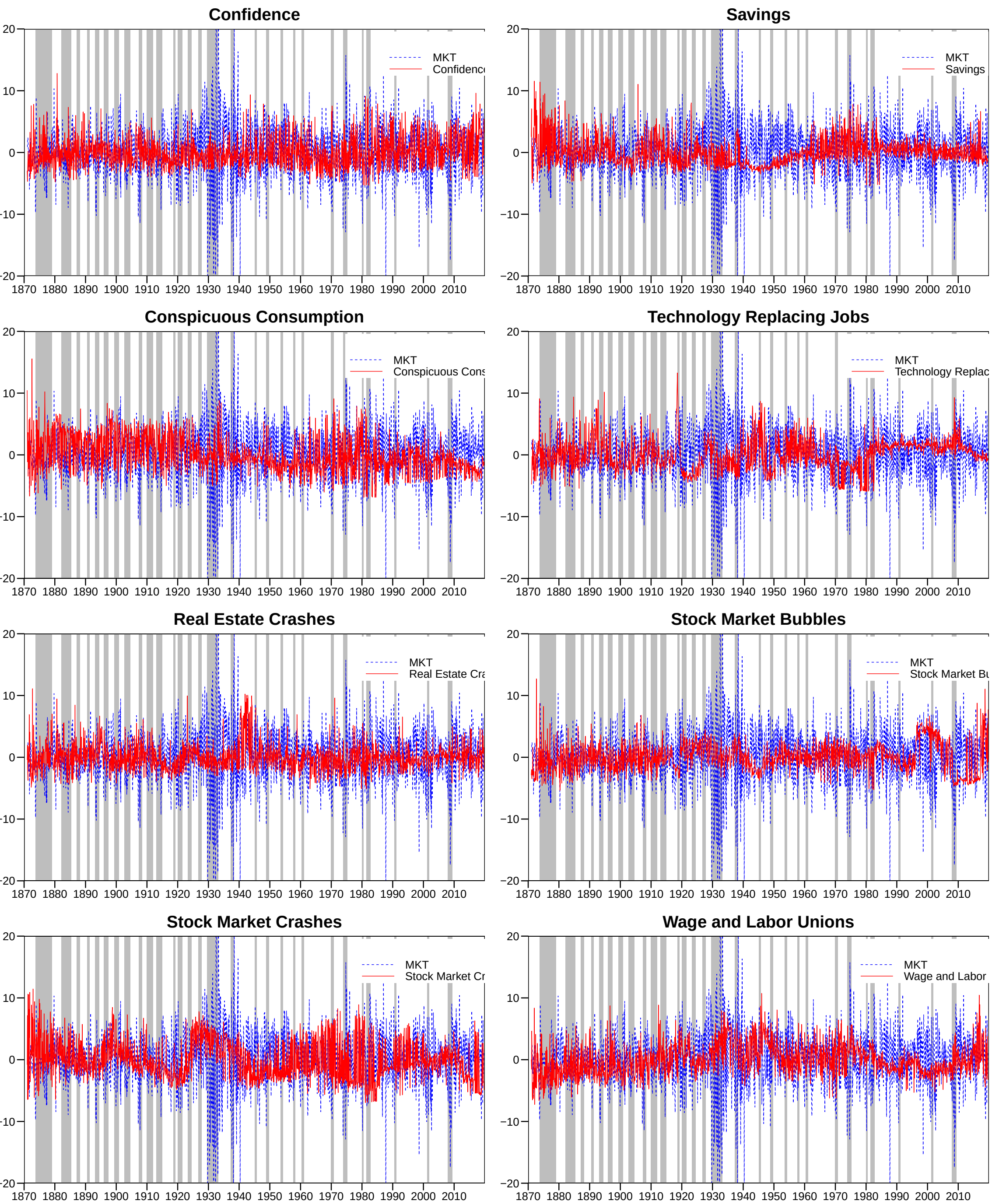


Figure 3.C.4. Historical Events

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.

