Forecasting Using Text-Based Uncertainty Measures

Kevin Benson* John Hull† Yoshio Nozawa†
Zissis Poulos† Vasily Strela* Yuntao Wu†

November 9, 2023

Abstract

We consider a new narrative-based measure of economic uncertainty derived from textual content in newspapers, and we compare it with existing narrative-based and market-based measures. We show that there is a strong relationship between narrative uncertainty and volatility in fixed income and commodity markets. We also show that narrative uncertainty has predictive power when it comes to forecasting the MOVE index and inflation.

Keywords: commodities, inflation, narrative uncertainty, machine learning, natural language processing

*RBC Capital Markets
†Joseph L. Rotman School of Management, University of Toronto

1 Introduction

In this project, we propose new measures of market uncertainty extracted from text in news articles. We quantify three aspects of uncertainty, namely uncertainty about monetary policy, inflation, and commodities. These are key towards determining the prices of fixed-income securities, and their importance has become more pronounced in today's market environment of high inflation and heightened economic uncertainty. Correctly capturing market uncertainty helps to assess the risk premiums of securities associated with these risks, as well as option prices that depend directly on uncertainty. It also complements studies[1] that focus primarily on the effects of first moments (*e.g.,* expectations).

From a practitioner's standpoint, developing news-based signals that can be used in forecasting models has been an area of active research within the investment community. Much of the effort, and many of the products offered by data vendors, have focused on applying NLP methods to news and social media in order to infer measures of sentiment. Such sentiment measures are taken as proxies for investor sentiment, and tend to be directional in nature; for instance, addressing the question of whether a company's press release is net positive or negative and predicting the impact on the issuer's stock price or credit spreads.

A common challenge in this type of work relates to the difficulty of isolating a clear signal between text-based sentiment measures and asset price movements, given the noise of overall market factors and other confounding factors. This raises the question of whether text-based sentiment is an informative predictor of volatility even if it is not a strong or consistent predictor of directional price movements. The research in this paper tackles this approach, by computing text-based measures of uncertainty and linking them to market-observed measures of uncertainty (i.e., volatility of asset prices).

Our uncertainty measures are extracted from major news outlets including the *New York Times* and *Wall Street Journal*. They are not only widely read by both professional and retail investors but are likely to have a high impact and capture new information in a timely fashion. Thus, the contents of the articles are likely to influence and reflect the uncertainty investors face in real time.

Unlike market-based measures of uncertainty, such as those derived from option prices (Bauer et al. [2021]), our measures directly capture how uncertain the news articles are. These measures are forward-looking, available at high frequency, and yet (unlike option prices) not affected by the risk premiums demanded by investors. Therefore, the use of our measure is an important step in advancing our knowledge of how to extract useful information from large amounts of underutilized resources (*i.e.,* news text) to correctly measure

[1] See for example Kuttner [2001], Bernanke and Kuttner [2005], Swanson [2021]

uncertainty. Specifically, we propose three measures of uncertainty: monetary policy uncertainty ('NMPU'), inflation uncertainty ('INFU') and commodity uncertainty ('GSCU'). We show that these new measures of uncertainty are strongly associated with option price-based measures of uncertainty and with the variation in the inflation of commodities, goods, and services. They suggest that studying the information contained in news articles is potentially useful in measuring the uncertainty faced by investors and in pricing options and a promising area for further work by practitioners.

2 Data and Methodology

In this section, we provide an overview of our methodology and explain our data to measure uncertainty. In addition, we explain the data sources used in our verification tests.

2.1 Extracting Uncertainty with Natural Language Processing

We employ a natural language processing (NLP) tool to construct an uncertainty measure. Full details are in Martineau et al. [2023]. Here we briefly explain the underlying idea.

NLP has recently witnessed significant advancements due to transformers, specifically the Bidirectional Encoder Representations from Transformers (BERT) architecture (Devlin et al. [2019]). BERT models undergo a "pre-training" phase using unlabeled text to predict missing words, ensuring they grasp the foundational structure and organization of language. Once pre-trained, these models can be slightly modified or "fine-tuned" to excel in specialized tasks. This ability to adapt a general pre-trained BERT for specific tasks, like uncertainty analysis, is known as transfer learning. For fine-tuning, only a minimal set of examples is needed, preventing the need to re-learn language properties.

Specifically, we use articles posted by the *Wall Street Journal* and *New York Times* obtained from ProQuest TDM Studio[2]. Then, we apply a publicly available BERT model[3] to each article, which is already fine-tuned for word-level uncertainty detection. The model is applied sequentially to every sentence of an article and it returns a list of all words that express uncertainty. We compute the article's uncertainty score as the number of words expressing uncertainty over the total number of words in the article. With the raw uncertainty score assigned to each available article from a single publication, we take the daily average to obtain a raw uncertainty score for day t, denoted UNC_t.

An attractive property of the method above is that it can identify uncertainty about

[2]https://tdmstudio.proquest.com/

[3]The model is available at https://huggingface.co/jeniakim/hedgehog

future outcomes and ignore uncertainty that is implied in the article but only existed in the past about outcomes that have already been resolved. The ability to discriminate between opinions about future economy/market states and current or past events has proven critical in other recent studies that design sentiment indices to predict GDP growth (see van Binsbergen et al. [2022]). Our measure quantifies only future uncertainty and is thus more relevant for capturing financial market uncertainty and option prices, which are inherently forward-looking.

The method we evaluate in this work belongs to the broader category of narrative-based measures of uncertainty, such as the ones proposed by Baker et al. [2016], Husted et al. [2020]. The advantage of the NLP approach is that uncertainty can be obtained at various frequencies, from daily to yearly, without the need for large article samples. Further, in traditional methodologies each article either contains uncertainty or not, whereas in the method we adopt each article obtains its own uncertainty score (ranging between 0.0 and 1.0). This allows for a fine-grained analysis that is necessary for capturing nuance in news text.

2.2 News Sample

The keywords used to identify relevant articles for each type of uncertainty index are listed in Table 1. We sample from newspapers only and exclude other source types such as magazines under the same publication. For the monetary policy uncertainty (NMPU), we use 70,356 *Wall Street Journal* articles and 21,101 *New York Times* articles from January 2, 1984 to March 22, 2023. These articles are mostly relevant to monetary policy, interest rates, and the Federal Open Market Committee. This dataset is the same dataset Martineau et al. [2023] uses. For the inflation uncertainty (INFU), we search over 60,413 *Wall Street Journal* articles and 42,294 *New York Times* articles from January 2, 1984 to June 30, 2023. For the commodity uncertainty (GSCU), the search terms are based on the major categories of the S&P GSCI components and some of the minor components with a weight of at least 4%. This results in 180,903 *Wall Street Journal* articles and 237,727 *New York Times* articles from January 2, 1984 to June 30, 2023.

2.3 Indices

In our analysis, we examine the correlation between our text-based uncertainty measures and the following measures of uncertainty:

1. Economic policy uncertainty (EPU): Daily from 1985-01 to 2023-02. This is a policy

uncertainty measure proposed in Baker et al. [2016]. (We scale down EPU by a factor of 100 in our analysis.)

2. Market-based monetary policy uncertainty (MPU-MKT): This is an uncertainty measure suggested by Bauer, Lakdawala, and Mueller [2021] for short-term rates based on Eurodollar futures options from 1990-01 to 2020-09.

3. Merrill Lynch Option Volatility Estimate (MOVE) monthly: The monthly MOVE is defined as the last valid MOVE of each month. The sample runs from 2002-11 to 2023-05.