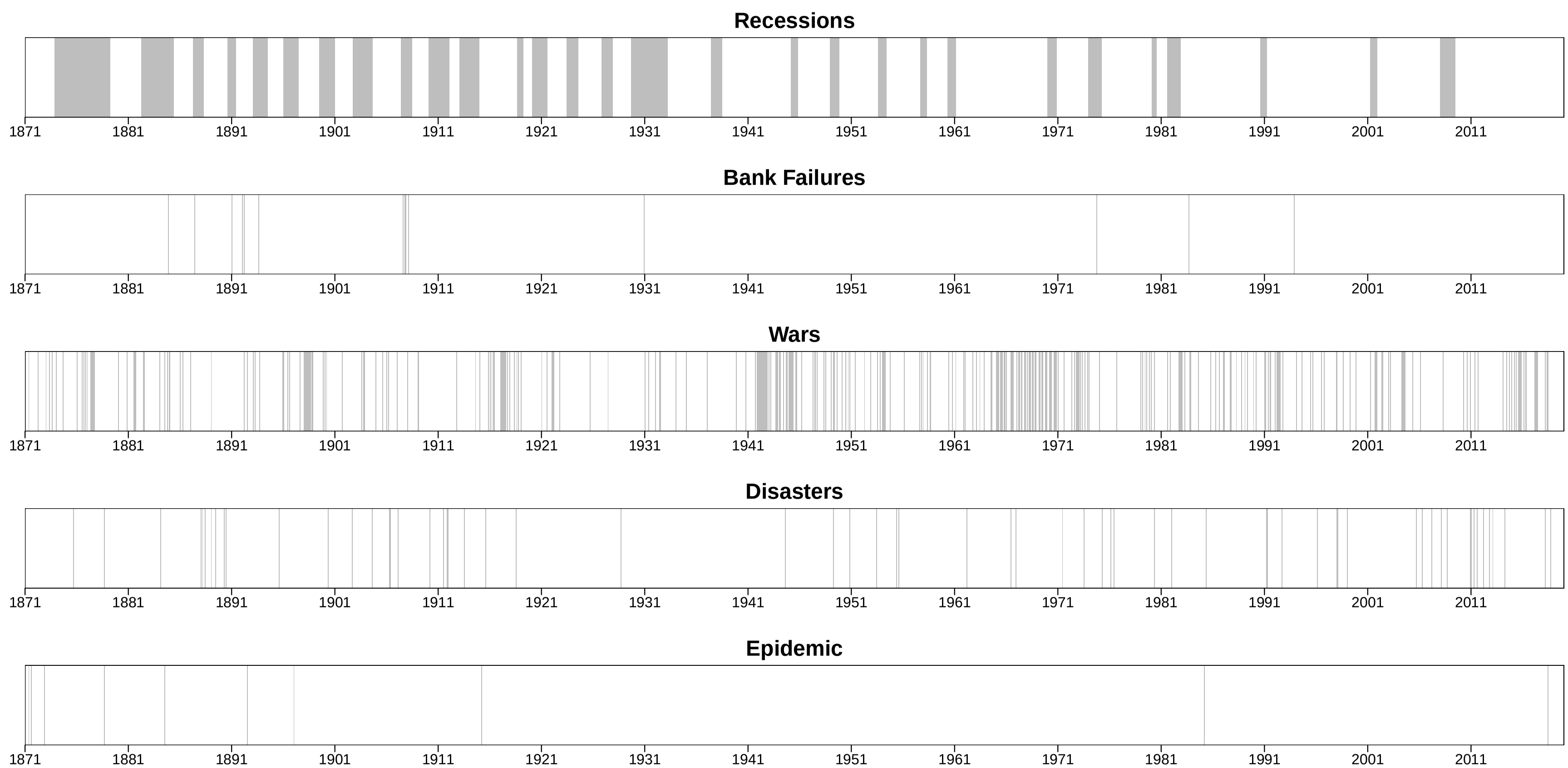


Table 3.C.1
Data Screening

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 3.C.2
Topic Correlations

This table presents the pairwise correlations among 14 monthly topic weights constructed according to the sLDA model described in Section 3.2 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 3.C.3
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 3.C.4
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):

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

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 3.C.5
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 3.C.6
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 1971 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 3.C.7
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

3.D Topic Weights Constructed by Raw Counts of Seed Words from the *NYT*

In this appendix, I conduct a robustness check for the main empirical results in the paper. Specifically, I 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 et al. (2017), Choudhury et al. (2019), Brown et al. (2020), and Bybee et al. (2021)). This paper blends the two branches by employing a semisupervised model in which I 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, I construct topic weights by simply counting the occurrences of seed words and scaling them by the total number of ngrams in the article.

Table 3.D.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).[45] This might seem low, but it shows the limitation of the raw 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

[45] On average, every month, I 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.

of 96%, much higher than the percentage (78%) obtained via the sLDA one. To remain consistent with the sLDA model, I also construct the PLS index from all topics.[46] Once again, the PLS index heavily loads on *War* and strongly correlates with this topic with a correlation coefficient of 99%.

To investigate whether manually constructed topics have the same market implications as the sLDA topics, I first use them to predict the monthly market returns in the sample. Table 3.D.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 Table 3.D.3, I 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 Table 3.D.4, 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 3.D.5, I report the out-of-sample R^2 computed with the frequency-based topics. Unsurprisingly, over the whole evaluation period of 1881–2019, the raw *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

[46] 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.

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, I need a statistical way to uncover and cluster them.

Table 3.D.1
Summary Statistics
Topic Weights Constructed by Raw 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 raw 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 3.D.2
Predicting One-Month Market Returns: Raw Topic Counts

This table presents the results of the following predictive regression:

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

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 raw 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. 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 3.D.3
Predicting Market Returns after Controlling for Economic Variables Topic Weights Constructed by Raw Counts of Seed Words

This table presents the results of the following predictive regression

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

in Panel A, and the following predictive regression

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

in Panel B, where R^e_{t+1} is the excess market return over the next month, x_t is the topic PLS index constructed by raw counts of seed words, and z_t is one of the 16 economic variables: 14 economic predictors from Goyal and Welch (2008), output gap from Cooper and Priestley (2009), and short interest from Rapach et al. (2016). The last row reports the results using PLS with the 16 economic variables. Returns are annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is in percentage; and t-stat is computed with the Newey and West (1987) standard errors. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

	Panel A: Univariate		Panel B: 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 3.D.4
Predicting Market Returns after Controlling for Uncertainty and Sentiment Variables Topic Weights Constructed by Raw Counts of Seed Words

This table presents the correlation between the topic PLS index with each of the uncertainty and sentiment variables in Panel A, the results of the following predictive regression

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

in Panel B, and the following predictive regression

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

in Panel C, where R^e_{t+1} is the excess market return over the next month, x_t is the topic PLS index constructed by raw counts of seed words, and z_t is one of the uncertainty variables—financial and macro uncertainty indexes from Jurado et al. (2015), economic policy uncertainty index from Baker et al. (2016), disagreement index from Huang et al. (2020), and implied volatility (VIX)—and 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). The last row reports the results using PLS with all uncertainty and sentiment variables. Returns are annualized percentages, and the independent variable is standardized to zero mean and unit variance. Adjusted R^2 is in percentage and t-stat is computed with the Newey and West (1987) standard errors. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

	Panel A: Correlations	Panel B: Univariate		Panel C: 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	185701-201910
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 3.D.5
Out-Of-Sample R^2
Topic Weights Constructed by Raw 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 raw 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
Topics	-0.26 *	-0.53 *	0.22 *	0.63 *
Shiller Topics	-0.69	-1.14	0.10	-0.59

3.E Discourse Topics from the *WSJ*

In this appendix, I check whether discourse topics extracted from 660 thousand *WSJ* articles predict stock market returns over the period 2000–2019.[47] WSJ is a mainstream media in the US, and readers are stock market participants. In contrast, *NYT*'s readers are educated and focus more on various news topics. In terms of politics, *NYT* and *WSJ* have polarized views. I apply sLDA to *WSJ* as an out-of-sample check. I focus on the past 20 years because this is the period where the *NYT* topics show the most robust predictability. I apply the same estimation method described in Section 3.2 to obtain the 14 time series of topic weights from the *WSJ* data.

Before extracting the 14 topics from the *WSJ* articles, I also conduct text-processing steps. Similar to the procedure applied to the *NYT* articles, I 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 Section 3.B for the description of text processing, cleaning, and converting into ngrams.

I plot the word clouds and time series of each topic in Figures 3.E.1 and 3.E.2. I report the summary statistics for these topics in Table 3.E.1.

Table 3.E.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

[47] Following the method in Section 3.2, I use the first 120 months from 1990 to 1999 to construct the first monthly topic weights.

prediction results (at the 5% level), although it is a negative predictor. Its R^2_{OS} is 0.89%, significant at the 10% level.

I also aggregate the topics with the PLS technique using all 14 topics (the "PLS" row) and 12 cases 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 estimate, it yields poor OOS results.

Table 3.E.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, I 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, I also document that *Stock Bubble* is a negative predictor for the *WSJ* articles. I 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.

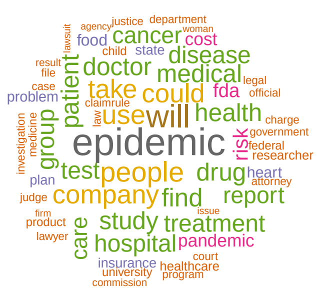
Figure 3.E.1. Discourse Topic Contents from the *WSJ*

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

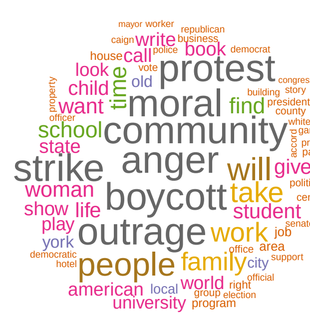
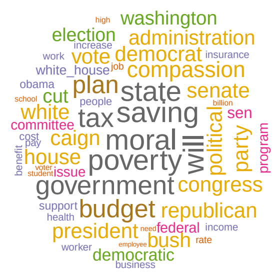


Figure 3.E.1. Discourse Topic Contents from the *WSJ* (Cont.)