4. Goldman Sachs Commodity Index Volatility (GSCI): We compute the daily log return for GSCI index and take the standard deviation over one-, three- and twelve-month window ending in month t. We denote the values by $\sigma^M_{t,\mathrm{GSCI}}$, $\sigma^Q_{t,\mathrm{GSCI}}$, $\sigma^Y_{t,\mathrm{GSCI}}$ respectively. The sample runs from 2002-11 to 2023-05.

5. CPI Inflation: We take the raw US CPI for all items in U.S. city average, all urban consumers, not seasonally adjusted, with base period in 1982-1984 from BLS[4], and compute the difference in the log CPI index over 12 months. The sample runs from 1984-01 to 2020-12.

For NMPU, GSCU, EPU, and Market-based MPU (MPU-MKT), we average trading day values over each month to get the monthly uncertainty scores.

For INFU, we count the number of times the word "inflation" appears in the article. We keep the articles when the word "inflation" appears once in the title or the first five sentences in the text, or when the word "inflation" appears at least twice in the article. Then we average the uncertainty of these articles over a month to get the monthly score for INFU. The screening leaves 34,479 articles in WSJ, and 22,264 articles in NYT.

In the following, we use UNC to denote one of NMPU, GSCU, INFU, EPU and MPU-MKT. Figure 1 shows the time-series data for the uncertainty indices we constructed based on WSJ news, while Figure 2 shows those based on NYT news. Figure 3 shows the time-series data for existing uncertainty measures, *i.e.*, EPU and MPU-MKT.

To assess the effect of changes in uncertainty, we consider the difference between the current month's uncertainty and last month's uncertainty. We also decompose changes in uncertainty into quarterly, yearly, and four-year horizons to incorporate the long-term time

[4]`https://www.bls.gov/cpi/data.htm`

effect:

$$\begin{aligned}
\mathrm{UNC}_t^{M-Q} &= \mathrm{UNC}_t - \frac{1}{3}\sum_{k=0}^{2}\mathrm{UNC}_{t-k}, \\
\mathrm{UNC}_t^{Q-Y} &= \frac{1}{3}\sum_{k=0}^{2}\mathrm{UNC}_{t-k} - \frac{1}{12}\sum_{k=0}^{11}\mathrm{UNC}_{t-k}, \\
\mathrm{UNC}_t^{Y-4Y} &= \frac{1}{12}\sum_{k=0}^{11}\mathrm{UNC}_{t-k} - \frac{1}{48}\sum_{k=0}^{47}\mathrm{UNC}_{t-k}.
\end{aligned} \tag{1}$$

Each term is a detrended moving average over the appropriate calendar interval (month, quarter, year, four years). Similar types of decomposition are used to capture variability in persistence of the fundamentals and long-cycle dependencies by Fisher et al. [2022], Ortu et al. [2013], Calvet and Fisher [2007], Martineau et al. [2023]. Here, we detrend UNC instead of the fundamentals, since the uncertainty values are now the exogenous predictive variable.

Panel A. of Table 2 shows the summary statistics of the indices. Panel B. of Table 2 shows the correlation between all the uncertainty measures. Our text-based uncertainty measures have a low correlation with the existing measures of economic policy uncertainty and market-based uncertainties, suggesting that there is potentially new information in our measures.

3 Results

3.1 MOVE Index

In this section, we analyze how our text-based measures of uncertainty relate to existing measures of uncertainty, such as the interest rate volatility.[5]

For trading, it is useful to forecast changes in option prices in the future. With this in mind, we regress next-month changes in MOVE on the current-month changes in the monetary policy uncertainty:

$$\mathrm{MOVE}_{t+1} - \mathrm{MOVE}_t = \alpha + \beta_1(\mathrm{UNC}_t - \mathrm{UNC}_{t-1}) + \epsilon_t. \tag{3}$$

[5] In the Appendix, we show that our uncertainty measure is strongly correlated with the MOVE index. To test this, we run a regression of changes in MOVE on changes in our monetary policy uncertainty:

$$\mathrm{MOVE}_t - \mathrm{MOVE}_{t-1} = \alpha + \beta_1(\mathrm{UNC}_t - \mathrm{UNC}_{t-1}) + \epsilon_t. \tag{2}$$

The results are reported in Tables A1 and A2.

Table 3 reports the results of this regression. Interestingly, we find significant slope coefficients on WSJ, EPU, MPU-MKT with negative signs. This suggests that when monetary policy uncertainty rises during a month, the MOVE index falls during the next month. Combined with the findings in Table A1, this suggests a mean reversion in uncertainty: higher-than-average uncertainty today predicts negative changes in the future.

Table 4 reports the regression estimates for the combined variables. Combining the uncertainty measures increases the adjusted R-squared from 0.040 to 0.068, suggesting that aggregating these measures helps to better predict the future. In addition, the sign of each variable is consistent with the individual regressions.

Taken together, our monetary-policy-based uncertainty measures capture the volatility priced in options on Treasury securities. Next, we examine the information content in our inflation uncertainty measure.

3.2 Uncertainty in CPI Inflation

For CPI inflation, we consider a regression of absolute CPI inflation on changes in the inflation uncertainty measure:

$$|\log \mathrm{CPI}_{t+12} - \log \mathrm{CPI}_t| = \alpha + \beta_1 \mathrm{UNC}_t^{M-Q} + \beta_2 \mathrm{UNC}_t^{Q-Y} + \beta_3 \mathrm{UNC}_t^{Y-4Y} + \epsilon_t. \tag{4}$$

This regression tests the effect of uncertainty changes on future variations in CPI.

Table 5 reports the results.[6] Many of the estimated coefficients are not significant. Interestingly, the variable defined as the average WSJ uncertainty this year minus the average over the last four years does have some predictive power: In particular, it is positively related to the change in inflation next year.

Table 6 reports the regression estimates for the combined variables. We only consider the UNC_t^{Y-4Y} variable because it is the only variable that has predictive power. In the forward prediction, adding the variables that have insignificant R^2 in the individual prediction decreases the overall R^2. Notably, WSJ-INFU remains significant in the multivariate forecast. However, EPU, MPU, and NYT-INFU don't help improve the predictions.

[6] In the Appendix, we report the results of the contemporaneous correlation by estimating a regression

$$|\log \mathrm{CPI}_t - \log \mathrm{CPI}_{t-12}| = \alpha + \beta_1 \mathrm{UNC}_t^{M-Q} + \beta_2 \mathrm{UNC}_t^{Q-Y} + \beta_3 \mathrm{UNC}_t^{Y-4Y} + \epsilon_t. \tag{5}$$

The results are reported in Tables A3 and A4.

3.3 Commodity Index Volatility

CPI inflation measures changes in the weighted average of the prices of all goods and services faced by consumers. However, the inflation uncertainty in news articles may be biased toward prices of newsworthy goods, such as gasoline. Therefore, it is possible that our measures are related more closely to the prices of those commodities rather than the general price index.

To examine this hypothesis, we study how the uncertainty scores correlate with the standard deviations in the Goldman Sachs commodity index. To this end, we consider the prediction model:

$$y_{t+h} - y_t = \alpha + \beta_1 \mathrm{UNC}_t + \epsilon_t, \tag{6}$$

where y_t is the monthly, quarterly, yearly standard deviation of log returns of GSCI indices ending in month t, *i.e.,* $\sigma^M_{t,\mathrm{GSCI}}$, $\sigma^Q_{t,\mathrm{GSCI}}$, $\sigma^Y_{t,\mathrm{GSCI}}$. The current month uncertainty level is used to predict how the standard deviation of log returns of GSCI will change after $h = 1, 3, 12$ months. Since the left-hand-side variable concerns commodity prices, we also use the text-based commodity uncertainty measures (WSJ-GSCU and NYT-GSCU) as an additional regressor.