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 3.E.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.

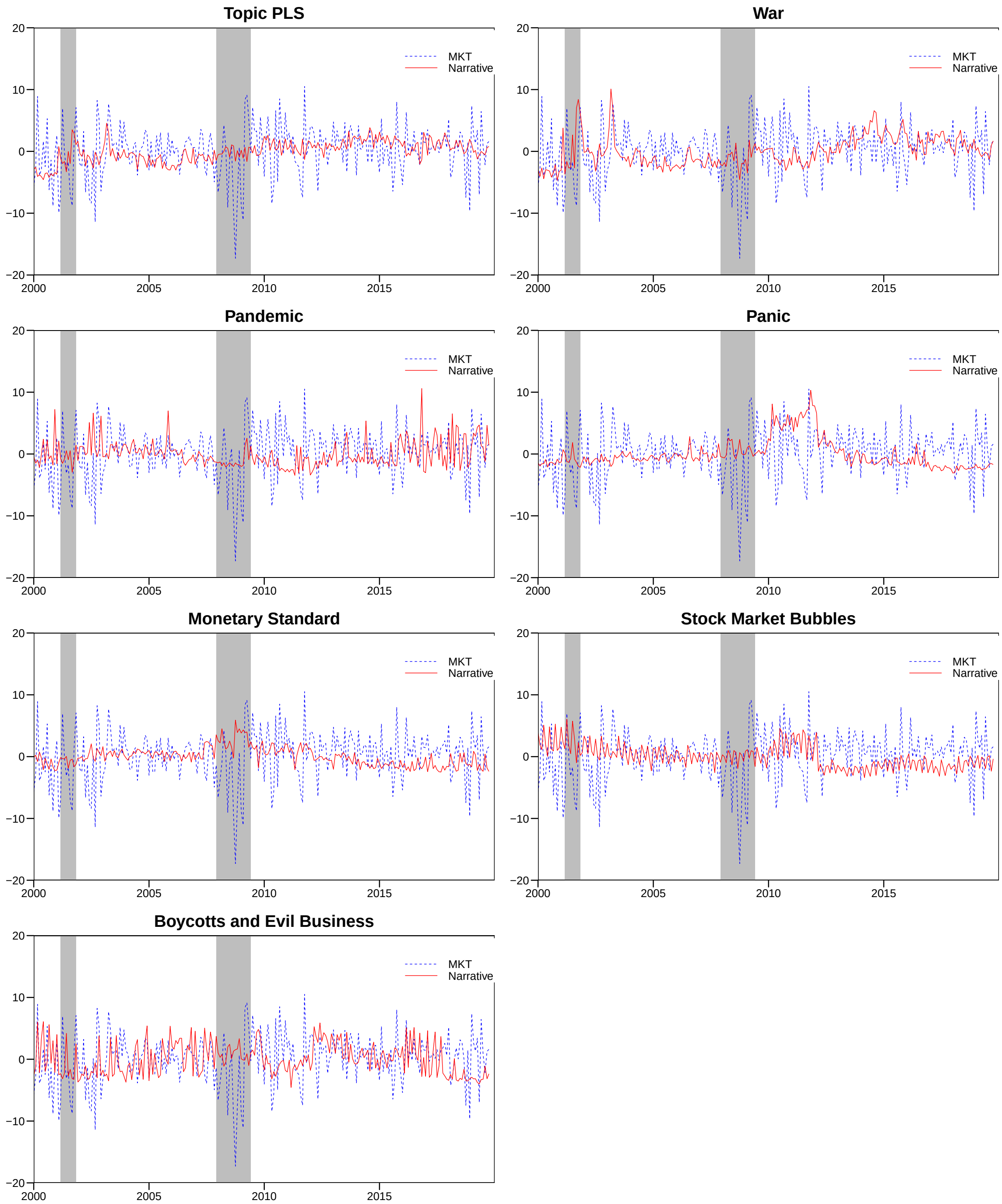


Figure 3.E.2. Time Series of Topic Weights from the *WSJ* (Cont.)

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.