Table 7 reports the estimation results. Overall, the uncertainty index we constructed and the EPU are negatively correlated to future changes in the volatility of GSCI. EPU is a strong indicator of how the volatility of GSCI will change over one-, three-, and twelve-month periods. Higher EPU of the current month means that the volatility of GSCI in the future will drop as compared to the current month. WSJ-GSCU is able to predict the change in volatility of GSCI over the next month, while NYT-GSCU is able to predict the changes over longer time periods.

Table 8 reports the combined regression of the variables. When the variables are combined to predict the changes in the volatility of GSCI, EPU remains significant across all regression specifications. On the other hand, the significance of WSJ-GSCU and NYT-NMPU are suppressed by EPU.[7] It is worth noting that, unlike our uncertainty measure, the EPU index includes components other than news narratives, such as tax code expiration dates and forecaster disagreement on macroeconomic variables. We leave it as an open question

[7] In Appendix, we also consider the following regression:

$$y_t = \alpha + \beta_1 \mathrm{UNC}_t + \epsilon_t, \tag{7}$$

where y_t is the monthly, quarterly, yearly standard deviation of log returns of GSCI indices ending in month t, *i.e.,* $\sigma^M_{t,\mathrm{GSCI}}$, $\sigma^Q_{t,\mathrm{GSCI}}$, $\sigma^Y_{t,\mathrm{GSCI}}$.

In Tables A5 and A6, we report the estimated coefficient β_1. In the table, the coefficient on the WSJ-based monetary policy and commodity uncertainty measures, as well as the economic policy uncertainty measure, are positively correlated with commodity volatility.

whether this is one of the reasons behind EPU's superior predictive power when it comes to the volatility of commodities prices.

4 Conclusion

In this paper, we demonstrate that uncertainty measures constructed by applying NLP to newspaper articles capture the uncertainty in the financial market well. The explanatory power is consistent across three broad types of uncertainty: namely, monetary policy, inflation, and commodity prices.

Since uncertainty can be traded via options, our results suggest that these NLP algorithms are potentially useful in predicting option prices. Our analysis shows that this is indeed the case for the interest-rate volatility measure (the MOVE index). In future research, one can easily extend this methodology to other asset classes such as equity index options (e.g., VIX), options on foreign exchanges, and international stocks.

The text-based measures are potentially useful in forming trading strategies because they are not contaminated by market microstructure noise as traded option prices are. To the extent those measures capture the fundamental values of options, it could predict future option prices.

Lastly, it will also be interesting to study how various text-based uncertainties move in response to salient events, including the news about the COVID-19 pandemic and ESG-related corporate incidents. While these events are known to significantly influence stock prices, the channel through which the event moves asset prices is not fully understood. The response of uncertainty can potentially help explain why asset prices move disproportionately to the magnitude of news and why liquidity tends to dry up simultaneously. These topics are left for future research projects.

Table 1: Search Words

This table reports the search terms used to select articles to construct the uncertainty indices from *Wall Street Journal* and *New York Times*. We retrieve the articles where the search words appear in either the headline or the article.

Uncertainty	Newspapers search words
Monetary Policy Uncertainty (NMPU)	('monetary policy' OR 'monetary policies' OR 'interest rate' OR 'interest rates' OR 'Federal fund rate' OR 'Federal funds rate' OR 'Fed fund rate' OR 'Fed funds rate') AND ('Federal Reserve' OR 'the Fed' OR 'Federal Open Market Committee' OR 'FOMC')
Inflation Uncertainty (INFU)	'Inflation'
Commodity Uncertainty (GSCU)	'Energy' OR 'Industrial Metals' OR 'Precious Metals' OR 'Agriculture' OR 'Livestock' OR 'Crude Oil' OR 'Gasoline' OR 'Gasoil' OR 'Heating Oil' OR 'Copper' OR 'Gold'

Table 2: Summary Statistics

Panel A. of this table reports the general statistics (mean, standard deviation, and the 1%, 50%, 99%-tiles) of the uncertainty indices we constructed (NMPU, INFU, GSCU), the existing uncertainty measures (EPU, MPU-MKT) and the existing indices relevant to market uncertainty, inflation and commodity volatility.
Panel B. of this table reports the correlation between all the constructed and existing uncertainty measures.

Panel A. Summary Statistics

	Mean	Std	1%	50%	99%
WSJ-NMPU	0.015	0.003	0.007	0.015	0.021
WSJ-INFU	0.017	0.002	0.012	0.017	0.024
WSJ-GSCU	0.011	0.001	0.008	0.011	0.014
NYT-NMPU	0.016	0.002	0.012	0.016	0.023
NYT-INFU	0.016	0.003	0.011	0.016	0.022
NYT-GSCU	0.011	0.001	0.009	0.011	0.012
EPU	1.038	0.542	0.406	0.891	2.762
MPU-MKT	0.902	0.338	0.364	0.941	1.559
MOVE	85.770	31.837	43.903	76.800	199.288
CPI	174.413	51.483	85.402	173.700	259.215
$\sigma^M_{t,\text{GSCI}}$	0.011	0.006	0.004	0.010	0.037
$\sigma^Q_{t,\text{GSCI}}$	0.012	0.005	0.005	0.011	0.032
$\sigma^Y_{t,\text{GSCI}}$	0.012	0.005	0.005	0.011	0.028

Panel B. Correlation

	WSJ-NMPU	WSJ-INFL	WSJ-GSCU	NYT-NMPU	NYT-INFU	NYT-GSCU	EPU	MPU-MKT
WSJ-NMPU	1.000	0.411	0.545	0.110	0.231	-0.027	0.114	-0.552
WSJ-INFL	0.411	1.000	0.282	0.062	0.137	-0.017	0.013	-0.166
WSJ-GSCU	0.545	0.282	1.000	0.070	0.052	0.225	0.144	-0.240
NYT-NMPU	0.110	0.062	0.070	1.000	0.441	0.135	0.187	-0.082
NYT-INFU	0.231	0.137	0.052	0.441	1.000	-0.048	0.082	-0.244
NYT-GSCU	-0.027	-0.017	0.225	0.135	-0.048	1.000	0.145	0.146
EPU	0.114	0.013	0.144	0.187	0.082	0.145	1.000	-0.067
MPU-MKT	-0.552	-0.166	-0.240	-0.082	-0.244	0.146	-0.067	1.000

Table 3: Forecast for MOVE Index

This table reports the results of the following forecast model:

$$\text{MOVE}_{t+1} - \text{MOVE}_t = \alpha + \beta_1(\text{UNC}_t - \text{UNC}_{t-1}) + \epsilon_t,$$

where UNC_t can be WSJ-NMPU, NYT-NMPU, EPU, and MPU-MKT. Here the time step for MOVE indices are shifted forward by 1 month to make it a prediction model. The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

	WSJ-NMPU	NYT-NMPU	EPU	MPU-MKT
	(1)	(2)	(3)	(4)
ΔUNC	-1074.8695**	447.2558*	-4.4964*	-41.3671***
	(528.89)	(235.61)	(2.68)	(15.15)
Intercept	-0.0154	-0.0149	-0.0030	-0.5035
	(0.76)	(0.77)	(0.77)	(0.87)
R^2	0.011	0.005	0.008	0.040
N	243	243	243	215

Table 4: Forecast for MOVE Index: Combining All Variables

This table reports the results of the following regression:

$$\text{MOVE}_{t+1} - \text{MOVE}_t = \alpha + \beta_1(\text{UNC}_t - \text{UNC}_{t-1}) + \epsilon_t,$$

where UNC_t can be WSJ-NMPU, NYT-NMPU, EPU, MPU-MKT and weighted combinations of the variables. The number of observations are set to be the minimum data range available (MPU-MKT). Here the time step for MOVE indices are shifted forward by 1 month to make it a prediction model. The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

	Combined 1 (1)	Combined 2 (2)	Combined 3 (3)
ΔMPU-MKT	-40.2094***	-41.7724***	-41.9443***
	(14.64)	(14.67)	(15.56)
ΔWSJ-NMPU	-898.9731	-1035.4571*	-998.5871*
	(571.21)	(568.27)	(581.44)
ΔNYT-NMPU		536.7384**	562.6752**
		(227.63)	(225.66)
ΔEPU			-5.3593**
			(2.67)
Intercept	-0.4875	-0.4896	-0.4623
	(0.86)	(0.86)	(0.87)
R^2	0.046	0.055	0.068
N	215	215	215

Table 5: Forecast for Absolute Changes in CPI

This table reports the results of the following forecast model:

$$|\log \mathrm{CPI}_{t+12} - \log \mathrm{CPI}_t| = \alpha + \beta_1 \mathrm{UNC}_t^{M-Q} + \beta_2 \mathrm{UNC}_t^{Q-Y} + \beta_3 \mathrm{UNC}_t^{Y-4Y} + \epsilon_t,$$

where UNC can be WSJ-INFU, NYT-INFU, EPU, and MPU-MKT. The definition of UNC_t^{M-Q}, UNC_t^{Q-Y} and UNC_t^{Y-4Y} can be found in Sec. 2.3. The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

	WSJ-INFU (1)	NYT-INFU (2)	EPU (3)	MPU-MKT (4)
UNC^{M-Q}	-0.1073	-0.0827	-0.0034	-0.0062
	(0.24)	(0.20)	(0.00)	(0.01)
UNC^{Q-Y}	-0.0016	-0.7927	-0.0004	-0.0118
	(0.75)	(0.71)	(0.00)	(0.01)
UNC^{Y-4Y}	3.2992**	-0.8668	-0.0065	0.0022
	(1.58)	(1.47)	(0.01)	(0.01)
Intercept	0.0259***	0.0258***	0.0254***	0.0233***
	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.025	0.004	0.014	0.012
N	432	432	420	360

Table 6: Forecast for Absolute Changes in CPI: Combining All Variables

This table reports the results of the following forecast model:

$$|\log \mathrm{CPI}_{t+12} - \log \mathrm{CPI}_t| = \alpha + \gamma_1 \mathrm{UNC}_t^{Y-4Y} + \epsilon_t,$$

where UNC can be WSJ-INFU, NYT-INFU, EPU, MPU-MKT and weighted combinations of the variables. The number of observations are set to be the minimum data range available (MPU-MKT). The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

	Combined 1 (1)	Combined 2 (2)	Combined 3 (3)
WSJ-INFU^{Y-4Y}	4.4129***	4.4111***	4.4026***
	(1.36)	(1.37)	(1.39)
EPU^{Y-4Y}	0.0015	0.0012	0.0013
	(0.00)	(0.00)	(0.00)
MPU^{Y-4Y}		0.0015	0.0015
		(0.00)	(0.00)
NYT-INFU^{Y-4Y}			0.0504
			(1.43)
Intercept	0.0232***	0.0233***	0.0233***
	(0.00)	(0.00)	(0.00)
R^2	0.077	0.075	0.072
N	360	360	360

Table 7: Forecast for Commodity Volatility

This table reports the results of the following regression:

$$y_{t+h} - y_t = \alpha + \beta_1 \text{UNC}_t + \epsilon_{t+h}, \quad h = 1, 3, 12.$$

where y_t is the monthly, quarterly, yearly standard deviation of log returns of GSCI indices ending in month t, *i.e.,* $\sigma^M_{t,\text{GSCI}}$, $\sigma^Q_{t,\text{GSCI}}$, $\sigma^Y_{t,\text{GSCI}}$. $h = 1$ for $\sigma^M_{t+1,\text{GSCI}}$, $h = 3$ for $\sigma^Q_{t+3,\text{GSCI}}$, and $h = 12$ for $\sigma^Y_{t+12,\text{GSCI}}$. UNC, UNC can be WSJ-NMPU, WSJ-GSCU, NYT-NMPU, NYT-GSCU, EPU, and MPU-MKT. The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

Panel A. $\sigma^M_{t+1,\text{GSCI}} - \sigma^M_{t,\text{GSCI}}$

	WSJ-NMPU (1)	WSJ-GSCU (2)	NYT-NMPU (3)	NYT-GSCU (4)	EPU (5)	MPU-MKT (6)
UNC	-0.0434	-0.2959**	0.0129	-0.2957	-0.0015***	0.0002
	(0.04)	(0.15)	(0.10)	(0.26)	(0.00)	(0.00)
Intercept	0.0007	0.0034**	-0.0002	0.0031	0.0016***	-0.0002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	-0.001	0.002	-0.002	-0.000	0.024	-0.003
N	470	470	470	470	458	369

Panel B. $\sigma^Q_{t+3,\text{GSCI}} - \sigma^Q_{t,\text{GSCI}}$

	WSJ-NMPU (1)	WSJ-GSCU (2)	NYT-NMPU (3)	NYT-GSCU (4)	EPU (5)	MPU-MKT (6)
UNC	-0.0609	-0.1952	-0.2167*	-0.5981*	-0.0024***	0.0008
	(0.07)	(0.22)	(0.13)	(0.35)	(0.00)	(0.00)
Intercept	0.0009	0.0022	0.0036*	0.0064*	0.0026***	-0.0007
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	-0.001	0.000	0.008	0.005	0.069	-0.000
N	469	469	469	469	457	369

Panel C. $\sigma^Y_{t+12,\text{GSCI}} - \sigma^Y_{t,\text{GSCI}}$

	WSJ-NMPU (1)	WSJ-GSCU (2)	NYT-NMPU (3)	NYT-GSCU (4)	EPU (5)	MPU-MKT (6)
UNC	-0.0672	-0.0170	-0.0735	-0.8597*	-0.0026***	0.0013
	(0.14)	(0.34)	(0.13)	(0.47)	(0.00)	(0.00)
Intercept	0.0013	0.0005	0.0015	0.0094*	0.0030***	-0.0010
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	-0.000	-0.002	-0.001	0.011	0.068	0.004
N	460	460	460	460	448	369

Table 8: Forecast for Commodity Volatility: Combining All Variables

This table reports the results of the following regression:

$$y_{t+h} - y_t = \alpha + \beta_1 \text{UNC}_t + \epsilon_{t+h}, \quad h = 1, 3, 12.$$

where y_t is the monthly, quarterly, yearly standard deviation of log returns of GSCI indices ending in month t, *i.e.,* $\sigma^M_{t,\text{GSCI}}$, $\sigma^Q_{t,\text{GSCI}}$, $\sigma^Y_{t,\text{GSCI}}$. $h = 1$ for $\sigma^M_{t+1,\text{GSCI}}$, $h = 3$ for $\sigma^Q_{t+3,\text{GSCI}}$, and $h = 12$ for $\sigma^Y_{t+12,\text{GSCI}}$. UNC can be WSJ-NMPU, WSJ-GSCU, NYT-NMPU, NYT-GSCU, EPU, MPU-MKT and weighted combinations of the variables. The number of observations are set to be the minimum data range available (MPU-MKT). The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

Panel A. $\sigma^M_{t+1,\text{GSCI}} - \sigma^M_{t,\text{GSCI}}$

	Combined 1 (1)	Combined 2 (2)	Combined 3 (3)	Combined 4 (4)	Combined 5 (5)
EPU	-0.0015**	-0.0015**	-0.0015**	-0.0015**	-0.0015**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
WSJ-GSCU	-0.1556	-0.1876	-0.1924	-0.1850	-0.1833
	(0.15)	(0.23)	(0.25)	(0.24)	(0.24)
WSJ-NMPU		0.0183	0.0196	0.0158	0.0132
		(0.07)	(0.08)	(0.07)	(0.08)
NYT-GSCIU			0.0272	0.0035	0.0058
			(0.36)	(0.34)	(0.36)
NYT-NMPU				0.0590	0.0587
				(0.13)	(0.13)
MPU-MKT					-0.0000
					(0.00)
Intercept	0.0033*	0.0034*	0.0031	0.0024	0.0025
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.023	0.021	0.018	0.016	0.013
N	369	369	369	369	369