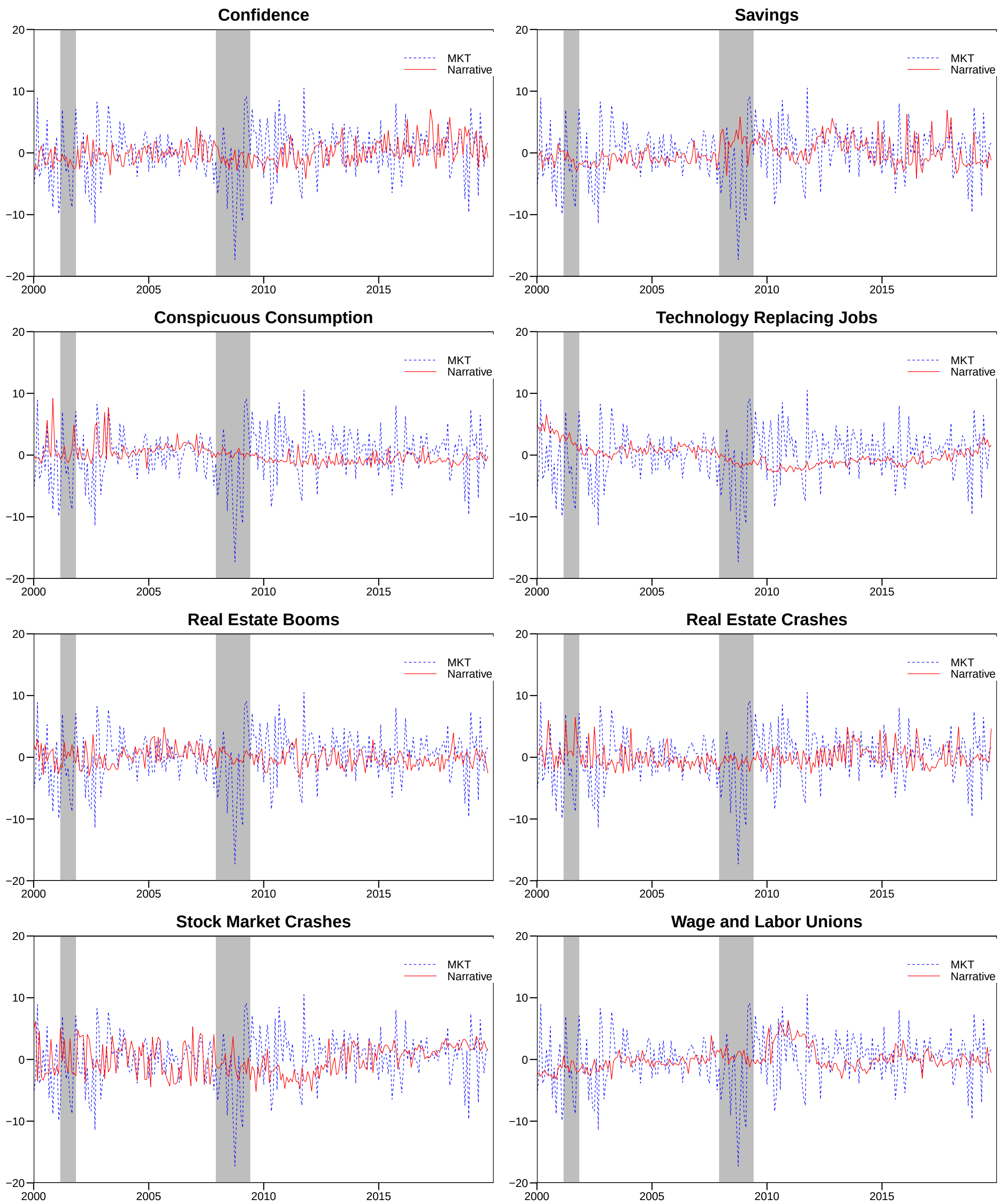


Table 3.E.1
Summary Statistics of Topics from the *WSJ*

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 3.2. 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 3.E.2
Predicting One-Month Market Returns with *WSJ* Topics

This table presents the results of the following predictive regression:

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

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 3.E.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

Adämmer, P. and Schüssler, R. A. (2020). Forecasting the Equity Premium: Mind the News! *Review of Finance*, 24(6):1313–1355.

Athey, S., Parashkevov, I., Sarukkai, V., and Xia, J. (2016). Bitcoin pricing, adoption, and usage: Theory and evidence.

Baker, M. and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61(4):1645–1680.

Baker, M. and Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21(2):129–152.

Baker, S. R., Bloom, N., and 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.

Berkman, H., Jacobsen, B., and Lee, J. B. (2011). Time-Varying Rare Disaster Risk and Stock Returns. *Journal of Financial Economics*, 101(2):313–332.