Panel B. $\sigma^{Q}_{t+3,\text{GSCI}} - \sigma^{Q}_{t,\text{GSCI}}$

	Combined 1 (1)	Combined 2 (2)	Combined 3 (3)	Combined 4 (4)	Combined 5 (5)
EPU	-0.0027***	-0.0027***	-0.0027***	-0.0025***	-0.0025***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
WSJ-GSCU	0.1114	0.2903	0.3555	0.3336	0.3214
	(0.26)	(0.39)	(0.38)	(0.38)	(0.36)
WSJ-NMPU		-0.1023	-0.1198	-0.1086	-0.0896
		(0.12)	(0.12)	(0.12)	(0.14)
NYT-GSCIU			-0.3693	-0.2992	-0.3164
			(0.40)	(0.39)	(0.42)
NYT-NMPU				-0.1740	-0.1716
				(0.12)	(0.12)
MPU-MKT					0.0003
					(0.00)
Intercept	0.0015	0.0010	0.0044	0.0064	0.0061
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
R^2	0.077	0.077	0.077	0.081	0.078
N	369	369	369	369	369

Panel C. $\sigma^{Y}_{t+12,\text{GSCI}} - \sigma^{Y}_{t,\text{GSCI}}$

	Combined 1 (1)	Combined 2 (2)	Combined 3 (3)	Combined 4 (4)	Combined 5 (5)
EPU	-0.0026***	-0.0027***	-0.0025***	-0.0026***	-0.0025***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
WSJ-GSCU	0.2399	0.5341	0.6518	0.6529	0.6109
	(0.39)	(0.55)	(0.56)	(0.56)	(0.51)
WSJ-NMPU		-0.1684	-0.2000	-0.2005	-0.1346
		(0.20)	(0.19)	(0.19)	(0.21)
NYT-GSCIU			-0.6662	-0.6699	-0.7294
			(0.56)	(0.54)	(0.59)
NYT-NMPU				0.0091	0.0176
				(0.12)	(0.12)
MPU-MKT					0.0010
					(0.00)
Intercept	0.0002	-0.0007	0.0054	0.0053	0.0043
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
R^2	0.064	0.068	0.072	0.070	0.070
N	369	369	369	369	369

Figure 1: WSJ Time-series

This figure shows the monthly WSJ-NMPU, WSJ-INFU, WSJ-GSCU from Jan 1984 together with the corresponding macroeconomic variables. The first row shows WSJ-NMPU with MOVE index. The second row shows WSJ-INFU with CPI. The third row shows WSJ-GSCU with σ^M_{GSCI}. All the variables are smoothed by a 12-month moving average. The gray vertical bars are NBER recessions.

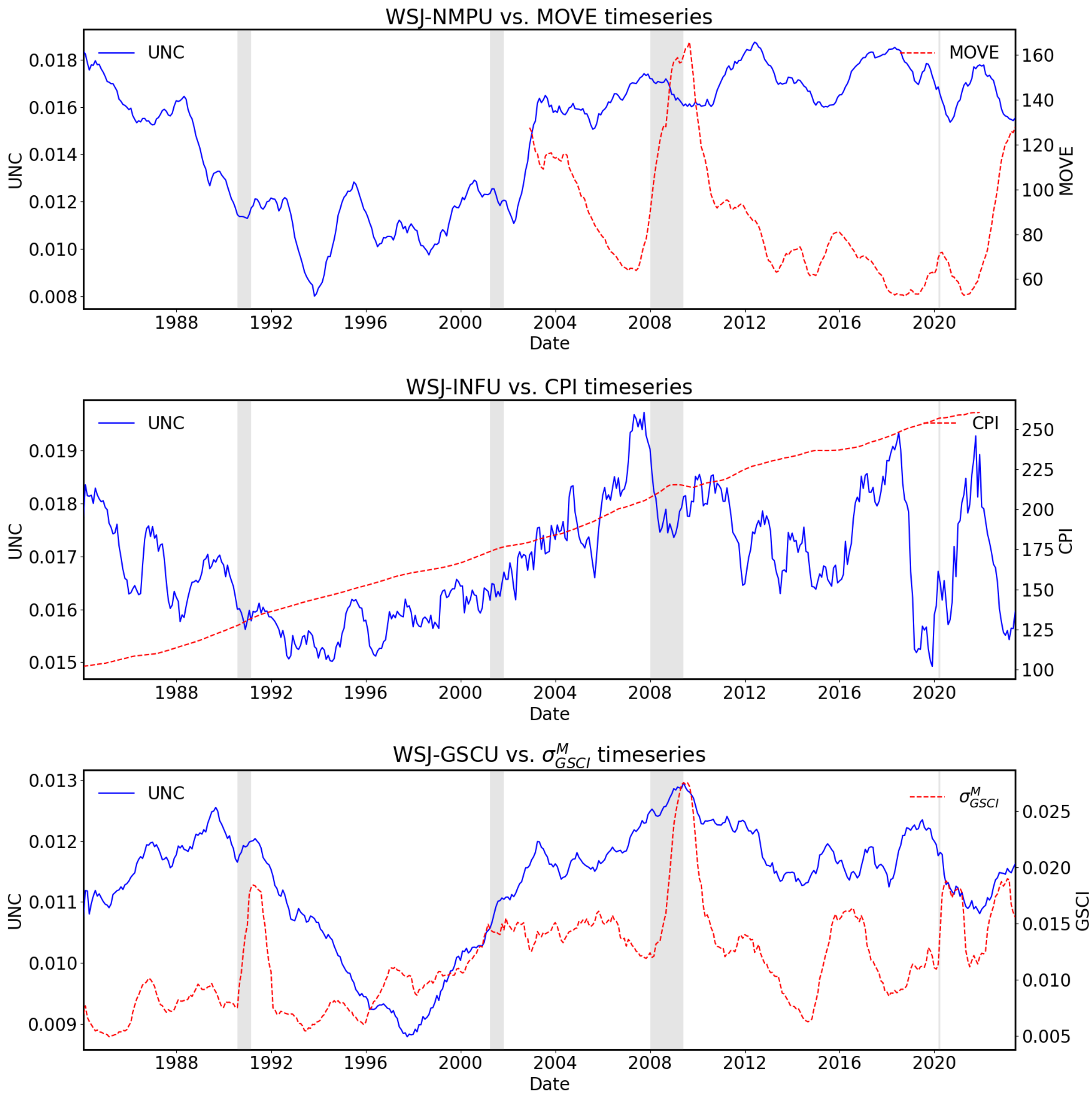
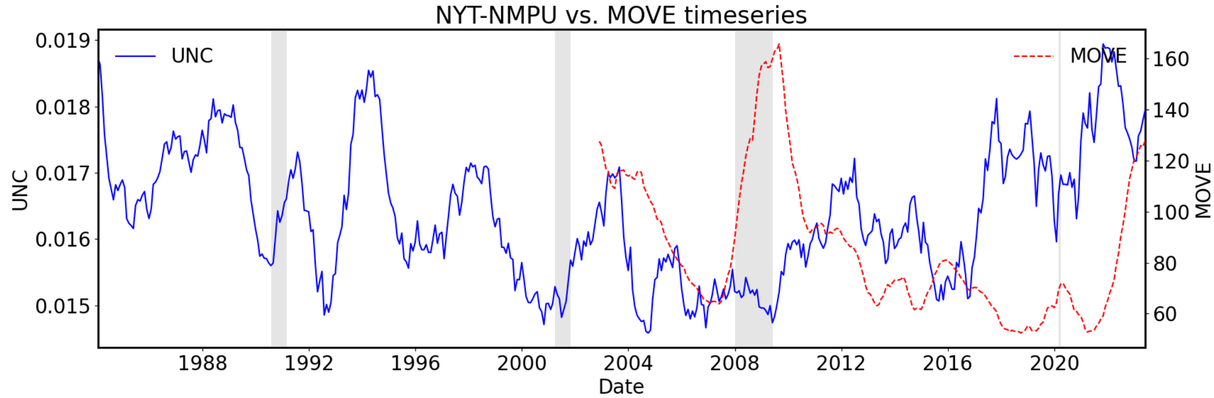
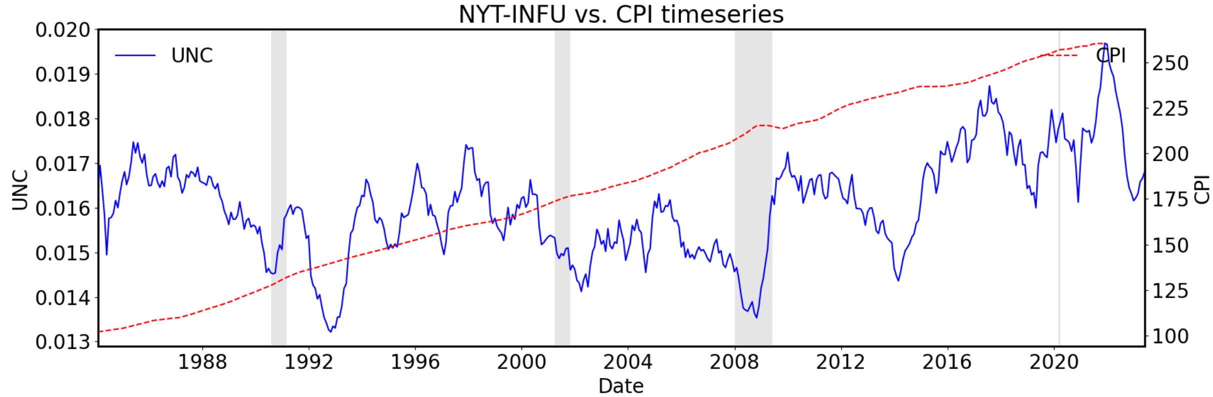