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

Blei, D. M., Ng, A. Y., and 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.

Bordalo, P., Gennaioli, N., La Porta, R., and Shleifer, A. (2020a). Expectations of fundamentals and stock market puzzles. Technical report, National Bureau of Economic Research.

Bordalo, P., Gennaioli, N., Ma, Y., and Shleifer, A. (2020b). Overreaction in macroeconomic expectations. *American Economic Review*, 110(9):2748–82.

Bordalo, P., Gennaioli, N., Porta, R. L., and Shleifer, A. (2019). Diagnostic expectations and stock returns. *Journal of Finance*, 74(6):2839–2874.

Boyd-Graber, J., Hu, Y., Mimno, D., and others, o. (2017). Applications of Topic Models. *Foundations and Trends in Information Retrieval*, 11(2-3):143–296.

Brock, W., Lakonishok, J., and LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *Journal of finance*, 47(5):1731–1764.

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

Brown, N. C., Crowley, R. M., and 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., and Xiu, D. (2021). Business News and Business Cycles. *Working Paper*.

Bybee, L., Kelly, B. T., and Su, Y. (2023). Narrative Asset Pricing: Interpretable Systematic Risk Factors from News Text. *Review of Financial Studies (forthcoming)*.

Caldara, D. and Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review*, 112(4):1194–1225.

Calomiris, C. W. and Mamaysky, H. (2019). How news and its context drive risk and returns around the world. *Journal of Financial Economics*, 133(2):299–336.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *Economic Journal*, 101(405):157–179.

Campbell, J. Y. and 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. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1(3):195–228.

Campbell, J. Y. and 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.

Chen, L., Da, Z., and Priestley, R. (2012). Dividend smoothing and predictability. *Management Science*, 58(10):1834–1853.

Chen, L., Da, Z., and Zhao, X. (2013). What drives stock price movements? *Review of Financial Studies*, 26(4):841–876.

Cherlin, A. J. (2010). *The Marriage-go-round: The State of Marriage and the Family in America Today*. Vintage, New York.

Chong, Y. Y. and Hendry, D. F. (1986). Econometric evaluation of linear macroeconomic models. *Review of Economic Studies*, 53(4):671–690.

Choudhury, P., Wang, D., Carlson, N. A., and 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. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.

Cochrane, J. (2009). *Asset pricing: Revised edition.* Princeton university press.

Cochrane, J. H. (1996). A cross-sectional test of an investment-based asset pricing model. *Journal of Political Economy*, 104(3):572–621.

Cochrane, J. H. (2005). *Asset pricing: Revised edition.* Princeton university press.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *Review of Financial Studies*, 21(4):1533–1575.

Cochrane, J. H. (2017). Macro-finance. *Review of Finance*, 21(3):945–985.

Cohen, R. B., Gompers, P. A., and Vuolteenaho, T. (2002). Who underreacts to cash-flow news? Evidence from trading between individuals and institutions. *Journal of financial Economics*, 66(2-3):409–462.

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

De La O, R. and Myers, S. (2021). Subjective cash flow and discount rate expectations. *Journal of Finance*, 76(3):1339–1387.

De La O, R. and Myers, S. (2022). When do subjective expectations explain asset prices? *Working paper.*

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial studies*, 22(5):1915–1953.

Detzel, A., Liu, H., Strauss, J., Zhou, G., and Zhu, Y. (2021). Learning and predictability via technical analysis: Evidence from bitcoin and stocks with hard-to-value fundamentals. *Financial Management*, 50(1):107–137.

Dictionary, O. E. (1993). Oxford English Dictionary. *Simpson, JA & Weiner, ESC.–1989.*

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

Dwyer, G. P. (2015). The economics of Bitcoin and similar private digital currencies. *Journal of Financial Stability*, 17:81–91.

Dyer, T., Lang, M., and 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. and Bollerslev, T. (1986). Modelling the Persistence of Conditional Variances. *Econometric Reviews*, 5(1):1–50.

Eshima, S., Imai, K., and Sasaki, T. (2020). Keyword Assisted Topic Models. *arXiv preprint arXiv:2004.05964*.

Fama, E. F. and French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy*, 96(2):246–273.

Farrell, J. (1995). Talk is cheap. *American Economic Review*, 85(2):186–190.

Farrell, J. and Rabin, M. (1996). Cheap talk. *Journal of Economic perspectives*, 10(3):103–118.

Ferguson, N. (2006). Political Risk and the International Bond Market between the 1848 Revolution and the Outbreak of the First World War. *Economic History Review*, 59(1):70–112.