Figure 2: NYT Time-series

This figure shows the monthly NYT-NMPU, NYT-INFU, NYT-GSCU from Jan 1984 together with the corresponding macroeconomic variables. The first row shows NYT-NMPU with MOVE index. The second row shows NYT-INFU with CPI. The third row shows NYT-GSCU with σ^M_{GSCI}. All the variables are smoothed by a 12-month moving average. The gray vertical bars are NBER recessions.





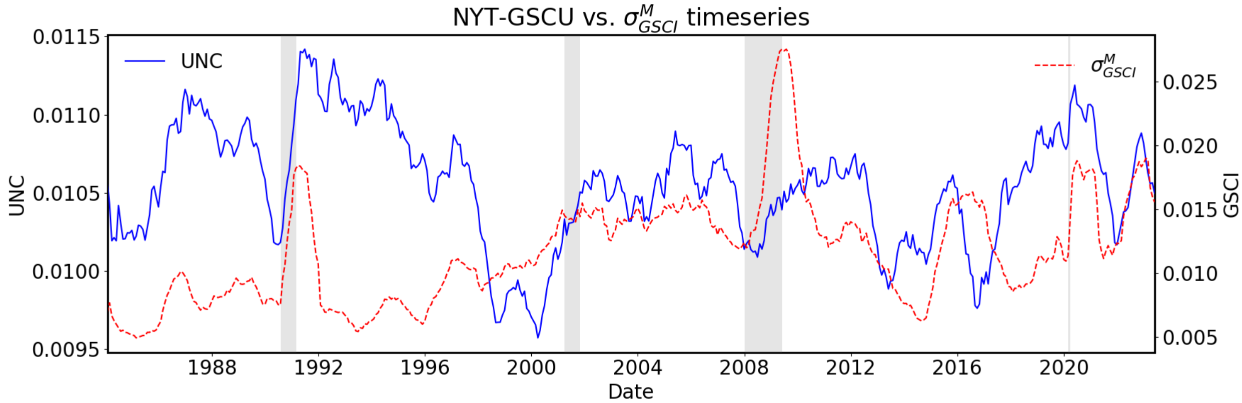
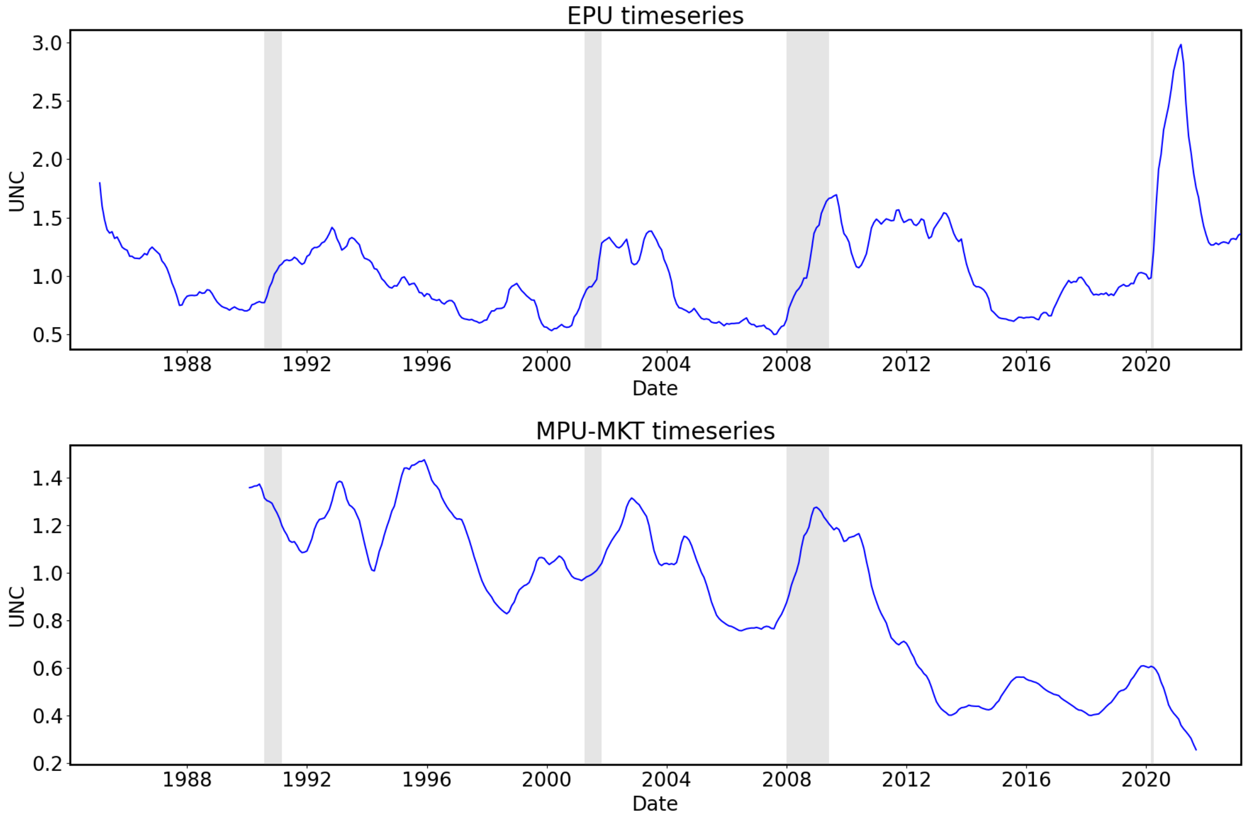


Figure 3: Existing Uncertainty Measures

This figure shows the monthly EPU and MPU-MKT from Jan 1984. Both the scores are smoothed by a 12-month moving average. The gray vertical bars are NBER recessions.