Fischhoff, B., Slovic, P., and Lichtenstein, S. (1977). Knowing with Certainty: The Appropriateness of Extreme Confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4):552.

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

Gao, C. and Martin, I. W. (2021). Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment. *Journal of Finance*, 76(6):3211–3254.

Garcia, D. (2013). Sentiment during recessions. *Journal of Finance*, 68(3):1267–1300.

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

Giglio, S., Maggiori, M., Stroebel, J., and Utkus, S. (2021). Five facts about beliefs and portfolios. *American Economic Review*, 111(5):1481–1522.

Golez, B. (2014). Expected returns and dividend growth rates implied by derivative markets. *Review of Financial Studies*, 27(3):790–822.

Golez, B. and Koudijs, P. (2018). Four centuries of return predictability. *Journal of Financial Economics*, 127(2):248–263.

Gómez-Cram, R. (2022). Late to Recessions: Stocks and the Business Cycle. *The Journal of Finance*, 77(2):923–966.

Goyal, A. and Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508.

Goyal, A., Welch, I., and Zafirov, A. (2021). A Comprehensive Look at the Empirical Performance of Equity Premium Prediction II. *SSRN*.

Goyal, A., Zafirov, A., and Welch, I. (2022). A comprehensive look at the empirical performance of equity premium prediction ii. *Working Paper*.

Greenwood, R. and Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies*, 27(3):714–746.

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

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and Deliberation within the FOMC: A Computational Linguistics Approach. *Quarterly Journal of Economics*, 133(2):801–870.

Harvey, D. I., Leybourne, S. J., and Newbold, P. (1998). Tests for forecast encompassing. *Journal of Business & Economic Statistics*, 16(2):254–259.

Hassan, T. A., Hollander, S., van Lent, L., and Tahoun, A. (2019). Firm-Level Political Risk: Measurement and Effects. *Quarterly Journal of Economics*, 134(4):2135–2202.

Hillert, A. and Ungeheuer, M. (2019). The Value of Visibility. *SSRN*.

Hirshleifer, D., Mai, D., and Pukthuanthong, K. (2023). War Discourse and the Cross-Section of Expected Stock Returns. *Working Paper*.

Homer, S. and Sylla, R. E. (1996). *A History of Interest Rates*. Rutgers University Press.

Hou, K., Xue, C., and Zhang, L. (2020). Replicating anomalies. *Review of Financial Studies*, 33(5):2019–2133.

Huang, D., Jiang, F., Tu, J., and Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *Review of Financial Studies*, 28(3):791–837.

Huang, D., Li, J., and Wang, L. (2020). Are Disagreements Agreeable? Evidence from Information Aggregation. *Journal of Financial Economics*.

Huang, H. and Lee, T.-H. (2010). To combine forecasts or to combine information? *Econometric Reviews*, 29(5-6):534–570.

Jagarlamudi, J., Daumé III, H., and Udupa, R. (2012). Incorporating Lexical Priors into Topic Models. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 204–213. Association for Computational Linguistics.

Jiang, F., Lee, J., Martin, X., and Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics*, 132(1):126–149.

Julliard, C. and Ghosh, A. (2012). Can Rare Events Explain the Equity Premium Puzzle? *Review of Financial Studies*, 25(10):3037–3076.

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

Kelly, B. and Pruitt, S. (2013). Market expectations in the cross-section of present values. *Journal of Finance*, 68(5):1721–1756.

Kelly, B. and Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, 186(2):294–316.

Koijen, R. S. and Van Nieuwerburgh, S. (2011). Predictability of returns and cash flows. *Annual Review of Financial Economics*, 3(1):467–491.

Kuchler, T. and Zafar, B. (2019). Personal experiences and expectations about aggregate outcomes. *Journal of Finance*, 74(5):2491–2542.

Landier, A. and Thesmar, D. (2020). Earnings expectations during the covid-19 crisis. *Review of Asset Pricing Studies*, 10(4):598–617.

Larsen, V. H. and Thorsrud, L. A. (2019). The Value of News for Economic Developments. *Journal of Econometrics*, 210(1):203–218.

Le Bris, D. (2012). Wars, inflation and stock market returns in france, 1870–1945. *Financial History Review*, 19(3):337–361.

Lettau, M. and Van Nieuwerburgh, S. (2008). Reconciling the return predictability evidence: Reconciling the return predictability evidence. *Review of Financial Studies*, 21(4):1607–1652.

Li, J. and Yu, J. (2012). Investor attention, psychological anchors, and stock return predictability. *Journal of Financial Economics*, 104(2):401–419.

Lochstoer, L. A. and Tetlock, P. C. (2020). What drives anomaly returns? *Journal of Finance*, 75(3):1417–1455.