Appendix

In this appendix, we report the estimates for the regression of the MOVE index, CPI inflation, and commodity volatility on the contemporaneous measure of the text-based uncertainties. This contrasts with the main results in the paper which focus on predicting future outcome.

Tables A1 and A2 report the results on the MOVE index. We show that WSJ-NMPU and MPU are positively correlated with the changes in the MOVE index.

Tables A3 and A4 report the results on the CPI inflation. We find that the long-term changes in the NYT-INFU are negatively related with the absolute changes in CPI inflation.

Tables A5 and A6 report the results on the GSCI volatility. We find that EPU and WSJ-GSCU are strongly positively correlated with the commodity volatility.

Overall, these results confirm that the text-based uncertainty measures capture an important variation in the uncertainty facing market participants.

Table A1: Contemporaneous Regression for MOVE Index

This table reports the results of the following regression:

$$\text{MOVE}_t - \text{MOVE}_{t-1} = \alpha + \beta_1(\text{UNC}_t - \text{UNC}_{t-1}) + \epsilon_t,$$

where UNC_t can be WSJ-NMPU, NYT-NMPU, EPU, and MPU-MKT. The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

	WSJ-NMPU (1)	NYT-NMPU (2)	EPU (3)	MPU-MKT (4)
ΔUNC	1409.5288**	127.5389	0.5900	51.3027***
	(571.79)	(349.79)	(5.58)	(17.94)
Intercept	0.0087	-0.0189	-0.0160	-0.1937
	(0.78)	(0.77)	(0.76)	(0.78)
R^2	0.021	-0.003	-0.004	0.064
N	243	243	243	214

Table A2: Contemporaneous Regression for MOVE Index: Combining All Variables

This table reports the results of the following regression:

$$\text{MOVE}_t - \text{MOVE}_{t-1} = \alpha + \beta_1(\text{UNC}_t - \text{UNC}_{t-1}) + \epsilon_t,$$

where UNC_t can be WSJ-NMPU, NYT-NMPU, EPU, MPU-MKT and weighted combinations of the variables. The number of observations are set to be the minimum data range available (MPU-MKT). The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

	Combined 1 (1)	Combined 2 (2)	Combined 3 (3)
ΔMPU-MKT	49.5146***	49.2829***	49.2792***
	(17.63)	(17.61)	(17.51)
ΔWSJ-NMPU	1104.8777**	1081.4274**	1074.9952**
	(542.34)	(531.80)	(541.18)
ΔNYT-NMPU		86.8129	79.2073
		(318.33)	(340.52)
ΔEPU			1.4495
			(5.47)
Intercept	-0.1950	-0.1963	-0.2014
	(0.79)	(0.79)	(0.77)
R^2	0.075	0.071	0.067
N	214	214	214

Table A3: Backward Regression for Absolute Changes in CPI

This table reports the results of the following regression:

$$|\log \mathrm{CPI}_t - \log \mathrm{CPI}_{t-12}| = \alpha + \beta_1 \mathrm{UNC}_t^{M-Q} + \beta_2 \mathrm{UNC}_t^{Q-Y} + \beta_3 \mathrm{UNC}_t^{Y-4Y} + \epsilon_t,$$

where UNC can be WSJ-INFU, NYT-INFU, EPU, and MPU-MKT. The definition of UNC_t^{M-Q}, UNC_t^{Q-Y} and UNC_t^{Y-4Y} can be found in Sec. 2.3. The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

	WSJ-INFU (1)	NYT-INFU (2)	EPU (3)	MPU-MKT (4)
UNC^{M-Q}	-0.0680	0.0107	0.0022	0.0053
	(0.17)	(0.19)	(0.00)	(0.01)
UNC^{Q-Y}	-0.8674	-0.3859	-0.0014	0.0014
	(0.64)	(0.69)	(0.00)	(0.01)
UNC^{Y-4Y}	0.4184	-4.7856***	-0.0056	-0.0012
	(1.38)	(1.38)	(0.00)	(0.01)
Intercept	0.0263***	0.0264***	0.0259***	0.0243***
	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.002	0.088	0.018	-0.007
N	444	444	432	369

Table A4: Backward Regression for Absolute Changes in CPI: Combining All Variables

This table reports the results of the following regression:

$$|\log \text{CPI}_t - \log \text{CPI}_{t-12}| = \alpha + \gamma_1 \text{UNC}_t^{Y-4Y} + \epsilon_t,$$

where UNC can be WSJ-INFU, NYT-INFU, EPU, MPU-MKT and weighted combinations of the variables. The number of observations are set to be the minimum data range available (MPU-MKT). The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

	Combined 1	Combined 2	Combined 3
	(1)	(2)	(3)
NYT-INFU^{Y-4Y}	-4.3011***	-4.4938***	-4.5007***
	(1.38)	(1.38)	(1.38)
WSJ-INFU^{Y-4Y}	1.9030	1.5856	1.5892
	(1.26)	(1.29)	(1.30)
EPU^{Y-4Y}		-0.0038	-0.0040
		(0.00)	(0.00)
MPU^{Y-4Y}			0.0012
			(0.01)
Intercept	0.0245***	0.0246***	0.0247***
	(0.00)	(0.00)	(0.00)
R^2	0.093	0.100	0.098
N	372	372	372

Table A5: Regression for Commodity Volatility

This table reports the results of the following regression:

$$y_t = \alpha + \beta_1 \text{UNC}_t + \epsilon_t,$$

where y_t is the monthly, quarterly, yearly standard deviation of log returns of GSCI indices ending in month t, *i.e.,* $\sigma^M_{t,\text{GSCI}}$, $\sigma^Q_{t,\text{GSCI}}$, $\sigma^Y_{t,\text{GSCI}}$. UNC can be WSJ-NMPU, WSJ-GSCU, NYT-NMPU, NYT-GSCU, EPU, and MPU-MKT. The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

Panel A. $\sigma^M_{t,\text{GSCI}}$

	WSJ-NMPU (1)	WSJ-GSCU (2)	NYT-NMPU (3)	NYT-GSCU (4)	EPU (5)	MPU-MKT (6)
UNC	0.2548**	1.5668***	-0.1856	0.7412	0.0041***	0.0002
	(0.11)	(0.36)	(0.15)	(0.60)	(0.00)	(0.00)
Intercept	0.0080***	-0.0060	0.0148***	0.0040	0.0077***	0.0122***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
R^2	0.016	0.088	0.003	0.005	0.129	-0.003
N	470	470	470	470	458	369

Panel B. $\sigma^Q_{t,\text{GSCI}}$

	WSJ-NMPU (1)	WSJ-GSCU (2)	NYT-NMPU (3)	NYT-GSCU (4)	EPU (5)	MPU-MKT (6)
UNC	0.2702**	1.3666***	-0.1101	0.6633	0.0040***	-0.0004
	(0.11)	(0.35)	(0.17)	(0.59)	(0.00)	(0.00)
Intercept	0.0080***	-0.0034	0.0139***	0.0051	0.0080***	0.0130***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
R^2	0.022	0.080	-0.000	0.005	0.153	-0.002
N	470	470	470	470	458	369

Panel C. $\sigma^Y_{t,\text{GSCI}}$

	WSJ-NMPU (1)	WSJ-GSCU (2)	NYT-NMPU (3)	NYT-GSCU (4)	EPU (5)	MPU-MKT (6)
UNC	0.3557***	1.2010***	-0.1435	0.1203	0.0030***	-0.0006
	(0.11)	(0.29)	(0.12)	(0.43)	(0.00)	(0.00)
Intercept	0.0070***	-0.0013	0.0146***	0.0110**	0.0094***	0.0134***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.055	0.084	0.003	-0.002	0.113	-0.001
N	470	470	470	470	458	369

Table A6: Regression for Commodity Volatility: Combining All Variables

This table reports the results of the following regression:

$$y_t = \alpha + \beta_1 \text{UNC}_t + \epsilon_t,$$

where y_t is the monthly, quarterly, yearly standard deviation of log returns of GSCI indices ending in month t, *i.e.*, $\sigma^M_{t,\text{GSCI}}$, $\sigma^Q_{t,\text{GSCI}}$, $\sigma^Y_{t,\text{GSCI}}$. UNC can be WSJ-NMPU, WSJ-GSCU, NYT-NMPU, NYT-GSCU, EPU, and MPU-MKT and weighted combinations of the variables. The number of observations are set to be the minimum data range available (MPU-MKT). The standard errors are calculated using Newy-west (HAC) with max lags set to $N^{\frac{1}{4}}$. *, **, *** denote the statistical significance at the 10%, 5%, 1% levels, respectively.

Panel A. $\sigma^M_{t,\text{GSCI}}$

	Combined 1 (1)	Combined 2 (2)	Combined 3 (3)	Combined 4 (4)	Combined 5 (5)
EPU	0.0035***	0.0035***	0.0036***	0.0038***	0.0039***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
WSJ-GSCU	1.5244***	1.6938***	1.7377***	1.6915***	1.6028***
	(0.33)	(0.50)	(0.53)	(0.50)	(0.47)
WSJ-NMPU		-0.0969	-0.1087	-0.0851	0.0539
		(0.19)	(0.19)	(0.18)	(0.17)
NYT-GSCIU			-0.2487	-0.1011	-0.2266
			(0.52)	(0.51)	(0.53)
NYT-NMPU				-0.3667**	-0.3488**
				(0.16)	(0.15)
MPU-MKT					0.0022
					(0.00)
Intercept	-0.0084**	-0.0089**	-0.0066	-0.0024	-0.0044
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
R^2	0.226	0.226	0.224	0.240	0.248
N	369	369	369	369	369

Panel B. $\sigma^{Q}_{t,\mathrm{GSCI}}$

	Combined 1 (1)	Combined 2 (2)	Combined 3 (3)	Combined 4 (4)	Combined 5 (5)
EPU	0.0036***	0.0036***	0.0037***	0.0038***	0.0039***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
WSJ-GSCU	1.3417***	1.3184***	1.3566***	1.3236***	1.2507***
	(0.33)	(0.49)	(0.51)	(0.49)	(0.47)
WSJ-NMPU		0.0133	0.0031	0.0199	0.1342
		(0.19)	(0.18)	(0.18)	(0.17)
NYT-GSCIU			-0.2165	-0.1109	-0.2140
			(0.56)	(0.54)	(0.56)
NYT-NMPU				-0.2623*	-0.2475*
				(0.15)	(0.14)
MPU-MKT					0.0018
					(0.00)
Intercept	-0.0062	-0.0061	-0.0041	-0.0011	-0.0028
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
R^2	0.253	0.251	0.250	0.259	0.265
N	369	369	369	369	369

Panel C. $\sigma^{Y}_{t,\mathrm{GSCI}}$

	Combined 1 (1)	Combined 2 (2)	Combined 3 (3)	Combined 4 (4)	Combined 5 (5)
EPU	0.0020***	0.0020***	0.0021***	0.0023***	0.0023***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
WSJ-GSCU	1.3960***	1.1387***	1.2111***	1.1770***	1.0922***
	(0.29)	(0.40)	(0.40)	(0.39)	(0.39)
WSJ-NMPU		0.1473	0.1279	0.1453	0.2782*
		(0.15)	(0.15)	(0.15)	(0.16)
NYT-GSCIU			-0.4098	-0.3008	-0.4207
			(0.42)	(0.40)	(0.41)
NYT-NMPU				-0.2707**	-0.2536**
				(0.11)	(0.10)
MPU-MKT					0.0021
					(0.00)
Intercept	-0.0049	-0.0042	-0.0005	0.0027	0.0007
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
R^2	0.227	0.233	0.234	0.250	0.265
N	369	369	369	369	369

References

Kenneth N Kuttner. Monetary policy surprises and interest rates: Evidence from the fed funds futures market. *Journal of Monetary Economics*, 47(3):523–544, 2001.

Ben S Bernanke and Kenneth N Kuttner. What explains the stock market's reaction to federal reserve policy? *Journal of Finance*, 60(3):1221–1257, 2005.

Eric T Swanson. Measuring the effects of federal reserve forward guidance and asset purchases on financial markets. *Journal of Monetary Economics*, 118:32–53, 2021.

Michael D Bauer, Aeimit Lakdawala, and Philippe Mueller. Market-Based Monetary Policy Uncertainty. *Economic Journal*, 132(644):1290–1308, 11 2021. ISSN 0013-0133. doi: 10.1093/ej/ueab086. URL `https://doi.org/10.1093/ej/ueab086`.

Charles Martineau, Zissis Poulos, Yuntao Wu, Cameron Thompson, Maryam Haghighi, Jun Yuan, and John Hull. Narrative monetary policy uncertainty. Sept 2023. Available at SSRN: `https://ssrn.com/abstract=4573829` or `http://dx.doi.org/10.2139/ssrn.4573829`.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv:1810.04805 [cs.CL]*, 2019. doi: 10.48550. URL `https://doi.org/10.48550/arXiv.1810.04805`.

Jules H. van Binsbergen, Svetlana Bryzgalova, Mayukh Mukhopadhyay, and Varun Sharma. (almost) 200 years of news-based economic sentiment. October 2022. Available at SSRN: `https://ssrn.com/abstract=4261249` or `http://dx.doi.org/10.2139/ssrn.4261249`.

Scott R. Baker, Nicholas Bloom, and Steven J. Davis. Measuring Economic Policy Uncertainty. *Quarterly Journal of Economics*, 131(4):1593–1636, 07 2016. ISSN 0033-5533. doi: 10.1093/qje/qjw024. URL `https://doi.org/10.1093/qje/qjw024`.

Lucas Husted, John Rogers, and Bo Sun. Monetary policy uncertainty. *Journal of Monetary Economics*, 115:20–36, 2020. ISSN 0304-3932. doi: https://doi.org/10.1016/j.jmoneco.2019.07.009. URL `https://www.sciencedirect.com/science/article/pii/S0304393218301661`.

Adlai Fisher, Charles Martineau, and Jinfei Sheng. Macroeconomic Attention and Announcement Risk Premia. *Review of Financial Studies*, 35(11):5057–5093, 02 2022. ISSN 0893-9454. doi: 10.1093/rfs/hhac011. URL `https://doi.org/10.1093/rfs/hhac011`.

Fulvio Ortu, Andrea Tamoni, and Claudio Tebaldi. Long-Run Risk and the Persistence of Consumption Shocks. *Review of Financial Studies*, 26(11):2876–2915, 11 2013. ISSN 0893-9454. doi: 10.1093/rfs/hht038. URL `https://doi.org/10.1093/rfs/hht038`.

Laurent E. Calvet and Adlai J. Fisher. Multifrequency news and stock returns. *Journal of Financial Economics*, 86(1):178–212, 2007. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2006.09.001. URL `https://www.sciencedirect.com/science/article/pii/S0304405X07000906`.