Loughran, T. and 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. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.

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

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

Ludvigson, S. C., Ma, S., and Ng, S. (2021). Uncertainty and business cycles: Exogenous impulse or endogenous response? *American Economic Journal: Macroeconomics*, 13(4):369–410.

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

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

Mcauliffe, J. and Blei, D. (2007). Supervised Topic Models. *Advances in Neural Information Processing Systems*, 20:121128.

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

Michaelides, A., Milidonis, A., and Nishiotis, G. P. (2019). Private information in currency markets. *Journal of Financial Economics*, 131(3):643–665.

Michaelides, A., Milidonis, A., Nishiotis, G. P., and Papakyriakou, P. (2015). The adverse effects of systematic leakage ahead of official sovereign debt rating announcements. *Journal of Financial Economics*, 116(3):526–547.

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

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

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Obaid, K. and Pukthuanthong, K. (2022). A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. *Journal of Financial Economics*.

Oosterlinck, K. and Landon-Lane, J. S. (2006). Hope Springs Eternal–French Bondholders and the Soviet Repudiation (1915–1919). *Review of Finance*, 10(4):507–535.

Pagnotta, E. and Buraschi, A. (2018). An equilibrium valuation of bitcoin and decentralized network assets. *Working Paper*.

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

Ramage, D., Hall, D., Nallapati, R., and Manning, C. D. (2009). Labeled lda: A supervised topic model for credit attribution in multi-labeled corpora. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 248–256.

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

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

Rhodes, R. (2012). *The Making of the Atomic Bomb*. Simon and Schuster, New York.

Rietz, T. A. (1988). The Equity Risk Premium: A Solution? *Journal of Monetary Economics*, 22(1):117–131.

Rutherford, E. (2012). The Scattering of α and β Particles by Matter and the Structure of the Atom. *Philosophical Magazine*, 92(4):379–398.

Schwert, G. W. (1990). Stock Market Volatility. *Financial Analysts Journal*, 46(3):23–34.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71(3):421–436.

Shiller, R. J. (2005). *Irrational exuberance*. Princeton University Press, Princeton, NJ.

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

Shiller, R. J. (2019). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

Smith, S. C. and Timmermann, A. (2021a). Break risk. *Review of Financial Studies*, 34(4):2045–2100.

Smith, S. C. and Timmermann, A. (2021b). Have risk premia vanished? *Journal of Financial Economics*.

Snowberg, E. and Wolfers, J. (2010). Explaining the Favorite–Long Shot Bias: Is It Risk-Love or Misperceptions? *Journal of Political Economy*, 118(4):723–746.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, 104(2):288–302.

Stein, J. C. (1989). Cheap talk and the Fed: A theory of imprecise policy announcements. *American Economic Review*, pages 32–42.

Steyvers, M. and Griffiths, T. (2007). Probabilistic topic models. In *Handbook of Latent Semantic Analysis*, pages 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.

Tversky, A. and Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5:297–323.

van Binsbergen, J. H., Bryzgalova, S., Mukhopadhyay, M., and Sharma, V. (2022). (Almost) 200 Years of News-Based Economic Sentiment. *Available at SSRN 4261249*.

Vuolteenaho, T. (2002). What drives firm-level stock returns? *Journal of Finance*, 57(1):233–264.

Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance*, 68(3):987–1035.

Walker, J. S. (2004). *Three Mile Island: A Nuclear Crisis in Historical Perspective*. University of California Press.

Watanabe, K. and Zhou, Y. (2020). Theory-Driven Analysis of Large Corpora: Semisupervised Topic Classification of the UN Speeches. *Social Science Computer Review*.

Williams, L. R. (1976). *How I made One million dollars in the commodity market last year*. Publikations-u. Handelsanst.

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

Zhou, G. (2018). Measuring investor sentiment. *Annual Review of Financial Economics*, 10:239–259.

Zhu, Y. and Zhou, G. (2009). Technical analysis: An asset allocation perspective on the use of moving averages. *Journal of Financial Economics*, 92(3):519–544.

VITA

Dat Y Mai was born in Ben Tre, Vietnam on March 5, 1993. He attended elementary and middle schools in Ben Tre City and graduated valedictorian from Ben Tre High School for the Gifted Students in May 2011. The following August, he entered Foreign Trade University-Vietnam and in May 2015 received a Bachelor of Economics with the first rank. He entered the University of Missouri-Columbia in August 2016 and received a Master of Business Administration in May 2018 and a Master of Arts in Quantitative Economics in December 2021. He is a Chartered Financial Analyst (CFA®) charter holder.